Journal of
Risk and Financial Management

Frontiers of Asset Pricing

Edited by
James W. Kolari and Seppo Pynnonen

Printed Edition of the Special Issue Published in *JRFM*

www.mdpi.com/journal/jrfm

MDPI

Frontiers of Asset Pricing

Frontiers of Asset Pricing

Editors

James W. Kolari
Seppo Pynnonen

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editors
James W. Kolari
Texas A&M University
USA

Seppo Pynnonen
University of Vaasa
Finland

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Journal of Risk and Financial Management* (ISSN 1911-8074) (available at: https://www.mdpi.com/journal/jrfm/special_issues/FAP).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-5845-5 (Hbk)
ISBN 978-3-0365-5846-2 (PDF)

Cover image courtesy of James W. Kolari

Contents

About the Editors . . . . **vii**

Preface to "Frontiers of Asset Pricing" . . . . **ix**

Seppo Pynnonen
Non-Parametric Statistic for Testing Cumulative Abnormal Stock Returns
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 149, doi:10.3390/jrfm15040149 . . . . **1**

James W. Kolari, Jianhua Z. Huang, Wei Liu and Huiling Liao
Further Tests of the ZCAPM Asset Pricing Model
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 137, doi:10.3390/jrfm15030137 . . . . **15**

James Nguyen, Wei-Xuan Li and Clara Chia-Sheng Chen
Mean Reversions in Major Developed Stock Markets: Recent Evidence from Unit Root, Spectral and Abnormal Return Studies
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 162, doi:10.3390/jrfm15040162 . . . . **39**

Klaus Grobys
On Survivor Stocks in the S&P 500 Stock Index
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 95, doi:10.3390/jrfm15020095 . . . . **59**

Ihsan Badshah and Hardjo Koerniadi
Net Buying Pressure and Informed Trading in the Options Market: Evidence from Earnings Announcements
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 53, doi:10.3390/jrfm15020053 . . . . **83**

Anupam Dutta and Elie Bouri
Outliers and Time-Varying Jumps in the Cryptocurrency Markets
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 128, doi:10.3390/jrfm15030128 . . . . **93**

Wei Liu and James W. Kolari
Multifactor Market Indexes
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 155, doi:10.3390/jrfm15040155 . . . . **101**

Haipeng Xing
A Singular Stochastic Control Approach for OptimalPairs Trading with Proportional Transaction Costs
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 147, doi:10.3390/jrfm15040147 . . . . **127**

Xinbo Hu and Paul J. Zak
A Procedure to Set Prices and Select Inventory in Thinly Traded Markets Using Data from eBay
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 297, doi:10.3390/jrfm15070297 . . . . **151**

Dean Leistikow, Yi Tang and Wei Zhang
Dynamic Conditional Bias-Adjusted Carry Cost Rate Futures Hedge Ratios
Reprinted from: *J. Risk Financial Manag.* **2022**, *15*, 12, doi:10.3390/jrfm15010012 . . . . **159**

Yasmeen Idilbi-Bayaa and Mahmoud Qadan
Forecasting Commodity Prices Using the Term Structure
Reprinted from: *J. Risk Financial Manag.* **2021**, *14*, 585, doi:10.3390/jrfm14120585 . . . . **177**

About the Editors

James W. Kolari

James W. Kolari is the JP Morgan Chase Professor of Finance and Academic Director of the Global Corporate Banking Program in the Department of Finance at Texas A&M University, College Station, Texas USA. He has taught money and capital markets as well as banking classes there since earning his PhD in Finance 1980. Over the years, various appointments have been held as a Visiting Scholar at the Federal Reserve Bank of Chicago, Fulbright Scholar to the Bank of Finland, and Senior Research Fellow at the Swedish School of Business and Economics (Hanken), Finland in addition to being a consultant to the U.S. Small Business Administration, U.S. Information Agency, and numerous banks and other organizations. With over 100 articles published in refereed journals, numerous other papers and monographs, 25 co-authored books, and over 200 competitive papers presented at academic conferences, he ranks in the top 1–2 percent of finance scholars in the United States. His papers have appeared in such domestic and international journals as the Journal of Finance, Journal of Business, Review of Financial Studies, Review of Economics and Statistics, Journal of Money, Credit and Banking, Journal of Banking and Finance, Critical Finance Review, Journal of Empirical Finance, Real Estate Economics, Journal of International Money and Finance, and the Scandinavian Journal of Economics. Papers in Chinese, Dutch, Finnish, Italian, Russian, and Spanish have appeared outside of the United States. He is a co-author of leading college textbooks in introductory business and commercial banking courses.

Seppo Pynnonen

Seppo Pynnonen is Professor of Statistics at the University of Vaasa and previously the Chairperson of the Department of Mathematics and Statistics. He has studied financial markets and taught various courses on statistical methodology, empirical finance, and mathematical finance covering undergraduate, graduate, and PhD levels since earning his PhD in mathematical statistics in 1988. He has published several papers in international finance and statistics journals including Review of Finance, Critical Finance Review, Journal of Empirical Finance, Journal of International Money and Finance, European Journal of Operational Research, Journal of Multivariate Analysis, and Communication in Statistics – Simulation and Computation.

Preface to "Frontiers of Asset Pricing"

The famed Capital Asset Pricing Model (CAPM) of Sharpe, Lintner, Mossin, and Black in the 1960s proposed an equilibrium theory wherein the expected return of an asset is a function of beta risk associated with the expected return of the market portfolio. The CAPM was a pathbreaking model derived from Markowitz portfolio theory and Tobin equilibrium pricing advances.

Unfortunately, in the 1990s, Fama and French published a series of widely-cited papers that documented little or no relation between beta risk and average U.S. stock returns. Concluding that the CAPM was dead, they proposed a number of empirically-based models incorporating long/short portfolio returns as multifactors. Their multifactor models supplanted the CAPM. Subsequently, researchers proposed similar models with different multifactors. However, as Cochrane has opined, a factor zoo has developed. Nowadays intense competition exists in terms of alternative multifactors and models.

James W. Kolari and Seppo Pynnonen

Editors

Journal of
Risk and Financial Management

MDPI

Article

Non-Parametric Statistic for Testing Cumulative Abnormal Stock Returns

Seppo Pynnonen

Department of Mathematics and Statistics, University of Vaasa, P.O. Box 700, FI-65101 Vaasa, Finland; sjp@uwasa.fi; Tel.:+358-21-449-8311

Abstract: Due to the non-normality of stock returns, nonparametric rank tests are gaining accceptance relative to parametric tests in financial economics event studies. In rank tests, financial assets' multiple day cumulative abnormal returns (CARs) are replaced by cumulated ranks. This paper proposes modifications to the existing approaches to improve robustness to cross-sectional correlation of returns arising from calendar time overlapping event windows. Simulations show that the proposed rank test is well specified in testing CARs and is robust towards both complete and partial overlapping event windows.

Keywords: finance; economics; event study; clustered event days; cross-sectional correlation; cumulated ranks; rank test; standardized abnormal returns

JEL Classification: G14; C10; C15

Citation: Pynnonen, Seppo. 2022. Non-Parametric Statistic for Testing Cumulative Abnormal Stock Returns. *Journal of Risk and Financial Management* 15: 149. https://doi.org/10.3390/jrfm15040149

Academic Editor: Ştefan Cristian Gherghina

Received: 8 December 2021
Accepted: 15 March 2022
Published: 23 March 2022

1. Introduction

Efficient markets has been and still is a cornerstone of asset pricing theory. Empirical work in this regard is largely concerned with the adjustment of security prices to relevant information. Fama (1970, 1991) refine relevant information into three hierarchical subsets of weak form, semi-strong form, and strong form Fama (1970), or equivalently, return predictability, event studies, and private information Fama (1991). Event studies investigate the effect of unexpected economic events on asset prices. Therefore, event studies can give the most direct evidence on market efficiency (c.f. Fama 1991, p. 1577). For this purpose, asset price data available from financial markets can be used with appropriate statistical testing methodology, reliability of which is central in inferences. In order to foster this, the current paper aims to fill the gap in existing (non-parametric) statistical testing by proposing non-parametric rank tests that are robust to cross-sectional dependency of asset returns in more general circumstances than the existing ones. Otherwise, refer to (Campbell et al. 1997, chp. 4) as an excellent introduction to event studies and related statistical methods.

Regarding methodology, standardizing returns by their respective standard deviations homogenizes data and has proven to improve testing performance. Because of this improvement, standardized return based tests by Patell (1976) and Boehmer et al. (1991) (BMP) have gained popularity over conventional non-standardized tests in testing event effects on mean security price performance. Harrington and Shrider (2007) found that in a short-horizon testing of abnormal returns (i.e., systematic deviation from expected behavior), one should always use methods that are robust to cross-sectional variation in the true abnormal returns.[1] They found that BMP is a good candidate for robust, parametric tests in conventional event studies.[2]

A major problem in statistical tests of returns is that the returns are not normally distributed (Fama 1976). Not surprisingly, non-parametric rank tests introduced by Corrado (1989, 2011); Corrado and Zivney (1992); Campbell and Wasley (1993) and Kolari and

Pynnonen (2011), among others, dominate parametric tests both in terms of better size and power (e.g., see Campbell and Wasley 1993; Corrado 1989; Corrado and Zivney 1992; Kolari and Pynnonen 2011; Kolari and Pynnönen 2010; Luoma 2011). Furthermore, rank tests by Corrado and Zivney (1992) and Kolari and Pynnonen (2011) that utilize event period re-standardized returns have proven to be robust to event-induced volatility (Kolari and Pynnonen (2011); Kolari and Pynnönen (2010)), cross-correlation due to event day clusterings (Kolari and Pynnönen 2010), and autocorrelation (Kolari and Pynnonen 2011). These studies are consistent with the view stated in the epilogue of Lehmann (2006): "Rank tests apply often to relatively simple solutions, such as one-, two-, and *s*-sample problems, and testing for independence and randomness, but for these situations they are often the method of choice". (Lehmann 2006, p. v). In addition, the results of rank tests are invariant to monotone transformations of the underlying returns; that is, whether the returns are defined as simple, continuously compounded log returns.

Existing rank based tests, however, are not robust to cross-sectional correlation if the event days are partially overlapping in calendar time. This *partial clustering* occurs when events are in calendar time scattered within an event window more or less randomly rather than clustered on the same calendar day (i.e., *complete clustering* as in Kolari and Pynnönen 2010). Figure 1 illustrates the various degrees of clustering in terms of three stocks. Panel A depict the non-clustered case, Panel B the partial clustering, and Panel C the complete clustering. In the complete clustering the event days are the same in calendar time, while in the partial clustering the event days may or may not be the same in calendar time but the event windows are more or less overlapping. In the non-clustered case the event windows are completely separate in calendar time. In this case all event effects can be investigated utilizing cross-sectional independence assumption of returns. In complete clustering cross-sectional correlation of returns must be fully accounted for. In the partial case the correlation can bias the results depending on the degree of overlapping. For example in the case of Panel C if the interest is only on the event day effect, as all the event days are different, there is no biasing effect by the correlation. On the other hand, if cumulative return effect over the whole event window is of interest, correlation of returns on the overlapping affects the joint behavior of the cumulative returns.

Jaffe (1974) is probably the first paper in event study testing to address the potential biasing effect of cross-sectional correlation due to clustered events. Table 2 of Kolari and Pynnönen (2010) explicitly addresses the issue by showing that already a virtually trivial cross-sectional correlation, such as 0.05, can severely bias testing for event effects towards material over-rejection. The present paper seeks to fill this gap of accounting for cross-sectional correlation in non-parametric even study testing also with partially clustered event days.

The paper is organized as follows. Section 2 reviews some related key literature. Section 3 defines the main concepts and derives some distributional properties of rank statistics. Section 4 introduces the new transformed rank test. Section 5 reports simulation results, and Section 6 concludes.

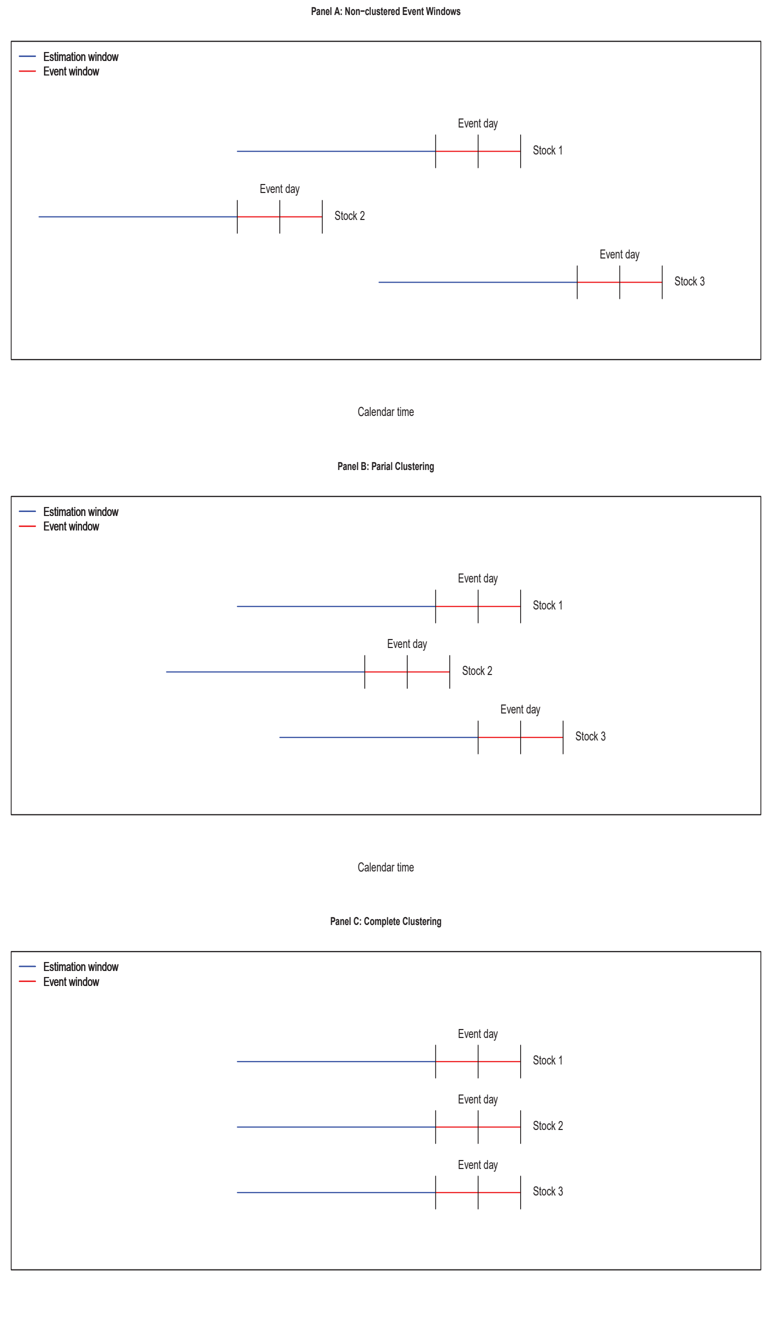

Figure 1. Event Windows Clustering.

2. Review of Related Literature

Patell and BMP parametric tests are straightforward tests of cumulative abnormal returns (CARs) over multiple day windows. With the correction suggested by Kolari and

Pynnönen (2010), these tests are useful in the case of completely clustered event days, and with the correction suggested by Kolari et al. (2018) when event days are either completely or partially clustered. By construction, the Corrado (1989) non-parametric rank test applies for testing single day event returns. Testing for CARs with the same logic implies the need of defining multiple-day returns that match the number of days in the CARs, (see (Corrado 1989, p. 395); (Campbell and Wasley 1993, footnote 4)). In practice this approach is carried out by dividing the estimation period and event period into intervals matching the number of days in the CAR. Unfortunately, this procedure is not useful for a number of reasons. Foremost among these is that it does not necessarily lead to a unique testing procedure. In addition, the abnormal return model should be re-estimated for each multiple-day CAR definition. Furthermore, for a fixed estimation period, as the number of days accumulated in a CAR increases, the number of multiple-day estimation period observations reduces quickly impractically low and thus would weaken the abnormal return model estimation (c.f., Kolari and Pynnönen 2010). Kolari and Pynnonen (2011) solve these issues in their generalized rank test approach.

On the other hand, Campbell and Wasley (1993) recommend using the Corrado (1989) rank test to test cumulative abnormal returns by simply accumulating the respective ranks to constitute cumulative ranks (see also Hagnäs and Pynnonen 2014). This practice is adopted in the Eventus® software Cowan (2007) and is probably, for the time being, the most popular procedure for multiple day applications of rank tests. An advantage is that this proceure implicitly accounts the cross-sectional correlation in the case of the complete clustering.

In spite of these attractive properties, the cumulative ranks test does not account for cross-sectional correlation due to calendar time partially overlapping event windows, i.e., the case of partial clustering. As referred above, even a small (positive) correlation biases the standard errors downwards leading to over-rejection of the null hypothesis of no event effect. Contributing to the event study literature, this paper proposes an adjustment for the standard errors that corrects the bias in non-parametric testing.

3. Distributional Properties of Ranks

We begin by fixing some notations and an underlying assumption to facilitate our theoretical discussion.

Assumption 1. *Stock returns r_{it} for firm i are weak white noise continuous random variables and are cross-sectionally independent over non-overlapping calendar days, or,*

$$\begin{array}{rcl} \mathbb{E}[r_{it}] & = & \mu_i \text{ for all } t \\ \text{var}[r_{it}] & = & \sigma_i^2 \text{ for all } t \\ \text{cov}[r_{it}, r_{iu}] & = & 0 \text{ for all } t \neq u \\ r_{it} \text{ and } & r_{ju} & \text{are independent whenever } i \neq j \text{ and } t \neq u. \end{array} \tag{1}$$

It is a stylized fact that the variances of the returns are time varying and that there is mild autocorrelation. The time varying volatility problem can be partially captured in terms of GARCH-modeling. However, typical GARCH-processes satisfy Assumption 1.

Let $\text{AR}_{it} = r_{it} - \mathbb{E}[r_{it}]$ denote the abnormal return of security i on day t, and following commonly used notations (e.g., Brown and Warner 1985, p. 6), let day $t = 0$ indicate the event day. Days from $t = T_0 + 1$ to $t = T_1$ represent the estimation period relative to the event day, and days from $t = T_1 + 1$ to $t = T_2$ represent the event window. The cumulative abnormal return (CAR) from τ_1 to τ_2 with $T_1 < \tau_1 \leq \tau_2 \leq T_2$, is defined as

$$\text{CAR}_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \text{AR}_{it}. \tag{2}$$

The time period from τ_1 to τ_2 is called in the following as a CAR window or CAR period.

Standardized abnormal returns are defined as

$$\mathrm{SAR}_{it} = \frac{\mathrm{AR}_{it}}{S_{\mathrm{AR}_i}}, \tag{3}$$

where

$$S_{\mathrm{AR}_i} = \sqrt{\frac{1}{T_1 - T_0 - 1}\sum_{t=T_0+1}^{T_1} \mathrm{AR}_{it}^2}. \tag{4}$$

Furthermore, for the purpose of accounting the possible event induced volatility, the re-standardized abnormal returns are defined as in Boehmer et al. (1991) (see also, Corrado and Zivney 1992), or

$$\mathrm{SAR}'_{it} = \begin{cases} \mathrm{SAR}_{it}/S_{\mathrm{SAR}_t}, & T_1 < t \le T_2 \\ \mathrm{SAR}_{it}, & \text{otherwise}, \end{cases} \tag{5}$$

where

$$S_{\mathrm{SAR}_t} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\mathrm{SAR}_{it} - \overline{\mathrm{SAR}}_t)^2} \tag{6}$$

is the time t cross-sectional standard deviation of SAR_{it}, $\overline{\mathrm{SAR}}_t = \frac{1}{n}\sum_{i=1}^{n}\mathrm{SAR}_{it}$, and n is the number of stocks in the portfolio. In addition, let K_{it} denote the rank numbers of abnormal returns, where $K_{it} \in \{1, \ldots, T\}$, $t = T_0 + 1, \ldots, T_2$, $T = T_2 - T_0$, and $i = 1, \ldots, n$.

If the available observations in the estimation period vary from one series to another, it is convenient to use *standardized ranks* with zero mean and unit variance. To do this, we compile the known results of rank statistics (e.g., Lehmann 2006, Appendix, Section 1) as described below.

Result 1. *Let K_{it} denote the rank numbers as defined above, then*

$$\begin{aligned} \mathbb{E}[K_{it}] &= (T+1)/2 && (7)\\ \mathrm{var}[K_{it}] &= (T^2-1)/12 && (8)\\ \mathrm{cov}[K_{is}, K_{it}] &= -(T+1)/12, (s \ne t). && (9) \end{aligned}$$

Definition 1. *Standardized ranks are defined as*

$$U_{it} = \frac{K_{it} - \frac{1}{2}(T+1)}{\sqrt{(T^2-1)/12}}. \tag{10}$$

(c.f., Hagnäs and Pynnonen 2014).

By Result 1, we obtain:

Result 2.

$$\begin{aligned} \mathbb{E}[U_{it}] &= 0 && (11)\\ \mathrm{var}[U_{it}] &= 1 && (12)\\ \mathrm{cov}[U_{is}, U_{it}] &= -1/(T-1). && (13) \end{aligned}$$

Next, we define *cumulative standardized ranks* for individual stocks.

Definition 2. *The cumulative standardized ranks of a stock i over the event days window form τ_1 to τ_2 are defined as*

$$U_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} U_{it}, \tag{14}$$

where $T_1 < \tau_1 \leq \tau_2 \leq T_2$.

From Result 2 and utilizing the variance-of-the-sum formula, $\mathrm{var}[U_i(\tau_1, \tau_2)] = \sum_{t=\tau_1}^{\tau_2} \mathrm{var}[U_{it}] + \sum_{s \neq t} \mathrm{cov}[U_{is}, U_{it}]$, we obtain:

Result 3.

$$\mathbb{E}[U_i(\tau_1, \tau_2)] = 0 \tag{15}$$

$$\mathrm{var}[U_i(\tau_1, \tau_2)] = \frac{\tau(T-\tau)}{T-1}, \tag{16}$$

where $i = 1, \ldots, n$*,* $T_1 < \tau_1 \leq \tau_2 \leq T_2$*, and* $\tau = \tau_2 - \tau_1 + 1$*.*

Rather than investigating individual (cumulative) returns, the practice in event studies is to aggregate individual returns into equally-weighted portfolios such that:

Definition 3. *The average cumulative standardized ranks are defined as the equally weighted portfolio of individual cumulative standardized ranks defined in* (14)*, i.e.,*

$$\bar{U}(\tau_1, \tau_2) = \frac{1}{n} \sum_{i=1}^{n} U_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \bar{U}_t, \tag{17}$$

where $T_1 < \tau_1 \leq \tau_2 \leq T_2$ *and*

$$\bar{U}_t = \frac{1}{n} \sum_{i=1}^{n} U_{it} \tag{18}$$

is the time t average of standardized ranks.

The expected value of $\bar{U}(\tau_1, \tau_2)$ is the same as that of the cumulative ranks of individual securities, or

$$\mathbb{E}[\bar{U}(\tau_1, \tau_2)] = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[U_i(\tau_1, \tau_2)] = 0.$$

If the event days are not clustered the cross-sectional correlations of the return series are zero (or at least negligible). Under the cross-sectional independence and by Equation (16), the variance of $\bar{U}(\tau_1, \tau_2)$ is

$$\sigma_\tau^2 = \mathrm{var}[\bar{U}(\tau_1, \tau_2)] = \frac{\tau(T-\tau)}{(T-1)n}. \tag{19}$$

Then by the central limit theorem

$$Z = \left(\frac{(T-1)n}{\tau(T-\tau)} \right)^{\frac{1}{2}} \bar{U}(\tau_1, \tau_2) \sim N(0, 1) \text{ as } n \to \infty. \tag{20}$$

The situation is more complicated if the event days are partially overlapping in calendar time which implies cross-sectional correlation. Recalling that the variances of $U_i(\tau_1, \tau_2)$ given in Equation (16) are constants (independent of i), we can write the cross-sectional covariance of $U_i(\tau_1, \tau_2)$ and $U_j(\tau_1, \tau_2)$ as

$$\mathrm{cov}\left[U_i(\tau_1, \tau_2), U_j(\tau_1, \tau_2)\right] = \frac{\tau(T-\tau)}{T-1} \rho_{ij}(\tau_1, \tau_2), \tag{21}$$

where $\rho_{ij}(\tau_1, \tau_2)$ is the cross-sectional correlation of $U_i(\tau_1, \tau_2)$, and $U_j(\tau_1, \tau_2)$, $i, j = 1, \ldots, n$. Utilizing this result and the variance-of-the-sum formula, the variance of $\bar{U}(\tau_1, \tau_2)$ in (17) becomes:

Result 4.

$$\text{var}[\bar{U}(\tau_1,\tau_2)] = \frac{1}{n^2}\sum_{i=1}^{n}\text{var}[U_i(\tau_1,\tau_2)] + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\text{cov}\left[U_i(\tau_1,\tau_2),U_j(\tau_1,\tau_2)\right]$$
$$= \frac{\tau(T-\tau)}{(T-1)n}(1+(n-1)\bar{\rho}_n(\tau_1,\tau_2)), \tag{22}$$

where

$$\bar{\rho}_n(\tau_1,\tau_2) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\rho_{ij}(\tau_1,\tau_2) \tag{23}$$

is the average cross-sectional correlation of cumulated ranks.

Cross-sectional dependence affects the asymptotic distribution of the statistic in Equation (20). However, as discussed in (Lehmann 1999, Scttion 2.8), it is frequently true that the asymptotic normality holds provided that the average correlation, $\bar{\rho}_n(\tau_1,\tau_2)$, tends to zero rapidly enough such that

$$\frac{1}{n}\sum_{i\neq j}^{n}\rho_{ij}(\tau_1,\tau_2) = (n-1)\bar{\rho}_n(\tau_1,\tau_2) \to \gamma \text{ as } n\to\infty, \tag{24}$$

where γ is some finite constant. Under this condition the limiting distribution of Z-statistic in (20) becomes $N(0,1+\gamma)$.

Otherwise, from practical point of view, the crucial result of Formula (22) is that the only unknown parameter to be estimated is the average cross-sectional correlation $\bar{\rho}_n(\tau_1,\tau_2)$. Hagnäs and Pynnonen (2014) discuss approaches to account implicitly for this average correlation in cumulated ranks tests when all events share the same calendar day, i.e., the case of complete clustering. These implicit approaches, however, do not work in the case of partial clustering. Therefore, by utilizing the procedure developed in Kolari et al. (2018), this paper proposes a method to estimate explicitly the cross-sectional correlation, $\bar{\rho}_n(\tau_1,\tau_2)$, and thereby solve the cross-sectional correlation problem in the case of the partial clustering.

4. Correlation Robust Test for Cumulated Ranks

Following Kolari et al. (2018), let τ_{ij}, $0\leq\tau_{ij}\leq\tau$ denote the number of calendar days stocks i and j share in common within the event windows. By independence in Assumption 1, the correlation, $\text{cor}\left[U_{iu},U_{jv}\right]$, of the standardized ranks U_{iu} and U_{jv} is zero whenever the underlying calendar days of the relative event days, u and v, differ and can be non-zero when the calendar days are the same. Denoting these non-zero correlations (which are also covariances) by ρ_{ij}, we get

$$\text{cov}\left[U_i(\tau_1,\tau_2),U_j(\tau_1,\tau_2)\right] = \sum_{u=\tau_1}^{\tau_2}\sum_{v=\tau_1}^{\tau_2}\text{cor}\left[U_{iu},U_{jv}\right] = \tau_{ij}\rho_{ij}.$$

Combining this with (21), we obtain

$$\rho_{ij}(\tau_1,\tau_2) = \left(\frac{T-1}{T-\tau}\right)\frac{\tau_{ij}}{\tau}\rho_{ij}. \tag{25}$$

We can assume that the overlapping window lengths, τ_{ij}, and the cross-sectional correlations, ρ_{ij}, are not dependent on each other so that $\sum_{i\neq j}\tau_{ij}\rho_{ij} = n(n-1)\bar{\tau}\bar{\rho}$, where $\bar{\tau}$

is the average number of overlapping calendar days, and $\bar{\rho}$ is the average cross-sectional correlation of U_i and U_j.[3] Consequently, we can rewrite (22) as

$$\text{var}[\bar{U}(\tau_1, \tau_2)] = \frac{\tau(T-\tau)}{(T-1)n}(1 + (n-1)\delta\bar{\rho}), \tag{26}$$

where $\delta = \bar{\tau}(T-1)/(\tau(T-\tau))$ adjusts the average correlation by the fraction of overlapping calendar days within the event window.

It is notable that, even though the average cross-sectional correlation, $\bar{\rho}$, in Equation (26) is based on $n(n-1)/2$ correlations, it can be computed without estimating individual correlations by utilizing the method introduced by Edgerton and Toops (1928). Instead of $n(n-1)/2$ individual correlations, it turns out that one needs to compute only $n+1$ variances, which is a computational problem of order n viz-a-viz of order n^2 with averaging the correlations. To illustrate the idea, consider n random variables x_j, $j = 1, \ldots, n$ and define the standardized variables $z_j = x_j/\sigma_j$. Next let $\bar{z} = \sum_j z_j/n$ denote the average of the variables. Then because $\text{var}\left[z_j\right] = 1$ and $\text{cov}\left[z_j, z_k\right] = \text{cor}\left[z_j, z_k\right] = \rho_{jk}$, variance of $\bar{z}$ becomes $\sigma_{\bar{z}}^2 = \text{var}[\bar{z}] = (1 + (n-1)\bar{\rho})/n$, we obtain

$$\bar{\rho} = (n\sigma_{\bar{z}}^2 - 1)/(n-1). \tag{27}$$

Hence, to estimate the average cross-sectional correlation, all we need are estimates of n standard deviations of the x-variables and the variance of $\bar{z}$. Finally, for large n, Equation (27) shows that $\bar{\rho} \approx \sigma_{\bar{z}}^2$.

Because in our case the calendar days of different stocks are only partially overlapping, we estimate the variance of the average utilizing the clustering robust estimation technique (e.g., see Cameron et al. 2011) suggested in Kolari et al. (2018).

Following Kolari et al. (2018), denote the calendar days of the returns in the combined estimation and event window as $t = 1, \ldots, L$, which implies that L becomes the number of clusters equaling the number of separate calendar days on which returns are available in the combined estimation and event windows. Let n_t denote the number of stocks having returns on calendar day t and define

$$U_t = \sum_{k=1}^{n_t} U_{kt}. \tag{28}$$

Then

$$U_t^2 = \sum_{k=1}^{n_t} U_{kt}^2 + \sum_{i \neq j}^{n_t} U_{it}U_{jt}, \tag{29}$$

so that

$$\sum_{i \neq j}^{n_t} U_{it}U_{jt} = U_t^2 - \sum_{k=1}^{n_t} U_{kt}^2. \tag{30}$$

Summing these up over the calendar days in the combined estimation and event window, we have

$$\sum_{t=1}^{L}\sum_{i \neq j}^{n_t} U_{it}U_{jt} = \sum_{t=1}^{L} U_t^2 - \sum_{t=1}^{L}\sum_{k=1}^{n_t} U_{kt}^2. \tag{31}$$

Taking the average, we get an estimator for the average correlation

$$\hat{\bar{\rho}} = \frac{1}{M}\sum_{t=1}^{L}\sum_{i \neq j}^{n_t} U_{it}U_{jt}, \tag{32}$$

where

$$M = \sum_{t=1}^{L} n_t(n_t - 1) \tag{33}$$

is the number of the cross-product terms. It is notable that days for which there is available only one return drop automatically out (if $n_t = 1$ for all t, then $\hat{\bar{\rho}} = 0$). The potentially tedious computation over all cross-products can be materially simplified by utilizing the right-hand- side of Equation (31). By Result 2 the variances of standardized ranks are all equal to one and means equal zero. Therefore, arranging the terms of the rightmost sum of Equation (31) to correspond to variance representations, the (double) sum becomes equal to $\sum_{t=1}^{L} n_t$, i.e., the total number of observations.[4] Thus, the only component we need to compute is the first sum of squares on the right-hand-side of (31). Therefore, similar to the illustration of computing the average correlation above, the computational effort of computing the average correlation is again of order n (rather than n^2). Finally, we get:

Result 5. *A computationally efficient form of the average correlation in* (32) *is*

$$\hat{\bar{\rho}} = \frac{N}{M}(s_U^2 - 1), \tag{34}$$

where $N = \sum_{t=1}^{L} n_t$ *is the total number of returns, M is given by* (33)*, and*

$$s_U^2 = \frac{1}{N}\sum_{t=1}^{L} U_t^2 \tag{35}$$

with U_t *given in Equation* (28)*. Variance,* s_U^2*, is a clustering robust variance estimator of standardized ranks in the presence of intra-cluster correlation (cf. e.g., Cameron et al. 2011).*

As noted earlier, $\hat{\bar{\rho}} = 0$ if all $n_t = 1$.

Given the estimator of the average cross-sectional correlation, $\bar{\rho}$, we can define an appropriate cross-sectional correlation robust test for the null hypothesis of zero cumulative abnormal returns

$$H_0 : \mu(\tau_1, \tau_2) = \mathbb{E}[\mathrm{CAR}(\tau_1, \tau_2)] = 0. \tag{36}$$

The test can be defined in terms of the cumulated ranks using the z-ratio

$$z_\tau = \frac{\bar{U}(\tau_1, \tau_2)}{\sigma_\tau \sqrt{1 + (n-1)\delta\hat{\bar{\rho}}}}, \tag{37}$$

where σ_τ is the square root of Equation (19), i.e., the variance

$$\sigma_\tau^2 = \frac{\tau(T-\tau)}{(T-1)n}$$

of $\bar{U}(\tau_1, \tau_2)$ for completely non-overlapping event windows in calendar time [i.e., when $\bar{\rho} = 0$ in Equation (26)], and $\tau = \tau_2 - \tau_1 + 1$ is the length of the window of cumulated abnormal returns.

In event studies, the combined length, T, of the estimation and event period remains fixed, while the number of event firms, n, defines the sample size, thereby being the dimension increased when dealing with the asymptotic distribution of associated test statistics.

Given that the condition in Equation (24) holds for $\hat{\bar{\rho}}$, the null distribution of z_τ is asymptotically normal with zero mean and unit variance.

Kolari and Pynnonen (2011) propose replacing the cumulative ranks in Definition 2 by a single rank number which is based on standardized cumulative abnormal returns (SCARs)

$$\mathrm{SCAR}_i(\tau_1, \tau_2) = \frac{\mathrm{CAR}_i(\tau_1, \tau_2)}{S_{\mathrm{CAR}_i(\tau_1,\tau_2)}} \tag{38}$$

in which $S_{\mathrm{CAR}_i(\tau_1,\tau_2)}$ is the standard deviation of $\mathrm{CAR}_i(\tau_1, \tau_2)$ (for details, see Kolari and Pynnonen 2011). Their approach again accounts implicitly for cross-sectional correlation due to completely overlapping event days. Here, we extend the approach to cover the partial overlapping case. Rather than using the scaled ranks defined in Kolari and Pynnonen (2011), we use the standardized ranks of Definition 1. Subsequently, denoting the standardized rank of $\text{SCAR}_i(\tau_1, \tau_2)$ by U_{i0}, we can base the rank test for testing the null hypothesis of zero cumulative abnormal returns in Equation (36) on the average ranks

$$\bar{U}_0 = \frac{1}{n}\sum_{i=1}^{n} U_{i0}. \tag{39}$$

If the event periods are completely non-overlapping, U_{i0}s are independent with zero mean and unit variance (see Definition 1), in which case the null distribution of $\bar{U}_0$ has zero mean and variance $1/n$. However, if the event days are partially overlapping, the components of $\bar{U}_0$ absorb the cross-sectional correlation over the CAR-window. The correlation that inflates the variance is inherited from the cross-sectional correlations of SCAR_is. Kolari et al. (2018) show that the variance inflation factor is of the form $(1 + (n-1)\nu\bar{\rho})$ as in Equation (26) with the exception that δ is replaced by $\nu = \bar{\tau}/\tau$, the ratio of the average number of overlapping calendar days within the CAR-window to the window length. With this correction the variance of $\bar{U}_0$ becomes $\text{var}[\bar{U}_0] = (1 + (n-1)\nu\bar{\rho})/n$. We can estimate the average cross-sectional correlation, $\bar{\rho}$, as in Equation (32) utilizing only the estimation period in computing $s_{\bar{U}}^2$. For this approach, the standardized ranks in Definition 1 are redefined for the estimation period abnormal returns. Alternatively one can estimate the cross-sectional correlation exactly as in Result 5. Both approaches will produce essentially the same result in most cases. With the estimated average correlation, we get a cross-sectional correlation robust generalized rank test statistic

$$z_{\tau,\text{grank}} = \frac{\sqrt{n}\,\bar{U}_0}{\sqrt{1 + (n-1)\nu\hat{\bar{\rho}}}}, \tag{40}$$

where $\nu = \bar{\tau}/\tau$. Again, given that the condition in Equation (24) holds for $\hat{\bar{\rho}}$, the null distribution of $z_{\tau,\text{grank}}$ is asymptotically normal with zero mean and unit variance.

5. Simulation Results

We generate artificial returns utilizing the Fama and French (2015) five-factor model (FF5),

$$(r_{it} - r_f)_t = \alpha_i + \beta_{i,\text{mkt}}(r_m - r_f)_t + \beta_{i,\text{smb}}\text{SMB}_t + \beta_{i,\text{hml}}\text{HML}_t + \beta_{\text{i,rmw}}\text{RMW}_t + \beta_{\text{i,cmw}}\text{CMW}_t + \epsilon_{it}, \tag{41}$$

where $r_m - r_f$ is the market excess return over the risk-free rate r_f, SMB, HML, RMW, and CMW are common market factors proposed by Fama and French. We utilize daily data from 2 January 1990 through 30 October 2020 (7770 daily returns) to generate 20,000 initial daily return series for this sample period. The regression coefficients for each stock are generated from multivariate normal distribution with mean vector $(0, 1, 0.5, 0.5, 0.5, 0.5)$ and covariance matrix $\sigma_i^2(X'X)^{-1}$, in which σ_i^2 is the variance of the error term ϵ. The stock specific σ_i^2 values are generated by drawing σ_is, the standard deviations, independently from a uniform distribution $U(1,3)$. This corresponds to a range of annual volatilities roughly from 10 percent to 48 percent. The $(X'X)$ matrix is the cross-product matrix of the Fama-French 5-factor regression model.[5] The (7770) error terms ϵ_{it} for stock i is generated independently from the normal distribution $N(0, \sigma_i^2)$.

In the simulations we define the abnormal returns with respect to the market model as

$$AR_{it} = (r_i - r_f)_t - (\hat{\alpha}_i + \hat{\beta}_i(r_m - r_f)_t), \tag{42}$$

where $\hat{\alpha}_i$ and $\hat{\beta}_i$ are ordinary least squares (OLS) estimates. Therefore, missing common factors introduce cross-sectional correlation between the abnormal returns. The event period is ±10 trading days around the event day $t = 0$, and the estimation period consists of 250 days prior the event periods, i.e., relative days $-260, \dots, -11$.

In forthcoming experiments we focus on the effect of cross-sectional correlation on the size of the test. Other issues, such as event induced volatility are well documented for example by Kolari and Pynnonen (2011); Kolari and Pynnönen (2010). Utilizing the base design initiated by Brown and Warner (1985), we generate 1000 samples of randomly selected 50 stocks (the returns of which are generated by the FF5 model in Equation (41)) with four over-lapping event days scenarios. In the first case of non-overlapping event days, the returns are cross-sectionally independent. In the second case of completely overlapping events, all firms share the same event day (calendar time), and in the third and fourth scenarios the event days are randomly scattered across 5 and 10 concecutive calendar days, i.e., one and two weeks of trading days, respectively.

We report two-tailed rejection rates for the null hypothesis of no event-effect across different event windows of ±1, ±2, ±5, and ±10 around the event day, i.e., window lengths $\tau = 1, 3, 5, 10$, and 21 days. In addition to statistic z_τ in Equation (37) we report results for the more traditional rank based test proposed by (Campbell and Wasley 1993, p. 85):

$$z_{cw} = \frac{\sum_{t=\tau_1}^{\tau_2} \bar{k}_t}{\sqrt{\tau} s_{\bar{k}}}, \tag{43}$$

where

$$\bar{k}_t = \frac{1}{n} \sum_{i=1} (K_{it} - \mathbb{E}[K_{it}]) \tag{44}$$

with $\mathbb{E}[K_{it}] = (T+1)/2$ and

$$s_{\bar{k}}^2 = \frac{1}{T} \sum_{t=T_0+1}^{T_2} \bar{k}_t^2. \tag{45}$$

Furthermore, we report results for traditional parametric (cross-sectional correlation non-robust) t-statistics popular in event studies (e.g., see (Campbell et al. 1997, chp. 4)),

$$t_\tau = \frac{\overline{\text{CAR}}(\tau_1, \tau_2)}{\text{s.e}(\text{CAR})}, \tag{46}$$

where $\overline{\text{CAR}}(\tau_1, \tau_2)$ is the sample average of $\text{CAR}_i(\tau_1, \tau_2)$ defined in (2), and s.e.(CAR) is the related standard error. Under independence, the null distribution of t_τ is asymptotically standard normal.

Table 1 summarizes the test statistics and their major features.

Table 2 reports the simulation results of the two-tailed rejection rates of the null hypothesis of no abnormal return at the 5% nominal rejection rate. The results are clear-cut. Panel A of the table reports the non-overlapping case with zero cross-sectional correlation. As expected, all statistics reject close to the nominal rate. Panel B reports results of complete overlapping. That is, all events share the same calendar day; hence, returns are prone to cross-sectional correlation. The new z_τ, $z_{\tau,\text{grank}}$, and the more traditional cumulative ranks statistic, z_{cw}, that account for cross-sectional correlation, reject reasonably close to the nominal rate up to event windows ±5 and exhibit some over-rejection on the longest event window ±10, i.e., 21 days. Not surprisingly, the parametric, non-cross-correlation robust statistic, t_τ, incrementally over-rejects as event windows increase in length. Panel C reports partial overlapping with events clustered randomly within 5 trading days (about a week). For event day testing also the a priori non-robust statistics perform well by rejecting at the nominal rate. However, they start to incrementally over-reject as the event window grows longer. The a priori partial overlapping robust statistics, z_τ and $z_{\tau,\text{grtank}}$, reject close to the nominal rate up to the event window lengths of 5 days and over-reject to some extent for the longest event windows of 11 and 21 days, albeit far less than the non-robust statistics of z_{cw} and t_τ. The results are pretty much similar with the decreased overlapping in Panel D.

Thus, we conclude that accounting for cross-sectional correlation is crucial to avoid biased inferences in statistical testing, not only due to complete overlapping of event windows, but also for partially overlapping cases. Regarding the latter, this paper has introduced two new test statistics that account for these cases.

Table 1. Test statistics and their key features.

Statistic	Type	Robustness Due to		
		Event	Correlation Caused by	
		Volatility	Complete Ovrlp	Partial Ovrlp
$z_\tau = \frac{\bar{U}(\tau_1,\tau_2)}{\sigma_\tau\sqrt{1+(n-1)\delta\hat{\bar{\rho}}}}$, Equation (37)	non-parametric	yes	yes	yes
$z_{\tau,\text{grank}} = \frac{\sqrt{n}\,\bar{U}_0}{\sqrt{1+(n-1)\nu\hat{\bar{\rho}}}}$, Equation (40)	non-parametric	yes	yes	yes
$z_{cw} = \frac{\sum_{t=\tau_1}^{\tau_2}\bar{k}_t}{\tau s_{\bar{k}}}$, Equation (43)	non-parametric	no	yes	no
$t_\tau = \frac{\text{CAR}(\tau_1,\tau_2)}{\text{s.e}(\text{CAR})}$, Equation (46)	parametric	yes	no	no

Table 2. Rejection rates of the null hypothesis of no event effect at the nominal 5% level when the events are no-overlapping, partially overlapping, and completely overlapping.

	CAR Window Length				
	1 Event Day	**3** $(-1,+1)$	**5** $(-2,+2)$	**11** $(-5,+5)$	**21** $(-10,+10)$
Panel A: Non-clustered events					
z_τ	0.048	0.054	0.050	0.052	0.064
$z_{\tau,\text{grank}}$	0.048	0.052	0.053	0.058	0.053
z_{cw}	0.052	0.050	0.051	0.052	0.063
t_τ	0.045	0.035	0.049	0.052	0.048
Panel B: Events clustered on the same trading day					
z_τ	0.059	0.051	0.059	0.064	0.072
$z_{\tau,\text{grank}}$	0.064	0.055	0.065	0.067	0.082
z_{cw}	0.059	0.052	0.061	0.064	0.075
t_τ	0.087	0.091	0.096	0.085	0.110
Panel C: Events clustered on 5 consecutive trading days					
z_τ	0.056	0.055	0.059	0.086	0.076
$z_{\tau,\text{grank}}$	0.056	0.057	0.066	0.083	0.075
z_{cw}	0.050	0.075	0.093	0.127	0.129
t_τ	0.045	0.063	0.077	0.112	0.102
Panel D: Events clustered on 10 consecutive trading days					
z_τ	0.056	0.055	0.059	0.086	0.076
$z_{\tau,\text{grank}}$	0.064	0.046	0.064	0.065	0.082
z_{cw}	0.059	0.062	0.091	0.116	0.133
t_τ	0.065	0.057	0.056	0.089	0.105

6. Summary and Conclusions

This paper proposed two variants of a new non-parametric rank based test statistic for testing cumulative abnormal returns in short-run event studies. The statistics are robust to event-induced volatility and cross-sectional correlation due to complete or partially overlapping event windows. This latter source of cross-sectional correlation is not taken into account by the existing non-parametric test statistics. Simulation results indicate that, unlike typically utilized test statistics, the proposed statistics reject the null hypothesis of no event effect close to the nominal significant level in the partially overlapping case. We conclude that accounting cross-sectional correlation is crucial to avoid biased inferences, not only due to complete overlapping of event windows but also for partial overlapping cases. The non-parametric test statistics proposed in this paper serve this purpose. A major limitation of utilizing non-parametric tests in financial economics is that they seem to play mainly side roles. For example, (Campbell et al. 1997, Sction 4.7) note that non-parametric tests are typically used in conjunction with parametric tests to check robustness of conclusions based on parametric tests. Even so, it should be noted that robustness checks are incrementally demanded in modern empirical financial research. Non-parametric methods can be the tools of choice in completing the task.

Funding: The research has not received external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available upon request from the author.

Acknowledgments: The author wants to thank James Kolari, Henk Snoo, Rudi Wietsma, and referees for many useful comments that improved the paper. All errors are the responsibility of the author.

Conflicts of Interest: The author declares no conflict of interest.

Notes

1. For discussion of true abnormal returns, see Harrington and Shrider (2007).
2. We define conventional event studies as those focusing only on mean stock price effects. Other types of event studies include (for example) the examination of return variance effects (Beaver (1968); Patell (1976)), trading volume (Beaver (1968); Cambell and Wasley (1996)), accounting performance (Barber and Lyon (1997)), and earnings management procedures (Dechow et al. (1995); Kothari et al. (2005)).
3. The equation follows by setting $\sum(x-\bar{x})(y-\bar{y}) = \sum xy - n\bar{x}\bar{y}$ to zero, so that $\sum xy = n\bar{x}\bar{y}$.
4. That is,

$$\sum_{t=1}^{L}\sum_{k=1}^{n_t} U_{kt}^2 = \sum_{t=t_1}^{L_1} U_{1t}^2 + \sum_{t=t_2}^{L_2} U_{2t}^2 + \cdots + \sum_{t=t_n}^{L_n} U_{nt}^2 = \sum_{i=1}^{n}\sum_{t=t_i}^{L_i} U_{it}^2,$$

where $t_i, t_i+1, \ldots, L_i$ indicate observations on stock i with $T_i = L_i - t_i + 1$, the number of observations. By Result 2 $\mathrm{var}[U_{it}] = 1$, so that $\sum_{t=t_i}^{L_i} U_{it}^2 = T_i$. Hence, $\sum_{t=1}^{L}\sum_{k=1}^{n_t} U_{kt}^2 = \sum_{i=1}^{n} T_i = N = \sum_{t=1}^{L} n_t$.

5. Factor returns have been downloaded from the French data library. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, accessed on 15 November 2021.

References

Barber, Brad M., and John D. Lyon. 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics* 43: 341–72. [CrossRef]

Beaver, William H. 1968. The information content of annual announcements. *Journal of Accounting Research Supplement* 6: 69–92. [CrossRef]

Boehmer, Ekkehart, Jim Musumeci, and Annette B. Poulsen. 1991. Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics* 30: 253–72. [CrossRef]

Brown, Stephen J., and Jerold B. Warner. 1985. Using daily stock returns. *Journal of Financial Economics* 14: 3–31. [CrossRef]

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. Robust inference with multiway clustering. *Journal of Business & Economic Statistics* 29: 238–49.

Campbell, Cynthia J., and Charles E. Wasley. 1993. Measuring security price performance using daily nasdaq returns. *Journal of Financial Economics* 33: 73–92. [CrossRef]

Campbell, Cynthia J., and Charles E. Wasley. 1996. Measuring abnormal daily trading volume for samples of NYSE/ASE and NASDAQ securities using parametric and nonparametric test statistics. *Review of Quantitative Finance and Accounting* 6: 309–26. [CrossRef]

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton: Princeton University Press.

Corrado, Charles J. 1989. A nonparametric test for abnormal security-price performance in event studies. *Journal of Financial Economics* 23: 385–95. [CrossRef]

Corrado, Charles J. 2011. Event studies: A methodology review. *Accounting & Finance* 51: 207–34.

Corrado, Charles J., and Terry L. Zivney. 1992. The specification and power of the sign test in event study hypothesis tests using daily stock returns. *Journal of Financial and Quantitative Analysis* 27: 465–78. [CrossRef]

Cowan, Arnold R. 2007. *Eventus 8.0 User's Guide*. Standard Edition 2.1. Ames: Cowan Research LC, pp. 88–89. Available online: www.eventstudy.com (accessed on 5 December 2021).

Dechow, Patricia M., Richard G. Sloan, and Amy P. Sweeney. 1995. Detecting earnings management. *Accounting Review* 70: 193–225.

Edgerton, Harold A., and Herbert A. Toops. 1928. A formula for finding the average intercorrelation coefficient of unranked raw scores without solving any of the individual correlations. *Journal of Educational Psychology* 19: 131–38. [CrossRef]

Fama, Eugene F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* 25: 383–417. [CrossRef]

Fama, Eugene F. 1976. *Foundations of Finance*. New York: Basic Books.

Fama, Eugene F. 1991. Efficient capital markets: Ii. *The Journal of Finance* 46: 1575–617. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116: 1–22. [CrossRef]

Hagnäs, Terhi, and Seppo Pynnonen. 2014. Testing for Cumulative Abnormal Returns in Event Studies with the Rank Test. Available online: https://ssrn.com/abstract=2479228 (accessed on 5 December 2021).

Harrington, Scott E., and David G. Shrider. 2007. All events induce variance: Analyzing abnormal returns when effects vary across firms. *Journal of Financial and Quantitative Analysis* 42: 229–56. [CrossRef]

Jaffe, Jeffrey F. 1974. Special information and insider trading. *The Journal of Business* 47: 410–28. [CrossRef]

Kolari, James W., Bernd Pape, and Seppo Pynnonen. 2018. Event Study Testing with Cross-Sectional Correlation Due to Partially Overlapping Event Windows. Available online: https://ssrn.com/abstract=3167271 (accessed on 5 December 2021).

Kolari, James W., and Seppo Pynnonen. 2011. Nonparametric rank tests for event studies. *Journal of Empirical Finance* 18: 953–71. [CrossRef]

Kolari, James W., and Seppo Pynnönen. 2010. Event study testing with cross-sectional correlation of abnormal returns. *The Review of Financial Studies* 23: 3996–4025. [CrossRef]

Kothari, Sagar P., Andrew J. Leone, and Charles E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39: 163–97. [CrossRef]

Lehmann, Eric L. 1999. *Elements of Large-Sample Theory*. Berlin and Heidelberg: Springer.

Lehmann, Eric L. 2006. *Nonparamterics: Statistical Mehtod Based on Ranks*, Revised 1st ed. Berlin and Heidelberg: Springer.

Luoma, Terhi. 2011. Nonparametric Event Study Tests for Testing Cumulative Abnormal Returns. Ph.D. dissertation, University of Vaasa, Department of Mathematics and Statistics, Vaasa, Finland. Available online: https://urn.fi/URN:ISBN:978-952-476-372-1 (accessed on 5 December 2021).

Patell, James M. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research* 14: 246–76. [CrossRef]

Journal of
Risk and Financial Management

MDPI

Article

Further Tests of the ZCAPM Asset Pricing Model

James W. Kolari [1,*], Jianhua Z. Huang [2], Wei Liu [3] and Huiling Liao [2]

[1] Department of Finance, Texas A&M University, College Station, TX 77843, USA
[2] Department of Statistics, Texas A&M University, College Station, TX 77843, USA; jianhua@stat.tamu.edu (J.Z.H.); huiling@stat.tamu.edu (H.L.)
[3] USAA Bank, San Antonio, TX 78288, USA; yuna0831@gmail.com
* Correspondence: j-kolari@tamu.edu; Tel.: +1-979-845-4803

Abstract: In a recent book, Kolari et al. developed a new theoretical capital asset pricing model dubbed the *ZCAPM*. Based on out-of-sample cross-sectional tests using U.S. stocks, the ZCAPM consistently outperformed well-known multifactor models popular in the finance literature. This paper presents further evidence that expands their sample period from 1927 to 2020. Results are provided for the subperiods 1927 to 1964 and 1965 to 2020. Our results corroborate those of KLH. In cross-sectional tests, the ZCAPM outperforms the CAPM as well as the Fama and French three-factor model and Carhart four-factor model. Outperformance is found in terms of both higher goodness of fit and the statistical significance of factor loadings. Interestingly, the earlier subperiod results highlight problems with the endogeneity of test assets in cross-sectional tests of multifactor models.

Keywords: asset pricing; zero-beta CAPM; return dispersion; expectation-maximization (EM) regression; latent variable

Citation: Kolari, James W., Jianhua Z. Huang, Wei Liu, and Huiling Liao. 2022. Further Tests of the ZCAPM Asset Pricing Model. *Journal of Risk and Financial Management* 15: 137. https://doi.org/10.3390/jrfm15030137

Academic Editors: Thanasis Stengos and Robert Brooks

Received: 3 December 2021
Accepted: 1 March 2022
Published: 15 March 2022

1. Introduction

This paper extends the recent work by Kolari et al. (2021) (hereafter KLH) in which they developed a new theoretical model of capital asset prices dubbed the *ZCAPM*. The authors derived the ZCAPM from Black's (1972) renowned zero-beta CAPM as a special case based on two unique efficient and inefficient orthogonal portfolios. This special case enabled the derivation of an alternative specification of the zero-beta CAPM. The ZCAPM is a parsimonious two-factor model comprised of beta risk associated with average market returns and zeta risk related to the cross-sectional return dispersion of assets in the market.[1] Based on the theoretical ZCAPM, an innovative empirical ZCAPM was developed using expectation–maximization (EM) regression methods.[2]

Subsequent empirical tests by KLH demonstrated that the ZCAPM is a superior asset pricing model that outperforms the CAPM as well as popular multifactor models, including the Fama and French (1992, 1993, 1995, 2015, 2018, 2020) three-, five-, and six-factor models in addition to the Carhart (1997), Hou et al. (2015), and Stambaugh and Yuan (2017) four-factor models. In their empirical tests of U.S. stock returns, out-of-sample Fama and MacBeth (1973) cross-sectional regression tests were conducted for the sample period 1965 to 2018. The ZCAPM consistently outperformed the aforementioned models in terms of both goodness-of-fit and statistical significance of zeta risk factor loadings. In some test asset portfolios, the empirical ZCAPM was able to achieve cross-sectional R^2 estimates as high as 95 percent and normally had values exceeding 70 percent: this goodness-of-fit is near perfect. It means that estimated risk parameters in an earlier period almost completely explain out-of-sample (next month) returns in the cross section of average stock returns. By comparison, other popular multifactor models typically had R^2 values noticeably lower than the ZCAPM in different test asset portfolios and sample periods. Regarding the statistical significance of factor loadings, zeta risk loadings associated with cross-sectional return dispersion in the ZCAPM almost always had t statistics in the range of

3 to 6. However, factors in popular multifactor models did not reach this level of statistical significance in cross-sectional tests. These findings are important in light of the recent work by Harvey et al. (2016) and Chordia et al. (2020), who found that factor loadings should attain t statistics of three or more to avoid false discoveries in asset pricing studies. The ZCAPM was the only model that passed the recommended validity tests.

Where does the ZCAPM fit into the prior literature? From a theoretical perspective, it is based on the general equilibrium framework of the capital asset pricing model (CAPM) of Treynor (1961, 1962), Sharpe (1964), Lintner (1965), and Mossin (1966). It applies the foundational mean–variance Markowitz (1959) portfolio theory and the Tobin (1958) equilibrium pricing methods to derive an alternative form of Black's zero-beta CAPM. Compared to extant asset pricing models that have been tested in the literature, the empirical ZCAPM is a new econometric model based on EM regression and mixture model methods. No previous asset pricing studies employ these methods in the estimation of factor models. The two factors in the empirical ZCAPM have precedent in the financial literature. Mean market returns are used to estimate beta risk. In addition, cross-sectional return dispersion is used to estimate zeta risk. Regarding this factor, a limited number of studies by Jiang (2010), Demirer and Jategaonkar (2013), Garcia et al. (2014), and Chichernea et al. (2015) augmented the market model form of the CAPM with a return dispersion factor. However, they used standard ordinary least squares (OLS) regression methods to estimate a coefficient related to return dispersion, rather than EM regression in a mixture model. The ZCAPM is different in its empirical estimation of this coefficient in that it explicitly models both positive and negative sensitivity to changes in return dispersion over time. A signal variable denoted $D_{jt} = +1, -1$ for asset j at time t (e.g., one day) is introduced to capture the potential two-sided effects of return dispersion on asset returns. As cross-sectional return dispersion increases in the population of assets at a point t in time, assets in the upper part of the distribution of returns experience increasing returns, and conversely those in the lower part of the distribution experience decreasing returns. If the return dispersion decreases, the opposite return effects occur for assets in the upper and lower parts of the distribution of returns. Since D_{jt} is a latent, unobservable variable, KLH estimated its probability using EM regression. This probability is multiplied by the coefficient on the return dispersion to obtain an estimate of the zeta risk, which is different from other previous studies that incorporated a return dispersion factor.

It is important to distinguish between cross-sectional return dispersion from time-series return dispersion. An example of the latter is the work of Bekaert et al. (2012), who employed the time-series standard deviation of returns for stocks as a factor in an asset pricing model. They used daily returns in a one-month period to compute monthly time-series standard deviations of returns for individual stocks and then averaged this idiosyncratic risk metric for N firms in the market to compute an aggregate idiosyncratic variance measure. Numerous studies have utilized a time-series market volatility factor, including those of Ang et al. (2006b, 2009), Adrian and Rosenberg (2008), Da and Schaumburg (2011), Chang et al. (2013), Bansal et al. (2014), Bollerslev et al. (2016), and Chen et al. (2021), among others.[3] Relevant to the ZCAPM, cross-sectional return dispersion is quite different from time-series dispersion. Earlier work by Jiang (2010) showed that, for U.S. stock returns in the period of 1963 to 2005, these two measures of return volatility are uncorrelated with one another in many sample periods. This evidence led Jiang to conclude that time-series and cross-sectional return dispersion are different market risk measures. Hence, the ZCAPM extends the small set of studies that incorporate cross-sectional return dispersion in an asset pricing model but has little or no connection to the larger body of time-series volatility studies.

The present study contributes further evidence on the ZCAPM. First, using U.S. stock return series available on Kenneth French's data library website[4], we extend the analysis period back to the 1928 to 1964 period. Second, we update their analyses to the period 1965 to 2020. ZCAPM results are benchmarked against the CAPM as well as the Fama and French (1992, 1993, 1995) three-factor model and Carhart (1997) four-factor model.

We do not test other multifactor models for which factors and test asset portfolios are not available on French's website. Test assets include 25 size and book-market equity ratio (BM) sorted, 25 size and momentum sorted, and 40 industry portfolios. We report results for out-of-sample, cross-sectional Fama and MacBeth tests. In general, our results support those of KLH. The empirical ZCAPM outperforms the CAPM as well as three- and four-factor models, in some cases by large margins. Zeta risk loadings are highly significant, with t statistics exceeding the recommended three hurdle rate in all cases. While multifactor loadings in the three- and four-factor models have t statistics exceeding 3.0 in the 1928 to 1964 subperiod, they generally do not in the more recent 1965 to 2020 subperiod. Additionally, we find that estimated zeta risk premiums are economically meaningful with a range from 0.47 percent to 1.29 percent per month per unit estimated zeta coefficient.

Graphical analyses of the ZCAPM, CAPM, and three- and four-factor models are also provided. In these cross-sectional analyses, fitted (or predicted) one-month-ahead excess stock returns are compared to realized (or actual) excess stock returns of test asset portfolios. Hence, these analyses are out-of-sample investable strategies. In general, we find that the ZCAPM outperforms other models. When industry portfolios are included in the test assets, the ZCAPM outperforms other models by considerable margins. These analyses demonstrate a major problem in testing the three- and four-factor models with endogenous test asset portfolios created from sorts on the same firm-level variables (i.e., size and BM) used to construct the size and value factors. In the earlier period of 1928 to 1964, this endogeneity problem worsened relative to the more recent 1965 to 2020 period due to smaller sample sizes of stocks, as the data go back in time. Our results support Lewellen et al. (2010), Daniel and Titman (2012), and others who have advocated for combining exogenous industry portfolios with other portfolios in asset pricing tests. In sum, our graphical analyses confirm the findings of KLH in support of the ZCAPM over the CAPM as well as three- and four-factor models, using long/short zero-investment portfolios as multifactors.

We conclude from these findings that the ZCAPM dominates other popular asset pricing models. Given that size, BM, and momentum sorted portfolios as test assets are exogenous to the ZCAPM's mean market return and return dispersion factors, this dominance is remarkable. Further research is recommended for applying the ZCAPM to different countries and asset classes (e.g., bonds, commodities, and real estate) to assess its performance relative to the existing asset pricing models. Additionally, applications to event studies, mutual and hedge funds, investment analysis, and other areas of finance are recommended.

The plan of this study is as follows. Section 2 overviews the ZCAPM. Section 3 describes our methodology, including data and empirical tests. Section 4 presents the empirical results. The last Section 5 gives the conclusion.

2. Overview of the ZCAPM

Here, we overview the theoretical ZCAPM and its companion empirical ZCAPM. Again, Kolari et al. (2021) (KLH) derived the ZCAPM as a special case of Black's zero-beta CAPM.[5] In their derivation, they focused on two orthogonal portfolios on the boundary of the mean–variance investment parabola—one that is efficient and one that is inefficient—with the same time-series variance of returns. Formulas of the expected returns for these two portfolios are written based on new insights concerning the mean–variance parabola. Upon substituting these expected returns into the zero-beta CAPM, the theoretical ZCAPM is obtained. Subsequently, the authors proposed a novel empirical ZCAPM for estimation purposes, using real world data. Unlike prior asset pricing models that use ordinary least squares (OLS) regression for estimation, the empirical ZCAPM utilizes expectation–maximization (EM) regression methods. In the forthcoming discussion, we abbreviate the derivations in the work of KLH to conserve space and highlight the main ideas of the theoretical and empirical ZCAPM. Readers interested in more details are referred to their book.

2.1. Theoretical ZCAPM

KLH mathematically proved two new insights about the Markowitz mean–variance investment parabola. First, they provided two mathematical proofs[6] to show that the width or span of the parabola is largely determined by the cross-sectional standard deviation of returns of all assets' returns. Second, given that this *return dispersion* defines the width of the parabola, the mean return of all assets should lie somewhere in the middle of the parabola on its axis of symmetry. The latter finding implies that the mean market portfolio used to proxy the market portfolio is inefficient. Regardless of whether all assets are equal- or value-weighted to form portfolios, the market portfolio in the CAPM, which lies on the efficient frontier, is far above the mean market portfolio that is located on the axis of symmetry. Consistent with the Roll (1977) critique, because the CAPM cannot be tested without an efficient portfolio, previous empirical tests of the CAPM using the mean market model returns to proxy market portfolio returns are invalid. The CAPM cannot be declared dead because it was never legitimately tested using efficient portfolios (see Fama and French 1996, 2004).

Figure 1 illustrates the return dispersion and mean market return characteristics of the mean–variance parabola. The x-axis is the time-series variance of returns for an asset or portfolio denoted as $\tilde{\sigma}^2_P$. In a one-day period of time, this variance can be measured by computing returns in, say, 10 min intervals during the day. On the y-axis is the expected returns of assets. The cross-sectional variance of returns of all assets in the market during the day is denoted as $\tilde{\sigma}_a$. Naturally, the mean market return denoted $E(\tilde{R}_a)$ must be located in the middle of the cross-sectional distribution of asset returns. In turn, it must be true that $E(\tilde{R}_a) \approx E(\tilde{R}_G)$, where the latter is the expected return on the global minimum variance portfolio G. Clearly, the mean market portfolio a is located far below the efficient frontier in Figure 1.

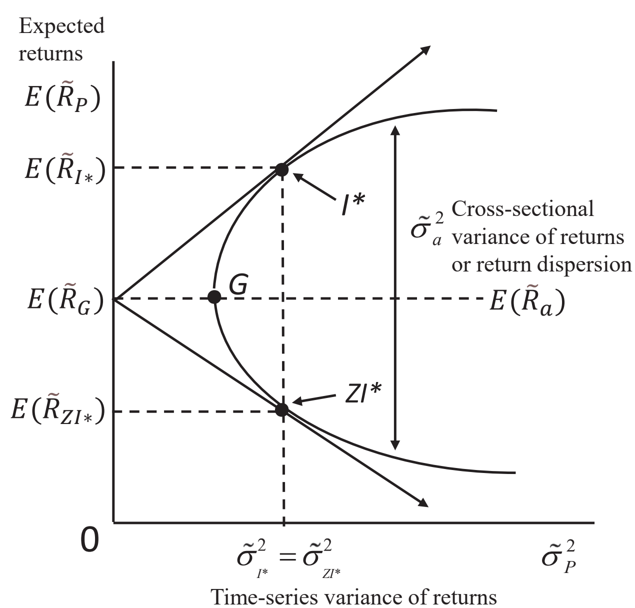

Figure 1. Geometric approach of the theoretical ZCAPM based on the Markowitz mean–variance investment parabola.

Next, KLH used this framework to identify two unique portfolios, I^* and ZI^*, that are uncorrelated with one another. These portfolios have the same time-series variance of returns or total risk, i.e., $\tilde{\sigma}^2_{I^*} = \tilde{\sigma}^2_{ZI^*}$.[7] Notice that portfolio I^* is on the efficient frontier, and portfolio ZI^* is inefficient on the parabola's lower boundary. A new geometry is introduced in this analysis. In the CAPM, the market portfolio M is geometrically located at the tangent point from a ray extending from the riskless rate to the efficient frontier. In the ZCAPM, portfolios I^* and ZI^* are located by moving along the axis of symmetry at the expected rate $E(\widetilde{R}_a)$ and then up or down, respectively, by the cross-sectional return dispersion $\widetilde{\sigma}_a$. Using this geometry, KLH defined the expected returns for portfolios I^* and ZI^* as follows:

$$E(\widetilde{R}_{I^*}) \approx E(\widetilde{R}_a) + f(\theta)\sigma_a \quad (1)$$

$$E(\widetilde{R}_{ZI^*}) \approx E(\widetilde{R}_a) - f(\theta)\sigma_a, \quad (2)$$

where $f(\theta)$ is a complex expression approximately equal to one (due to almost completely random risky asset returns[8]).

Assuming $f(\theta) = 1$, KLH substituted the expected returns for portfolios I^* and ZI^* into Black's zero-beta CAPM to derive the theoretical ZCAPM without a riskless asset. The zero-beta CAPM specifies the expected return for the ith asset as

$$E(\widetilde{R}_i) = E(\widetilde{R}_{ZM}) + \beta_{i,M}[E(\widetilde{R}_M) - E(\widetilde{R}_{ZM})] \quad (3)$$

$$E(\widetilde{R}_i) = \beta_{i,M}E(\widetilde{R}_M) + (1 - \beta_{i,M})[E(\widetilde{R}_{ZM})], \quad (4)$$

where $\beta_{i,M}$ is the sensitivity or beta risk of asset i's return with respect to the excess return of the expected market portfolio return, $E(\widetilde{R}_M)$ and its zero-beta (uncorrelated) portfolio expected return, or $E(\widetilde{R}_{ZM})$. The latter is the borrowing rate in Black's model, unlike the riskless rate R_f in the CAPM.[9]

For portfolios I^* and ZI^*, their expected returns are

$$E(\widetilde{R}_{I^*}) = \beta_{I^*,M}E(\widetilde{R}_M) + (1 - \beta_{I^*,M})E(\widetilde{R}_{ZM}) \quad (5)$$

$$E(\widetilde{R}_{ZI^*}) = \beta_{ZI^*,M}E(\widetilde{R}_M) + (1 - \beta_{ZI^*,M})E(\widetilde{R}_{ZM}), \quad (6)$$

where $\beta_{I^*,M}$ and $\beta_{ZI^*,M}$ are beta risks of portfolios I^* and ZI^* associated with market portfolio M, respectively. Solving these equations[10], we obtain the general form of the zero-beta CAPM:

$$E(\widetilde{R}_i) = \beta_{i,I^*}E(\widetilde{R}_{I^*}) + (1 - \beta_{i,I^*})E(\widetilde{R}_{ZI^*}), \quad (7)$$

where $\beta_{i,I^*} = (\beta_{i,M} - \beta_{ZI^*,M})/(\beta_{I^*,M} - \beta_{ZI^*,M})$. As observed by Roll (1980), the above expression shows that the zero-beta CAPM can be specified in terms of any efficient portfolio and its orthogonal zero-beta (inefficient) counterpart on the mean–variance parabola. Here, KLH re-wrote the zero-beta CAPM using the unique portfolios I^* and ZI^*.

Upon substituting $E(\widetilde{R}_{I^*})$ and $E(\widetilde{R}_{ZI^*})$ in Equations (1) and (2) into Equation (7), the theoretical ZCAPM can be specified as follows:

$$\begin{aligned} E(\widetilde{R}_i) &= \beta_{i,I^*}E(\widetilde{R}_{I^*}) + (1 - \beta_{i,I^*})E(\widetilde{R}_{ZI^*}) \\ &= E(\widetilde{R}_{ZI^*}) + \beta_{i,I^*}[E(\widetilde{R}_{I^*}) - E(\widetilde{R}_{ZI^*})] \\ &= E(\widetilde{R}_a) - \sigma_a + \beta_{i,I^*}\{[E(\widetilde{R}_a) + \sigma_a] - [E(\widetilde{R}_a) - \sigma_a]\} \\ &= E(\widetilde{R}_a) + (2\beta_{i,I^*} - 1)\sigma_a \\ E(\widetilde{R}_i) &= E(\widetilde{R}_a) + Z^*_{i,a}\sigma_a, \end{aligned} \quad (8)$$

where $Z^*_{i,a} = 2\beta_{i,I^*} - 1$.

Adding a third riskless asset rate R_f, and again using the definitions of $E(\widetilde{R}_{I^*})$ and $E(\widetilde{R}_{ZI^*})$ in Equations (1) and (2), the expected return of the ith asset is

$$\begin{aligned} E(\widetilde{R}_i) &= w_{I^*}E(\widetilde{R}_{I^*}) + w_{ZI^*}E(\widetilde{R}_{ZI^*}) + w_f R_f \\ &= w_{I^*}[E(\widetilde{R}_a) + \sigma_a] + w_{ZI^*}[E(\widetilde{R}_a) - \sigma_a] + w_f R_f \\ &= (w_{I^*} + w_{ZI^*})E(\widetilde{R}_a) + (w_{I^*} - w_{ZI^*})\sigma_a + w_f R_f, \end{aligned} \tag{9}$$

where I^*, ZI^*, and f are orthogonal assets with corresponding weights w_{I^*}, w_{ZI^*}, and w_f that sum to one with both long and short positions in the assets allowed. By rearranging terms and using Equation (9), the final form of the theoretical ZCAPM becomes

$$E(\widetilde{R}_i) - R_f = (w_{I^*} + w_{ZI^*})[E(\widetilde{R}_a) - R_f] + (w_{I^*} - w_{ZI^*})\sigma_a \tag{10}$$

$$E(\widetilde{R}_i) - R_f = \beta_{i,a}[E(\widetilde{R}_a) - R_f] + Z^*_{i,a}\sigma_a, \tag{11}$$

where beta risk coefficient $\beta_{i,a} = w_{I^*} + w_{ZI^*}$ measures the sensitivity of the ith asset's excess returns to average market excess returns of all assets, and zeta risk coefficient $Z^*_{i,a} = w_{I^*} - w_{ZI^*}$ measures the sensitivity of an asset's excess returns to the market return dispersion of all assets.[11] KLM used the notation $\beta_{i,a}$ to denote beta risk with respect to the average returns on the portfolio of n assets in the market. This beta is distinguished from CAPM market beta $\beta_{i,M}$ with respect to the market portfolio (typically denoted simply as β_i).

Returning to the mean–variance parabola, it is interesting that beta risk and zeta risk in the theoretical ZCAPM can be used to describe its architecture, including not only boundary portfolios, but locations of assets and portfolios within the parabola. Assets and portfolios with positive (negative) zeta risk lie in the upper (lower) portion of the investment parabola. On any zeta risk curve in the parabola, as beta risk increases, the time-series variance of returns increases. As shown in Kolari et al. (2021, Figure 10.1, p. 272 and Figure 10.2, p. 274), an interlocking web of beta and zeta risks result that shape the parabola with zeta risk increasing vertically and beta risk increasing horizontally. Hence, the parabola contains a risk structure based on the systematic risks of assets with respect to average market returns and market return dispersion. Interestingly, in Chapter 10 of their book, KLH confirmed this architecture using out-of-sample (next month) empirical evidence for U.S. stock portfolios. Portfolios along the highest zeta risk curve comprise the efficient frontier.[12] Additionally, the mean market portfolio a lies approximately on the axis of symmetry of the parabola. Supporting this conjecture, in Chapter 9 of their book, KLH constructed relatively efficient portfolios and showed that the CRSP market index lies along the axis of symmetry of the parabola.

2.2. Empirical ZCAPM

Figure 2 shows how assets in the upper and lower portions of the mean–variance parabola are affected over time in response to changes in the mean market returns and cross-sectional return dispersion of all assets in the market.[13] Comparing $t = 1$ to $t = 2$, when the average market returns do not change, we see that asset returns in the upper (lower) portion of the parabola experience increasing (decreasing) returns as the return dispersion increases. At $t = 3$, the return dispersion decreases, which tends to decrease (increase) the asset returns in the upper (lower) portion of the parabola. Of course, as mean market returns decrease in this period, all asset returns decrease in concert with lower mean market returns. In period $t = 4$, mean market returns increase but the return dispersion changes little, if at all. In this period, all assets' returns tend to increase under these conditions.

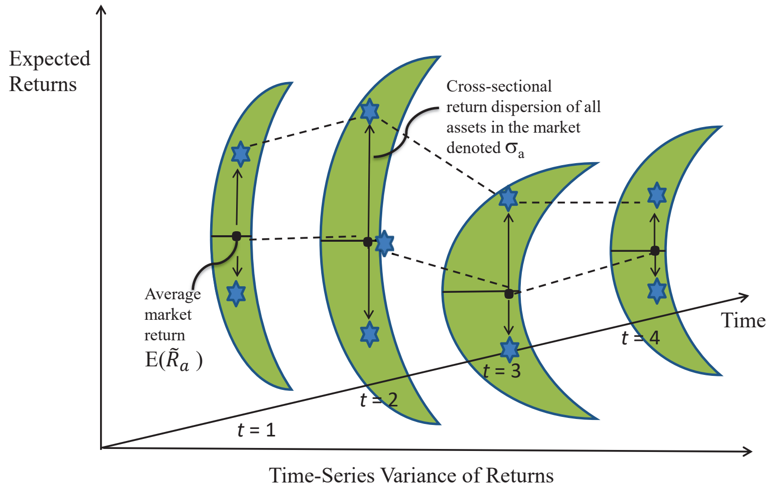

Figure 2. As the investment parabola moves over time t, its level and width change. The level changes with average market returns, and the width changes with cross-sectional return dispersion of all assets in the market. Assets in the upper (lower) half of the parabola experience opposite return effects of changing return dispersion, whereas as all assets' returns move up and down in concert with the average market returns.

How did KLH empirically model the time-series behavior of the mean–variance parabola depicted in Figure 2? The positive and negative effects of return dispersion on asset returns in the upper and lower portions of the parabola need to be taken into account. To solve this problem, they introduced a dummy signal variable denoted D_{it} for each ith asset. The following novel empirical ZCAPM is proposed:

$$\widetilde{R}_{it} - R_{ft} = \alpha_i + \beta_i(\widetilde{R}_{at} - R_{ft}) + Z_i D_{it}\widetilde{\sigma}_{at} + \widetilde{u}_{it}, \quad t = 1, \cdots, T \tag{12}$$

where $R_{it} - R_{ft}$ is the excess return for the ith asset over the riskless rate at time t, β_i measures sensitivity to excess average market returns equal to $R_{at} - R_{ft}$, Z_i measures sensitivity to return dispersion σ_{at}, D_{it} is a signal variable with values $+1$ and -1 representing positive and negative return dispersion effects on stock returns at time t, respectively, and $u_{it} \sim$ iid $N(0, \sigma_i^2)$. No previous studies modeled two-sided return dispersion risk using positive and negative risk loadings. A previous study by Ang et al. (2006a) estimated downside and upside market betas, i.e., β^- and β^+, using excess market returns over time, below and above the mean market return, but did not introduce a dummy variable in their analyses. Similarly, Lettau et al. (2014) found that the cross section of currency returns can be explained by downside market beta risk. More recently, Bollerslev et al. (2016) proxied good and bad stock return volatility by utilizing a relative difference in the semi-variance measure but again did not use a dummy variable approach to simultaneously model their effects to predict returns.

Departing from previous literature, signal variable D_{it} is modeled by KLH as an unknown or latent (hidden) variable. They defined D_{it} as an independent random variable with the following two-point distribution:

$$D_{it} = \begin{cases} +1 & \text{with probability } p_i \\ -1 & \text{with probability } 1 - p_i, \end{cases} \tag{13}$$

where p_i (or $1 - p_i$) is the probability of a positive (or negative) return dispersion effect, and D_{it} are independent of u_{it}.[14]

To estimate the empirical ZCAPM's parameters $\theta_i = (\beta_i, Z_i, p_i)$, KLH employed the expectation–maximization (EM) algorithm of Dempster et al. (1977) (See also Jones and McLachlan 1990; McLachlan and Peel 2000; McLachlan and Krishnan 2008). Their book gives detailed step-by-step estimation procedures. Unlike any previous asset pricing model, the empirical ZCAPM can be characterized as a probabilistic mixture model with two mixture components. Each component itself is a two-factor regression model (see Equations in Note 14). Hidden dummy variable D_{it} determines the operative regression model.

Notice that the coefficient of the return dispersion in regression Equation (12) is a random variable $Z_{i,a}D_{i,t}$ with two possible values, $+Z_{i,a}$ or $-Z_{i,a}$, based on the sign of signal variable $D_{i,t}$. Here, the signal variable has mean $E(D_{it}) = 2p_i - 1$ and variance $\text{Var}(D_{it}) = 4p_i(1 - p_i)$. They separate the mean from the random coefficient $Z_{i,a}D_{it}$ associated with σ_{at} as follows:

$$Z_{i,a}D_{it} = Z_{i,a}(2p_i - 1) + Z_{i,a}[D_{it} - (2p_i - 1)]. \tag{14}$$

Thus, using definitions $Z^*_{i,a} = Z_{i,a}(2p_i - 1)$ and $u^*_{it} = Z_{i,a}[D_{it} - (2p_i - 1)]\sigma_{at} + u_{it}$, the marginal form of the empirical ZCAPM relation (12) becomes

$$R_{it} - R_{ft} = \beta_{i,a}(R_{at} - R_{ft}) + Z^*_{i,a}\sigma_{at} + u^*_{it}, \quad t = 1, \cdots, T. \tag{15}$$

where the term $Z^*_{i,a}\sigma_{at}$ results from integrating out the probability distribution of the unobservable signal variable in the term $Z_{i,a}D_{it}\sigma_{at}$ in model (12). Regression parameter $Z^*_{i,a}$ represents the zeta risk loading in the theoretical ZCAPM as specified in Equation (11).

It should be mentioned that there is no mispricing error term (i.e., $\alpha_i = 0$) in empirical ZCAPM relation (15). In tests using U.S. stock returns, KLH found that introducing an α_i term did not lower the residual variance and therefore did not improve in-sample data fitting. In the present study, the α term is not needed, as we test the empirical ZCAPM using standard Fama and MacBeth (1973) cross-sectional regression analyses, to be discussed shortly.[15]

The positive or negative sign of zeta risk loading $Z^*_{i,a}$ is determined by the probability p_i of signal variable D_{it} in sample period $t = 1, \cdots, T$. If $p_i > 1/2$ (or $< 1/2$), $Z^*_{i,a}$ has a positive (or negative) sign. By way of interpretation, $Z^*_{i,a}$ measures the average increase or decrease in asset returns in response to a one unit change in market return dispersion σ_{at}.

Setting the empirical ZCAPM apart from other studies that include a return dispersion factor cited earlier in Section 1, the variance of the error term u^*_{it} in relation (15) is not constant. This heterogeneity of error variance can be defined as follows:

$$\text{Var}(u^*_{it}) = 4\, p_i(1 - p_i)\, Z^2_{i,a}\, \sigma^2_{at} + \text{Var}(u_{it}). \tag{16}$$

Due to this property, other studies incorporating return dispersion as a factor are misspecified.

KLH provided Matlab, R, and Python programs for EM estimation of the empirical ZCAPM at GitHub (https://github.com/zcapm (accessed on 1 September 2021)): Programs to run cross-sectional Fama and MacBeth regression tests are provided also. In this study, we employ their R programs due to the faster estimation speed relative to the Matlab and Python programs.

3. Cross-Sectional Tests

3.1. Data

U.S. stock returns for all common stocks on the Center for Research in Security Prices (CRSP) database are used. Daily stock returns are gathered for two subperiods: (1) January 1928 to December 1964, and (2) January 1965 to December 2020. CRSP value-weighted index returns and 30-day U.S. Treasury bill rates, in addition to size, value, and momentum factors, are downloaded from Kenneth French's online database website.

We compute the return dispersion factor for the ZCAPM as the daily cross-sectional standard deviation of returns of all stocks in the market:

$$\sigma_{at} = \sqrt{\frac{n}{n-1}\sum_{i=1}^{n} w_{it-1}(R_{it} - R_{at})^2}, \quad (17)$$

where n is the total number of stocks, w_{it-1} is the previous day's market vale weight for the ith stock, R_{it} is the return of the ith stock on day t, and R_{at} is the value-weighted average return of all available stocks in the country on day t.

To benchmark the performance of the ZCAPM in our empirical tests, we employ the following asset pricing models:

- CAPM in market model form (See Sharpe 1963; Fama 1968). with an excess market return factor (*MKT-RF*) defined as the value-weighted CRSP return minus the U.S. Treasury bill rate;
- The Fama and French (1992, 1993, 1995) three-factor model based on augmenting the CAPM with a size factor (viz. small minus large firms' stock returns denoted as *SMB*) and a value factor (viz. high value minus low value firms' stock returns denoted *HML*);
- The Carhart (1997) four-factor model based on augmenting the three-factor model with a momentum factor (viz. stocks with high past returns minus stocks with low past returns denoted *MOM*).

French's website contains construction details for the multifactors *SMB*, *HML*, and *MOM*. As defined there, based on portfolio deciles, *SMB* is the average return on the three small portfolios minus the average return on the three big portfolios. *HML* is the average return on the two value portfolios minus the average return on the two growth portfolios. Additionally, *MOM* is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios.

Descriptive statistics for our data are provided in Table 1. Compared to the market, size, value, and momentum factors, the magnitude of the cross-sectional return dispersion is much larger at 1.56 percent, compared to a range of only 0.003 percent to 0.03 percent for the other factors. In addition, with the exception of return dispersion, notice that the standard deviations of factors are quite large relative to their mean values; hence, these factors can fluctuate widely over time.

Table 1. Descriptive statistics for U.S. stock returns in the sample period of January 1928 to December 2020.

Panel A. 1928 to 1964						
	90 Portfolios	***MKT-RF***	***SMB***	***HML***	***MOM***	***Ret Disp***
Mean	0.09	0.03	0.003	0.02	0.02	1.56
Std dev	1.80	1.16	0.66	0.68	0.79	0.86
Panel B. 1965 to 2020						
	90 Portfolios	***MKT-RF***	***SMB***	***HML***	***MOM***	***Ret Disp***
Mean	0.07	0.03	0.007	0.01	0.03	1.79
Std dev	1.21	1.03	0.54	0.55	0.76	0.60

Table 1 gives the means and standard deviations of returns for test asset portfolios and asset pricing factors also. The 90 test assets are formed by combining 25 size and book-to-market ratio (BM) portfolios, 25 size and momentum portfolios, and 40 industry portfolios. Two subperiods are used: (1) January 1928 to December 1964, and (2) January 1965 to December 2020.

3.2. Cross-Sectional Regression Tests

We provide two different cross-sectional regression tests. The first test is based on the standard two-step Fama and MacBeth (1973) regression analyses. Step one estimates the time-series regression equation for the asset pricing model, using daily returns in a one-year period for each of the test asset portfolios. Estimated factor loadings for beta and zeta coefficients are retained for use in the next step. Step two is a cross-sectional regression with one-month-ahead returns for test asset portfolios as the dependent variable and beta and zeta risk factor loadings from the previous year as the independent variables. This procedure is rolled forward one month at a time until the end of the sample period. This procedure represents an investable strategy in the sense that an investor could implement it in the real world. Out-of-sample returns are related to prior risk parameter estimates to assess the validity of models. No tampering or manipulation is possible in this setup. In this regard, Simin (2008, p. 356) commented that the use of step-ahead (e.g., one-month-ahead) returns in this procedure to assess the predictive ability of asset pricing models mitigates a number of evaluation problems, including data snooping, the use of R^2 as a measure of goodness-of-fit, and efficiency issues. Likewise, Ferson et al. (2013) argued that the practical value of asset pricing models should be assessed using out-of-sample tests as in the two-step Fama and MacBeth procedure discussed above.

In the second step of the Fama–MacBeth procedure, we run the following cross-sectional regression to test the empirical ZCAPM based on estimates of beta and zeta risk coefficients (or loadings) from time-series regression (15):

$$R_{i,T+1} - R_{fT+1} = \lambda_0 + \lambda_\beta \hat{\beta}_i + \lambda_{Z^*} \hat{Z}_i^* + u_{it}, i = 1, ..., N, \quad (18)$$

where λ_β and λ_{Z^*} are coefficient estimates of the market price of beta risk (associated with sensitivity to mean market returns) and the market price of zeta risk (associated with sensitivity to cross-sectional market volatility or return dispersion) in percent terms, respectively, and the other notation is as before. According to Ferson (2019), estimated risk premiums λ_β and λ_{Z^*} approximate mimicking portfolio returns that are long stocks with higher betas or zetas and short stocks with lower betas or zetas. As observed by Cochrane (2005, pp. 250–51), t-statistics associated with estimated factor prices $\hat{\lambda}_k$ using the monthly rolling approach are corrected for cross-sectional correlation of residual errors (and therefore, are similar to Shanken (1992) corrected OLS standard errors).

It should be noted that beta loadings ($\hat{\beta}_i$) are time invariant for the most part with similar values, using daily or monthly returns. The reason for this invariance is that they are benchmarked to one corresponding to the beta risk of the average market return of all assets. By contrast, KLH noted that zeta risk loadings ($\hat{Z}_i^*$) are time variant (i.e., the holding period can affect their estimated values) due to not being benchmarked to one. By way of interpretation, the estimated market price λ_{Z^*} related to the return dispersion measures the risk premium per unit zeta risk. Given that time-series regression (15) is used to estimate risk parameters with daily returns, and the cross-sectional regression Equation (18) uses one-month-ahead excess returns as the dependent variable, $\hat{Z}_i^*$ can be rescaled from a daily to monthly basis as follows:

$$R_{i,T+1} - R_{fT+1} = \lambda_0 + \lambda_\beta \hat{\beta}_i + \lambda_{Z^*} \hat{Z}_i^* N_{T+1} + u_{it}, i = 1, ..., N, \quad (19)$$

where N_{T+1} is the number of trading days in month $T+1$ (i.e., 21 days), $Z_i^* N_{T+1}$ is the monthly estimated zeta risk, and λ_{Z^*} is the monthly risk premium associated with zeta risk. This rescaling does not change the risk premium $\hat{\lambda}_{Z^*}$ per unit zeta risk.[16]

Another important statistic in the cross-sectional regressions is the R^2 estimate. Following Jagannathan and Wang (1996) and Lettau and Ludvigson (2001, footnote 17, p. 1254), we compute this goodness-of-fit measure by using the R^2 statistic from a single regression approach. Using the 1928 to 1964 (1965 to 2020) subperiod, we obtain 444 (672) monthly estimates of $\hat{\lambda}_k$ for each test asset portfolio as we roll forward month by month to the end of the analysis subperiod. We also have the same number of one-month-ahead realized excess returns for each portfolio. After taking the averages of the $\hat{\lambda}_k$s and realized excess returns for each portfolio, the average realized excess returns for the n portfolios are regressed on the average $\hat{\lambda}_k$s to obtain an estimate of R^2.

The above discussion of R^2 estimation leads to a second cross-sectional test. In this test, we compute the one-month-ahead average realized excess returns for the n portfolios as before. Additionally, we compute the one-month-ahead average fitted excess returns for each portfolio. To do this, for each portfolio, the empirical ZCAPM is estimated, and β_i and Z_i^* risk parameters are retained. In the next month $T+1$, these risk parameters are multiplied by estimated factor prices of risk, or λ_ks to compute the fitted excess return for each portfolio. Rolling forward month by month to the end of the subperiod, a series of fitted excess returns are available to compute the average fitted excess return. Finally, plots of average realized excess returns (x-axis) and average fitted excess returns (y-axis) for the n portfolios are created. If the model works perfectly, all points will lie on a 45-degree line from the origin.

4. Cross-Sectional Regression Results

Tables 2 and 3 report the results for the out-of-sample Fama and MacBeth cross-sectional regression tests in subperiods 1928 to 1964 and 1965 to 2020, respectively.

Table 2. Out-of-sample Fama–MacBeth cross-sectional regressions for U.S. stocks: January 1928 to December 1964.

Panel A: 25 Size and BM Sorted Portfolios							
Model	$\hat{\lambda}_0$	$\hat{\lambda}_\beta$	$\hat{\lambda}_{Z^*}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	R^2
CAPM	2.57	−0.73					0.00
	(4.89)	(−2.12)					
Three-factor	1.86	−1.05		1.37	1.24		0.79
	(4.71)	(−2.65)		(5.44)	(5.19)		
Four-factor	1.82	−1.12		1.25	1.37	−1.44	0.92
	(5.20)	(−2.88)		(5.07)	(5.25)	(−2.84)	
ZCAPM	1.42	−0.18	0.92				0.98
	(3.80)	(−0.54)	(6.19)				
Panel B: 25 Size and Momentum Sorted Portfolios							
Model	$\hat{\lambda}_0$	$\hat{\lambda}_\beta$	$\hat{\lambda}_{Z^*}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	R^2
CAPM	4.64	−2.68					0.07
	(4.60)	(−3.38)					
Three-factor	3.47	−3.27		1.57	3.28		0.85
	(5.07)	(−4.86)		(4.62)	(3.68)		
Four-factor	4.46	−4.17		1.44	3.22	0.22	0.94
	(5.89)	(−5.68)		(4.32)	(3.93)	(0.76)	
ZCAPM	1.73	−0.39	1.29				0.98
	(3.44)	(−0.93)	(10.22)				
Panel C: 90 Total Portfolios							
Model	$\hat{\lambda}_0$	$\hat{\lambda}_\beta$	$\hat{\lambda}_{Z^*}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	R^2
CAPM	2.86	−0.95					0.00
	(5.62)	(−2.88)					
Three-factor	1.92	−1.27		1.65	1.33		0.59
	(5.45)	(−3.70)		(6.80)	(4.48)		
Four-factor	1.97	−1.33		1.57	1.41	−0.05	0.59
	(5.36)	(−3.73)		(6.62)	(4.83)	(−0.20)	
ZCAPM	1.65	−0.28	1.00				0.95
	(5.18)	(−1.06)	(9.53)				

The portfolio *MKT* in the ZCAPM is the value-weighted mean market portfolio rather than a proxy for the market portfolio *M* as in the CAPM.

Table 3. Out-of-sample Fama–MacBeth cross-sectional regressions for U.S. stocks: January 1965 to December 2020.

Panel A: 25 Size and BM Sorted Portfolios							
Model	$\hat{\lambda}_0$	$\hat{\lambda}_\beta$	$\hat{\lambda}_{Z^*}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	R^2
CCAPM	2.18	−1.41					0.72
	(8.82)	(−6.11)					
Three-factor	2.45	−1.75		0.36	0.26		0.83
	(10.78)	(−7.93)		(2.46)	(1.93)		
Four-factor	2.45	−1.77		0.36	0.23	−0.43	0.86
	(10.43)	(−7.79)		(2.50)	(1.74)	(−1.47)	
ZCAPM	1.48	−0.77	0.47				0.98
	(5.82)	(−3.25)	(4.00)				
Panel B: 25 Size and Momentum Sorted Portfolios							
Model	$\hat{\lambda}_0$	$\hat{\lambda}_\beta$	$\hat{\lambda}_{Z^*}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	R^2
CCAPM	1.53	−0.78					0.38
	(6.90)	(−3.52)					
Three-factor	1.81	−1.23		0.54	0.03		0.76
	(7.77)	(−5.41)		(3.75)	(0.11)		
Four-factor	2.06	−1.42		0.52	−0.20	0.47	0.80
	(9.05)	(−6.70)		(3.61)	(−0.91)	(2.48)	
ZCAPM	0.87	−0.19	0.69				0.93
	(4.06)	(−0.85)	(6.07)				
Panel C: 90 Total Portfolios							
Model	$\hat{\lambda}_0$	$\hat{\lambda}_\beta$	$\hat{\lambda}_{Z^*}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{MOM}$	R^2
CAPM	1.82	−0.94					0.26
	(9.06)	(−4.45)					
Three-factor	1.52	−0.93		0.57	0.26		0.48
	(9.42)	(−5.04)		(3.72)	(1.85)		
Four-factor	1.58	−0.97		0.53	0.23	0.26	0.54
	(9.91)	(−5.37)		(3.37)	(1.82)	(1.39)	
ZCAPM	1.22	−0.43	0.63				0.87
	(6.81)	(−2.29)	(8.50)				

The portfolio *MKT* in the ZCAPM is the value-weighted mean market portfolio rather than a proxy for the market portfolio *M* as in the CAPM.

4.1. Subperiod 1928 to 1964

The cross-sectional regression results for the 1928 to 1964 subperiod based on three different test asset portfolios are shown in Panels A to C in Table 2. In Panel A for the 25 size and BM portfolios, we see that the CAPM performed the worst with virtually no explanatory power at $R^2 = 0$ and a marginally significant negative market price for beta $\hat{\lambda}_\beta = -0.73$ ($t = -2.12$). Since beta risk should be positively priced in the CAPM, our results confirm those of Fama and French (1992, 1993, 1995) and many others that do not support the CAPM. By contrast, the best performing models are the four-factor model and ZCAPM with estimated R^2 values of 92 percent and 98 percent, respectively. In assessing the relative performance of different models, it is important to recognize that the three- and four-factor models use endogenous test asset portfolios; that is, size, BM, and momentum test asset portfolios are constructed from the same firm-level characteristics as the respective factors. Notably, these firm-level characteristics are exogenous to the ZCAPM, which makes the almost perfect goodness-of-fit of the ZCAPM quite remarkable.

In terms of the significance of factor loadings, again the ZCAPM outperforms the other models. The $\hat{\lambda}_{Z^*}$ market price of zeta risk loadings is 1.29 percent per month with a very high t-statistic of 6.19. The t-statistics for $\hat{\lambda}_{SMB}$ and $\hat{\lambda}_{HML}$ are very high also in the range of 5.07 to 5.44. As discussed earlier, Harvey et al. (2016) and Chordia et al. (2020) recommended that asset pricing factors have t statistics greater than 3.0 to avoid false discoveries. Our findings suggest that the size and value factors of Fama and French are not false discoveries. This inference holds for the return dispersion factor in the ZCAPM also. The estimated market price of momentum risk loadings $\hat{\lambda}_{MOM}$ with $t = -2.98$, which is borderline significant relative to the threshold hurdle rate. However, its market price has a negative sign that is difficult to explain (i.e., higher risk should imply higher risk premiums). Even so, adding momentum to the three-factor model noticeably boosts its goodness-of-fit from 79 percent to 92 percent in the four-factor model.

The results in Panel B for the 25 size and momentum portfolios are similar to those in Panel A. Again, the ZCAPM has almost perfect goodness-of-fit at a 98 percent R^2 estimate, and $\hat{\lambda}_{Z^*}$ has the highest t statistic of all factors tested. Regarding the latter, the t statistic equals 10.22, which is extremely high. No previous studies to our knowledge have reported a t statistic this high in cross-sectional regression tests. As before, the four-factor model outperforms the three-factor model, and the CAPM does the worst in terms of very low goodness-of-fit.

Lastly, Panel C contains the results for 90 combined portfolios including 40 industry portfolios. These results mitigate endogeneity problems by incorporating exogenous industry test assets. Upon doing so, the three- and four-factor models' performance diminishes substantially compared to Panels A and B. Now their R^2 values only reach 59 percent, which is far below that of the ZCAPM with near perfect goodness-of-fit at 95 percent. While the t statistics for $\hat{\lambda}_{SMB}$ and $\hat{\lambda}_{HML}$ are high in the range of 4.48 to 6.80, they are well below that of the market price of zeta risk $\hat{\lambda}_{Z^*}$ at 9.53.

Another finding in Panels A to C of Table 2 is that the intercept term $\hat{\lambda}_0$ is somewhat lower for the ZCAPM compared to the other models. This pattern is most clearly seen in Panel B, wherein $\hat{\lambda}_0 = 1.73$ percent per month ($t = 3.44$) for the ZCAPM, compared to estimates in the range of 3.47 percent to 4.64 percent for the other models. This lower mispricing error further supports the ZCAPM.

In sum, the ZCAPM outperforms the popular three- and four-factor models, even when endogenous test assets are used (which are exogenous to the ZCAPM factors). Consistent with earlier studies, the CAPM performs poorly in cross-sectional tests. When exogenous industry portfolios are added to the test assets, the ZCAPM outperforms other models by a large margin. The latter results are the most reliable and highlight the dominance of the ZCAPM, compared to often-used multifactor models.

4.2. Subperiod 1965 to 2020

In Table 3, we repeat the cross-sectional analyses in Table 2 for the subperiod 1965 to 2020. The results are similar to those in the earlier subperiod, with the exception that the three- and four-factor models' performance diminishes noticeably. For example, these models now have R^2 values of 83 percent and 86 percent, compared to 72 percent and 92 percent in Panel A of Table 2. None of the t statistics for these multifactor models breaks the recommended 3.0 threshold—namely, they range from 1.74 to 2.50. In addition, the market price of momentum loadings $\hat{\lambda}_{MOM}$ is insignificant with a negative sign. By contrast, the ZCAPM has a near perfect R^2 value of 98 percent, and the market price of zeta risk loadings $\hat{\lambda}_{Z^*}$ is highly significant with $t = 4.00$. The CAPM performs better in this subperiod with R^2 = 72 percent, but the market price of beta risk is again significantly negative at $\hat{\lambda}_{\beta} = -1.41$ percent ($t = -6.11$).

In Panel B, the results using 25 size and momentum portfolios are little changed. The ZCAPM continues to outperform the multifactor models, even with endogenous assets (that are exogenous to the ZCAPM). Now the market prices of size loadings $\hat{\lambda}_{SMB}$ have t statistics exceeding the 3.0 threshold at 3.75 and 3.61 in the three- and four-factor models. The market price of momentum loadings $\hat{\lambda}_{MOM}$ is positive and significant at $t = 2.48$, but falls below the 3.0 threshold. Recall that it was negative and significant in the earlier subperiod for the 25 size and BM portfolios. So here, we see some instability in the momentum factor results over time. By comparison, the ZCAPM's $\hat{\lambda}_{Z^*} = 0.47$ percent has $t = 6.07$, which is much higher than the size loadings. As in Panel A, the goodness-of-fit of the ZCAPM surpasses the other models by a larger margin than in the earlier subperiod. These results suggest that the earlier subperiod has greater endogeneity problems than the later subperiod due to smaller sample sizes as you go back in time before 1965. Hence, the earlier subperiod findings underscore the endogeneity problem in the three- and four factor models.

Finally, Panel C provides the results for the 90 combined portfolios with industry portfolios included. As before, the inclusion of exogenous test assets reduces the performance of the multifactor models. The ZCAPM has $R^2 = 87$ percent and $t = 8.50$ with respect to the market price of zeta risk $\hat{\lambda}_{Z^*} = 0.63$ percent, which exceeds the R^2 values of the three- and four-factor models at 0.48 percent and 0.54 percent, respectively, and t-statistics of 3.72 and 3.37 for $\hat{\lambda}_{SMB}$ at 0.57 percent and 0.53 percent. The market prices of value and momentum loadings are insignificant at the 5 percent level in these test assets. The inability of these factors to consistently be significant from over time and across test assets suggests that they are false discoveries. Only the size factor continues to pass the 3.0 threshold, even when exogenous assets are included in the test assets.

We should mention that previously cited studies that incorporated a return dispersion factor in an OLS time-series regression model obtain much weaker and ambiguous findings than the EM ZCAPM regression model with a dummy latent variable. Unlike the present study, Verousis and Voukelators (2015) found that return dispersion loadings are negatively priced. Other studies by Jiang (2010), Demirer and Jategaonkar (2013), Garcia et al. (2014), and Chichernea et al. (2015) reported positive prices of return dispersion loadings, but the significance levels did not consistently exceed the 3.0 threshold. In this regard, because they used in-sample cross-sectional tests rather than out-of-sample tests as in the present study, their results cannot be directly compared to our results.

Another noteworthy finding is that, as in the earlier subperiod, the intercept terms $\hat{\lambda}_0$ for the ZCAPM are lower than those for the other models in Panels A to C in Table 3. We infer that this is likely due to the better goodness-of-fit of the ZCAPM compared to the other models.

In sum, the ZCAPM outperforms the CAPM and commonly used multifactor models in terms of both goodness-of-fit and significance of return dispersion factor loadings. Differences in performance are greater, using exogenous industry assets, which call attention to the endogeneity problem in using tests assets sorted on firm-level characteristics that are used to construct long-short, zero-investment factors. We infer that our results corroborate those in KLH—that is, the ZCAPM consistently dominates multifactor models in out-of-sample cross-sectional regression tests. According to KLH, the return dispersion factor outperforms multifactors due to the fact that the latter are actually rough return dispersion measures that capture different slices within the total return dispersion. The size factor is long small stocks' returns and short big stocks' returns, and so captures a portion of the total return dispersion. Sometimes, multifactors switch from positive to negative market prices of risk in cross-sectional tests; for example, the market prices of momentum factor loadings switch from negative in the earlier subperiod to negative in the later subperiod in Tables 2 and 3, respectively. The reason for this erratic pricing behavior is that momentum was capturing negative zeta risk in the earlier subperiod and positive zeta risk in the later subperiod. Surely, momentum is a return dispersion measure, as it is defined as past winner stocks' turns minus past loser stocks' returns. These multifactors can shift around within the total return dispersion of stocks over time and at times become insignificant, which is what we find in our results in Tables 2 and 3. Because multifactors are proxies for different slices of return dispersion, they are related to the ZCAPM. As the theoretical ZCAPM posits, return dispersion is needed to span the return and risk dimensions of the mean–variance Markowitz investment parabola. To locate efficient and orthogonal inefficient portfolios per Black's zero-beta CAPM, the return dispersion is a critical asset pricing factor that is needed to augment the mean market return factor.

4.3. Cross-Sectional Fitted and Realized Excess Returns

Fama and MacBeth (1973, p. 613) observed that: "As a normative theory the model only has content if there is some relation between future returns and estimates of risk that can be made on the basis of current information". Following this logic, Lettau and Ludvigson (2001) and other researchers typically generated graphs of cross-sectional fitted excess returns and realized excess returns for different test asset portfolios. We discussed details for computing these out-of-sample returns in the previous section. We next display in Figures 3–10 illustrations of the relation between actual and fitted excess returns.

In Figures 3–6 corresponding to the earlier subperiod, we provide the cross-sectional results for the CAPM, three-factor model, four-factor model, and ZCAPM. In Figure 3, the CAPM demonstrates no relation between past beta estimates (used to compute future fitted excess returns) and future realized excess returns. The three- and four-factor models in Figures 4 and 5, respectively, do a much better job. Even so, they have difficulties with portfolios with average realized excess returns greater than about 4 percent per month. For these higher risk portfolios, fitted excess returns tend to underestimate the realized excess returns. We infer that some portion of risk is left out of these models which explains this downward bias. By contrast, the ZCAPM in Figure 6 correctly prices these higher risk portfolios. Hence, the ZCAPM more completely measures risk than the multifactor models. Additionally, and of major importance as a normative theory, based on the ZCAPM, portfolios fall fairly close to the 45-degree line from the origin for fitted and realized excess returns.

Turning to Figures 7–10 related to the later subperiod, we find similar patterns in fitted versus realized excess returns. In Figure 7, the CAPM does better than in the previous subperiod but a markedly flat relation is obvious that fails to capture a linear relation between risk and return. The three- and four-factor models in Figures 8 and 9, respectively, do a much better job than the CAPM but again have difficulties with underestimating the fitted excess returns of high return (risk) portfolios. In Figure 10, the ZCAPM clearly demonstrates a closer fit between fitted and realized excess returns than the other models in the cross section of average stock returns. Additionally, the ZCAPM better prices high return (risk) portfolios.

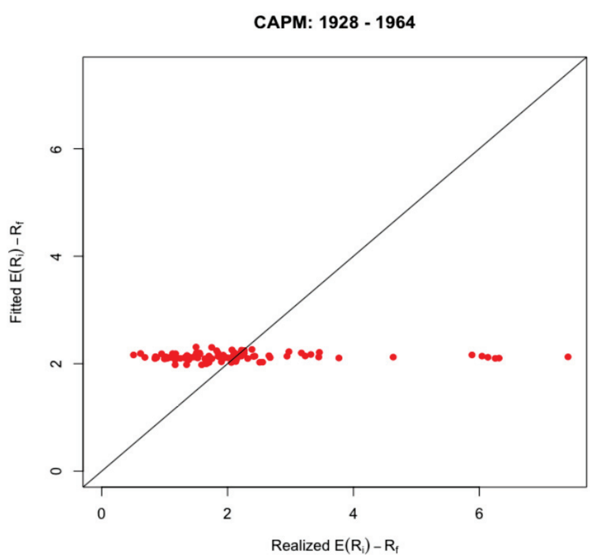

Figure 3. Out-of-sample cross-sectional relationships for the CAPM between average one-month-ahead fitted (predicted) excess returns in percent (y-axis) and average one-month-ahead realized excess returns in percent (x-axis): January 1928 to December 1964.

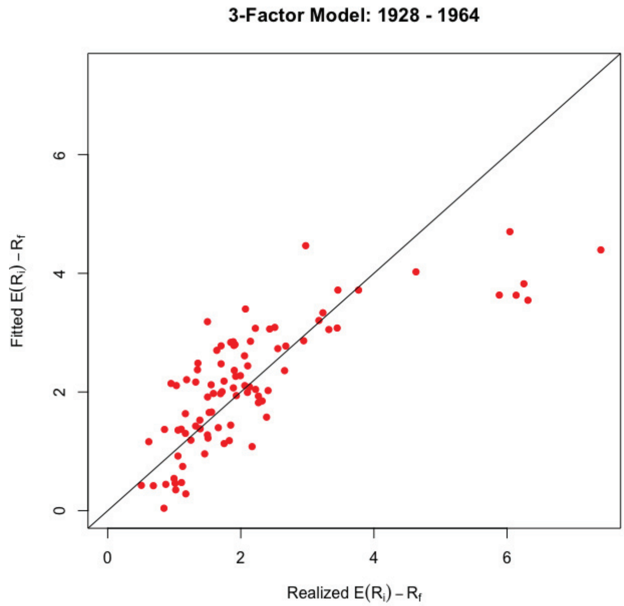

Figure 4. Out-of-sample cross-sectional relationships for the Fama and French three-factor model between average one-month-ahead fitted (predicted) excess returns in percent (*y*-axis) and average one-month-ahead realized excess returns in percent (*x*-axis): January 1928 to December 1964.

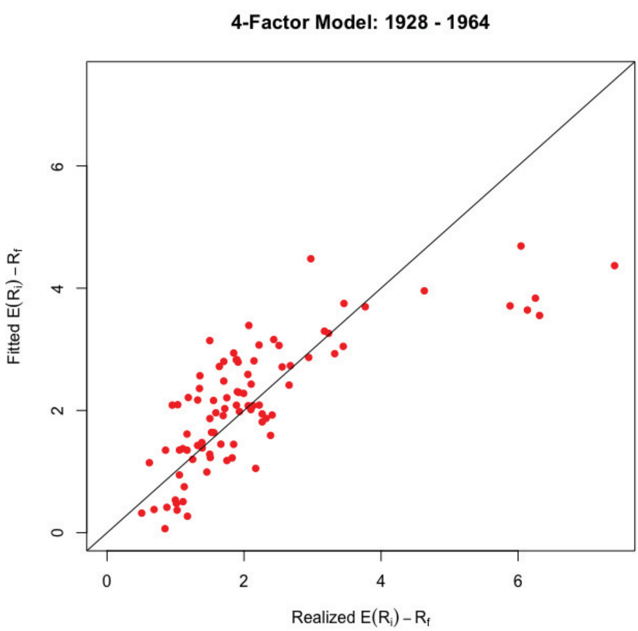

Figure 5. Out-of-sample cross-sectional relationships for the Fama and French four-factor model between average one-month-ahead fitted (predicted) excess returns in percent (*y*-axis) and average one-month-ahead realized excess returns in percent (*x*-axis): January 1928 to December 1964.

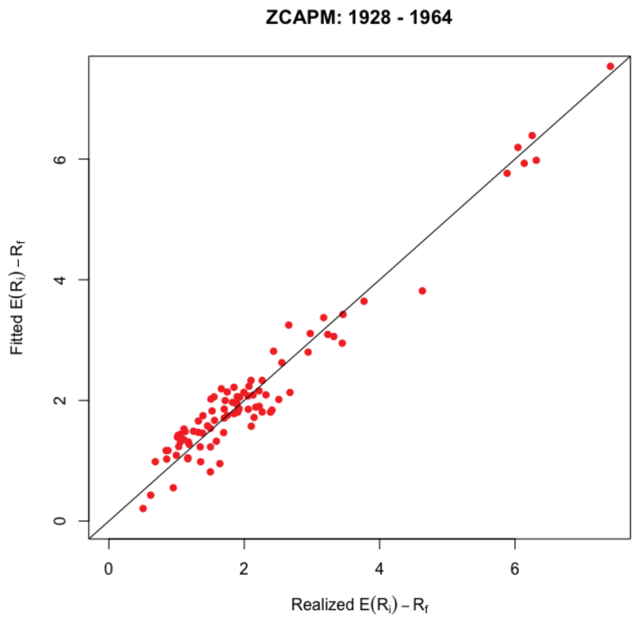

Figure 6. Out-of-sample cross-sectional relationships for the ZCAPM between average one-month-ahead fitted (predicted) excess returns in percent (*y*-axis) and average one-month-ahead realized excess returns in percent (*x*-axis): January 1928 to December 1964.

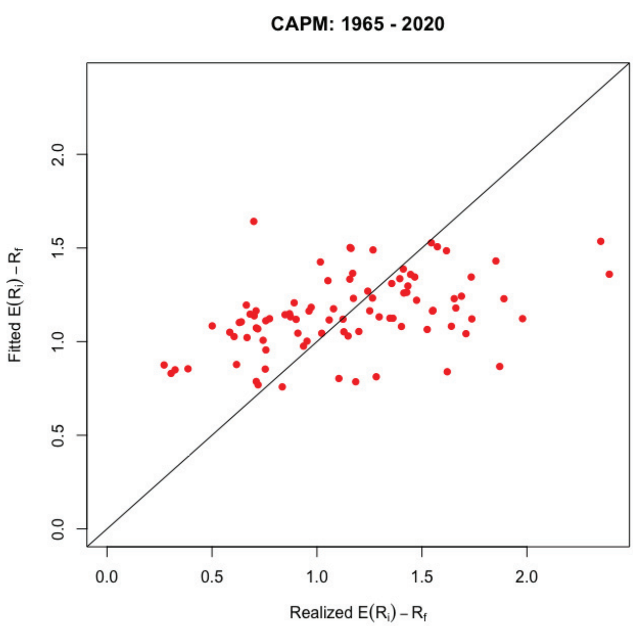

Figure 7. Out-of-sample cross-sectional relationships for the CAPM between average one-month-ahead fitted (predicted) excess returns in percent (*y*-axis) and average one-month-ahead realized excess returns in percent (*x*-axis): January 1965 to December 2020.

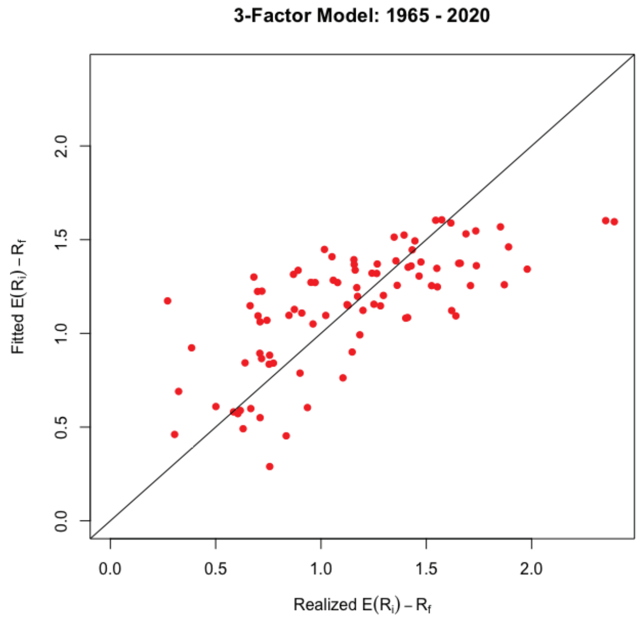

Figure 8. Out-of-sample cross-sectional relationships for the Fama and French three-factor model between average one-month-ahead fitted (predicted) excess returns in percent (*y*-axis) and average one-month-ahead realized excess returns in percent (*x*-axis): January 1965 to December 2020.

Figure 9. Out-of-sample cross-sectional relationships for the Fama and French four-factor model between average one-month-ahead fitted (predicted) excess returns in percent (*y*-axis) and average one-month-ahead realized excess returns in percent (*x*-axis): January 1965 to December 2020.

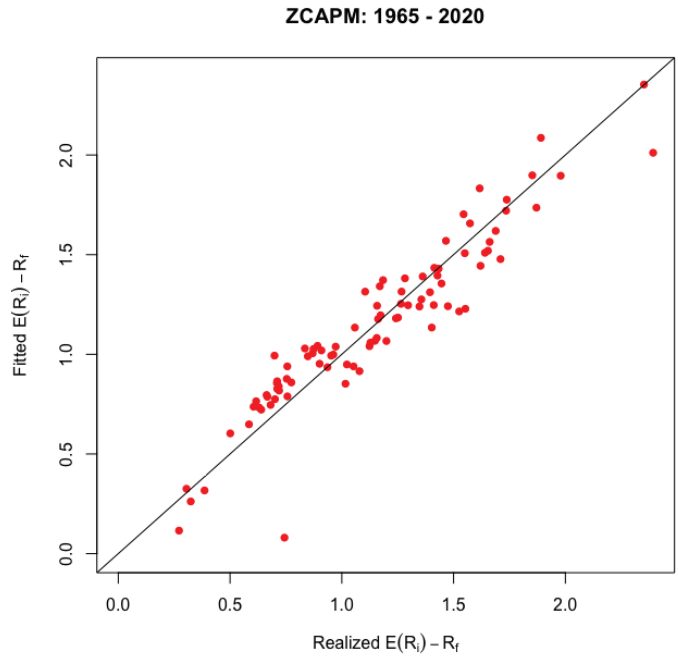

Figure 10. Out-of-sample cross-sectional relationships for the ZCAPM between average one-month-ahead fitted (predicted) excess returns in percent (y-axis) and average one-month-ahead realized excess returns in percent (x-axis): January 1965 to December 2020.

In sum, confirming Fama and French (1992, 1993, 1995) and others, graphical results for the CAPM suggest that there is no relation between the one-month-ahead fitted excess returns based on market beta risk and realized excess returns in the cross section of the average stock returns. Fama and French's three-factor model noticeably boosts the goodness-of-fit compared to the CAPM. Additionally, Carharts' four-factor model further improves the goodness-of-fit, especially in the earlier subperiod of 1928 to 1964. When industry portfolios are added to the test assets, the performance of the three- and four-factor models decreases considerably, whereas the ZCAPM continues to perform quite well. Graphs of average fitted and realized excess returns show that test asset portfolio returns fall closer to the 45-degree line from the origin for the ZCAPM, compared to the multifactor models. Unlike the ZCAPM, the latter multifactor models have difficulty in pricing higher return (risk) portfolios. Thus, the ZCAPM outperforms other models in cross-sectional analyses of stock returns.

5. Conclusions

This study extended the previous work by Kolari et al. (2021) (KLH) on tests of a new asset pricing model derived as a special case of Black's (1972) zero-beta CAPM, dubbed the *ZCAPM*. KLH investigated U.S. stock returns in the sample period 1965 to 2018. After reviewing the theoretical and empirical versions of the ZCAPM, we expanded their analyses by taking into account the earlier subperiod 1928 to 1964 as well as the later subperiod of 1965 to 2020. Standard out-of-sample Fama and MacBeth (1973) cross-sectional regression analyses were applied to a variety of test asset portfolios, including 25 size and BM portfolios, 25 size and momentum portfolios, and 90 combined portfolios with 40 industry portfolios. We benchmarked the ZCAPM results against the CAPM with a single market factor, the Fama and French (1992, 1993, 1995) three-factor model augmented with size and value factors, and the Carhart (1997) four-factor model augmented with a momentum factor.

Our results corroborate those in KLH that the ZCAPM consistently dominates multifactor models, especially when using exogenous industry portfolios. Our CAPM results are similar to previous authors that find little or no support for the hypothesized positive relation between beta and average returns. The three- and four-factor models did much better than the CAPM but primarily showed strength using endogenous test asset portfolios based on size and momentum characteristics that are contained in the size and momentum factors. Even so, the ZCAPM outperformed these popular multifactor models. Interestingly, multifactor models did better in the earlier subperiod in all likelihood due to the smaller sample sizes of stocks relative to the later subperiod; that is, smaller sample sizes exacerbate the endogeneity problem in cross-sectional tests. When using exogenous industry portfolios, the multifactor models' performance declined substantially, whereas the ZCAPM continued to perform quite well in both earlier and later subperiods.

We conclude that, similar to the findings of KLH, the ZCAPM consistently outperformed multifactor models. A key reason for this outperformance is that the cross-sectional standard deviation of all stock's returns more comprehensively captures the return dispersion than the selected multifactors that are themselves rough measures of the return dispersion. While popular multifactors show significance in our tests, and can surpass the recommended 3.0 *t*-statistic thresholds in terms of the market prices of factor loadings in some test asset portfolios, their results tend to be inconsistent across subperiods and test asset portfolios. By contrast, *t* statistics associated with the market price of return dispersion loadings in the ZCAPM always exceed 3.0 in different subperiods and test asset portfolios. Additionally, near perfect goodness-of-fit was achieved by the ZCAPM for portfolios sorted on firm-level characteristics, even though these characteristics are exogenous to the mean market return and cross-sectional return dispersion factors of the ZCAPM.

Cochrane (2011, p. 1061) observed that, "...the world would be much simpler if betas on only a few factors, important in the covariance matrix of returns, accounted for a larger number of mean characteristics". The ZCAPM embodies a parsimonious two-factor model with mean market return and return dispersion factors. Regarding the latter return dispersion factor, the ZCAPM takes into account long/short, zero-investment factors based on firm-level characteristics that themselves are rough measures of total return dispersion. Future research is recommended on different countries[17] as well as other asset classes, including bonds, commodities, real estate, etc. Moreover, further work on applications to other areas of finance is recommended, such as investment analysis, the cost of equity, event studies, etc.

Author Contributions: J.W.K. conceptualized the main research problem, coordinated research activities, and worked on the final manuscript writing and editing. W.L. derived the theoretical model and helped in empirical tests of the model. J.Z.H. developed the empirical model. H.L. ran the empirical analyses and provided programming assistance. All authors have read and agreed to the published version of the manuscript.

Funding: The research received external funding from the Teachers Retirement System of Texas.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Data is available upon request from the authors.

Acknowledgments: The authors gratefully acknowledge helpful comments from conference participants at the Financial Management Association International 2012 in Atlanta, Georgia, Multinational Finance Society 2012 in Kracaw, Poland, Midwest Finance Association 2012 in New Orleans, Louisiana, and the Southern Finance Association 2020 (virtual). We received the Best Paper Award in Investments at the former conference. Financial and professional investment support from the Teachers Retirement System of Texas is appreciated. A number of individuals at academic institutions and investment firms have shared useful comments and suggestions with us over the years, including Ali Anari, Saurabh Biswas, Yong Chen, Brett Cornwell, Wayne Ferson, Wesley Gray, Klaus Grobys, Britt Harris, Hagen Kim, Johan Knif, Qi Li, Liqian Ren, Seppo Pynnönen, William Smith, Mikhail Sokolov, Sorin Sorescu, Ahmet Tuncez, Ivo Welch, Mark Westerfield, and David Veal. Yao Han provided earlier assistance with cross-sectional statistical tests. Zhao Tang Luo conducted replication exercises. Mohsin Sadaqat assisted with data collection.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1 Numerous authors link return dispersion to economic fundamentals, including the business cycle, economic uncertainty, and macroeconomic shocks, including Loungani et al. (1990), Christie and Huang (1994), Bekaert and Harvey (1997, 2000), Connolly and Stivers (2003), Stivers (2003), Pastor and Veronesi (2009), Angelidis et al. (2015), and others.

2 See seminal work by Dempster et al. (1977) on the development of EM regression as well as applications in other areas of finance by Harvey and Liu (2016) and Chen et al. (2017). Wikipedia provides an excellent overview of EM regression and further citations to statistics literature.

3 These studies compute a variety of market volatility factors, including the time-series volatility index (VIX) of the Chicago Board of Options Exchange (CBOE), time-series variance of market returns, and volatility-of-volatility metrics. See (Ferson 2019, chp. 34) for an excellent discussion of studies using time-series volatility factors in the asset pricing literature.

4 See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library (accessed on 1 September 2021)

5 See also their earlier work in Liu et al. (2012) and Liu (2013).

6 One proof uses random matrix mathematics, and the second proof is based on Markowitz mathematical methods.

7 This result was proved by KLH by means of two different geometric methods, including the Roll's (1980) well-known geometric approach.

8 By contrast, the riskless rate is constant and therefore nonrandom.

9 In early CAPM studies, such as that of Black et al. (1972), it was found that $E(\widetilde{R}_{ZM}) > R_f$, which implied that security market line (SML) had a higher intercept and lower slope than the theoretical CAPM.

10 To do this, KLH use the latent variable approach in conditional asset pricing (see Gibbons and Ferson 1985; Ferson and Locke 1998).

11 The following conditions hold: (1) assuming all funds are invested in either I^* or ZI^*, then $\beta_{I^*,a} = \beta_{ZI^*,a} = 1$ and $Z^*_{I^*,a} = 1$ or $Z^*_{ZI^*,a} = -1$, respectively; (2) assuming no riskless asset, Equation (11) reduces to Equation (8) (i.e., $\beta_{i,a} \equiv w_{I^*} + w_{ZI^*} = 1$); and (3) assuming the restriction $w_f > 0$ (i.e., no borrowing at the riskless rate is allowed), then $\beta_{i,a} < 1$.

12 The market portfolio M lies on the efficient frontier at the tangent point of a ray from the riskless rate.

13 This diagram is based on Figure 3.3 in Kolari et al. (2021, p. 68).

14 More specifically, KLH defined $\mathcal{T}_+ = \{t : 1 \le t \le T, D_{it} = +1\}$ and $\mathcal{T}_- = \{t : 1 \le t \le T, D_{it} = -1\}$ as sets of time indices associated with positive and negative signs of the signal variable. As such, the empirical ZCAPM Equation (12) becomes a two equation model:

$$R_{it} - R_{ft} = \beta_{i,a}(R_{at} - R_{ft}) + Z_{i,a}\sigma_{at} + u_{it}, \;\; t \in \mathcal{T}_+$$
$$R_{it} - R_{ft} = \beta_{i,a}(R_{at} - R_{ft}) - Z_{i,a}\sigma_{at} + u_{it}, \;\; t \in \mathcal{T}_-,$$

where first equation has probability p_i, and the second equation has probability $1 - p_i$.

15 As we will see, the goodness-of-fit of the empirical ZCAPM was exceptional, with estimated adjusted R^2 values as high as 98 percent in some tests. These results imply that the absence of a time-series α_i term in the empirical ZCAPM did not affect the cross-sectional results.

16 Without recaling Z_i^* to a monthly basis, estimates of λ_{Z^*} would be much larger and not comparable to λ_β estimates related to beta loadings.

17 See the working paper by Kolari et al. (2021) on international stock market tests of the ZCAPM in Canada, France, Germany, Japan, the United Kingdom, and the United States.

References

Adrian, Tobias, and Joshua Rosenberg. 2008. Stock returns and volatility: Pricing the short-run and long-run components of market risk. *Journal of Finance* 63: 2997–3030. [CrossRef]

Ang, Andrew, Joseph Chen, and Yuhang Xing. 2006a. Downside risk. *Review of Financial Studies* 19: 1191–239. [CrossRef]

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2006b. The cross-section of volatility and expected returns. *Journal of Finance* 61: 259–99. [CrossRef]

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2009. High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics* 91: 1–23. [CrossRef]

Angelidis, Tinotheos, Athanasios Sakkas, and Nikolaos Tessaromatis. 2015. Stock market dispersion, the business cycle and expected factor returns. *Journal of Banking and Finance* 59: 256–79. [CrossRef]

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron. 2014. *Volatility, the Macroeconomy, and Asset Prices*. Working Paper. Durham: Duke University; Philadelphia: University of Pennsylvania.
Bekaert, Geert, and Campbell Harvey. 1997. Emerging equity market volatility. *Journal of Financial Economics* 43: 29–77. [CrossRef]
Bekaert, Geert, and Campbell Harvey. 2000. Foreign speculators and emerging equity markets. *Journal of Finance* 55: 565–613. [CrossRef]
Bekaert, Geert, Robert J. Hodrick, and Xiaoyan Zhang. 2012. Aggregate idiosyncratic volatility. *Journal of Financial and Quantitative Analysis* 47: 1155–85. [CrossRef]
Black, Fischer. 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45: 444–54. [CrossRef]
Black, Fischer, Michael C. Jensen, and Myron Scholes. 1972. The Capital Asset Pricing Model: Some Empirical Tests. In *Studies in the Theory of Capital Markets*. Edited by Michael C. Jensen. New York: Praeger.
Bollerslev, Tim, Sophia Zhengzi Li, and Bingzhi Zhao. 2016. *Good Volatility, Bad Volatility, and the Cross-Section of Stock Returns*. Working Paper. Durham: Duke University; East Lansing: Michigan State University.
Carhart, Mark M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52: 57–82. [CrossRef]
Chang, Bo Young, Peter Christoffersen, and Kris Jacobs. 2013. Market skewness risk and the cross-section of stock returns. *Journal of Financial Economics* 107: 46–68. [CrossRef]
Chen, Te-Feng, Tarun Chordia, San-Lin Chung, and Ji-Chai Lin. 2021. Volatility-of-volatility risk in asset pricing. *Review of Asset Pricing Studies* 12: 289–335. [CrossRef]
Chen, Yong, Michael Cliff, and Haibei Zhao. 2017. Hedge funds: The good, the bad, and the lucky. *Journal of Financial and Quantitative Analysis* 52: 1081–109. [CrossRef]
Chichernea, Doina C., Anthony D. Holder, and Alex Petkevich. 2015. Does return dispersion explain the accrual and investment anomalies? *Journal of Accounting and Economics* 60: 133–48. [CrossRef]
Chordia, Tarun, Amit Goyal, and Alessio Saretto. 2020. Anomalies and false rejections. *Review of Financial Studies* 33: 2134–79. [CrossRef]
Christie, William Gary, and R. D. Huang. 1994. The changing functional relation between stock returns and dividend yields. *Journal of Empirical Finance* 1: 161–91. [CrossRef]
Cochrane, John H. 2005. *Asset Pricing: Revised Edition*. Princeton: Princeton University Press.
Cochrane, John H. 2011. Presidential address: Discount rates. *Journal of Finance* 56: 1047–108. [CrossRef]
Connolly, Robert, and Chris Stivers. 2003. Momentum and reversals in equity index returns during periods of abnormal turnover and return dispersion. *Journal of Finance* 58: 1521–56. [CrossRef]
Da, Zhi, and Ernst Schaumburg. 2011. *The Pricing of Volatility Risk across Asset Classes*. Working Paper. Notre Dame: University of Notre Dame; New York: Federal Reserve Bank of New York.
Daniel, Kent, and Sheridan Titman. 2012. Testing factor-model explanations of market anomalies. *Critical Finance Review* 1: 103–39. [CrossRef]
Demirer, Riza, and Shrikant P. Jategaonkar. 2013. The conditional relation between dispersion and return. *Review of Financial Economics* 22: 125–34. [CrossRef]
Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 39: 1–38.
Fama, Eugene F. 1968. Risk, return, and equilibrium: Some clarifying comments. *Journal of Finance* 23: 29–40. [CrossRef]
Fama, Eugene F., and James D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81: 607–36. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47: 427–65. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 1995. Size and book-to-market factors in earnings and returns. *Journal of Finance* 50: 131–56. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 1996. The CAPM is wanted, dead or alive. *Journal of Finance* 51: 1947–58. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 2004. The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives* 18: 25–46. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116: 1–22. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 2018. Choosing factors. *Journal of Financial Economics* 128: 234–52. [CrossRef]
Fama, Eugene F., and Kenneth R. French. 2020. Comparing cross-section and time-series factor models. *Review of Financial Studies* 33: 1892–926. [CrossRef]
Ferson, Wayne E. 2019. *Empirical Asset Pricing: Models and Methods*. Cambridge: The MIT Press.
Ferson, Wayne E., and Dennis H. Locke. 1998. Estimating the cost of capital through time: Analysis of sources of error. *Management Science* 44: 485–500. [CrossRef]
Ferson, Wayne E., Suresh K. Nallareddy, and Biqin Xie. 2013. The "out-of-sample" performance of long-run risk models. *Journal of Financial Economics* 107: 537–56. [CrossRef]
Garcia, René, Daniel Mantilla-Garcia, and Lionel Martellini. 2014. A model-free measure of aggregate idiosyncractic volatility and the prediction of market returns. *Journal of Financial and Quantitative Analysis* 49: 1133–65. [CrossRef]

Gibbons, Michael R., and Wayne E. Ferson. 1985. Testing asset pricing models with changing expectations and an unobservable market portfolio. *Journal of Financial Economics* 14: 217–36. [CrossRef]
Harvey, Campbell R., and Yan Liu. 2016. *Rethinking Performance Evaluation*. Working Paper No. 22134. Cambridge: National Bureau of Economic Research.
Harvey, Campbll R., Yan Liu, and Heqing Zhu. 2016. And the cross-section of expected returns. *Review of Financial Studies* 29: 5–68. [CrossRef]
Hou, Kewei, Chen Xue, and Lu Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28: 650–705. [CrossRef]
Jagannathan, Ravi, and Zhenyu Wang. 1996. The conditional CAPM and the cross-section of asset returns. *Journal of Finance* 51: 3–53. [CrossRef]
Jiang, Xiaoquan. 2010. Return dispersion and expected returns. *Financial Markets and Portfolio Management* 24: 107–35. [CrossRef]
Jones, P. N., and G. J. McLachlan. 1990. Algorithm AS 254: Maximum likelihood estimation from grouped and truncated data with finite normal mixture models. *Applied Statistics* 39: 273–82. [CrossRef]
Kolari, James W., Jianhua Z. Huang, Hilal Anwar Butt, and H. Liao. 2021. *International Tests of the ZCAPM Asset Pricingmodel*. Working Paper. College Station: Texas A&M University; Karachi: Institute of Business Administration.
Kolari, James W., Wei Liu, and Jianhua Z. Huang. 2021. *A New Model of Capital Asset Prices: Theory and Evidence*. Cham: Palgrave Macmillan.
Lettau, Martin, Matteo Maggiori, and Michael Weber. 2014. Conditional risk premia in currency markets and other asset classes. *Journal of Financial Economics* 114: 197–225. [CrossRef]
Lettau, Martin, and Sydney Ludvigson. 2001. Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy* 109: 1238–87. [CrossRef]
Lewellen, Jonathan, Stefan Nagel, and Jay Shanken. 2010. A skeptical appraisal of asset pricing tests. *Journal of Financial Economics* 96: 175–94. [CrossRef]
Lintner, John. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47: 13–37. [CrossRef]
Liu, Wei. 2013. A New Asset Pricing Model Based on the Zero-Beta CAPM: Theory and Evidence. Ph.D. thesis, Texas A&M University, College Station, TX, USA.
Liu, Wei, James W. Kolari, and Jianhua Z. Huang. 2012. A new asset pricing model based on the zero-beta CAPM. Paper presented at Annual Meetings of the Financial Management Association, Atlanta, GA, USA, October 17–20.
Loungani, Prakash, Mark Rush, and William Tave. 1990. Stock market dispersion and unemployment. *Journal of Monetary Economics* 25: 367–88. [CrossRef]
Markowitz, Harry M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.
McLachlan, Geoffrey J., and Thriyambakam Krishnan. 2008. *The EM Algorithm and Extensions*, 2nd ed. New York: John Wiley & Sons.
McLachlan, Geoffrey, and David Peel. 2000. *Finite Mixture Models*. New York: Wiley Interscience.
Mossin, Jan. 1966. Equilibrium in a capital asset market. *Econometrica* 34: 768–83. [CrossRef]
Pastor, Lubos, and Pietro Veronesi. 2009. Technological revolutions and stock prices. *American Economic Review* 99: 1451–83. [CrossRef]
Roll, Richard. 1977. A critique of the asset pricing theory's tests, part I: On past and potential future testability of the theory. *Journal of Financial Economics* 4: 129–76. [CrossRef]
Roll, Richard. 1980. Orthogonal portfolios. *Journal of Financial and Quantitative Analysis* 15: 1005–12. [CrossRef]
Shanken, Jay. 1992. On the estimation of beta pricing models. *Review of Financial Studies* 5: 1–34. [CrossRef]
Sharpe, William F. 1963. A simplified model for portfolio analysis. *Management Science* 9: 277–93. [CrossRef]
Sharpe, William F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19: 425–42.
Simin, Timothy. 2008. The poor performance of asset pricing models. *Journal of Financial and Quantitative Analysis* 43: 335–80. [CrossRef]
Stambaugh, Roger F., and Yu Yuan. 2017. Mispricing factors. *Review of Financial Studies* 30: 1270–315. [CrossRef]
Stivers, Chris. 2003. Firm-level return dispersion and the future volatility of aggregate stock market returns. *Journal of Financial Markets* 6: 389–411. [CrossRef]
Tobin, James. 1958. Liquidity preference as behavior toward risk. *Review of Economic Studies* 4: 65–86. [CrossRef]
Treynor, Jack L. 1961. Market Value, Time, and Risk. Unpublished work.
Treynor, Jack L. 1962. Toward a Theory of Market Value of Risky Assets. Unpublished work.
Verousis, Thanos, and Nikolaos Voukelators. 2015. Cross-sectional dispersion an expected returns. Paper presented at 5th Annual Conference of the Financial Engineering and Banking Society, Nantes, France, June 11–13.

Journal of
Risk and Financial Management

MDPI

Article

Mean Reversions in Major Developed Stock Markets: Recent Evidence from Unit Root, Spectral and Abnormal Return Studies

James Nguyen [1,*], Wei-Xuan Li [2] and Clara Chia-Sheng Chen [3]

[1] College of Business, Engineering and Technology, Texas A&M University Texarkana, Texarkana, TX 75503, USA
[2] School of Business, Stockton University, Galloway, NJ 08205, USA; liw1@stockton.edu
[3] School of Business, Madonna University, Livonia, MI 48150, USA; ccchen@madonna.edu
* Correspondence: james.nguyen@tamut.edu

Abstract: We revisited the issue of return predictability in three major developed markets (USA, UK and Japan) using a unique dataset from the Wharton Research Data Services database and a comprehensive set of traditional and recent statistical methods. We specifically employed a variety of traditional linear and nonlinear tests, latest multiple-break unit root tests and spectral analysis to test the efficient market hypothesis. Our results show that these stock markets generally are inefficient. We further explored whether the departure from market efficiency can be used to generate profitable trades and found that abnormal returns exist in all three markets. We found evidence of abnormal returns associated with the break dates identified in the models which are correlated with major historical events around the world. Our findings have important implications for investors and policymakers.

Keywords: efficient market hypothesis; unit root; spectral analysis; abnormal returns

Citation: Nguyen, James, Wei-Xuan Li, and Clara Chia-Sheng Chen. 2022. Mean Reversions in Major Developed Stock Markets: Recent Evidence from Unit Root, Spectral and Abnormal Return Studies. *Journal of Risk and Financial Management* 15: 162. https://doi.org/10.3390/jrfm15040162

Academic Editors: James W. Kolari and Seppo Pynnonen

Received: 19 January 2022
Accepted: 24 March 2022
Published: 1 April 2022

1. Introduction

The efficient market hypothesis (EMH), introduced by Eugene Fama in 1970, states that financial asset prices entirely reflect all available information, making it impossible for investors to beat the market. The EMH posits that stock prices are sensitive to every bit of information in the market and that movements of stock prices are unpredictable. Therefore, there should not be a momentous difference between the optimal forecast and actual stock prices, and the probability of making abnormal profits in the stock market is asymptotically zero. The theory has attracted many supporters as well as critics. Shiller (1981) documented that stock price variation should not be explained by fundamentals. Some of the results which show little alpha (risk-adjusted return) and no persistence were published by Carhart (1997), Lettau and Van Nieuwerburgh (2008), Fama and French (2010), Busse et al. (2010), Bertone et al. (2015), etc. Richard Thaler, a Nobel laureate in Economics in 2017, has helped reignite this debate. Thaler, one of the founders of "behavioral finance", has put the notion of the EMH in doubt and provided scientific explanations for the existence of irrational market behaviors. The empirical evidence is mixed, and the research community is "torn" between the EMH and behavioral finance camps (Verheyden et al. 2015).

A review of the EMH in developed markets reveals a widespread but not definitive consensus that markets tend toward efficiency, although there are periods of informational inefficiency and periods of speculative bubbles (behavioral finance) (e.g., French and Roll 1986; De Long et al. 1990). Carhart (1997) showed that the performance of mutual funds does not reflect superior stock-picking skills. Fama and French (2010) showed that few mutual funds produce returns sufficient to cover their costs. Busse et al. (2010) found that an investment manager's superior risk-adjusted returns are indistinguishable from zero.

Finally, Bertone et al. (2015) showed that the US market had become significantly more efficient even during very short-term intervals. More recently, Durusu-Ciftci et al. (2017) argued that the evidence for the EMH is mixed. One reason is that traditional tests ignore the presence of structural breaks, leading to invalid statistical inferences. Another potential issue is that traditional unit root tests only allow for one of two breaks in the data—a problem that can be overcome by some of the multiple-break unit root tests employed in our study.

Our research contributes to the literature by testing market efficiency in three major developed markets, the USA, the UK and Japan, for the first time—to our knowledge—using unique authoritative stock price indices provided by the WRDS. Our study also complements those that examine this topic for major stock markets, especially the study of the US, UK and Japanese stock markets by Urquhart and McGroarty (2016), Urquhart and Hudson (2013), Borges (2010) and Narayan and Smyth (2007). However, we employed a number of recent and powerful statistical tests to study this issue. Specifically, in this paper, we utilized highly regarded tests such as those used by Elliott et al. (1996), Ng and Perron (2001) and Brock et al. (1996, BDS) which had not been widely used in this line of research in addition to the highly popular traditional statistical tests such as the BDS and variance ratios. Further, we took advantage of the latest multiple-break unit root tests by Lumsdaine and Papell (1997, LP), Lee and Strazicich (2003, LS), Narayan and Popp (2010, NP) and Ender and Lee (2012, EL).[1] To increase the robustness of our results, we adopted recent spectral tests commonly found in the electrical engineering literature to further assess the EMH in the three developed markets in question. The final novelty of our study is the analysis of abnormal returns. Specifically, we explored whether the departure from market efficiency can be used to generate profitable trades.

By way of preview, we found that the three stock market indices in our study exhibit mean reversions. The rather surprising finding of market inefficiency (contradicting many prior findings of market efficiency for highly developed markets) may indicate more pronounced information asymmetry, limited competition and not fully developed financial and banking systems within these countries. The paper is organized as follows. Section 2 presents a brief review of the related studies. Section 3 discusses the data and the methodology. Section 4 discusses the empirical results. Section 5 provides some discussions of the findings. Finally, Section 6 concludes the study with some remarks.

2. Brief Literature Review

Numerous studies have explored the predictability of equity returns. Early studies documented that macroeconomic and financial variables are useful predictors of equity returns. For example, Fama and Schwert (1977) found a positive relationship between inflation and expected returns. Chen et al. (1986) showed that term spread, expected and unexpected inflation, industrial production and credit spread can explain the variations of equity returns in the US dividend yields (or dividend/price ratios) and also demonstrate the strong predictive power of equity returns (e.g., Shiller 1982; Bekaert and Hodrick 1992; Campbell and Hamao 1992; Solnik 1993; Campbell and Shiller 1988; Fama and French 1988; Ang and Bekaert 2007; Golez and Koudijs 2018). Interest rates, documented by Ang and Bekaert (2007) and Rapach et al. (2013), are reliable predictors of equity returns. Size and book-to-market ratio along with the market factor, presented by Fama and French (1992, 1993), are also important variables to predict equity returns. Examining firms' fundamentals and equity prices in the USA, Bhargava (2014) found that the following variables were important predictors: earnings per share, total assets, long-term debt, dividends per share and unemployment and interest rates.

Other studies incorporate liquidity to explore its relationship with equity returns (e.g., Amihud 2002; Bekaert et al. 2007). Amihud (2002) found a positive relationship between expected returns and contemporaneous unexpected illiquidity. Bekaert et al. (2007) documented that local market liquidity is an important determinant of equity returns in emerging markets. Another line of research examines the effect of investor sentiment on equity returns (e.g., Baker and Wurgler 2006, 2007). Baker and Wurgler (2006, 2007) documented a negative relationship between investor sentiment and subsequent equity returns. Nyberg and Pönkä (2016) documented the predictability of other equity market returns with the information from the US market.

A number of studies most related to our current research include the following studies. Golez and Koudijs (2018) combined the annual stock market data for the Netherlands/UK (1629–1812), the UK (1813–1870) and the USA (1871–2015) and showed that dividend yields are stationary and consistently forecast returns over both short and long horizons. Goetzmann et al. (2001) estimated a new index for the New York stock market between 1815 and 1925. They found little evidence for return predictability, but data limitations forced them to approximate dividends for the period before 1870. Mitra et al. (2017) examined the efficiency of 31 stock index series spanning 26 countries across the world. They found periods of departure from the martingale difference hypothesis among the stock index series around the world. The results are consistent with the adaptive market hypothesis whereby stock markets remain efficient most of the time but there are periods when markets become inefficient. Urquhart and Hudson (2013) also empirically investigated the adaptive market hypothesis for the US, UK and Japanese markets using very long-run data. Daily data were divided into five-yearly subsamples and subjected to linear and nonlinear tests to determine how the independence of stock returns had behaved over time. Their results from the linear autocorrelation, runs and variance ratio tests reveal that each market shows evidence of being an adaptive market, with returns going through periods of independence and dependence. However, results from nonlinear tests show strong dependence for every subsample in each market. Urquhart and McGroarty (2016) examined the adaptive market hypothesis in S&P 500, FTSE 100, NIKKEI 225 and EURO STOXX 50 by testing stock return predictability using daily data from January 1990 to May 2014. Their results show that there are periods of statistically significant return predictability, but also periods of no statistically significant predictability in stock returns. Narayan and Smyth (2007) showed evidence on the random walk hypothesis in G7 stock price indices using unit root tests which allow for one and two structural breaks in the trend. Evidence of mean reversion only exists for the stock price index of Japan. In short, no consensus has been reached.

3. Data and Methodologies

Our dataset was obtained from the Wharton Research Data Services (WRDS) country price index database. A major advantage of using this database is that all price series have a consistent data format. Our sample contained daily data for 23 major stock market indices in the USA, the UK and Japan. The indices were market capitalization-weighted, adjusted for stock splits and dividends. Compustat Global—Security Daily was used to construct the indices. The portfolio was rebalanced annually at the end of the last trading day of June for each country. Observations were removed if the market capitalization was not positive or if the exchange information was missing. For firms with multiple issues, the issue with the largest market capitalization was chosen. Additionally, a security had to be in the top 50% of the market capitalization of that country and traded at the stock exchanges located within the country in question. The currency of the security price had to be consistent with its ISO currency code. Lastly, only common ordinary shares were included in the indices. We extracted the country price indices for the USA, the UK and Japan in monthly frequency. In this study, our sample period for the USA spanned from January 1926 to December 2016, while the sample period for the UK and Japan started in December 1989, ending in December 2015. We employed a battery of tests typically used in the literature as well as a number of recent methods.[2]

Among the most important tests for market efficiency (i.e., random walk) are unit root tests. The weak-form efficient market hypothesis states that stock prices move in a random walk fashion, or that past prices cannot be used to predict future prices. The random walk model is commonly specified as follows:[3]

$$y_t = \mu + y_{t-1} + \varepsilon_t$$

where y_t is the log of price or stock index return in a number of studies, μ is the drift term and ε_t is the random disturbance term. To evaluate this hypothesis, we examined the returns on the country price indices and tested for independence of their return series. To test the random walk hypothesis it is necessary to examine the existence of a unit root in a return series. More specifically, we conducted traditional, highly regarded unit root tests and more recent single- as well as multiple-break unit root tests.[4] We first used the following highly popular unit root tests in this study: augmented Dickey–Fuller (ADF), Phillips–Perron, Elliott–Rothenberg–Stock and Ng–Perron tests. We then employed the Zivot and Andrews (1992) test as a single-break unit root test. For multiple-break unit root tests, we utilized the models developed by Lumsdaine and Papell (1997), Lee and Strazicich (2003), Narayan and Popp (2010), Ender and Lee (2012). Finally, we computed the abnormal returns for each price index using the structural break information found in those tests.

3.1. Unit Root Tests

ADF test (1981): This is our baseline test which evaluates if a series is stationary or random-walk (unit root), mainly for comparison purposes. The null hypothesis of a unit root is rejected if the test statistic is less (or more negative) than their associated critical values:

$$y_t = a_0 + \gamma_1 y_{t-1} + \theta t + \sum_{i=2}^{k} \beta_i y_{t-i} + \varepsilon_t$$

where y_t in our setting is the stock index return in month t. The null hypothesis is $\gamma_1 = 1$, a unit root in the return series. One problem with the ADF method is the selection of lag length (Schwert 1989). We, therefore, used Akaike's information criterion to select the optimal lag length (to ensure that the residual was white noise) to mitigate this issue. We also performed the Phillips and Perron (1988, PP) test, a more powerful test than the ADF test (Dickey and Fuller 1981), but with better size distortions.

ERS test (1996): This is basically a modified ADF test where Elliot, Rothenberg and Stock (ERS) show that their DF-GLS test has the power function close to the point optimal test which has better power properties. This test not only provides a higher power than the ADF and PP tests, but can also distinguish persistent stationary processes from nonstationary processes. The test has the same null hypothesis as the ADF test, and its results are interpreted similarly. To our knowledge, this was the first time the ERS tests were used to examine the market efficiency hypothesis, at least for our sample of countries.

Ng–Perron test (2001): Using the procedure in the ERS test to create efficient versions of the modified PP tests of Perron and Ng (1996), Ng and Perron (2001) showed that these tests do not have the same serious size distortions as the PP tests (used in many studies reviewed in the paper) for errors with large autoregressive and moving average roots. As a result, they can give a much higher power than the PP tests. Ng and Perron constructed four test statistics which are based on the PP tests (MZ_α and MZ_t statistics), the Bhargava (1986) (MSB) statistic and the ERS point optimal statistic (MP_T). We used the modified AIC for lag selection as suggested by the authors to maximize the power. Interpretations of results for these tests are similar to those of the ADF tests discussed above. To our knowledge, this was the first time these tests were used to examine the market efficiency hypothesis for our sample of countries.

3.2. Multiple-Break Unit Root Tests

Perron (1989) showed that structural change and unit roots are intimately related, and it is important to note that conventional unit root tests (as performed in many of the reviewed studies) are biased toward a false unit root null when the data are trend-stationary with a structural break. This observation has led to the development of a large amount of literature with unit root tests that remain valid in the presence of a break. One of the novel contributions of our study is the inclusion of multiple-break unit root tests by Lumsdaine and Papell (1997, LP), Lee and Strazicich (2003, LS), Narayan and Popp (2010, NP) and Ender and Lee (2012, EL). The main limitation of a unit root test, according to Zivot and Andrews (1992), is that it allows only for one break in the data and has a lower power than the tests described below. While Perron (1989) specified an *a priori* fixed break date, the ZA tests can endogenously determine a break date from the data.

Lumsdaine and Papell (1997): Improving on ZA, the LP multiple unit root tests allow for more than one (unknown) breakpoint in either the trend, the intercept or both the trend and the intercept of the data. We used two, four and six lags for the base model as well as automatic lag selections using the AIC and the BIC.[5] The null hypothesis is that there is a unit root in the data. Thus, if the null hypothesis is rejected, the return series is predictable, and vice versa. It should be noted that these are computationally intensive methods when two or more breaks are selected if the dataset is fairly large (more than 500).

Lee and Strazicich (2003) showed that their model outperforms that of Lumsdaine and Papell (1997, LP) in simulations and that, unlike the LP unit root test, rejection of the null unambiguously implies a stationary trend or return predictability in our case. They also showed that the power of the tests increases substantially when two or more breaks are taken into account. It is a minimum Lagrange multiplier test for testing the presence of a unit root with two structural breaks. We employed both the "Crash" model to allow for a sudden change in level but no change in the trend and the "Break" model to account for simultaneous changes in the level and the trend. The location of breakpoints is determined endogenously by conducting a grid search to locate the minimum t-statistics. We used a 10% trimming of data points at each end of the series. The critical values for the test were provided by Lee and Strazicich (2003). It is important to note that the critical values for the model with breaks in the intercept and the trend are dependent on break locations.

Narayan and Popp (2010): This has been one of the most cited tests in recent years. Narayan and Popp showed that their model outperforms those of LP and LS. Furthermore, NP possesses a more stable power and correct size. Further, the NP test accurately recognizes the break date. Since break dates are endogenously determined within the model, this test requires no prior knowledge for possible timings of structural breaks. In our study, we considered two different models, with the first model allowing two structural breaks (level) and the second model allowing two structural breaks (level and trend). The interpretation of the model is similar to those of LP and LS. To our knowledge, the NP test has not been used to study the stock market indices of our three countries.

Ender and Lee (2012): This test, also known as the Fourier unit root test, is one of the latest tests within this class. EL surpasses the aforementioned multiple-break unit root test by reducing specification errors about break dates and their forms (gradual or sharp), leading to an increase in the power of tests. The test uses trigonometric functions to capture deviations greater than the average of the dependent variable and takes into account multiple structural breaks. A major advantage of these tests is that there is no need to know a priori the break dates, the exact number of breaks and the form of breaks. EL utilizes a dynamic (time variant) deterministic intercept term consisting of sine and cosine functions to determine the essence of the process or whether there is a breakpoint or nonlinear trend. EL employs a specific data-generating regression model with the smallest residual sums of square at the most appropriate frequency, as well as a more precise approximation including multiple frequencies.

3.3. Spectral Analysis

A series of tests employed in this study are the recently available spectral tests commonly used in electrical engineering.[6] We first examined the periodogram for each country to help identify the dominant periods, cyclical properties or periodicities across different frequencies (high and low) in a series. We looked for peaks or hidden periodic components in the data. If a series seems very smooth, for example, then the values of the periodogram for low frequencies will be large relative to its other values, and vice versa. For a random walk series, all sinusoids should be of similar importance, and the periodogram will vary randomly around a constant. On the other hand, if a series exhibits very pronounced spectra at higher frequencies, this may indicate that the series is driven by dynamics or transient features that frequently come and go. In this case, we would typically consider this time series as stationary (we would typically classify it as nonstationary if spectra are more prominent near zero frequency). Further, we employed Fisher's G-test to check for the proportion of intensity represented at each specific frequency to determine if the observed peak at that frequency is random or not. Particularly, this test reveals if the series in question is white noise (i.e., a stationary process) in the sense that its maximum ordinate is not significant enough. Finally, utilizing a normalized integrated spectrum, we tested the hypothesis if observations from each of series follow a white noise process.

3.4. Abnormal Returns

Another novel feature of our work is the analysis of abnormal returns. We explored in this section whether a departure from market efficiency can be used to generate profitable trades. Since the stock markets in our study were found to be inefficient, it was interesting to explore their abnormal returns. To do this, we split the sample period by the multiple structural breaks identified in these tests into subsample periods. The random walk model and a rolling 36-month estimation period were used to compute the 1-month-ahead predicted return ($\hat{y}_{t+1}$):

$$y_t = c + \varepsilon_t$$

$$\hat{y}_{t+1} = \hat{c} = \frac{1}{36}\sum(y_t + y_{t-1} + \cdots + y_{t-35})$$

where y_t is the return in month t, c is the constant and $\hat{y}_{t+1}$ is the predicted return in month $t+1$.

We then subtracted the predicted return from the realized return in each month to calculate the abnormal return (AR):

$$AR_{t+1} = y_{t+1} - \hat{y}_{t+1}.$$

Summing up the monthly abnormal returns is the cumulative abnormal return (CAR) in a subsample period:

$$CAR = \sum_{t=36}^{T} AR_{t+1}.$$

The importance of a structural break and its impact on abnormal profits should not be overlooked as the existence of significant abnormal returns may suggest that the market in question is inefficient.

3.5. Other Tests

Variance ratio tests: This test, after Lo and MacKinlay (1988), has been shown to be more powerful and reliable than the ADF tests and is robust to heteroscedasticity. It is based on the notion that if a series follows a random walk process, then the variance of its qth period difference should be q times the variance of its 1-period difference. If the variance ratio test statistic is greater than 1, then the series is positively correlated. We chose two, four, eight and 16 periods as these periods are typically chosen in the literature. The variance ratio of the Lo and MacKinlay tests whether the variance ratio is equal to 1 for a particular holding period. For each country, we presented its variance ratio, its Chow and Denning (1993) joint maximum z-statistic (since we chose more than one period) and its associated *p*-values (we did not report the individual test statistics as they are qualitatively similar). The null hypothesis of random walk is rejected if the *p*-value for the z-statistic is small (i.e., less than 0.05 for a 5% significance level). We noted that for a given set of test statistics, the random walk hypothesis is rejected if any one of the variance ratios is considerably dissimilar to one. The results of this test are not reported to conserve space as they are similar to those obtained using the Lo and MacKinlay tests.

BDS test (1996): This is perhaps the most popular (nonlinear) test for detecting serial dependence in time series data, after Brock et al. (1996). A number of studies have found evidence of the movement of asset returns. The BDS tests the null hypothesis of independent and identically distributed (IID) process against an unknown alternative. The test is estimated for different embedding dimensions (m) and distances (e). The null hypothesis of randomness is rejected if the BDS statistic exceeds 2 for a 95% confidence and 3 for a 99% confidence. For ease of interpretation, we presented results using different dimensions (m = 2–6) and e = 0.5. The distance e was selected to make sure a certain fraction of the total number of pairs of points in the sample lie within e of each other as this approach is most invariant to the distribution of the series in question. Furthermore, we let e vary from 0.50 to 2 (the higher this value, the lower the power of the test). The results for the tests where e was higher than 0.5 are not reported as they are similar to those of the baseline case. As a further robustness test, especially when dealing with shorter series, we also chose the option of calculating bootstrapped *p*-values for the test statistic using various repetitions to increase the accuracy of the *p*-values (the results are not shown as they are qualitatively similar to those from the standard tests).

4. Empirical Results

Table 1 presents the summary statistics of the data[7]. As found in many prior studies, all the return series for the USA, the UK and Japan were not normally distributed, based on their associated Jarque–Bera statistics. We also examined the correlation matrix (results not shown) and observed that these return series are positively (and statistically significant) related,[8] similar to those found in other developed markets in several prior studies. The rather high kurtosis numbers suggest the higher likelihood of extreme returns in the data for all the three countries. The skewness numbers indicate high volatility, with some extreme gains for the USA and losses for the UK and Japan.

Table 1. Descriptive statistics.

	USA	UK	Japan
	y_{US}	y_{UK}	y_{JP}
Mean	0.0062	0.0070	0.0013
Median	0.0091	0.0116	0.0011
Max.	0.4222	0.1129	0.1843
Min.	−0.2994	−0.1306	−0.2012
Std. dev.	0.0542	0.0405	0.0567
Skewness	0.2928	−0.4591	−0.0535
Kurtosis	12.4360	3.6832	3.8240
Jarque–Bera	4063.0700 ***	17.0834 ***	9.0045 **
p	0.0000	0.0002	0.0111
NOB	1092	313	313

Notes: This table reports the descriptive statistics of monthly returns, y, on the country price indices for the USA, the UK and Japan. The sample period for the USA spanned from January 1926 to December 2016, while the sample period for the UK and Japan started in December 1989 to December 2015. Notations ** and *** indicate 5% and 1% significance levels, respectively.

Table 2 shows the results for simple unit root tests. At the 1% level of significance, the ADF and Phillip–Perron tests unanimously rejected the random walk hypothesis. Table 3 displays the results for the ERS tests which also rejected the null hypothesis. It is interesting to note that the random walk hypothesis was rejected by only two of the four tests for the USA, the UK and Japan indicated in Table 4 (Ng–Perron). Table 5 reports the Zivot–Andrews test results. Again, the unit root or the random walk hypothesis was rejected at the 1% significance level. These results are in line with some of the prior reviewed studies but are in stark contrast to those obtained by Narayan and Smyth (2007), except for Japan whose price series was found to be stationary. It is possible that their tests (LS, LP, Perron, Zivot and Andrews) suffer from the same problems as those discussed by Narayan and Popp (2010) and Ender and Lee (2012) which are performed in our study. While the test found a structural break in April 2000 for the US, in March 2009 for the UK and in March 2007 for Japan, these results should be interpreted with extreme caution (Perron 1989).[9]

Table 2. Unit root tests: Augmented Dickey–Fuller and Phillip–Perron Tests.

	USA		UK		Japan	
	ADF Test	**Phillips–Perron Test**	**ADF Test**	**Phillips–Perron Test**	**ADF Test**	**Phillips–Perron Test**
a_0	0.0046	0.0046	0.008795 *	0.008795 *	−0.0077	−0.0077
	(0.1585)	(0.1585)	(0.0585)	(0.0585)	(0.2334)	(0.2334)
γ_1	−0.9163 ***	−0.9163 ***	−0.9444 ***	−0.9444 ***	−0.9034 ***	−0.9034 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
θ	0.0000	0.0000	−0.0000	−0.0000	0.0000	0.0000
	(0.7117)	(0.7117)	(0.5508)	(0.5508)	(0.1158)	(0.1158)
Adj. R^2	0.4572	0.4572	0.4696	0.4696	0.4480	0.4480
NOB	1092	1092	313	313	313	313

Notes: ADF denotes the augmented Dickey–Fuller test. For details, please see Dickey and Fuller (1979) and Phillips and Perron (1988). The model specification for the ADF and Phillips–Perron tests is: $y_t = a_0 + \gamma_1 y_{t-1} + \theta t + \sum_{i=2}^{p} \beta_i y_{t-i} + \varepsilon_t$, where y_t is the stock index return in month t. The null hypothesis is $\gamma_1 = 1$, a unit root in the return series. p-values are in the parentheses. Notations *, ** and *** denote 10%, 5% and 1% significance levels, respectively.

Table 3. Unit root tests: Elliott–Rothenberg–Stock test.

	USA	**UKA**	**Japan**
Elliott–Rothenberg–Stock test statistic	0.6490 ***	0.7920 ***	−15.398 ***
Test critical values: 1% level	3.9600	3.9915	−3.4712
Test critical values: 5% level	5.6200	5.6374	−2.9076
Test critical values: 10% level	6.8900	6.8770	−2.6008
NOB	1092	313	313

Notes: The equations of unit root testing by Elliott et al. (1996) are specified as follows: $y_t = d_t + U_t$, $U_t = \alpha U_{t-1} + v_t$, where y_t is the stock index return, d_t is a deterministic component, v_t is an unobserved stationary error with zero mean, and its spectral density at frequency of zero is a positive value. In the GLS-detrended series, $\widetilde{y_t} \equiv y_t - \hat{\varphi}' Z_t$, $\hat{\varphi}$ minimizes $S(\overline{\alpha}, \varphi) = \left(y^{\overline{\alpha}} - \varphi' Z^{\overline{\alpha}}\right)' \left(y^{\overline{\alpha}} - \varphi' Z^{\overline{\alpha}}\right)$, where Z_t is a set of deterministic components and $\overline{\alpha} = \left(1 + \frac{\overline{c}}{T}\right)$. The null hypothesis of a unit root is $\alpha = 1$, while the alternative hypothesis is $\alpha = \overline{\alpha}$. The likelihood ratio statistic is defined as $L = S(\overline{\alpha}) - S(1)$, where $S(\overline{\alpha}) = min_{\varphi} S(\overline{\alpha}, \varphi)$. The statistic of a feasible point optimal test is $P_T = [S(\overline{\alpha}) - S(1)]/S^2_{AR}$. S^2_{AR} is the autoregressive estimate of the spectral density at zero frequency of v_t. $S^2_{AR} = \hat{\sigma}_k / \left(1 - \hat{\beta}(1)\right)^2$. In an augmented Dickey–Fuller equation, $y_t = d_t + \gamma_1 y_{t-1} + \sum_{i=2}^{k} \beta_i y_{t-i} + \varepsilon_{tk}$, $\hat{\beta}(1) = \sum_{i=2}^{k} \hat{\beta}_i$ and $\hat{\sigma}^2_k = (T-k)^{-1} \sum_{t=k+1}^{T} \hat{\varepsilon}^2_{tk}$, where T is the total of time periods and k is the lag length. Notation *** denotes a 1% significance level.

Table 4. Unit root tests: Ng–Perron test.

	USA				UK				Japan			
	MZ_α	MZ_t	MSB	MP_T	MZ_α	MZ_t	MSB	MP_T	MZ_α	MZ_t	MSB	MP_T
Ng–Perron test statistics	−96.1995 [a]	−6.9339 [a]	0.0721	0.9531	−8.1361	−2.0149	0.2476 [a]	11.2068 [a]	−153.1710 [a]	−8.7497 [a]	0.0571	0.6000
Asym. critical values: 1% level	−23.8000	−3.4200	0.1430	4.0300	−23.8000	−3.4200	0.1430	4.0300	−23.8000	−3.4200	0.1430	4.0300
Asym. critical values: 5% level	−17.3000	−2.9100	0.1680	5.4800	−17.3000	−2.9100	0.1680	5.4800	−17.3000	−2.9100	0.1680	5.4800
Asym. critical values: 10% level	−14.2000	−2.6200	0.1850	6.6700	−14.2000	−2.6200	0.1850	6.6700	−14.2000	−2.6200	0.1850	6.6700
NOB	1092	1092	1092	1092	313	313	313	313	313	313	313	313

Notes: The equations of unit root testing by Ng and Perron (2001) are specified as follows: $y_t = d_t + U_t$, $U_t = \alpha U_{t-1} + v_t$, where y_t is the stock index return, d_t is a deterministic component, v_t is an unobserved stationary error with zero mean, and its spectral density at zero frequency is a positive value. $d_t = \sum_{i=0}^{p} \varphi t^i$. The analysis by Ng and Perron (2001) focused on $p = 0$, 1, but it remains valid in general cases. The null hypothesis of a unit root is $\alpha = 1$, while the alternative hypothesis is $\alpha < 1$. In an augmented Dickey–Fuller equation, $y_t = d_t + \gamma_1 y_{t-1} + \sum_{i=2}^{k} \beta_i y_{t-i} + \varepsilon_{tk}$, $\hat{\beta}(1) = \sum_{i=2}^{k} \hat{\beta}_i$ and $\hat{\sigma}_k^2 = (T-k)^{-1} \sum_{t=k+1}^{T} \hat{\varepsilon}_{tk}^2$. $S_{AR}^2 = \hat{\sigma}_k / (1-\hat{\beta}(1))^2$. $MZ_\alpha = \left(T^{-1} y_T^2 - S_{AR}^2\right)\left(2T^{-2} \sum_{t=1}^{T} y_{t-1}^2\right)^{-1}$, $MSB = \left(T^{-2} \sum_{t=1}^{T} y_{t-1}^2 / S_{AR}^2\right)^{(1/2)}$. $MZ_t = MZ_\alpha \times MSB$. The statistic for the modified feasible point optimal test by Ng and Perron (2001) is as follows: when $p = 0$, $MP_T^{GLS} = \left[c^{-2} T^{-2} \sum_{t=1}^{T} \tilde{y}_{t-1}^2 - \bar{c} T^{-1} \tilde{y}_T^2\right] / S_{AR}^2$. When $p = 1$, $MP_T^{GLS} = \left[c^{-2} T^{-2} \sum_{t=1}^{T} \tilde{y}_{t-1}^2 + (1-\bar{c}) T^{-1} \tilde{y}_T^2\right] / S_{AR}^2$. Notation [a] denotes a 1% significance level.

Table 5. Single-break unit root tests: Zivot–Andrews test.

	USA	UK	Japan
Zivot–Andrews test statistic	−14.17325 ***	−16.8966 ***	−16.26222 ***
1% critical value	−5.57	−5.34	−5.57
5% critical value	−5.08	−4.93	−5.08
10% critical value	−4.82	−4.58	−4.82
Breakpoint	April 2000	March 2009	March 2007
NOB	1092	313	313

Notes: Zivot and Andrews (1992) modified three models developed by Perron (1989), the crash model (model A), the changing growth model (model B) and the changes in the level and slope of the trend function (model C), to endogenously determine a breakpoint from the data. The following are the modified models: Model A: $y_t = \hat{\mu}^A + \theta^A DU_t(\hat{\lambda}) + \hat{\beta}^A t + \hat{\alpha}^A\ y_{t-1} + \sum_{j=2}^{k} \hat{c}_j^A y_{t-j} + \hat{\varepsilon}_t$, model B: $y_t = \hat{\mu}^B + \hat{\beta}^B t + \hat{\gamma}^B DT_t^*(\hat{\lambda}) + \hat{\alpha}^B\ y_{t-1} + \sum_{j=2}^{k} \hat{c}_j^B y_{t-j} + \hat{\varepsilon}_t$, model C: $y_t = \hat{\mu}^C + \hat{\theta}^C DU_t(\hat{\lambda}) + \hat{\beta}^C t + \hat{\gamma}^C DT_t^*(\hat{\lambda}) + \hat{\alpha}^C\ y_{t-1} + \sum_{j=2}^{k} \hat{c}_j^C y_{t-j} + \hat{\varepsilon}_t$, where y_t in our setting is the stock index return in month t, $\lambda = T_B/T$, T_B is the breakpoint, T is the total of time periods, $DU_t(\lambda) = 1$ if $t > T\lambda$ and zero otherwise and $DT_t^*(\lambda) = t - T\lambda$ if $t > T\lambda$ and zero otherwise. Notation $\wedge$ is the estimated value of the break function. The null hypothesis of a unit root is $\alpha = 1$. The test statistic is $t_{\hat{\alpha}^i}(\lambda)$, and $i = A,\ B,\ C.$ λ was chosen to minimize the one-sided t-statistic for testing the unit root (i.e., $\alpha^i = 1$). Notation *** denotes a 1% significance level.

Table 6 presents the results of the LP multiple-break unit tests with two lags and two breaks as typically suggested in the econometric literature. First, the null hypothesis of a unit root (with two or more breaks) was rejected by both tests at the 1% significance level for the USA, the UK and Japan.[10] Table 9 reports the findings for Narayan and Popp. Again, the unit root or the random walk hypothesis was rejected at the 1% significance level. The test also found two breaks. Similarly, the LS test rejected the random walk hypothesis, as shown in Tables 7 and 8. It is interesting to note that LS only found one break for the US and two breaks for the UK and Japan and that the break dates in LS and LP are quite different—a well-documented phenomenon in the literature. The NP test (Table 9) appears to do a better job in capturing breaks in all the three series which occurred around the financial crisis starting in 2007. The NP results also unambiguously rejected the null hypothesis, based on model 1 (break in the level, not reported) and model 2 allowing for breaks in both the level and the trend (shown in the table). These break dates found in the LP and NP tests were later used in the final part of our paper to study the associated abnormal returns in these countries. Finally, the findings for EL presented in Table 10 are similar to those of LP, LS and NP. Note that EL, while allowing for an unknown number of breaks, does not report the number of breaks. The optimal lags chosen to minimize the residual sum of squares were six, seven and two for the US, the UK and Japan, respectively.

Table 6. Multiple-break unit root tests: Lumsdaine–Papell test.

	USA	UK	Japan
μ	0.0037	0.0161	−0.0204
	(0.7498)	(2.4993)	(−1.9811)
β	0.0000	−0.0001	0.0003
	(0.6149)	(−1.8426)	(2.2255)
θ	−0.0108	0.0395	−0.0505
	(−1.4678)	(3.4249)	(−2.992)
γ	0.0000	−0.0007	0.0006
	(1.4546)	(−3.0054)	(1.7457)
ω	−0.025	0.0499	−0.0508
	(−2.6197)	(3.8389)	(−3.0221)
ψ	0.0000	0.0005	−0.0006
	(0.5755)	(1.7308)	(−1.7839)
α	−1.0284 ***	−1.0022 ***	−0.9501 ***
	(−14.094)	(17.655)	(−16.6473)
NOB	1092	313	313
Number of breaks	2	2	2
First break	March 1968	March 2003	February 2000
Second break	April 2000	February 2009	January 2006

Notes: The model specification for the Lumsdaine and Papell (1997) test is as follows: $y_t = \mu + \beta t + \theta DU1_t + \gamma DT1_t + \omega DU2_t + \psi DT2_t + \alpha y_{t-1} + \sum_{i=2}^{k} c y_{t-i} + \varepsilon_t$, where y_t in our setting is the stock index return in month t, $DU1_t$ ($DU2_t$) is an indicator dummy for a mean shift at $TB1$ ($TB2$), the time breakpoint, and $DT1$ ($DT2$) is the corresponding trend shift variable. The null hypothesis is $\alpha = 1$, a unit root in the return series. Given that $\delta_1 = TB1/T$ and $\delta_2 = TB2/T$, the test statistic is defined as $\hat{t}(\delta_1, \delta_2) \Longrightarrow \int_0^1 w^*(s)dw(s) / \left[\int_0^1 w^*(s)^2 ds\right]^{(\frac{1}{2})}$, where $w(s)$ is a Wiener process. T-statistics are in brackets. Notation *** denotes a 1% significance level.

Table 7. Multiple-break unit root tests: Lee and Strazicich test: the crash model.

	USA	UK	Japan
μ	−0.0120 ***	0.0010	0.0023
	(−6.1079)	(0.4163)	(0.7139)
δ_{DU1}	−0.0445	0.0655	0.1313
	(−0.8311)	(1.5284)	(2.2769)
δ_{DU2}		0.0630	0.0629
		(1.4687)	(1.1020)
ϕ	−0.9113 ***	−0.5163 ***	−0.7294 ***
	(−10.9841)	(−5.9937)	(−8.4737)
Minimum test stat. (tau)	−10.9841	−5.9937	−8.4737
Test critical values: 1% level	−3.7980	−4.2264	−4.2264
Test critical values: 5% level	−3.2300	−3.6356	−3.6356
Test critical values: 10% level	−2.9250	−3.2995	−3.2995
Breakpoint	June 1981	September 2003	March 1993
		February 2010	April 2003
NOB	1092	313	313

Notes: The specification for the crash model in the Lee and Strazicich (2003) test is as follows: $y_t = \delta' Z_t + \phi \widetilde{S}_{t-1} + \mu_t$, $\widetilde{S}_t = y_t - \widetilde{\psi}_x - Z_t\widetilde{\delta}$, $\widetilde{\psi}_x = y_1 - Z_1\widetilde{\delta}$, where Z_t is a set of exogenous variables, $Z_t' = [1, t, DU1_t, DU2_t]$ and δ' is a set of coefficients $[\delta_1, \delta_1, \delta_{DU1}, \delta_{DU2}]$. The null hypothesis is $\phi = 1$, a unit root in the return series. T-statistics are in brackets. Notation *** denotes a 1% significance level.

Table 8. Multiple-break unit root tests: Lee and Strazicich test: the break model.

	USA	UK	Japan
μ	−0.0161 ***	−0.0508 ***	−0.0339 ***
	(−7.0990)	(−11.2039)	(−6.5624)
δ_{DU1}	−0.0192	−0.2457 ***	0.0995
	(−0.3569)	(−5.9904)	(1.7233)
δ_{DU2}		−0.0983	−0.1006
		(−2.5119)	(−1.7423)
δ_{DT1}	0.0124	0.1138	−0.0481 *
	(2.4237)	(8.4136)	(−4.8467)
δ_{DT2}		−0.0273	0.0965
		(−2.2535)	(7.0711)
ϕ	−0.9139 ***	−1.0956 ***	−0.8416 ***
	(−10.9961)	(−13.7221)	(−11.0827)
Minimum test stat. (tau)	−10.9961	−13.7221	−11.0827
Test critical values: 1% level	−4.4612	−5.6458	−5.5177
Test critical values: 5% level	−3.9240	−4.9246	−5.0260
Test critical values: 10% level	−3.6492	−4.6474	−4.7586
Breakpoint	November 2005	August 2008	November 2005
		September 2009	March 2010
NOB	1092	313	313

Notes: The specification for the break model in the Lee and Strazicich test is as follows: $y_t = \delta' Z_t + \phi \widetilde{S}_{t-1} + \mu_t$, $\widetilde{S}_t = y_t - \widetilde{\psi}_x - Z_t\widetilde{\delta}$, $\widetilde{\psi}_x = y_1 - Z_t\widetilde{\delta}$, where Z_t is a set of exogenous variables, $Z_t' = [1, t, DU1_t, DU2_t, DT1_t, DT2_t]$ and δ' is a set of coefficients $[\delta_1, \delta_1, \delta_{DU1}, \delta_{DU2}, \delta_{DT1}, \delta_{DT2}]$. The null hypothesis is $\phi = 1$, a unit root in the return series. T-statistics are in brackets. Notations *** and * denote 1% and 10% significance levels.

Table 9. Multiple-break unit root tests: Narayan and Popp test.

	USA	UK	Japan
Narayan and Popp test statistic	12.666 ***	17.534 ***	16.397 ***
1% critical value	5.287	5.318	5.318
5% critical value	4.692	4.741	4.741
10% critical value	4.396	4.430	4.430
Breakpoint	July 2007	June 2008	August 2008
	January 2009	September 2008	April 2009
NOB	1092	313	313

Notes: This table reports the test statistic of the model with a break and a trend in the paper by Narayan and Popp (2010). The null hypothesis is a unit root in the return series. The test is based on the following process: $y_t = d_t + U_t$, $U_t = t_{t-1} + \varepsilon_t$, $\varepsilon_t = \psi^*(L)\varepsilon_t = A^*(L)B(L)^{-1}e_t$, where y_t is the return series with a deterministic component d_t and a schochastic component U_t, e_t is iid $(0, \sigma^2)$ with $A^*(L)$ and $B(L)$ being polynomial lags of order p and q lying outside the unit circle. Model 1 in the paper by Narayan and Popp (2010) allows for two breaks in the level. Model 2 (shown) allows for two breaks in the level and the trend. Notation *** denotes a 1% significance level.

Table 10. Multiple-break unit root tests: Ender and Lee test.

	USA	UK	Japan
Ender and Lee test statistic	10.299 ***	4.175 **	8.144 ***
1% critical value	4.560	4.610	3.730
5% critical value	4.030	4.070	3.120
10% critical value	3.770	3.790	2.830
Chosen lag	6	7	2
Frequency	1	1	5
NOB	1092	313	313

Notes: Ender and Lee test (2012) is a modification of the DF test in which $d(t)$ or the time-dependent deterministic term is added to the test regression: $Y(t) = d(t) + \alpha Y_{t-1} + e_t$ and e_t is iid $\left(0, \sigma^2\right)$, where Y is the stock return. The unit root null hypothesis of $\alpha = 1$ is tested by approximating $d(t)$ with the following Fourier function : $d(t) = \phi_0 + \phi_{sin} \cdot sin(2\pi kt/T) + \phi_{cos} \cdot cos(2\pi kt/T) + \varepsilon_t$, where $\varepsilon_t = \alpha e_{t-1} + u_t$, k is the single frequency component and measures the amplitude and displacement of the sinusoidal component of $d(t)$, $t = 1, 2, \ldots, T$. The above equation is estimated for all integer values of k which lie between the interval [1, 5] and selecting the estimation which produces the lowest residual sum of squares. Notations *** and ** denote 1% and 5% significance levels, respectively.

The results from the normalized integrated spectrum tests are shown in Figure 1 (USA), Figure 2 (UK) and Figure 3 (Japan). The null of stationarity was rejected at the 5% significance level in all the three return series as the statistics fell within the two bands. Fisher's G-tests and periodograms for each country, not shown to save space, are qualitatively similar.

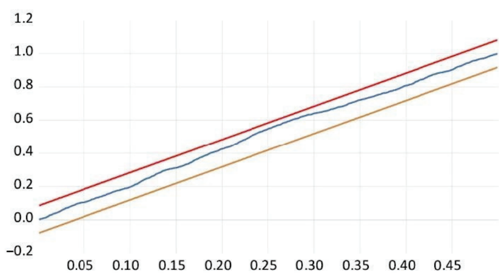

Figure 1. Spectral tests: Normalized integrated spectrum for the USA. Note: This test evaluates the null hypothesis that the data are stationary (white noise). It is based on the normalized integrated spectrum with the following statistic: $U_p = \frac{\sum_{k=1}^{p} I\left(w_q\right)}{\sum_{k=1}^{n/2} I\left(w_q\right)}$, where $I\left(\omega(i)\right)$ is the ith maximum. The test statistic (vertical axis) is plotted against its frequency (horizontal axis). If the deviations of the statistic from the $p/(n/2)$ line do not exceed $\pm a\sqrt{2/n}$, the null will not be rejected, where a is set equal to 1.36 for 95% confidence.

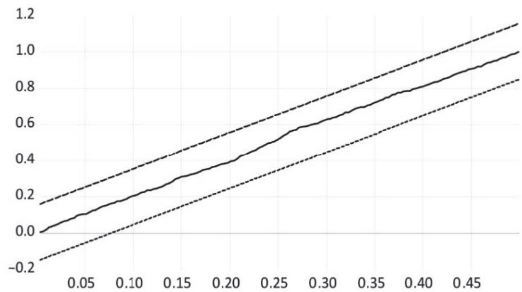

Figure 2. Spectral tests: Normalized integrated spectrum for the UK. Note: This test evaluates the null hypothesis that the data are stationary (white noise). It is based on the normalized integrated spectrum with the following statistic: $U_p = \frac{\sum_{k=1}^{p} I\left(w_q\right)}{\sum_{k=1}^{n/2} I\left(w_q\right)}$, where $I\left(\omega(i)\right)$ is the ith maximum. The test statistic (vertical axis) is plotted against its frequency (horizontal axis). If the deviations of the statistic from the $p/(n/2)$ line do not exceed $\pm a\sqrt{2/n}$, the null will not be rejected, where a is set equal to 1.36 for 95% confidence.

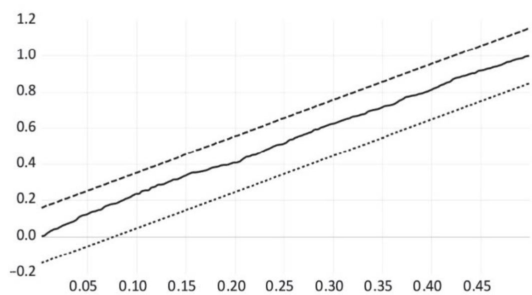

Figure 3. Spectral tests: Normalized integrated spectrum for Japan. Note: This test evaluates the null hypothesis that the data are stationary (white noise). It is based on the normalized integrated spectrum with the following statistic: $U_p = \frac{\sum_{k=1}^{p} I\left(w_q\right)}{\sum_{k=1}^{n/2} I\left(w_q\right)}$, where $I\left(\omega(i)\right)$ is the ith maximum. The test statistic (vertical axis) is plotted against its frequency (horizontal axis). If the deviations of the statistic from the $p/(n/2)$ line do not exceed $\pm a\sqrt{2/n}$, the null will not be rejected, where a is set equal to 1.36 for 95% confidence.

Table 11 reports the mean abnormal returns and the cumulative abnormal returns for the USA, the UK and Japan. To conserve space, we presented these statistics for a sample period with two structural breaks identified by the Lumsdaine–Papell test and the Narayan and Popp test. The two structural breaks identified using the Narayan and Popp test coincide with the recent global financial crisis period. Table 11a,b show that the mean abnormal return in most of subsample periods for the USA, the UK and Japan is close to zero. However, significant cumulative abnormal returns are found for the USA, the UK and Japan, lending support for market inefficiency. Interestingly, the cumulative abnormal returns were all positive (negative) for Japan (UK) in these subsample periods. The cumulative abnormal returns in the subsample periods ranged from 14.64% to 81.02% for Japan, whereas they were between −4.05 and −64.51% for the UK. The positive (negative) cumulative abnormal returns for Japan indicated that the stock market in Japan (UK) consistently outperformed (underperformed) the random walk model.

Table 11. Abnormal and cumulative abnormal returns for the USA, the UK and Japan. (a) Sample period split by breakpoints identified by the Lumsdaine–Papell test. (b) Sample period split by breakpoints identified using the Narayan and Popp test.

(a)					
USA	Subsample period 1	First breakpoint	Subsample period 2	Second breakpoint	Subsample period 3
	Jan. 1926–Feb 1968	Mar. 1968	Apr. 1968–Mar. 2000	Apr. 2000	May 2000–Dec. 2016
Mean Ab. Ret	−0.0010		0.0007		0.0018
	(0.0671)		(0.0447)		(0.0405)
Cum. Ab. Ret	−0.4511		0.2281		0.2924
NOB	470		348		164
UK	Subsample period 1	First breakpoint	Subsample period 2	Second breakpoint	Subsample period 3
	Dec. 1989–Feb. 2003	Mar. 2003	Apr. 2003–Jan. 2009	Feb. 2009	Mar. 2009–Dec. 2015
Mean Ab. Ret	−0.0042		−0.0190		−0.0029
	(0.0393)		(0.0441)		(0.0313)
Cum. Ab. Ret	−0.5166		−0.6451		−0.1334
NOB	123		34		46
Japan	Subsample period 1	First breakpoint	subsample period 2	Second breakpoint	Subsample period 3
	Dec. 1989–Jan. 2000	Feb. 2000	Mar. 2000–Dec. 2005	Jan. 2006	Feb. 2006–Dec. 2015
Mean Ab. Ret	0.0077		0.0238		0.0093
	(0.0572)		(0.0407)		(0.0508)
Cum. Ab. Ret	0.6635		0.8102		0.7698
NOB	86		34		83
(b)					
US	Subsample period 1	First breakpoint	Subsample period 2	Second breakpoint	Subsample period 3
	Jan. 1926–Jun. 2007	Jul. 2007	Aug. 2007–Dec. 2008	Jan. 2009	Feb. 2009–Dec. 2016
Mean Ab. Ret	−0.0003		N/A		−0.0014
	(0.0567)		N/A		(0.0304)
Cum. Ab. Ret	−0.2935		N/A		−0.0800
NOB	942		N/A		59
UK	Subsample period 1	First breakpoint	Subsample period 2	Second breakpoint	Subsample period 3
	Dec. 1989–May 2008	June. 2008	Jul. 2008–Aug. 2008	Sep. 2008	Oct. 2008–Dec. 2015
Mean Ab. Ret	−0.0008		N/A		−0.0008
	(0.0368)		N/A		(0.0321)
Cum. Ab. Ret	−0.1555		N/A		−0.040538
NOB	186		N/A		51
Japan	Subsample period 1	First breakpoint	Subsample period 2	Second breakpoint	Subsample period 3
	Dec. 1989–Jul. 2008	Aug. 2008	Sep. 2008–Mar. 2009	Apr. 2009	May 2009–Dec. 2015
Mean Ab. Ret	0.0008		N/A		0.0051
	(0.0525)		N/A		(0.0513)
Cum. Ab. Ret	0.146402		N/A		0.2248
NOB	188		N/A		44

Notes: A rolling 36-month estimation period was used to compute the 1-month-ahead predicted return from the random walk model. Each month, the predicted return is subtracted from the realized return to obtain an abnormal return. The cumulative abnormal return is the sum of abnormal returns in a subsample period. The following are the specifications of the random walk model, predicted return ($\hat{y}$), abnormal return (AR) and cumulative abnormal return (CAR): $y_t = c + \varepsilon_t$, $\hat{y}_{t+1} = \hat{c} = \frac{1}{36}\sum\left(y_t + y_{t-1} + \cdots + y_{t-35}\right)$, $AR_{t+1} = y_{t+1} - \hat{y}_{t+1}$, $CAR = \sum_{t=36}^{T} AR_{t+1}$. If a subsample period is shorter than 36 months, predicted return, abnormal return and cumulative abnormal return are not computed. The standard deviation of abnormal returns is in parentheses.

Table 11a also shows that the cumulative abnormal return for the USA was −45.11%, 22.81%, and 29.24% in the periods between January 1926 and February 1968, April 1968 and March 2000, May 2000 and December 2016, respectively. In Table 11b, we find cumulative abnormal returns of −29.35% and −8% for the USA during the periods of January 1926 to June 2007 and February 2009 to December 2016, respectively. The presence of significant cumulative abnormal returns again suggests that the US stock market is not efficient. Overall, our empirical evidence implies that abnormal profits can be exploited if structural breaks are correctly identified and appropriate trading strategies are implemented. The importance of a structural break and its impact on abnormal profits cannot be overlooked.

It is important to note that the structural breaks found correspond to major historical economic events. For example, 1968 is the year of economic crisis in the USA (Collins 1996): the Bretton Woods Agreement caused the balance of payments deficit in the USA. In March 1968, foreign investors started selling US dollars to buy gold, which led to the crack of the Bretton Woods Agreement. In April 2000, the NASDAQ Composite Index plummeted 10% (Johansen and Sornette 2020). When the UK joined the Iraq War in March 2003, the FTSE 100 Index hit bottom at 3272. In January 2009, the UK entered the recession, and the unemployment rate rose in February 2009. For Japan, the recession of the Japanese economy started in the 1990s and continued to 2002. The Nikkei 225 Index rose above 20,000 yen in March 2000 because of the dot.com spillover effect from the USA. In the same month, news that Japan had entered a recession led to a global selloff which adversely affected technology stocks. In January 2006, Japan continued its expansion which started in 2002.

5. Discussion

The overall findings of mean reversions in our study may suggest that stock index prices behave in an ergodic manner. Horst and Wenzelburger (2008) showed in a theoretical model that financial market dynamics is ergodic if the interaction between households is sufficiently weak. In this case, market shares settle down to a unique equilibrium. However, when ergodicity no longer holds (if the interactive complementarities in the financial market are "too powerful"), "history matters" and the long-run market shares of competing financial mediators are path-dependent.

Our results also lend support to the existence of "market anomalies" or "behavioral finance" as discussed in earlier sections of the paper. Even in an imperfectly efficient market, Grossman and Stiglitz (1980) showed that there still exist opportunities for abnormal investment returns due to superior information gathering by some analysts. Lo and MacKinlay (1988) demonstrated that the serial correlation of share prices is significantly significant. Therefore, there is a possibility of short-term returns on share prices when investors realize that share prices move consequently in the same direction. Studying the American market with high-frequency data for the S&P 500 index, Peters (1994) found a persistent time series with strong autocorrelation. Findings from other recent studies, discussed in the literature review, are also consistent with our present results.

What may be the reasons for the mixed empirical evidence for the efficient market hypothesis?[11]

We do not have a solid answer but believe that the conflicting findings may be a result of discrepancies in the datasets used in prior studies. From our prior experience, estimation results using data from the same stock indices obtained from different databases can sometimes be quite dissimilar, perhaps due to various methodologies used in constructing the data series. Econometric methods employed in a given study can also play a role. Bhargava (2014) demonstrated that certain approaches in testing for random walk (such as those by Lo and Andrew's variance ratio and related tests) can lead to erroneous results. Our study, we believe, mitigated some of these shortcomings by employing a comprehensive battery of highly regarded tests on an authoritative database. It is surprising that this was the first time, to our knowledge, data from the WRDS stock price indices were used to examine this issue.

6. Concluding Remarks

The main rationale for our research was that previous studies had found mixed results with regard to the efficient market hypothesis. We set out to explore this topic for the USA, the UK and Japan with a recent dataset and improved statistical methods. We contributed to the existing literature by employing a comprehensive battery of tests including several high-power multiple-break unit root and novel spectral tests. We further computed the abnormal returns using the break dates captured in the models. We then linked those abnormal profits to their associated economic events. We found that stock market indices in the USA, the UK and Japan are generally not efficient. While our results are in line with a number of recent studies, they do not support the findings of several earlier studies reviewed in the paper. Therefore, definitive empirical evidence for mean reversions in highly developed markets remains elusive. It will be interesting to extend the present study to include market indices in other advanced countries in future studies. Finally, based on the findings in this study, it may be concluded that investors could possibly be able to earn arbitrage profits due to market inefficiency even in highly developed stock markets.

Author Contributions: Conceptualization, J.N.; methodology, J.N. and W.-X.L.; software, J.N.; validation, J.N., W.-X.L. and C.C.-S.C.; formal analysis, J.N.; investigation, J.N. and W.-X.L.; resources, J.N.; data curation, J.N.; writing—original draft preparation, J.N.; writing—review and editing, J.N., W.-X.L. and C.C.-S.C.; visualization, J.N.; supervision, J.N.; project administration, J.N.; funding acquisition, J.N. and W.-X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available upon request from the authors.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. These tests allow for more than one structural break in the data and, if not accounted for, can lead to misleading results (Lumsdaine and Papell 1997; Lee and Strazicich 2003; Narayan and Popp 2010; Ender and Lee 2012).
2. We strongly suggest the readers refer to the original papers for detailed derivations of the models and test statistics. Due to space limitations and the large number of tests examined in this study, it is not practical to discuss each of them in detail.
3. There is another model based on the ergodic theorem stating that past and present probability distributions define the probability distribution, which will help forecast future market prices. The ergodic principle posits that the future is predetermined by the existing variables such as market fundamentals. Therefore, it is possible to forecast the future by analyzing the present and past data. If the system is nonergodic, on the other hand, the probability distributions of past and present do not provide a statistically reliable estimate for the probability of future events. A reviewer commented that stock prices appear to be random, yet they are "chaotic" in reality. This presents a challenge for the random walk model. Klinkova and Grabinski (2017) and Grabinski and Klinkova (2019) showed that using arithmetic means in chaotically varying quantities does not always yield results to random variations and that the "ultimate" financial model is not possible. Furthermore, ergodicity can be assumed in random variations but, generally, not in chaotic ones.
4. We selected high-impact and widely cited tests (most of which were originally published in elite journals in the fields of econometrics, statistics, finance and economics) to be used in our study to avoid the "kitchen-sink" approach.
5. To conserve space, we reported the results for two lags since the results were essentially the same for any of these methods.
6. Please refer to Wei (2018) and Ronderos (2014) for detailed discussions of the tests in this section.
7. To conserve space, we did not report the results from all the tests conducted in this study discussed in the Data and Methodologies section, especially when the vast majority of the findings were similar. Rather, we focused on the more interesting and important test results. In addition to the reported tests, we completed a variety of older random walk tests such as the Brock et al. (1996), various versions of variance ratio, runs and autocorrelation tests as in several of the reviewed articles and found the results were essentially unchanged (and did not report them in the Results section). The complete results are available from the authors upon request.

[8] The correlation coefficients between the WRDS indices and those of Compustat are between 0.95 and 0.98 for the countries in our sample.

[9] An anonymous reviewer noted that one typically wants to show that the measured results are stronger with a statistical significance when there is a null hypothesis or placebo. In many cases, the null hypothesis is also a result of observation. As such, it has a distribution. Including both distributions, consequently, changes the way one proves statistical significance. In a recent study, Tormählen et al. (2021) showed that in order to obtain identical significance, it may be necessary to perform twice as many experiments than in a setting where the placebo distribution is ignored. They also showed that statistical significance may be inaccurate due to "misuse" of the central limit theorem.

[10] The specification with three and more structural breaks was tested. However, our statistical software only found two breaks. Furthermore, the results remained similar regardless of the number of lags employed.

[11] We thank an anonymous referee for his/her many stimulating questions, including this one.

References

Amihud, Yakov. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Market* 5: 31–56. [CrossRef]

Ang, Andrew, and Geert Bekaert. 2007. Return predictability: Is it there? *Review of Financial Studies* 20: 651–707. [CrossRef]

Baker, Malcolm, and Jeffrey Wurgler. 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61: 1645–80. [CrossRef]

Baker, Malcolm, and Jeffrey Wurgler. 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21: 129–51. [CrossRef]

Bekaert, Geert, and Robert James Hodrick. 1992. Characterizing predictable components in excess returns on equity and foreign exchange markets. *Journal of Finance* 47: 467–509. [CrossRef]

Bekaert, Geert, Campbell Russell Harvey, and Christian Lundblad. 2007. Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies* 20: 1783–831. [CrossRef]

Bertone, Stephen, Imants Paeglis, and Rahul Ravi. 2015. (How) has the market come more efficient? *Journal of Banking & Finance* 54: 72–86.

Bhargava, Alok. 1986. On the theory of testing for unit roots in observed time series. *Review of Economic Studies* 53: 369–84. [CrossRef]

Bhargava, Alok. 2014. Firms' fundamentals, macroeconomic variables and quarterly stock prices in the US. *Journal of Econometrics* 183: 241–50. [CrossRef]

Borges, Maria Rosa. 2010. Efficient market hypothesis in European stock markets. *The European Journal of Finance* 16: 711–26. [CrossRef]

Brock, William, W. Davis Dechert, Blake Lebaron, and Jose Scheinkman. 1996. A test for independence based on a correlation dimension. *Econometric Review* 15: 197–235. [CrossRef]

Busse, Jeffrey A., Amit Goyal, and Sunil Wahal. 2010. Performance and persistence in institutional investment management. *Journal of Finance* 65: 765–90. [CrossRef]

Campbell, John Young, and Yasushi Hamao. 1992. Predictable stock returns in the United States and Japan: A study of long-term capital market integration. *Journal of Finance* 47: 43–69. [CrossRef]

Campbell, John Young, and Robert James Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1: 195–228. [CrossRef]

Carhart, Mark. 1997. On persistence in mutual fund performance. *Journal of Finance* 52: 57–82. [CrossRef]

Chen, Nai-Fu, Richard Roll, and Stephen Ross. 1986. Economic forces and the stock market. *Journal of Business* 59: 383–403. [CrossRef]

Chow, K. Victor, and Karen C. Denning. 1993. A simple multiple variance ratio test. *Journal of Econometrics* 58: 385–401. [CrossRef]

Collins, Robert M. 1996. The Economic Crisis of 1968 and the Waning of the "American Century". *The American Historical Review* 101: 396–422. [CrossRef]

De Long, James Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert James Waldmann. 1990. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance* 45: 379–95. [CrossRef]

Dickey, David Alan, and Wayne Arthur Fuller. 1979. Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74: 427–31.

Dickey, David Alan, and Wayne Arthur Fuller. 1981. Likelihood rtio statistics for autoregressive time series with a unit root. *Econometrica* 49: 1057–72. [CrossRef]

Durusu-Ciftci, Dilek, Mustafa Serdar Ispir, and Hakan Yetkiner. 2017. Financial development and economic growth: Some theory and more evidence. *Journal of Policy Modeling* 39: 290–306. [CrossRef]

Elliott, Graham, Thomas J. Rothernberg, and James Harold Stock. 1996. Efficient tests for an autoregressive unit root. *Econometrica* 64: 831–36. [CrossRef]

Ender, Walter, and Junsoo Lee. 2012. A unit root test using a fourier series to approximate smooth breaks. *Oxford Bulletin of Economics and Statistics* 74: 574–99. [CrossRef]

Fama, Eugene Francis, and G. William Schwert. 1977. Asset returns and inflation. *Journal of Financial Economics* 5: 115–46. [CrossRef]

Fama, Eugene Francis, and Kenneth Ronald French. 1988. Permanent and temporary components of stock prices. *Journal of Political Economy* 96: 246–73. [CrossRef]

Fama, Eugene Francis, and Kenneth Ronald French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47: 427–65. [CrossRef]

Fama, Eugene Francis, and Kenneth Ronald French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56. [CrossRef]

Fama, Eugene Francis, and Kenneth Ronald French. 2010. Luck versus skill in the cross section of mutual fund returns. *Journal of Finance* 65: 1915–47. [CrossRef]

French, Kenneth Ronald, and Richard Roll. 1986. Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics* 17: 5–26. [CrossRef]

Goetzmann, William, Roger Ibbotson, and Liang Peng. 2001. A new historical database for the NYSE 1815 to 1925: Performance and predictability. *Journal of Financial Markets* 4: 1–32. [CrossRef]

Golez, Benjamin, and Peter Koudijs. 2018. Four centuries of return predictability. *Journal of Financial Economics* 127: 248–63. [CrossRef]

Grabinski, Michael, and Galiya Klinkova. 2019. Wrong use of average implies wrong results from many heuristic models. *Applied Mathematics* 10: 605–18. [CrossRef]

Grossman, Sanford, and Joseph Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70: 393–408.

Horst, Ulrich, and Jan Wenzelburger. 2008. On non-ergodic asset prices. *Economic Thoery* 34: 207–34. [CrossRef]

Johansen, Anders, and Didier Sornette. 2020. Condensed Matter and Complex Systems. *European Physical Journal B* 17: 319–28. [CrossRef]

Klinkova, Galiya, and Michael Grabinski. 2017. Conservation laws derived from systemic approach and symmetry. *International Journal of Latest Trends in Finance and Economics* 7: 1307–12.

Lee, Junsoo, and Mark C. Strazicich. 2003. Minimum lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics* 85: 1082–89. [CrossRef]

Lettau, Martin, and Stijn Van Nieuwerburgh. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* 21: 1607–52. [CrossRef]

Lo, Andrew Wen-Chuan, and A. Craig MacKinlay. 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies* 1: 41–66. [CrossRef]

Lumsdaine, Robin L., and David H. Papell. 1997. Multiple trend breaks and the unit-root hypothesis. *Review of Economics and Statistics* 79: 212–18. [CrossRef]

Mitra, Subrata, Manojit Chattopadhyay, Charan Parikshit, and Bawa Jaslene. 2017. Identifying periods of market inefficiency for return predictability. *Applied Economics Letters* 24: 668–71. [CrossRef]

Narayan, Paresh, and Stephan Popp. 2010. A new unit root test with two structural breaks in level and slope at unknown time. *Journal of Applied Statistics* 37: 1425–38. [CrossRef]

Narayan, Paresh, and Russell Smyth. 2007. Are shocks to energy consumption permanent or temporary? Evidence from 182 countries. *Energy Policy* 35: 333–41. [CrossRef]

Ng, Serena, and Pierre Perron. 2001. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69: 1519–54. [CrossRef]

Nyberg, Henri, and Harri Pönkä. 2016. International sign predictability of stock returns: The role of the United States. *Economic Modelling* 58: 323–38. [CrossRef]

Perron, Pierre. 1989. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57: 1361–401. [CrossRef]

Peters, Edgar E. 1994. *Fractal Market Analysis: Applying Chaos Theory to Investment and Economics*. New York: John Wiley & Sons Inc.

Phillips, Peter C. B., and Pierre Perron. 1988. Testing for a unit root in time series regression. *Biometrika* 75: 335–46. [CrossRef]

Rapach, David E., Jack K. Strauss, and Guofu Zhou. 2013. International stock return predictability: What is the role of the United States? *Journal of Finance* 68: 1633–62. [CrossRef]

Ronderos, Nicolas. 2014. Spectral Analysis Using Eviews. Available online: https://www.eviews.com/Addins/SpectralAnalysis.aipz (accessed on 18 January 2022).

Schwert, G. William. 1989. Tests for unit roots: A monte carlo investigation. *Journal of Business and Economic Statistics* 7: 147–60.

Shiller, Robert James. 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71: 421–36.

Shiller, Robert James. 1982. Consumption, asset markets, and macroeconomic fluctuations. *Carnegie-Rochester Conference Series on Public Policy* 17: 203–38. [CrossRef]

Solnik, Bruno. 1993. The performance of international asset allocation strategies using conditioning information. *Journal of Empirical Finance* 1: 33–55. [CrossRef]

Tormählen, Maike, Galiya Klinkova, and Michael Grabinski. 2021. Statistial Significance Revisited. *Mathematics* 9: 958. [CrossRef]

Urquhart, Andrew, and Robert Hudson. 2013. Efficient or adaptive markets? Evidence from major stock markets using very long run historic data. *International Review of Financial Analysis* 28: 130–42. [CrossRef]

Urquhart, Andrew, and Frank McGroarty. 2016. Are stock markets really efficient? Evidence of the adaptive market hypothesis. *International Review of Financial Analysis* 47: 39–49. [CrossRef]

Verheyden, Tim, Lieven De Moor, and Filip Van den Bossche. 2015. Towards a new framework on efficient markets. *Research in International Business and Finance* 34: 294–308. [CrossRef]

Wei, William W. S. 2018. *Time Series Analysis*, 2nd ed. Boston: Addison Wesley.
Zivot, Eric, and Donald Wilfrid Kao Andrews. 1992. Further evidence on the great crash, the oil-price shock, and the unit-orot hypothesis. *Journal of Business & Economic Statistics* 10: 251–70.

Journal of
Risk and Financial Management

MDPI

Article

On Survivor Stocks in the S&P 500 Stock Index

Klaus Grobys [1,2]

[1] Finance Research Group, School of Accounting and Finance, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland; klaus.grobys@uwasa.fi
[2] Innovation and Entrepreneurship (InnoLab), University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland

Abstract: This paper investigates the performance and characteristics of survivor stocks in the S&P 500 index. Using both in-sample and out-of-sample comparisons, survivor stocks outperformed this market index by a considerable margin. Relative to other S&P 500 index companies, survivor stocks tend to be small-value stocks that exhibit high profitability and invest conservatively. Surprisingly, survivor stocks tend to be loser stocks with negative exposure to the momentum factor. Further analyses show that the volatility of the survivor stocks portfolio is less exposed to tail risks and responds less to shocks in the innovation process.

Keywords: asset pricing; S&P 500 index; survivor stocks; risk factors; momentum

JEL Classification: G10; G12; G15; G19

Citation: Grobys, Klaus. 2022. On Survivor Stocks in the S&P 500 Stock Index. *Journal of Risk and Financial Management* 15: 95. https://doi.org/10.3390/jrfm15020095

Academic Editor: Robert Brooks

Received: 27 January 2022
Accepted: 16 February 2022
Published: 21 February 2022

1. Introduction

The purpose of this paper is to investigate the performance of long-run survivor stocks in the S&P 500 index and their characteristics. We make use of Standard & Poor's 2 March 2007 announcement and use the CRSP database to retrieve data for all survivor firms that exist until December 2019 but may have dropped out of the S&P 500 index in the ex post 2 March 2007 announcement period. We refer to this stock portfolio as *all survivors*. The statistical properties of this survivor portfolio are compared to the S&P 500 index. Additionally, we examine the survivor portfolios' outperformance relative to the index in general as well as risk-adjusted performance. Treating the ex ante March 2007 period as in sample and the ex post March 2007 period as out of sample, we further investigate whether a structural change occurred in the performance of survivor stocks in the ex post announcement period. In addition, we replicate our analysis using publicly available data retrieved from Yahoo. Lastly, in an effort to gain a deeper understanding of the performance of our survivor stocks portfolio, we explore the dynamics of factor exposures across time.

Our paper contributes to the academic literature in a number of ways. The S&P 500 index is widely considered to be an important gauge of the U.S. equity market and is prominently quoted in stock markets around the world (Gnabo et al. 2014). Being elected to join the constituent companies in the S&P 500 index is quite a feat: while there are more than 4000 listed companies in the U.S. stock market[1], only approximately 10% of these companies achieve membership in the S&P 500 stock index. Moreover, a company must pass the following battery of criteria to be selected by the Index Committee:[2] (1) primarily U.S. based, (2) market capitalization exceeding $8.2 billion, (3) highly liquid shares[3], (4) public trading of 50% or more of its outstanding shares, (5), positive earnings in the most recent quarter, and (6) a positive sum for the previous four quarters' earnings. Only very successful companies can fulfill these strict requirements. According to Chen and Lin (2018), member companies benefit from reductions in financial constraints, and a lower cost of equity, among other advantages.[4] Unfortunately, over time, most companies eventually do not pass these criteria and are dropped from the index. Particularly relevant to the present study, on 2 March 2007, Standard & Poor published the list of companies that have been in the S&P 500 index since March 1957. Remarkably, only 17% of the original constituent companies survived over 50 years. These long-run survivors represent less than 2% of all listed stocks in the U.S. To our knowledge, no previous study investigates both the performance and characteristics of these exceptionally successful companies.

Long-run survivors in this well-known market index are exceptional in terms of fulfilling strict criteria for membership. Survivor companies remained profitable despite many economic shocks that periodically occurred over time. The closest published study to the present work is Siegel and Schwartz (2006), who investigated the long-term returns of the original S&P 500 constituent companies from March 1957 to December 2003. The authors found that the buy-and-hold returns of the original 500 companies outperformed the returns on the continually updated S&P 500 index used by investment professionals. Their study contradicted earlier research by McKinsey & Company's Foster and Kaplan (2001), who documented that new companies added to the S&P 500 index generated higher returns than the original companies. In their study, the performance of three portfolios was examined: (1) a survivor portfolio of 125 original companies that remained intact (except possibly for a name change) from 1957 to 2003, (2) a portfolio of direct descendants consisting of the shares of companies in the survivors portfolio plus the shares issued by companies that acquired an original S&P 500 company, and (3) a portfolio of total descendants including all companies in the direct descendants portfolio and all the spin-offs and other stocks distributed by the companies in the portfolio of direct descendants. Their results indicated that differences in average returns between these portfolios were negligible.[5] Our approach to identifying survivor stocks differs from theirs by using those companies announced as survivor companies by Standard & Poor on 2 March 2007. Additionally, we expand their sample period beyond 2003 to encompass the 2008–2010 global financial crisis. In doing so, as proposed in Harvey et al. (2016), our analyses take into account multiple testing hurdles.

Numerous studies have investigated companies in the S&P 500 index. Chan et al. (2013) explored the long-term effects of S&P 500 index additions and deletions on sample stocks from 1962 to 2003. The authors documented significant long-term price increases for both categories of stocks, with deleted stocks outperforming added stocks. As already noted, they found that firms added to the S&P 500 index gain a competitive edge in terms of reductions in financial constraints and the cost of equity. Platikanova (2016) examined revisions in cash holdings and the market valuation of investment opportunities of 475 firms added to the S&P 500 in the 1980–2010 period. They found a larger decrease in cash for index inclusions in sectors with high financial dependence. Shankar and Miller (2006) investigated market reactions to S&P 600 index inclusions and deletions. They observed significant announcement effects in terms of price increases, trading volume, and institutional ownership. Afego (2017) provided an excellent literature survey on the effects of changes in stock index composition. He argued that the vast majority of studies in this research area focused on price and volume effects for S&P 500 index firms due to the enormous value of investment assets directly benchmarked to the index. The survey revealed that S&P 500 stocks face significant short-term price pressures due to exceptionally high trading volumes by tracker funds with an estimated $2.2 trillion in directly-linked funds. Finally, Chen and Lin (2018) documented that companies in the S&P 500 index gain a competitive advantage over non-S&P 500 industry competitors in terms of positive stock valuation effects at the expense of competitors. Index inclusion is associated with both a decrease in the cost of equity and an increase in capital investment for newly added firms. Our study contributes to this literature by focusing on what we can learn from those companies that survived in the S&P 500 over a long period of time.

Interestingly, we find that the risk-adjusted average excess return of our portfolio of survivor stocks is 5.16% per annum after controlling for the excess returns of the S&P 500 index. This finding supports Siegel and Schwartz (2006), who documented that the original S&P 500 constituent stocks outperformed the index. Our findings indicate that this outperformance is even more pronounced after controlling for market risk. Relative to the S&P 500 index, we find that survivors tend to be on average small-value stocks that exhibit high profitability and conservative capital investment. Moreover, survivor stocks' returns are negatively correlated with momentum returns, which suggests that their returns more closely mimic losers rather than winners in momentum portfolios. During the financial crisis of 2008 and 2009, survivor stocks earned higher profits, invested more aggressively in capital, and decreased in size over time relative to the S&P 500 index. Additionally, the value characteristic of survivor stocks appears to be sample specific to the post-financial crisis period. We infer that survivor companies were better able to withstand the stresses of economic downturns than other S&P 500 index firms.

Using Standard & Poor's announcement on March 2007 as a structural break, we further explore whether survivor companies thereafter experienced a decrease in performance as measured by their average excess returns until December 2019. Since the evidence does not support this pattern, our results again support Siegel and Schwartz (2006). We further investigate the volatility process of the survivor stock portfolio as opposed to the S&P 500 index. We find that the volatility of the survivor stocks portfolio responds less to shocks in the return-generating innovation process than the index. This finding is surprising given the small fraction of survivor stocks in the index. Moreover, we find that the portfolio of survivor stocks is less exposed to fat tails than the index, such that investors are less exposed to extreme events. Finally, replicating the analyses using publicly available data for survivors on Yahoo Finance, our results using CRSP data to construct the survivors portfolio are corroborated. Based on the empirical evidence, we conclude that survivor stocks are different from other stocks in the S&P 500 index with remarkable resilience to withstand economic downturns and coincident stock market collapses.

The next section describes the data. Section 3 discusses our methodological approach. Section 4 provides the empirical results. The last section concludes.

2. Data

On 2 March 2007, Standard & Poor, the world's leading index provider, released the list of survivor companies in the S&P 500 index from March 1957 to March 2007. The list is publicly available on the internet.[6] Interestingly, only 86 original constitute firms of this well-known market index survived over the past 50 years, which corresponds to 17.20% of the 500 original constitute firms.[7] We begin our data collection as follows: (i) use the survivor list of company names; (ii) search the corresponding stock ticker; and (iii) employ the CRSP database to match the stock ticker with corresponding stock returns.[8] For these 86 survivor companies, we identified available data associated with 92 stocks. Table A1 in the Appendix A lists the firm names. The number of stocks over time is plotted in Figure 1. Using survivor stocks, we compute an equal-weighted average portfolio denoted as the *all survivors portfolio* ($RET^{ALL}_{SURVIVOR}$).[9]

This figure illustrates the number of available survivor stock observations over time using the CRSP database. Additionally, we retrieve monthly data for the Fama and French (2018) risk factors (viz., six-factor model) and Treasury bill rate from Kenneth French's website. Since data for the profitability factor (RMW) and investment factor (CMA) are not available before July 1963, we download data series for the size (SMB) and value (HML) factors as well as RMW and CMA factors from July 1963 to November 2019. Table 1 provides descriptive statistics for portfolios $RET^{ALL}_{SURVIVOR}$, SMB, HML, RMW, CMA, and the S&P 500 index. As shown there, the average gross return of $RET^{ALL}_{SURVIVOR}$ is 40 basis points per month higher than the average gross return of the S&P 500 index. The survivor stock portfolio exhibits a monthly standard deviation of returns equal to 3.94%, which is slightly lower that of the S&P 500 index at 4.27%. Relevant to these comparisons, it is important to bear in mind that the number of stocks in $RET^{ALL}_{SURVIVOR}$ (viz., 67) is considerably lower than the S&P 500 index.

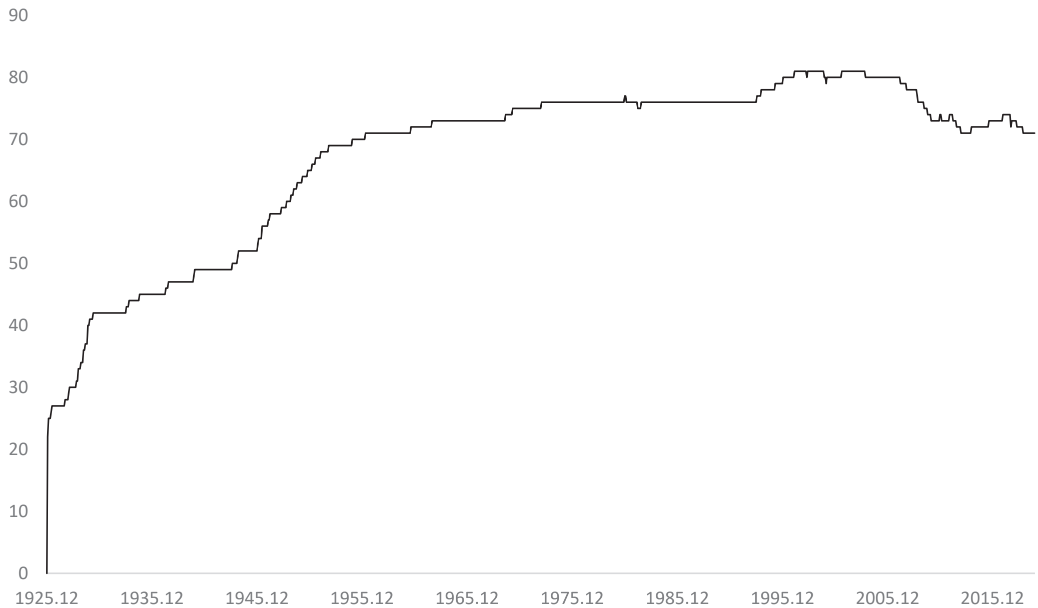

Figure 1. Evolution of survivor stocks in the sample period.

Table 1. Descriptive portfolio statistics.

	$RET^{ALL}_{SURVIVOR}$	*S&P 500*	*SMB*	*HML*	*RMW*	*CMA*	*UMD*
Mean	1.05	0.65	0.21	0.26	0.26	0.26	0.66
Median	1.28	0.91	0.09	0.25	0.22	0.11	0.71
Maximum	15.25	16.30	18.05	12.60	13.38	9.56	18.36
Minimum	−18.67	−21.76	−14.86	−14.11	−18.48	−6.86	−34.39
Std. dev.	3.94	4.27	3.02	2.87	2.17	1.99	4.19
Skewness	−0.36	−0.44	0.33	0.01	−0.33	0.32	−1.28
Kurtosis	5.22	4.87	6.02	5.39	15.38	4.61	13.19

This table reports the descriptive statistics of the survivor stocks portfolio, the S&P 500 index, and the Fama and French (2018) risk factors. The sample period is from July 1963 to December 2019. The figures are given in terms of percentages.

3. Statistical Analysis

3.1. Risk Adjustments and Survivor Stock Portfolio Characteristics

Here, we investigate the outperformance of the *all survivors portfolio* relative to the S&P 500 index. For this purpose, we regress $RET^{ALL,excess}_{SURVIVOR,t}$ on the excess returns of the S&P 500 index denoted $RET^{excess}_{S\&P500,t}$ as follows:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta \cdot RET^{excess}_{S\&P500,t} + u_t. \quad (1)$$

Table 2 reports the regression estimates, which confirm that survivor stocks outperformed the S&P 500 index by a large margin. $RET^{ALL,excess}_{SURVIVOR,t}$ generated an average return of 5.16% per annum in excess of $RET^{excess}_{S\&P500,t}$, with a t-statistic equal to 7.11 that is significant at any level.[10] The loading on $RET^{excess}_{S\&P500,t}$ is slightly less than unity, such that on average survivor stocks do not exhibit higher betas than the S&P 500 index.

Table 2. Regression estimates for all survivors using different asset pricing models.

Alpha	***S&P 500***	***SMB***	***HML***	***RMW***	***CMA***	***UMD***	R^2
0.43 *** (7.11)	0.86 *** (60.11)						0.84
0.32 *** (5.97)	0.89 *** (68.02)	0.08 *** (4.63)	0.26 *** (13.39)				0.88
0.21 *** (4.10)	0.92 *** (72.51)	0.15 *** (8.39)	0.16 *** (6.61)	0.26 *** (10.41)	0.20 *** (5.36)		0.90
0.26 *** (5.16)	0.91 *** (72.56)	0.15 *** (8.75)	0.13 *** (5.09)	0.27 *** (11.31)	0.22 *** (6.14)	−0.07 *** (−6.11)	0.90

*** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET^{ALL,excess}_{SURVIVOR,t}$ on the excess returns of the S&P 500 index as well and different asset pricing models. Ordinary *t*-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Can the outperformance of the survivor stocks be explained by exposures to the Fama and French (1993, 2015, 2018) risk factors? To address this question, we regress $RET^{ALL,excess}_{SURVIVOR,t}$ successively on the Fama and French (1993, 2015, 2018) three-, five-, and six-factor models defined, respectively, as follows:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + u_t, \tag{2}$$

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t \ldots + \beta_4 \cdot RMW_t + \beta_5 \cdot RMW_t + u_t \tag{3}$$

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t \ldots + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t. \tag{4}$$

The estimated regression results for these models are reported in rows two to four in Table 2. Several findings are worth noting. First, regardless of the asset pricing model, survivor stocks outperform the S&P 500 index. The economic magnitude of risk-adjusted returns, as measured by the regression intercepts, varies from 21 to 32 basis points per month with *t*-statistics between 4.10 and 5.97, indicating significance at any level. Predictably, the variation in the excess returns of the S&P 500 index explains 84% of the variation in the excess returns of the survivor stock portfolio. Controlling for various risk factors only marginally increases the *R*-squared value. Second, the positive loading on the size factor implies that the survivor stocks tend to be smaller stocks. However, this finding needs to be interpreted relative to the S&P 500 index; that is, survivor stocks are relatively smaller than the average index stock. Third, statistically significant exposures with respect to the value, profitability, and investment factors imply that survivor stocks tend to be value stocks that are profitable and invest conservatively. Fourth, and last, an unexpected finding is that the statistically significant loading on the momentum factor is negative in sign. Consequently, survivor stocks tend to have returns more correlated on average with loser than winner stocks. In view of the previously discussed S&P 500 index listing requirements, this finding is surprising.

To further investigate survivor stocks' characteristics relative to S&P 500 index companies, we employ a simultaneous equation model wherein $RET^{ALL,excess}_{SURVIVOR,t}$ and $RET^{excess}_{S\&P500,t}$ are modeled in the following system of equations:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha_1 + \beta_{1,1} \cdot RET^{excess}_{CRSP,t} + \beta_{1,2} \cdot SMB_t + \beta_{1,3} \cdot HML_t \ldots + \beta_{1,4} \cdot RMW_t + \beta_{1,5} \cdot CMA_t + \beta_{1,6} \cdot UMD_t + u_{1,t} \tag{5}$$

$$RET^{excess}_{S\&P500,t} = \alpha_2 + \beta_{2,1} \cdot RET^{excess}_{CRSP,t} + \beta_{2,2} \cdot SMB_t + \beta_{2,3} \cdot HML_t \ldots + \beta_{2,4} \cdot RMW_t + \beta_{2,5} \cdot CMA_t + \beta_{2,6} \cdot UMD_t + u_{2,t}. \tag{6}$$

Due to the high contemporaneous correlation between $RET^{ALL,excess}_{SURVIVOR,t}$ and $RET^{excess}_{S\&P500,t}$, we use seemingly unrelated regression (SUR) to estimate system (5) and (6). If a set of equations has contemporaneous cross-equation error correlation (i.e., the error terms in the regression equations are correlated), SUR addresses this issue by using a two-step estimation procedure that explicitly models the cross-equation error correlation. Since the correlation between $RET^{ALL,excess}_{SURVIVOR,t}$ and $RET^{excess}_{S\&P500,t}$ is manifested in $COV(u_{1,t},\ u_{2,t}) \neq 0$, SUR appears to be an adequate econometric model. In the previous analysis, we examined the average characteristics of survivor stocks relative to the underlying S&P 500 index. Using these equations, we conduct similar analyses but employ an overall market index proxied by the excess returns of the value-weighted CRSP index. The latter index is typically used in tests of the Fama and French (1993, 2015, 2018) asset pricing models.

The results are reported in Table 3. First, while the t-statistic associated with $\hat{\alpha}_1$ is statistically not different from zero, the t-statistic corresponding to $\hat{\alpha}_2$ is significantly negative at any statistical level. Hence, this evidence suggests that the outperformance of survivor stocks is driven by the underperformance of the S&P 500 index relative to the more general CRSP index. Second, the point estimator $\hat{\beta}_{2,2} = 0.16$ with a corresponding t-statistic of -30.46 confirms that stocks in the S&P 500 index tend to be large relative to those in the CRSP index. Since the t-statistic of $\hat{\beta}_{1,2}$ corresponding to -0.00 suggests that survivor stocks are not small stocks relative to the CRSP index, our evidence can only be interpreted to mean that survivor stocks are smaller relative to the average stock in the S&P 500. Third, even though the positive exposures with respect to *HML*, *RMW*, and *CMA* suggest that the average stock in the S&P 500 index tends to be a value firm that is profitable and invests conservatively, the exposures to these risk factors are very low in the range of only 0.02 to 0.06. By contrast, survivor stocks exhibit exposure with respect to these risk factors that are considerably larger in terms of their economic magnitudes with a range from 0.15 to 0.35 and t-statistics significant at any level.[11] Survivor stocks appear to perform better on all of these metrics. Fourth, and last, a surprising finding is that survivor stocks are, on average, considerably more exposed to loser stocks than the S&P 500. The exposure of the survivor stocks portfolio to the momentum factor is -0.09 as opposed to -0.02 for the S&P 500 index with respect to the momentum factor.

Table 3. Further asset pricing regression tests of all survivors.

Dependent var.	Alpha	$CRSP^{excess}$	*SMB*	*HML*	*RMW*	*CMA*	*UMD*	R^2
$RET^{ALL,excess}_{SURVIVOR}$	0.03	0.91 ***	−0.00	0.15 ***	0.33 ***	0.26 ***	−0.09 ***	0.91
	(0.60)	(74.79)	(−0.00)	(6.04)	(13.98)	(7.49)	(−7.70)	
$RET^{excess}_{S\&P500}$	−0.25 ***	1.00 ***	−0.16 ***	0.02 ***	0.06 ***	0.04 ***	−0.02 ***	0.99
	(−15.81)	(262.94)	(−30.46)	(3.08)	(7.89)	(3.48)	(−5.16)	

*** Statistically significant on a 1% level.

This table reports regresses $RET^{ALL,excess}_{SURVIVOR,t}$ and $RET^{excess}_{S\&P500}$ on the excess returns of the CRSP index as well as other risk factors in Fama and French's (2018) six-factor model. Ordinary t-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

3.2. Out-of-Sample Performance

It is important to recognize that our analysis incorporates information that the naïve investor did not know before March 2007 when Standard & Poor released the list of survivor S&P 500 index companies since March 1957. Here, we consider the out-of-sample question: What has been the performance of the survivor portfolio since March 2007? To explore whether survivors continued to outperform the S&P 500 index in the ex post announcement period, we add a binary dummy variable (denoted as x_t) to the regression models formulated in Equations (2)–(6) as follows:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + d \cdot x_t + \beta \cdot RET^{excess}_{S\&P500,t} + u_t \tag{7}$$

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + d \cdot x_t + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + u_t \tag{8}$$

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + d \cdot x_t + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t \ldots + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + u_t \tag{9}$$

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + d \cdot x_t + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t \ldots + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t \tag{10}$$

where x_t is a binary dummy variable with value equal to 0 in the pre-announcement period July 1963–March 2007 and 1 in the post-announcement period April 2007–November 2020. If risk-adjusted returns in the ex post March 2007 period, as measured by the sum $\alpha + d$, are statistically significantly lower, we expect that the t-statistic for parameter d will be significantly negative. The results in Table 4 show that the parameter estimate $\hat{d}$ is negative in all model specifications but with an economic magnitude close to zero and, in most model specifications, statistically not different from zero. These findings suggest that, even in the post-announcement period, survivors continued to outperform the S&P 500 index.

Table 4. Out-of-sample performance of all survivors.

Alpha	Dummy	*S&P 500*	*SMB*	*HML*	*RMW*	*CMA*	*UMD*	R^2
0.50 ***	−0.32 **	0.86 ***						0.84
(7.33)	(−2.22)	(60.33)						
0.35 ***	−0.12	0.89 ***	0.08 ***	0.25 ***				0.88
(5.67)	(−0.91)	(68.01)	(4.58)	(13.22)				
0.23 ***	−0.12	0.92 ***	0.15 ***	0.16 ***	0.26 ***	0.20 ***		0.90
(4.08)	(−1.00)	(72.52)	(8.34)	(6.51)	(10.41)	(5.37)		
0.30 ***	−0.19 *	0.91 ***	0.15 ***	0.12 ***	0.27 ***	0.22 ***	−0.08 ***	0.90
(5.33)	(−1.65)	(72.66)	(8.68)	(4.90)	(11.33)	(6.17)	(−6.25)	

* Statistically significant on a 10% level. ** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET^{ALL,excess}_{SURVIVOR,t}$ on the excess returns of the S&P 500 index as well as different asset pricing models. The regression models include a dummy variable denoted d with a value of 0 in the period from July 1963 to March 2007 and a value of 1 in the period April 2007–December 2019. Ordinary t-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

3.3. Time-Varying Factor Exposures

To better understand the risk determinants of survivor stocks' returns, we estimate the Fama and French (2018) six-factor model:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t, \quad (11)$$

where a 60 month window is used to estimate the parameter vector $\hat{\boldsymbol{\beta}}_t = (\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_6)$, and the estimation window is rolled forward one month at a time to the end of our sample period. This approach enables observation of trends over time in the estimated parameters.

In Figures 2–7, we report the time-varying factor exposures based on Fama and French's six-factor model from July 1968 to November 2020. From casual inspection of Figure 2, we see that the survivor portfolio's exposure to excess S&P 500 index returns is stable over time with beta close to unity. By contrast, Figure 3 shows that exposure to the size factor increases over time and exhibits noticeable volatility. We infer that, while survivor stocks in the S&P 500 index are large companies, as time passes, these companies become smaller with respect to this market index. This intertemporal pattern is consistent with Taleb (2012), who observed that " … in spite of what is studied in business schools concerning 'economics of scale', size hurts you at times of stress; it is not a good idea to be large during difficult times." (Taleb 2012, p. 279). Even though survivor companies engaged in mergers and acquisitions over time (Siegel and Schwartz 2006), our findings suggest that survivor companies grew smaller in size relative to the S&P 500 index in general. Visual inspection of Figure 2 shows a clear linear trend of the survivor stock portfolio's exposure against the size factor, which reaches its peak in May 2007. This peak occurred shortly before the early phase of the financial crisis starting in the beginning of August 2007 with the seizure in the banking system precipitated by BNP Paribas announcing that it was ceasing activity in three hedge funds operating with U.S. mortgage debt. When stock prices collapsed in the wake of the financial crisis, the market capitalization of those firms remained relatively stable. This pattern is implied by the sharp decrease from a positive exposure against the size factor in May 2007 until reaching its minimum with the economically largest negative exposure against the size factor in May 2012. A similar pattern can be seen after the stock market crises of 1972, 1987, and 1997. From an investment point of view, this finding suggests that survivor stocks may serve as safe havens in times of turmoil because their market capitalizations increase relative to the S&P 500.

Figure 2. Dynamic evolution of market factor beta exposure of survivors over time.

This figure plots the dynamic evolution of excess returns of the survivor stock portfolio's time-varying market beta exposure (i.e., β_1) based on the following regression equation:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t$$

where $RET^{excess}_{S\&P500,t}$ is the excess return on the S&P 500 index, and *SMB*, *HML*, *RMW*, *CMA*, and *UMD* are the risk factors in the Fama and French (2018) six-factor model. This model is estimated iteratively on a monthly basis using a rolling time window of 60 months. The sample is from July 1968 to December 2019.

This figure plots the dynamic evolution of excess returns of the survivor stock portfolio's time-varying size beta exposure (i.e., β_2) based on the following regression equation:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t$$

where $RET^{excess}_{S\&P500,t}$ is the excess return on the S&P 500 index, and *SMB*, *HML*, *RMW*, *CMA*, and *UMD* are the risk factors of the Fama and French (2018) six-factor model. This model is estimated iteratively on a monthly basis using a rolling time window of 60 months. The sample is from July 1968 to December 2019.

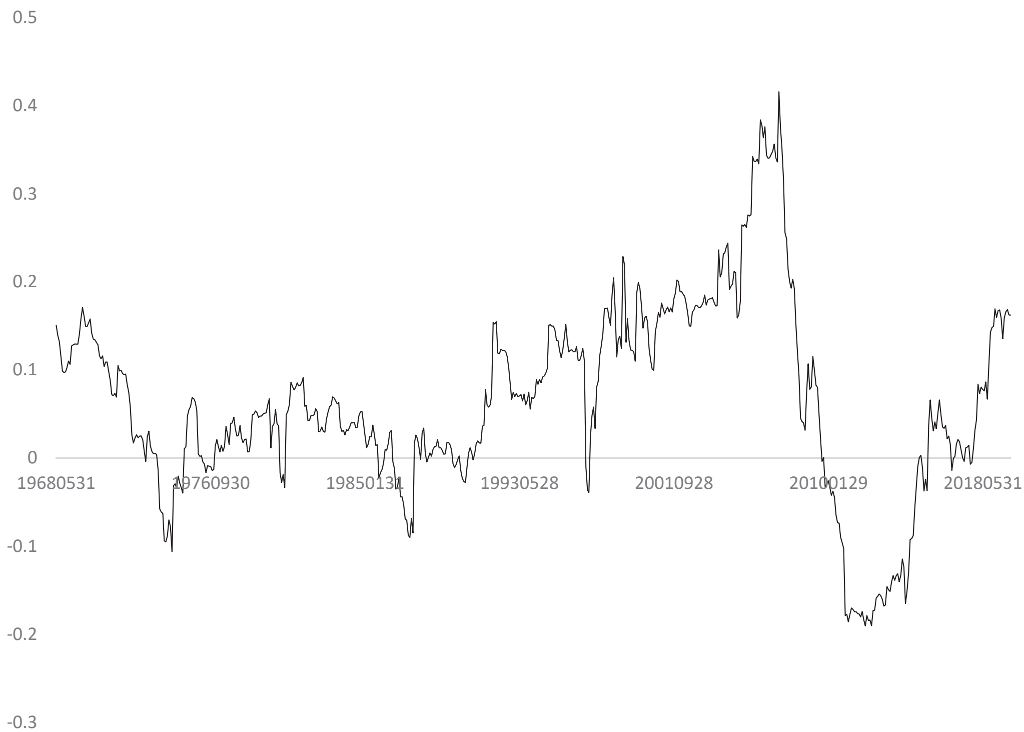

Figure 3. Dynamic evolution of size factor beta exposure of survivors over time.

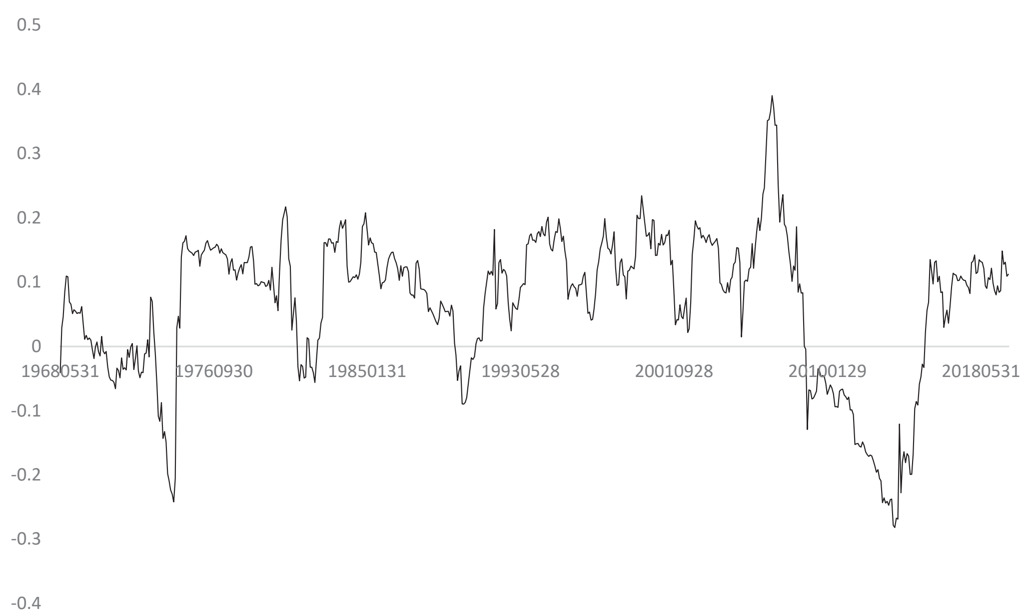

Figure 4. Dynamic evolution of value factor beta exposure for survivors over time.

Figure 5. Dynamic evolution of profitability factor beta exposure for survivors over time.

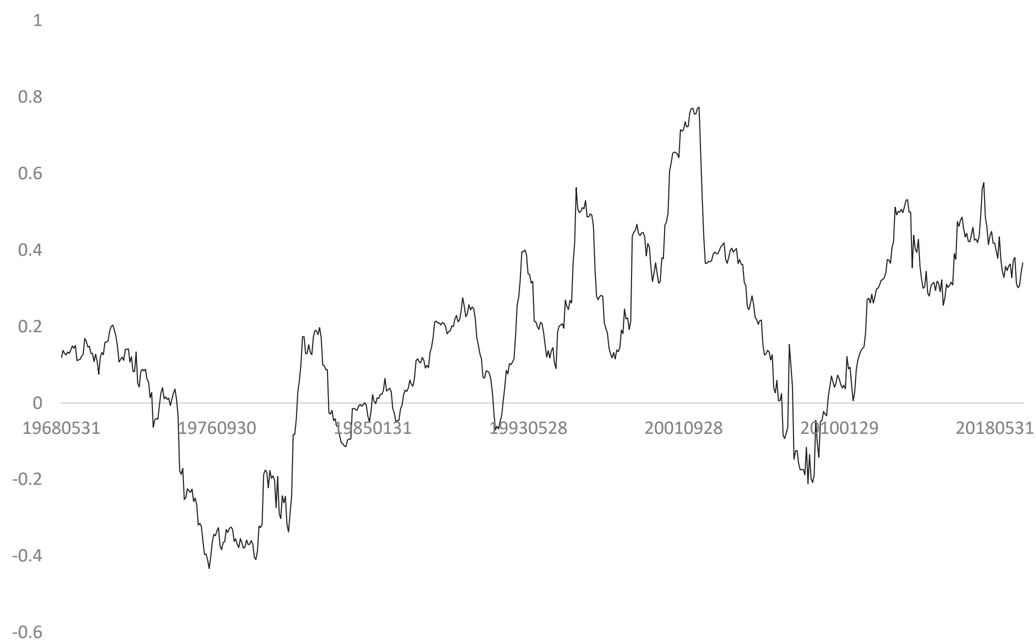

Figure 6. Dynamic evolution of investment factor beta exposure for survivors over time.

Figure 7. Dynamic evolution of momentum factor beta exposure for survivors over time.

Regardless of our findings in Table 2 that survivors tend to be value stocks on average, as shown in Figure 4, survivor companies experience dynamic changes in their value factor exposures. For example, from the early 1990s to 2007, survivor companies were on average value companies relative to the S&P 500 index.[12] However, in 2007 the loading on the value factor drops dramatically and thereafter continues to decrease throughout the financial crisis. From 2007 to 2015, survivor stocks were on average growth stocks. We observe a similar pattern in the mid-to-late 1980s. In sum, our findings suggest that survivor companies may be a safe haven in times of economic stress. These stocks benefit from long-term growth trends that are independent of economic cycles and tend to perform well in periods of economic downturns.

This figure plots the dynamic evolution of the excess returns of the survivor stock portfolio's time-varying value beta exposure (i.e., β_3) based on the following regression equation:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t$$

where $RET^{excess}_{S\&P500,t}$ is the excess return on the S&P 500 index, and *SMB*, *HML*, *RMW*, *CMA*, and *UMD* are the risk factors of the Fama and French (2018) six-factor model. This model is estimated iteratively on a monthly basis using a rolling time window of 60 months. The sample is from July 1968 to December 2019.

Extending our analyses, in Figure 5, we show that survivor companies become more profitable relative to S&P 500 index companies over time. The time-varying profitability factor exposure is on average negative until the end of the 1990s and thereafter increases and becomes positive on average.

This figure plots the dynamic evolution of excess returns of the survivor stock portfolio's time-varying profit beta exposure (i.e., β_4) based on the following regression equation:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t$$

where $RET^{excess}_{S\&P500,t}$ is the excess return on the S&P 500 index, and *SMB*, *HML*, *RMW*, *CMA*, and *UMD* are the risk factors of the Fama and French (2018) six-factor model. This model is estimated iteratively on a monthly basis using a rolling time window of 60 months. The sample is from July 1968 to December 2019.

Unlike the profitability factor, the time-varying exposure against the investment factor plotted in Figure 6 exhibits a pattern similar to the time-varying value factor. While survivor stocks invested more aggressively relative to S&P 500 index companies in the beginning of the sample period, they tended to invest more conservatively relative to S&P 500 index companies in the later years. In this regard, during the financial crisis starting in 2008, the average investment factor exposure was −0.11 from June 2007 to June 2009; by comparison, the average exposure of the S&P 500 companies was 0.36 in this period. We infer that survivor companies were profitable firms in good financial position, which enabled them to invest more aggressively than the average company in the S&P 500 index. Companies that are profitable in times of economic stress and increase investment are attractive from the perspective of investors. While these stocks were on average value stocks in the crisis period, they became growth companies after June 2009.

This figure plots the dynamic evolution of excess returns of the survivors stock portfolio's time-varying investment beta exposure (i.e., β_5) based on the following regression equation:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t$$

where $RET^{excess}_{S\&P500,t}$ is the excess return on the S&P 500 index, and *SMB*, *HML*, *RMW*, *CMA*, and *UMD* are the risk factors of the Fama and French (2018) six-factor model. This model is estimated iteratively on a monthly basis using a rolling time window of 60 months. The sample is from July 1968 to December 2019.

Lastly, in Figure 7, we plot the time-varying momentum factor exposure of the survivor portfolio. Surprisingly, we find that, for 69% of the sample observations, the exposure against the momentum factor is negative. This negative sign in conjunction with the increasingly positive loading for profitability suggests that survivor companies are stocks that generate returns mimicking stocks with low cumulative returns in the past 12 months despite high profitability levels. This finding is interesting in view of the fact that the profitability factor is positively correlated with momentum, which implies that profitable firms tend to be winner stocks. Our findings indicate that survivor firms are the exceptions as their returns co-move with profitable loser stocks.

This figure plots the dynamic evolution of excess returns of the survivor stock portfolio's time-varying momentum beta exposure (i.e., β_6) based on the following regression equation:

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha + \beta_1 \cdot RET^{excess}_{S\&P500,t} + \beta_2 \cdot SMB_t + \beta_3 \cdot HML_t + \beta_4 \cdot RMW_t + \beta_5 \cdot CMA_t + \beta_6 \cdot UMD_t + u_t$$

where $RET^{excess}_{S\&P500,t}$ is the excess return on the S&P 500 index, and *SMB*, *HML*, *RMW*, *CMA*, and *UMD* are the risk factors of the Fama and French (2018) six-factor model. This model is estimated iteratively on a monthly basis using a rolling time window of 60 months. The sample is from July 1968 to December 2019.

3.4. Conditional Volatility

Does uncertainty in the survivor stocks portfolio differ from S&P 500 stocks? Since the survivor stocks portfolio contains relatively few stocks that are small compared to S&P 500 stocks, one might expect that the survivor stocks exhibit more pronounced responses to volatility shocks and are more exposed to tail risks. To explore this issue, we estimate Exponential Generalized Conditional Heteroskedasticity (EGARCH) models for both the excess returns of S&P 500 and survivor stocks as follows:

$$R^{excess}_{i,t} = \mu_i + \epsilon_{i,t},$$

$$\epsilon_{i,t} = \zeta_{i,t}\sigma_{i,t},$$

where $R_{i,t}^{excess}$ is the excess return at time t, $i = \{S\&P\ 500_t,\ all\ survivors_t\}$, μ_i denotes the intercept term of the mean equation, and $\epsilon_{i,t}$ is the residual term at time t. The equation for the variance is:

$$ln\left(\sigma_{i,t}^2\right) = c_i + \alpha_i\left|\frac{\epsilon_{i,t-1}}{\sigma_{i,t-1}}\right| + \beta_i ln\left(\sigma_{i,t-1}^2\right) + \gamma_i\frac{\epsilon_{i,t-1}}{\sigma_{i,t-1}},$$

where $\sigma_{i,t}^2$ is the conditional variance at time t, and the parameter vector $\boldsymbol{\theta}_{BTC} = (\mu_i, c_i, \alpha_i, \beta_i, \gamma_i)$ is estimated using maximum-likelihood estimation. As observed earlier from Table 1, given that both return series exhibit high kurtosis, we assume that the innovation process follows a fat-tailed t-distribution (i.e., $\zeta_{BTC,t}|\boldsymbol{\Omega}_{t-1} \sim t(v)$ with v degrees of freedom).[13] Our sample period is from July 1963 to November 2020.

Table 5 reports our findings. First, using maximum-likelihood estimation accounting for fat-tailed data via the t-distribution supports our earlier finding—namely, the average excess returns of the survivor portfolio are economically larger than the average excess returns of the S&P 500 index. Second, the estimated alpha in the conditional variance equation of the S&P 500 index equals 0.20, which is almost twice as high as the corresponding alpha for the survivor portfolio. Both alpha estimates are statistically significant at the 5% level. This finding suggests that, despite their smaller size and numbers, the volatility of survivor stocks responds less than S&P 500 stocks to shocks in the data generating innovation process. Third, beta and gamma estimates in the variance equations are very close to each other for both portfolios, which suggests that both volatilities respond similarly to bad news in the data generating innovation process and to the long-run conditional variance. Fourth, we observe that portfolio returns of survivor stocks are less exposed to fat tails than other S&P 500 stocks, as the estimated degrees of freedom for Student's t-distribution is higher for the former portfolio. We infer that survivor stocks are less exposed to extreme events.

Table 5. Estimating volatility processes.

	μ	c	α	β	γ	v
S&P 500	0.39 *** (2.77)	0.11 (1.28)	0.20 *** (2.97)	0.90 *** (28.26)	−0.16 *** (−3.99)	9.53
All survivors	0.69 *** (5.23)	0.13 * (1.92)	0.11 ** (2.09)	0.92 *** (36.31)	−0.19 *** (−5.39)	10.95

* Statistically significant on a 10% level. ** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

This table reports the estimates for the EGARCH model with mean equation:

$$R_{i,t}^{excess} = \mu_i + \epsilon_{i,t}$$

$$\epsilon_{i,t} = \zeta_{i,t}\sigma_{i,t},$$

where $R_{i,t}^{excess}$ is the excess return of at time t, $i = \{S\&P\ 500_t,\ all\ survivors_t\}$, μ_i denotes the intercept term of the mean equation, and $\epsilon_{i,t}$ is the residual term at time t. The equation for the variance is:

$$ln(\sigma_{i,t}^2) = c_i + \alpha_i\left|\frac{\epsilon_{i,t-1}}{\sigma_{i,t-1}}\right| + \beta_i ln\left(\sigma_{i,t-1}^2\right) + \gamma_i\frac{\epsilon_{i,t-1}}{\sigma_{i,t-1}},$$

where $\sigma_{i,t}^2$ is the conditional variance at time t, and the parameter vector $\boldsymbol{\theta}_{BTC} = (\mu_i, c_i, \alpha_i, \beta_i, \gamma_i)$ is estimated using maximum-likelihood estimation. The models assume that the innovation process follows a fat-tailed t-distribution (i.e., $\zeta_{BTC,t}|\boldsymbol{\Omega}_{t-1} \sim t(v)$ with v degrees of freedom). The z-statistics are given in parentheses. The sample period is from July 1963 to December 2019.

4. Robustness Checks

4.1. The Multiple Testing Problem

Using a multiple testing framework to derive threshold levels for testing statistical significance, Harvey et al. (2016) re-evaluated 296 cross-sectional asset pricing phenomena. Their findings showed that 27% to 53% are likely false discoveries. Following these authors, we use the higher cut-off corresponding to 3.39 for testing statistical significance. Our main results in Tables 2–4 remain statistically significant. For instance, irrespective of which asset pricing model is used for risk adjusting the survivor portfolio, the regression intercepts reported in Table 2 exceed 3.39 by a large margin. In this respect, the lowest *t*-statistic of 4.10 is generated when using the Fama and French (2015) five-factor model.[14]

4.2. Replication Using Publicly Available Data

Hou et al. (2020), who conducted an extensive replication of 452 asset pricing anomalies, found that approximately 80% of these anomalies fail scientific replication. Subsequently, the authors recommended scientific replications of test results. To address this issue, we replicate our analyses using publicly available data from Yahoo.[15] Matching the data from Standard & Poor's announcement on 2 March 2007 with the database provided by Yahoo, we find data available for 71 stocks. Among these stocks, we excluded Raytheon Technologies Corporation (RTX) due to extreme outliers in the sample period.[16] Descriptive statistics for the final sample of 70 survivor stocks are shown in Appendix A Table A2. In Appendix A Figure A1, we plot the evolution of available survivor stocks over time (i.e., survivors are added as Yahoo Finance stock data becomes available from the earliest date of February 1962). Among these stocks, 14 survivor stocks had complete return series available from February 1962 to November 2020.[17] The final sample of 70 observations are used to form the replicated *all survivors portfolio* (denoted $RET^{ALL}_{SURVIVOR}$) which is equal weighted. As before, we retrieved data for the Fama and French (2018) risk factors (viz., six-factor model) and Treasury bill rate from Kenneth French's website. Since data for the profitability factor (*RMW*) and investment factor (*CMA*) are not available before July 1963, we retrieve data for the size factor (*SMB*), value factor (*HML*), *RMW*, and *CMA* from July 1963 to December 2019. Descriptive statistics for portfolio $RET^{ALL}_{SURVIVOR}$, *SMB*, *HML*, *RMW*, *CMA*, and the S&P 500 index are provided in Table A3. Next, we run the same regressions as in Equations (1)–(4). The results are reported in Table A4. Again survivor stocks outperformed the S&P 500 index by a considerable margin. $RET^{ALL,excess}_{SURVIVOR,t}$ generated an average return of 5.88% per annum in excess of $RET^{excess}_{S\&P500,t}$, with a *t*-statistic equal to 7.01 that is significant at any level.

The loading on $RET^{excess}_{S\&P500,t}$ is slightly less than unity, such that on average survivor stocks do not exhibit higher betas than the S&P 500 index. Employing different variations of the Fama and French (1993, 2015, 2018) does not change our results. Here, the economic magnitudes of risk-adjusted returns, as measured by the regression intercepts, varies from 27 to 40 basis points per month with *t*-statistics between 4.50 and 6.31 indicating statistical significance at any level. Note that variation in the excess returns of the S&P 500 index explains 83% of the variation in the excess returns of the survivor stock portfolio. Controlling for various risk factors negligibly increases the *R*-squared value. Second, the positive loading on the size factor implies that the survivor stocks tend to be smaller stocks. Third, statistically significant exposures with respect to the value, profitability, and investment factors imply that our replicated portfolio of survivor stocks is, on average, exposed to value stocks that are profitable and invest conservatively. Finally, we find a statistically significantly negative loading on the momentum factor. As a consequence, survivor stocks tend to have returns more correlated on average with loser than winner stocks. In sum, our results strongly confirm the key findings of our previous analysis based on CRSP data.

4.3. Survivor Stocks' Characteristics Relative to S&P 500 Index

Next, to further investigate survivor stocks' characteristics relative to S&P 500 index companies, we employ the simultaneous equation model in Equations (5) and (6). Based on SUR econometric estimation, the results are reported in Table A5. First, while the *t*-statistic associated with $\hat{\alpha}_1$ is statistically not different from zero, and the *t*-statistic corresponding to $\hat{\alpha}_2$ is significantly negative at any statistical level. This result implies that the outperformance of survivor stocks is driven by the underperformance of the S&P 500 index relative to the more general CRSP index. Second, the *t*-statistic of $\hat{\beta}_{1,2}$ equal to -1.81 suggests that survivor stocks are not small relative to the CRSP index, implying that survivor stocks are smaller relative to the average stock in the S&P 500. Third, survivor stocks exhibit exposures with respect to *HML*, *RMW*, and *CMA* that are considerably larger than the ones of the S&P 500′s in terms of their economic magnitudes with a range from 0.15 to 0.35 with *t*-statistics significant at any level. We infer survivor stocks appear to perform better on all of these metrics. Fourth, and last, survivor stocks are, on average, considerably more exposed to loser stocks than the S&P 500. The exposure of the survivor stocks portfolio to the momentum factor is -0.12 as opposed to -0.02 for the S&P 500 index with respect to the momentum factor. Again, our replicated portfolio of survivor stocks strongly supports the key results of our main analysis.

Further, we address the question: What has been the performance of the survivor portfolio since March 2007? To investigate if our replicated portfolio of survivors continued to outperform the S&P 500 index in the ex post announcement period, we again employ Equations (7) to (10) using our replicated survivor portfolio drawn from the Yahoo database. The results are reported in Table A6. Once again, the parameter estimate $\hat{d}$ is negative in all model specifications but with an economic magnitude close to zero and statistically not different from zero. These findings suggest that, even in the post-announcement period, our replicated portfolio of survivors continued to outperform the S&P 500 index. Finally, we explore the conditional volatility of our replicated portfolio of survivor stocks. The results, as reported in Table A7, clearly support earlier evidence in Section 3.4.

4.4. Equal-Weighted Market Factor

As mentioned earlier, we have valid reasons to use equal-weighted portfolios in our current study. Value weighting would distort the overall portfolio return distribution because market capitalization as a financial variable is pareto distributed, implying that if value-weighted portfolios were used, a very small number of stocks would receive extraordinarily high weights. In our study, we are interested in the revealing potential common links among survivor stocks. A valid question that may arise is, however, could the outperformance of survivor stocks be an artefact of using equal-weighted stocks in the portfolio? To explore this issue, we download 49 equal-weighted industrial portfolios from Kenneth French website, compute the simple average return and subtract the U.S. risk free rate.[18] We use this portfolio as proxy for an equal-weighted U.S. market factor in excess form. Again, we make use of a multiple equation model as in Section 3, that is, we estimate

$$RET^{ALL,excess}_{SURVIVOR,t} = \alpha_1 + \beta_{1,1}\cdot RET^{excess}_{EQUAL,t} + \beta_{1,2}\cdot SMB_t + \beta_{1,3}\cdot HML_t + \beta_{1,4}\cdot RMW_t + \beta_{1,5}\cdot CMA_t + \beta_{1,6}\cdot UMD_t + u_{1,t} \quad (12)$$

$$RET^{excess}_{S\&P500,t} = \alpha_2 + \beta_{2,1}\cdot RET^{excess}_{EQUAL,t} + \beta_{2,2}\cdot SMB_t + \beta_{2,3}\cdot HML_t + \beta_{2,4}\cdot RMW_t + \beta_{2,5}\cdot CMA_t + \beta_{2,6}\cdot UMD_t + u_{2,t} \quad (13)$$

where $RET^{excess}_{EQUAL,t}$ is our proxy for an equal-weighted U.S. market factor in excess form and all other notation is as before. The results are reported in Table A8. We observe from Table A8. that the survivor stocks portfolio generates a risk-adjusted payoff of 27 basis points per month, whereas the S&P 500 underperforms the equal-weighted portfolio by 28 basis points per month. Testing the parameter difference $(\hat{\alpha}_1 - \hat{\alpha}_2) = 0.56$ for statistical significance, gives us a value of 57.03 for the estimated test statistic. Since the test statistic is under the null hypothesis distributed as chi-square with one degree of freedom with corresponding critical value of 3.84 for a 5% significance level, we can reject the null hypotheses (*p*-value 0.0000). Hence, the outperformance of the survivor stock portfolio is not driven by equal-weighting the stocks in the survivor portfolio. Another interesting issue which we observe from Table A8 is that the loading against the size factor is less negative for the survivor stocks portfolio even after controlling for the equal-weighted excess market factor. Next, testing the parameter difference $(\hat{\beta}_{1,6} - \hat{\beta}_{2,6}) = -0.10$ for statistical significance gives us a value of 33.12 for the estimated test statistic. Since the test statistic is under the null hypothesis distributed as chi-square with one degree of freedom with corresponding critical value of 3.84 for a 5% significance level, we can reject the null hypotheses (p-value 0.0000). Hence, the survivor stock portfolio is relatively less exposed to winner stocks than the average S&P 500 firm which confirms our earlier findings.

4.5. Additional Robustness Checks

In our analysis, we followed the mainstream literature in using ordinary t-statistics (e.g., Fama and French 2015, 2017, 2018). One may wonder whether our results hold when accounting for heteroskedasticity and autocorrelation consistent (HAC) t-statistics.[19] To address this issue, we employ the HAC covariance matrix estimator proposed from Newey and West (1987) accounting for a lag order of $l = 1$ and replicate the main results from Table 4. The results are reported in Table A9. We observe from Table A9 that the results do not change. Hence, we infer that our results are robust with respect to potential autocorrelation and heteroskedasticity in the data.

5. Conclusions

On 2 March 2007, Standard & Poor released a list of companies that had been in the S&P 500 index since March 1957. Over this 50 year period, only 86 companies survived index membership requirements. Companies listed in the S&P 500 index are special in the sense that they are leading companies influential to the U.S. and global economies. A number advantages accrue to members, including reductions in financial constraints, the cost of equity, and other shadow costs, among others. Due to relatively high hurdles for membership, most companies drop out of the index over time. This study sought to investigate the performance and characteristics of survivor stocks in the S&P 500 index. Due to data availability, our survivor stocks covered the period from July 1963 to December 2019.

We found that survivor stocks outperformed the S&P 500 index by a large margin in this sample period. Their outperformance was unchanged after taking into account checks revealed that this phenomenon is not sample period specific. Relative to S&P 500 companies, survivor stocks tend to be, on average, small-value stocks that exhibit high profitability and invest conservatively. A surprising finding was that survivor stocks also tend to be loser stocks with negative exposure to the momentum factor. Further analyses revealed that survivor stocks decreased in size over time relative to other S&P 500 companies. Additionally, the value characteristic of survivor stocks shifted to be consistent with growth in periods of economic distress. Unlike other index stocks, survivors were relatively profitable and increased capital investments in times of economic stress. In this regard, survivors' returns were less exposed to fat tails than other S&P 500 stocks. Further analyses revealed that the survivor stock portfolio outperformed the S&P 500 index even in the post-March 2007 period after the public announcement by Standard & Poor's list of 50 year survivor companies. Additionally, replicating the survivor portfolio using publicly available data corroborated our findings. We conclude that survivor stocks are different from other stocks in the S&P 500 index, with remarkable resilience to withstand economic downturns and coincident stock market collapses. Comparative research is recommended on survivor stocks in other major stock markets around the world. Are survivor characteristics local or global in nature? Moreover, future research is encouraged to explore the return evolution for firms exiting the S&P 500. Since this is beyond the scope of this paper, this issue is left for future research.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

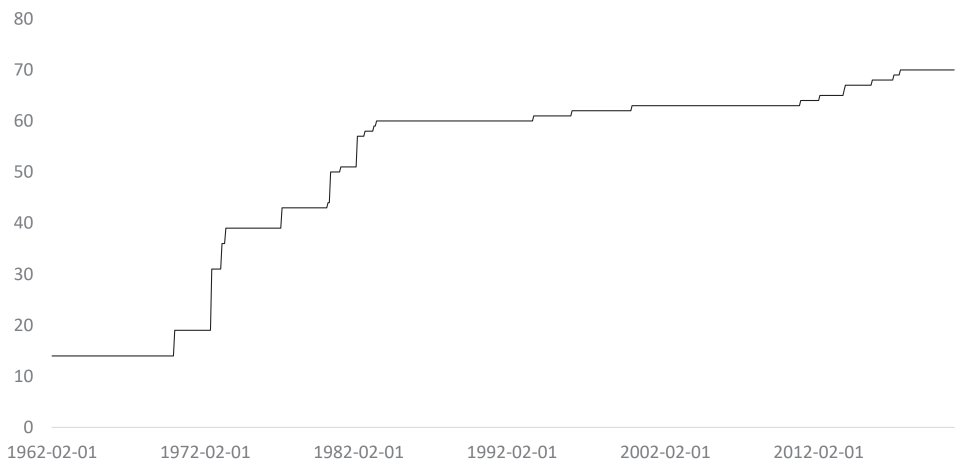

Figure A1. Evolution of survivor stocks in the sample period.

This figure illustrates the number of available survivor stock observations over time using the Yahoo database.

Table A1. Survivor firms based on the CRSP database.

1	AMERICAN WATER WORKS & ELEC INC	56	AMERICAN TYPE FOUNDERS INC
2	WEST PENN ELECTRIC CO	57	ATF INC
3	ALLEGHENY POWER SYSTEMS INC	58	DAYSTROM INC
4	ALLEGHENY ENERGY INC	59	SCHLUMBERGER LTD
5	ALLIED CHEMICAL & DYE CORP	60	STANDARD OIL CO CALIFORNIA
6	ALLIED CHEMICAL CORP	61	CHEVRON CORP
7	ALLIED CORP	62	CHEVRONTEXACO CORP
8	ALLIED SIGNAL INC	63	CHEVRON CORP NEW
9	HONEYWELL INTERNATIONAL INC	64	UNION TANK CAR CO
10	ARCHER DANIELS MIDLAND CO	65	TRANS UNION CORP
11	BURROUGHS ADDING MACH CO	66	UNITED STATES STEEL CORP
12	BURROUGHS CORP	67	USX CORP
13	UNISYS CORP	68	U S X MARATHON GROUP
14	COCA COLA CO	69	MARATHON OIL CORP
15	CONSOLIDATED GAS CO NY	70	KRAFT HEINZ CO
16	CONSOLIDATED EDISON CO NY INC	71	WRIGLEY WILLIAM JR CO
17	CONSOLIDATED EDISON INC	72	AMERICAN HOME PRODUCTS CORP
18	DETROIT EDISON CO	73	WYETH
19	D T E ENERGY CO	74	SOUTHERN CALIFORNIA EDISON CO
20	DU PONT E I DE NEMOURS & CO	75	SCE CORP
21	EATON AXLE & SPRING CO	76	EDISON INTERNATIONAL
22	EATON MFG CO	77	ALCOA CORP
23	EATON YALE & TOWNE INC	78	GOODYEAR TIRE & RUBBER CO
24	EATON CORP	79	HERSHEY CHOCOLATE CORP
25	EATON CORP PLC	80	HERSHEY FOODS CORP
26	STANDARD OIL CO N J	81	HERSHEY CO
27	EXXON CORP	82	KROGER GROCERY & BAKING CO
28	EXXON MOBIL CORP	83	KROGER COMPANY
29	ELECTRIC BOAT CO	84	DOWDUPONT INC

Table A1. *Cont.*

30	GENERAL DYNAMICS CORP	85	DUPONT DE NEMOURS INC
31	GENERAL ELECTRIC CO	86	MELVILLE SHOE CORP
32	GENERAL MOTORS CORP	87	MELVILLE CORP
33	GENERAL MOTORS CO	88	CVS CORP
34	INGERSOLL RAND CO	89	CVS CAREMARK CORP
35	INGERSOLL RAND CO LTD	90	CVS HEALTH CORP
36	INGERSOLL RAND PLC	91	GENERAL MILLS INC
37	TRANE TECHNOLOGIES PLC	92	MCGRAW HILL PUBLISHING INC
38	INTERNATIONAL BUSINESS MACHS COR	93	MCGRAW HILL INC
39	FORTUNE BRANDS HOME & SECUR INC	94	MCGRAW HILL COS INC
40	TRANSCONTINENTAL OIL CO	95	MCGRAW HILL FINANCIAL INC
41	OHIO OIL CO	96	S&P GLOBAL INC
42	MARATHON OIL CO	97	KIMBERLY CLARK CORP
43	PACIFIC GAS & ELEC CO	98	PHELPS DODGE CORP
44	PG & E CORP	99	HERCULES POWDER CO
45	LOFT INC	100	HERCULES INC
46	PEPSI COLA CO	101	MINNEAPOLIS HONEYWELL REGULATOR
47	PEPSICO INC	102	HONEYWELL INC
48	PHILIP MORRIS & CO LTD	103	PENNEY J C INC
49	PHILIP MORRIS INC	104	PENNEY J C CO INC
50	PHILIP MORRIS COS INC	105	COMMONWEALTH & SOUTHERN CORP
51	ALTRIA GROUP INC	106	SOUTHERN CO
52	PHILLIPS PETROLEUM CO	107	CATERPILLAR TRACTOR INC
53	CONOCOPHILLIPS	108	CATERPILLAR INC
54	EASTMAN KODAK CO	109	COLGATE PALMOLIVE PEET CO
55	AMERICAN TYPE FOUNDERS CO	110	COLGATE PALMOLIVE CO
111	DEERE & CO IL	168	PITNEY BOWES INC
112	DEERE & CO DEL	169	TEXAS UTILITIES CO
113	DEERE & CO	170	TXU CORP
114	BRISTOL MYERS CO	171	ALUMINUM COMPANY AMER
115	BRISTOL MYERS SQUIBB CO	172	ALCOA INC
116	BOEING AIRPLANE CO	173	ARCONIC INC
117	BOEING CO	174	HOWMET AEROSPACE INC
118	ABBOTT LABS	175	NORTHROP AIRCRAFT INC
119	ABBOTT LABORATORIES	176	NORTHROP CORP
120	DOW CHEMICAL CO	177	NORTHROP GRUMMAN CORP
121	LOCKHEED AIRCRAFT CORP	178	RAYTHEON MANUFACTURING CO
122	LOCKHEED CORP	179	RAYTHEON CO
123	LOCKHEED MARTIN CORP	180	CAMPBELL SOUP CO
124	WEST VA PULP & PAPER CO	181	FORD MOTOR CO
125	WESTVACO CORP	182	FORD MOTOR CO DEL
126	MEADWESTVACO CORP	183	COOPER TIRE & RUBBER CO
127	WESTROCK CO	184	OCCIDENTAL PETROLEUM CORP
128	INTERNATIONAL PAPER & PWR CO	185	UNION PACIFIC CORP
129	INTERNATIONAL PAPER CO	186	BURLINGTON NORTHERN INC
130	PHILADELPHIA ELECTRIC CO	187	BURLINGTON NORTHERN SANTA FE CP
131	P E C O ENERGY CO	188	SEALED AIR CORP
132	EXELON CORP	189	CSX CORP
133	PFIZER CHAS & CO INC	190	NORFOLK SOUTHERN CORP
134	PFIZER INC	191	ALLSTATE CORP
135	COOPER BESSEMER CORP	192	SANTA FE FINANCIAL CORP
136	COOPER INDUSTRIES INC	193	NGC CORP
137	COOPER INDUSTRIES LTD	194	DYNEGY INC
138	COOPER INDUSTRIES PLC	195	ITT HARTFORD GROUP INC
139	PITTSBURGH PLATE GLASS CO	196	HARTFORD FINANCIAL SVCS GRP INC
140	P P G INDUSTRIES INC	197	QUEST DIAGNOSTICS INC
141	MINNESOTA MINING & MFG CO	198	SEALED AIR CORP NEW
142	3M CO	199	ROCKWELL COLLINS INC

Table A1. *Cont.*

143	MERCK & CO INC	200	DYNEGY INC NEW
144	MERCK & CO INC NEW	201	DYNEGY INC DEL
145	GALVIN MANUFACTURING CO	202	DYNEGY INC NEW DEL
146	MOTOROLA INC		
147	MOTOROLA SOLUTIONS INC		
148	CAROLINA POWER & LIGHT CO		
149	CP & L ENERGY INC		
150	PROGRESS ENERGY INC		
151	CONSUMERS PWR CO		
152	CONSUMERS POWER CO		
153	C M S ENERGY CORP		
154	PUBLIC SERVICE ELECTRIC & GAS CO		
155	PUBLIC SERVICE ENTERPRISE GP INC		
156	HALLIBURTON OIL WELL CEMENTING		
157	HALLIBURTON COMPANY		
158	NORTHERN STATES POWER CO MN		
159	XCEL ENERGY INC		
160	MIDDLE SOUTH UTILITIES INC		
161	ENTERGY CORP		
162	ENTERGY CORP NEW		
163	AMERICAN GAS & ELECTRIC CO		
164	AMERICAN ELECTRIC POWER CO INC		
165	CONSOLIDATED GAS ELEC LT & PWR		
166	BALTIMORE GAS & ELECTRIC CO		
167	CONSTELLATION ENERGY GROUP INC		

This table reports the firms of the corresponding survivor stocks.

Table A2. Descriptive statistics for survivor stocks.

Ticker/Metric	AA	ABT	ADM	AEP	ALL	ATI	BA	BMY	BURL	CAT
Mean	0.90	1.47	1.09	0.92	1.11	1.29	1.44	1.15	2.97	1.36
Median	0.73	1.36	1.09	1.10	1.05	0.36	1.45	1.11	3.63	1.17
Maximum	54.02	22.12	32.08	28.70	30.97	62.50	48.44	43.72	25.24	40.14
Minimum	−55.59	−20.74	−27.36	−17.77	−42.78	−50.26	−45.47	−28.87	−26.73	−35.91
Std. dev.	9.98	5.97	7.89	5.73	7.59	16.38	9.59	6.97	8.51	8.40
Skewness	−0.02	−0.15	0.12	0.03	−0.72	0.60	0.23	0.08	−0.49	0.03
Kurtosis	7.32	3.65	3.98	4.09	8.28	4.71	5.60	5.86	4.08	4.65
Sample start	1962-02	1980-04	1980-04	1970-02	1993-07	1999-12	1962-02	1972-07	2013-11	1962-02
Ticker/Metric	**CL**	**CMS**	**COO**	**COP**	**CPB**	**CSX**	**CVS**	**CVX**	**DD**	**DE**
Mean	1.15	0.65	1.46	1.10	0.94	1.45	1.18	1.28	1.33	1.41
Median	1.21	0.90	1.05	1.36	0.94	1.55	0.75	1.34	0.66	1.26
Maximum	49.25	41.77	88.45	39.92	33.00	29.11	56.86	36.30	182.16	45.30
Minimum	−21.59	−44.17	−52.59	−35.94	−18.76	−31.41	−36.36	−21.46	−67.77	−29.86
Std. dev.	6.74	8.30	13.69	8.29	6.75	7.87	8.51	6.74	12.27	8.44
Skewness	0.76	−0.12	0.56	0.20	0.26	−0.14	0.41	0.35	5.96	0.08
Kurtosis	8.89	9.87	8.26	6.00	4.35	3.89	7.96	4.99	89.80	4.53
Sample start	1973-06	1973-03	1983-02	1982-01	1973-03	1980-12	1973-02	1962-02	1972-07	1972-07
Ticker/Metric	**DTE**	**ED**	**EIX**	**ETN**	**ETR**	**EXC**	**F**	**GD**	**GDP**	**GE**
Mean	0.95	0.98	1.10	1.96	0.98	1.04	1.17	1.44	0.94	0.90
Median	0.95	0.97	1.30	1.98	0.96	1.10	0.53	1.17	0.64	0.37
Maximum	54.18	45.00	25.92	72.89	39.24	30.69	127.38	34.00	99.77	37.20
Minimum	−22.41	−52.50	−36.90	−30.33	−24.54	−24.14	−57.88	−27.95	−33.17	−29.84
Std. dev.	5.70	6.07	6.65	8.40	6.84	6.47	11.04	8.10	18.54	7.39
Skewness	1.03	−0.17	−0.69	0.97	0.53	0.08	2.59	0.23	2.98	0.20
Kurtosis	13.92	16.22	6.72	12.15	6.58	4.78	33.33	4.74	18.24	5.39
Sample start	1962-02	1962-02	1973-06	1972-07	1972-07	1973-06	1972-07	1977-02	2017-01	1962-02

Table A2. *Cont.*

Ticker/Metric	GIS	GM	GT	HAL	HIG	HON	HSY	IBM	IP	IR
Mean	1.24	0.78	0.81	1.08	1.29	1.27	1.48	0.84	0.91	2.72
Median	1.12	0.33	0.37	1.24	1.39	1.40	1.11	0.36	0.50	1.85
Maximum	19.75	28.10	75.56	54.90	103.57	51.05	27.60	35.38	79.83	26.70
Minimum	−24.08	−31.87	−41.74	−59.61	−74.82	−38.19	−24.91	−24.86	−37.61	−24.37
Std. dev.	5.65	9.06	10.77	10.83	12.94	8.12	6.45	6.94	8.60	10.65
Skewness	0.08	0.18	0.60	−0.15	1.43	0.06	0.20	0.19	1.04	0.12
Kurtosis	3.99	4.26	8.13	6.24	22.84	6.97	4.85	4.83	13.67	3.31
Sample start	1980-04	2010-12	1962-02	1972-07	1996-01	1970-02	1983-04	1962-02	1962-02	2017-06
Ticker/Metric	**JCPNQ**	**KHC**	**KMB**	**KO**	**KODK**	**KR**	**LMT**	**MMM**	**MO**	**MRK**
Mean	−0.15	−0.49	1.22	1.56	9.41	2.04	1.70	0.98	1.81	1.14
Median	−0.26	0.12	0.85	1.33	−1.60	1.62	1.33	1.10	1.91	1.05
Maximum	52.91	24.70	33.21	33.22	879.82	214.90	48.26	25.80	195.19	31.34
Minimum	−47.83	−30.94	−17.10	−29.55	−72.63	−67.84	−38.81	−27.83	−69.65	−26.62
Std. dev.	11.83	9.02	5.86	6.38	97.98	13.48	8.89	6.06	10.71	6.89
Skewness	0.05	−0.45	0.82	0.18	8.23	7.80	0.22	−0.01	7.94	−0.05
Kurtosis	5.44	4.51	6.31	5.59	73.33	124.80	6.38	4.68	154.67	4.00
Sample start	1973-03	2015-08	1980-04	1962-02	2013-10	1977-02	1977-02	1970-02	1962-02	1970-02
Ticker/Metric	**MRO**	**MSI**	**NOC**	**NSG**	**OXY**	**PBI**	**PCG**	**PEP**	**PFE**	**PG**
Mean	0.92	1.25	1.53	1.36	0.98	1.06	0.85	1.20	1.15	1.04
Median	0.44	1.10	1.67	1.31	0.76	0.54	1.13	0.93	1.05	0.79
Maximum	86.02	30.73	33.88	25.53	72.62	73.04	45.71	36.89	39.67	24.69
Minimum	−60.08	−33.49	−35.56	−31.52	−64.63	−48.66	−45.26	−28.41	−24.01	−35.42
Std. dev.	10.55	9.56	8.30	7.55	9.63	10.06	8.45	6.34	7.01	5.48
Skewness	0.91	−0.14	−0.04	−0.07	0.64	0.64	−0.39	0.07	0.28	−0.34
Kurtosis	12.41	3.64	5.16	4.20	14.90	11.64	11.37	6.76	4.77	6.10
Sample start	1970-02	1977-02	1982-02	1982-07	1982-02	1972-07	1972-07	1972-07	1972-07	1962-02
Ticker/Metric	**PPG**	**PREX**	**ROK**	**SEE**	**SLB**	**SO**	**UIS**	**UNP**	**XEL**	**XOM**
Mean	1.39	0.89	1.84	1.69	0.77	1.24	0.81	1.35	0.91	0.98
Median	1.42	0.00	1.93	2.00	0.50	1.28	0.06	1.48	1.18	0.93
Maximum	26.71	52.50	160.92	149.35	39.16	22.57	130.19	34.44	42.13	22.69
Minimum	−32.32	−63.83	−57.90	−63.97	−49.47	−14.26	−55.92	−33.43	−58.50	−25.13
Std. dev.	7.29	10.88	11.83	11.41	9.28	4.99	16.10	7.49	6.07	5.31
Skewness	0.06	−0.53	4.96	3.98	−0.20	0.02	1.30	0.05	−0.96	0.03
Kurtosis	4.58	18.46	73.04	61.54	5.76	3.89	12.92	4.77	22.91	4.51
Sample start	1980-04	2012-03	1982-01	1980-04	1982-01	1982-01	1972-08	1980-02	1973-03	1962-02

This table reports the descriptive statistics for all available data on survivor stocks. The data are downloaded from Yahoo.com and sorted in alphabetical order.

Table A3. Descriptive portfolio statistics for the scientific replication.

	$RET^{ALL}_{SURVIVOR}$	*S&P 500*	*SMB*	*HML*	*RMW*	*CMA*	*UMD*
Mean	1.13	0.65	0.21	0.26	0.26	0.26	0.66
Median	1.23	0.91	0.09	0.25	0.22	0.11	0.71
Maximum	18.21	16.30	18.05	12.60	13.38	9.56	18.36
Minimum	−20.41	−21.76	−14.86	−14.11	−18.48	−6.86	−34.39
Std. dev.	4.38	4.27	3.02	2.87	2.17	1.99	4.19
Skewness	−0.31	−0.44	0.33	0.01	−0.33	0.32	−1.28
Kurtosis	5.80	4.87	6.02	5.39	15.38	4.61	13.19

This table reports the descriptive statistics of the survivor stock portfolio, S&P 500 index, and Fama and French (2018) risk factors. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Table A4. Regression estimates for the replicated survivor portfolio using different asset pricing models.

Alpha	***S&P 500***	***SMB***	***HML***	***RMW***	***CMA***	***UMD***	***R*²**
0.49 *** (7.01)	0.92 *** (56.88)						0.83
0.40 *** (6.31)	0.96 *** (64.32)	0.05 ** (2.46)	0.30 *** (13.56)				0.86
0.27 *** (4.50)	1.00 *** (67.75)	0.12 *** (5.87)	0.18 *** (6.44)	0.27 *** (9.37)	0.24 *** (5.65)		0.88
0.34 *** (5.72)	0.98 *** (68.03)	0.12 *** (6.17)	0.13 *** (4.69)	0.29 *** (10.37)	0.27 *** (6.58)	−0.10 *** (−6.87)	0.89

** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET^{ALL,excess}_{SURVIVOR,t}$ based on the Yahoo database on the excess returns of the S&P 500 index as well as different asset pricing models. Ordinary *t*-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Table A5. Multiple equation model analysis of the replicated survivor portfolio.

Dependent var.	**Alpha**	***CRSP*^*excess*^**	***SMB***	***HML***	***RMW***	***CMA***	***UMD***	***R*²**
$RET^{ALL,excess}_{SURVIVOR}$	0.09 (1.54)	0.98 *** (67.50)	−0.04 * (−1.81)	0.15 *** (5.45)	0.35 *** (12.32)	0.31 *** (7.47)	−0.12 *** (−8.14)	0.89
$RET^{excess}_{S\&P500}$	−0.25 *** (−15.81)	1.00 *** (262.94)	−0.16 *** (−30.46)	0.02 *** (3.08)	0.06 *** (7.89)	0.04 *** (3.48)	−0.02 *** (−5.16)	0.99

* Statistically significant on a 10% level. *** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET^{ALL,excess}_{SURVIVOR,t}$ based on the Yahoo database on the excess returns of the S&P 500 index as well as other risk factors in Fama and French's (2018) six-factor model. Ordinary *t*-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Table A6. Out-of-sample performance of the replicated survivor portfolio.

Alpha	**Dummy**	***S&P 500***	***SMB***	***HML***	***RMW***	***CMA***	***UMD***	***R*²**
0.56 *** (7.05)	−0.31 * (−1.88)	0.93 *** (57.01)						0.83
0.40 *** (5.55)	−0.02 (−0.16)	0.96 *** (64.27)	0.05 ** (2.45)	0.30 *** (13.41)				0.86
0.28 *** (4.03)	−0.03 (−0.21)	1.00 *** (67.70)	0.12 *** (5.85)	0.18 *** (6.38)	0.27 *** (9.37)	0.24 *** (5.65)		0.88
0.37 *** (5.43)	−0.12 (−0.92)	0.98 *** (68.02)	0.12 *** (6.12)	0.13 *** (4.54)	0.29 *** (10.39)	0.27 *** (6.59)	−0.10 *** (−6.92)	0.89

* Statistically significant on a 10% level. ** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET^{ALL,excess}_{SURVIVOR,t}$ based on the Yahoo database on the excess returns of the S&P 500 index as well as different asset pricing models. The regression models include a dummy variable denoted d with a value of 0 in the period from July 1963 to March 2007 and a value of 1 in the period April 2007–December 2019. Ordinary *t*-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Table A7. Estimating volatility processes for the replicated survivor portfolio.

	μ	c	α	β	γ	v
S&P 500	0.39 *** (2.77)	0.11 (1.28)	0.20 *** (2.97)	0.90 *** (28.26)	−0.16 *** (−3.99)	9.53
All survivors	0.76 *** (5.44)	0.14 * (1.79)	0.14 ** (2.40)	0.91 *** (34.06)	−0.20 *** (−5.43)	10.69

* Statistically significant on a 10% level. ** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

Here, we use Yahoo data and replicate the portfolio of survivor stocks. This table reports the estimates for the EGARCH model with mean equation:

$$R_{i,t}^{excess} = \mu_i + \epsilon_{i,t}$$

$$\epsilon_{i,t} = \zeta_{i,t}\sigma_{i,t},$$

where $R_{i,t}^{excess}$ is the excess return of at time t, $i = \{S\&P\ 500_t,\ all\ survivors_t\}$, μ_i denotes the intercept term of the mean equation, and $\epsilon_{i,t}$ is the residual term at time t. The equation for the variance is:

$$ln(\sigma_{i,t}^2) = c_i + \alpha_i \left| \frac{\epsilon_{i,t-1}}{\sigma_{i,t-1}} \right| + \beta_i ln\left(\sigma_{i,t-1}^2\right) + \gamma_i \frac{\epsilon_{i,t-1}}{\sigma_{i,t-1}},$$

where $\sigma_{i,t}^2$ is the conditional variance at time t, and the parameter vector $\boldsymbol{\theta}_{BTC} = (\mu_i, c_i, \alpha_i, \beta_i, \gamma_i)$ is estimated using maximum-likelihood estimation. The models assume that the innovation process follows a fat-tailed t-distribution (i.e., $\zeta_{BTC,t}|\boldsymbol{\Omega}_{t-1} \sim t(v)$ with v degrees of freedom). The z-statistics are given in parentheses. The sample period is from July 1963 to December 2019.

Table A8. Multiple equation model analysis of the replicated survivor portfolio and equal-weighted U.S. equity index.

Dependent var.	Alpha	$CRSP^{excess}$	SMB	HML	RMW	CMA	UMD	R^2
$RET_{SURVIVOR}^{ALL,excess}$	0.27 *** (3.02)	0.87 *** (40.75)	−0.80 *** (−20.27)	0.10 ** (2.41)	0.19 *** (4.35)	0.05 (0.77)	0.00 (0.08)	0.75
$RET_{S\&P500}^{excess}$	−0.28 *** (−5.05)	0.90 *** (68.29)	−0.85 *** (−36.06)	−0.05 * (−1.75)	0.05 * (1.92)	−0.11 ** (−2.78)	0.11 *** (7.77)	0.90

* Statistically significant on a 10% level. ** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET_{SURVIVOR,t}^{ALL,excess}$ based on the Yahoo database on the excess returns of the S&P 500 index as well as other risk factors in Fama and French's (2018) six-factor model. The factor model specification employs the average excess returns of 49 equal-weighted Fama and French U.S. industrial portfolios as proxy for the market factor. Ordinary t-statistics are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Table A9. Out-of-sample performance of all survivors with robust t-statistics.

Alpha	Dummy	*S&P 500*	*SMB*	*HML*	*RMW*	*CMA*	*UMD*	R^2
0.50 *** (5.99)	−0.32 ** (−2.47)	0.86 *** (33.76)						0.84
0.35 *** (4.96)	−0.12 (−0.94)	0.89 *** (62.46)	0.08 ** (2.36)	0.25 *** (4.38)				0.88
0.23 *** (3.30)	−0.12 (−0.99)	0.92 *** (63.69)	0.15 *** (6.89)	0.16 *** (4.99)	0.26 *** (3.63)	0.20 *** (5.30)		0.90
0.30 *** (4.20)	−0.19 * (−1.71)	0.91 *** (62.52)	0.15 *** (7.30)	0.12 *** (3.86)	0.27 *** (3.70)	0.22 *** (6.13)	−0.08 *** (−2.75)	0.90

* Statistically significant on a 10% level. ** Statistically significant on a 5% level. *** Statistically significant on a 1% level.

This table reports the results of regressing portfolio $RET^{ALL,excess}_{SURVIVOR,t}$ on the excess returns of the S&P 500 index as well as different asset pricing models. The regression models include a dummy variable denoted d with a value of 0 in the period from July 1963 to March 2007 and a value of 1 in the period April 2007–December 2019. Robust t-statistics using the covariance matrix estimator proposed from Newey and West (1987) with lag order $l = 1$ are reported in parentheses. The figures are given in terms of percentages. The sample period is from July 1963 to December 2019.

Notes

1 See https://fred.stlouisfed.org/series/DDOM01USA644NWDB (accessed on 14 January 2021).

2 As of February 2019 guidance.

3 The value of a stock's market capitalization traded annually should be at least a quarter million dollars of its shares in each of the previous six months.

4 Additional possible advantages include reduced information asymmetry due to greater scrutiny by investors, increased investor recognition as an industry leader, and a decline in shadow costs. See studies by Denis et al. (2003), Chen et al. (2004), Baran and King (2012), and Chan et al. (2013).

5 For instance, the annualized sample average return varies between 13.63% per annum and 13.75% per annum for the arithmetic return and equal-weighted portfolios, respectively.

6 See https://www.globalpapermoney.com/s-p-releases-list-of-86-companies-in-the-s-p-500-since-1957-cms-1023 (accessed on 31 January 2022).

7 The question arises how does the non-survivorship manifest itself over time? The so-called Lindy law could explain this phenomenon. In this regard, Taleb points out that the Lindy effect (or law) corresponds to situations where the conditional expectation of additional life expectancy increases with time, which requires the survival function of survival time to be that of a power law. A discussion of this issue is provided in Taleb's study "Lindy as a Distance from an Absorbing Barrier", which is available at https://www.academia.edu/44944654 (accessed on 31 January 2022). Future research could elaborate on this issue and model the survival, respectively, non-survival functions for companies in the S&P 500. This issue is, however, beyond the scope of this study and therefore left for future research.

8 As a last resort, the stock name was used to find stock return data in the CRSP database. In this regard, a company could change names or the same company could have different stocks. It is important to note that companies could have similar names and one stock could be changed to another one as successor in the CRSP's dataset. Additionally, ticker symbols for companies can change. Hence, we used the output produced from the CRSP database for tickers associated with corresponding company names. Finally, one stock does not necessarily mean one firm in the CRSP database. For instance, a firm could change its stock to be a different one. Moreover, a stock could also belong to different firms. As an example, firm A spins off into X, Y, and Z different firms. The original stock (in terms of its permno in CRSP) stays with firm X. However, the core business of firm A is actually in firm Z. Now firm Z is assigned a new stock (permno). When we have firm Z's name, and we expand its history, we include the original stock for firm A. In the dataset, we used (to be more inclusive) the stock for firm A in the past as well as firm Z's stock.

9 There are good reasons to use equal-weighted portfolios in the present study. Most importantly, market capitalization as a financial variable is pareto distributed, which means that if value-weighted portfolios were used, a very small number of stocks would receive extraordinarily high weights. Hence, value weighting would distort the overall portfolio return distribution. This distortion occurs when variables deviate from the normal distribution. Our sample stocks share one commonality—namely, survivorship. We are mainly interested in this common link, rather than potential size effects.

10 Because the kurtosis of the regression residuals is 9.87, one could argue that standard t-statistics are not valid for making statistical inference. If we assume a t-distribution with v = 4.5 degrees of freedom, the corresponding kurtosis will be 15, which is much larger than 9.74. Using a 5% significance level, the critical value of this distribution is 2.66. Since the t-statistic of 7.01 well exceeds 2.66, we can safely deduce that our statistical inference is valid.

11 As an example, the survivor stock portfolio's loading against the profitability factor exceeds the S&P 500′s loading by a factor of 6.5 implying that, on average, survivor stocks are considerably more profitable than the average S&P 500.

12 Peak exposure to the value factor was reached in February 2007 at an economic magnitude of 0.37.

13 Note that excess kurtosis is a stylized fact of financial market data. The use of the Gaussian distribution for modeling the conditional volatilities may lead to misleading results. For this reason, we employ t-distributions to model the innovation processes, which explicitly takes into account the fat-tailed data observed here.

14 Importantly, our higher cut-off of 3.39 decreases the likelihood that the performance of the survivor stock portfolio diminished in the ex post March 2007 period. Note that the t-statistic of -2.22 for our dummy variable in the CAPM model specification indicated a significant structural break on a common 5% level using standard critical values.

15 As precedent, Alexander and Dimitriu (2005) used Yahoo Finance. Data providers such as CRSP impose relatively high charges for data, whereas Yahoo Finance is freely available, thereby expanding research replicability to a larger audience of scholars. Many universities around the world do not subscribe to CRSP due to costs; in such cases, Yahoo Finance is available.

16 We found that 7.38% of *RTX* returns exceeded 100% in the sample period from February 1970 to December 2019.

17 These 14 survivor stocks with the longest available data are: Alcoa Corporation (AA), the Boeing Company (BA), Caterpillar Inc. (CAT), Chevron Corporation (CVX), DTE Energy Company (DTE), Consolidated Edison, Inc. (ED), General Electric Company (GE), Goodyear Tire & Rubber Company (GT), International Business Machines Corporation (IBM), International Paper Company (IP), Coca-Cola Company (KO), Altria Group, Inc. (MO), Procter & Gamble Company (PG), and the Exxon Mobil Corporation (XOM). These companies are very old and were originally established between 1823 and 1925. Ten of these 14 companies were originally founded before 1900.

18 We would like to thank an anonymous reviewer for suggesting this additional robustness check.

19 We thank an anonymous reviewer for suggesting this additional robustness check.

References

Afego, Pyemo N. 2017. Effects of changes in stock index compositions: A literature survey. *International Review of Financial Analysis* 52: 228–39. [CrossRef]

Alexander, Carol, and Anca Dimitriu. 2005. Indexing and Statistical Arbitrage. *Journal of Portfolio Management* 31: 50–63. [CrossRef]

Baran, Lindsay, and Tao-Hsien King. 2012. Cost of equity and S&P 500 index revisions. *Financial Management* 41: 457–81.

Chan, Kalok, Hung Wan Kot, and Gordon Y. N. Tang. 2013. A comprehensive long-term analysis of S&P 500 index additions and deletions. *Journal of Banking and Finance* 37: 4920–30.

Chen, Honghui, Gregory Noronha, and Vijay Singal. 2004. The price response to S&P 500 index additions and deletions: Evidence of asymmetry and a new explanation. *Journal of Finance* 59: 1901–29.

Chen, Sheng-Syan, and Yueh-Hsiang Lin. 2018. The competitive effects of S&P 500 index revisions. *Journal of Business Financial Accounting* 45: 997–1027.

Denis, Diane K., John J. McConnell, Alexei V. Ovtchinnikov, and Yun Yu. 2003. S&P 500 index additions and earnings expectations. *Journal of Finance* 58: 1821–40.

Fama, Eugene F., and Kenneth R. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2015. A Five-Factor Asset Pricing Model. *Journal of Financial Economics* 116: 1–22. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2018. Choosing Factors. *Journal of Financial Economics* 128: 234–52. [CrossRef]

Foster, Richard, and Sarah Kaplan. 2001. *Creative Destruction: Why Companies That Are Built to Last Underperform the Market and How to Successfully Transform Them*. New York: Random House.

Gnabo, Jean-Yves, Lyudmyla Hvozdyk, and Jérôme Lahaye. 2014. System-wide tail comovements: A bootstrap test for cojump identification on the S&P 500, US bonds and currencies. *Journal of International Money and Finance* 48: 147–74.

Harvey, Campbell R., Yan Liu, and Heqing Zhu. 2016. . . . and the cross-section of expected returns. *Review of Financial Studies* 29: 5–68. [CrossRef]

Hou, Kewei, Chen Xue, and Lu Zhang. 2020. Replicating Anomalies. *Review of Financial Studies* 33: 2019–133. [CrossRef]

Newey, Whitney K., and Kenneth D. West. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55: 703–8. [CrossRef]

Platikanova, Petya. 2016. S&P 500 index addition, liquidity management and Tobin's Q. *Accounting and Finance* 56: 479–508.

Shankar, Gowri S., and James M. Miller. 2006. Market Reaction to Changes in the S&P SmallCap 600 Index. *Financial Review* 41: 339–60.

Siegel, Jeremy J., and Jeremy D. Schwartz. 2006. Long-term returns on the original S&P 500 companies. *Financial Analysts Journal* 62: 18–31.

Taleb, Nassim N. 2012. *Antifragile*. New York: Penguin Books.

Journal of
Risk and Financial Management

Article

Net Buying Pressure and Informed Trading in the Options Market: Evidence from Earnings Announcements

Ihsan Badshah * and Hardjo Koerniadi

Department of Finance, Business School, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand; hardjo.koerniadi@aut.ac.nz
* Correspondence: ihsan.badshah@aut.ac.nz

Abstract: By employing the modified net buying pressure as a measure of informed option trading, this study tested whether option trading around quarterly earnings announcements is either directionally motivated and/or volatility motivated. We found evidence that is consistent with the idea that option investors have private information prior to positive earnings announcements and use at-the-money options to exploit their informational advantage. In the post-event period, however, informed option investors trade by using deep-out-of-the-money and out-of-the-money options. We documented limited evidence on the volatility-motivated option trading, and our results suggest that this type of option trading could be motivated by hedging purposes only.

Keywords: earnings; announcements; options; informed trading; net buying pressure; volatility; direction; at-the-money; out-of-the-money; deep-out-of-the-money

JEL Classification: G10; G13; G14

Citation: Badshah, Ihsan, and Hardjo Koerniadi. 2022. Net Buying Pressure and Informed Trading in the Options Market: Evidence from Earnings Announcements. *Journal of Risk and Financial Management* 15: 53. https://doi.org/10.3390/jrfm15020053

Academic Editors: James W. Kolari and Seppo Pynnonen

Received: 15 December 2021
Accepted: 21 January 2022
Published: 24 January 2022

1. Introduction

The literature reports that option investors trade on the volatility of underlying stock returns (e.g., Patell and Wolfson 1981; Gharghori et al. 2017; Chen and Wang 2016). Other studies, however, document that option investors are able to predict the direction of stock prices prior to major corporate events (e.g., Jin et al. 2012; Atilgan 2014; Chan et al. 2015).

Kang and Park (2008) propose direction-learning hypothesis and volatility-learning hypothesis and provide notable evidence (for KOSPI 200 index options) supporting direction-learning hypothesis while rejecting the volatility-learning hypothesis. Later Chen and Wang (2016) extend the study of Kang and Park (2008) by arguing that option investors betting on the direction of stock returns trade differently from those betting on the volatility of stock returns. When positive news is expected to increase both stock returns and volatility, directional (volatility) traders will sell (buy) put options. By contrast, when negative news is expected to decrease stock returns but increase volatility, directional (volatility) traders will sell (buy) call options. This exogenous shock, consequently, can have simultaneous, but offsetting effects on option informed trading measures. Given that the directional and volatility option investors may trade differently on the same impending news, it is important to distinguish these two types of option trades.

Except for Kang and Park (2008) and Chen and Wang (2016), prior studies do not differentiate trading strategies of these two types of option investors.[1] To fill this gap in the literature, this paper examines US option investors' trading prior to quarterly earnings announcements with respect to both expected changes in stock returns and the volatility of stock returns related to the event. We focus on earnings announcements due to the fact that the two types of option investors are likely to trade around such an event (Patell and Wolfson 1981; Jin et al. 2012; Atilgan 2014). To separate these two types of option trading from each other, we employ net buying pressure (NBP), which was initially developed by Bollen and Whaley (2004) and then modified further by Chen and Wang (2016), as the informed option trading measure in our study. Our study also examines at which option moneyness informed option trading occurs. To investigate this issue, we split our sample based on the moneyness of the options in an effort to better understand how moneyness affects option informed trading.

Additionally, Kim and Verrecchia's (1991, 1994) information-based trading theory suggests that informed investors trade, not only on private information prior to an event, but could also trade in the post-event period due to their superior ability in processing publicly disclosed information from a corporate announcement. Therefore, to test this conjecture, we also examine the relation between net buying pressures and both stock returns and stock return volatility during the post-event window.

We found that directional informed trading prior to earnings announcements occurs in at-the-money (ATM) options. This finding could be due to the relatively higher liquidity and lower transactions costs of ATM options (Chakravarty et al. 2004). We also found evidence that option investors trade on stock-return volatility prior to earnings announcements, particularly in OTM options. A further analysis showed that option investors trade by using deep-out-of-the-money (DOTM) and out-of-the-money (OTM) call and put options in the post-event window period.

Our study contributes to the literature in several ways. First, we provide empirical evidence whether option investors trade on the expected changes in stock returns and/or on the volatility of stock returns around quarterly earnings announcements. Second, we provide evidence that option investors have private information on good announcements prior to this event and have superior ability to process publicly disclosed information. Third, our study provides a deeper insight by providing empirical evidence about in which option moneyness these informed option transactions occur.

The remainder of this paper proceeds as follows. The next section reviews the research design and option informed trading measures. Section 3 describes the sample selection process. Section 4 presents the empirical results. Section 5 concludes.

2. Research Design

To proxy for demand in the underlying stock equivalent, Bollen and Whaley (2004) propose net buying pressure (NBP), measured as the difference between the number of buyer-motivated contracts and the number of seller-motivated contracts, multiplied by the absolute value of option delta. Bollen and Whaley (2004) identify buyer-motivated options as trades executed at the price above the midpoint of prevailing bid and ask prices. We collected option-related data from the OptionMetrics database. This database, however, does not provide transaction prices of options. Therefore, we used the current midpoint of bid and ask prices as the proxy for transaction price of an option. We identified an option trading as a buyer-(seller-) motivated option trading if a current midpoint price of an option is higher (lower) than its previous midpoint price. This procedure was repeated for the entire universe of call and put options for US equities. Option net buying pressure (NBP) measure is calculated as the difference between the number of buyer-motivated contracts and seller-motivated contracts, multiplied by the absolute value of the option's delta. Following Bollen and Whaley (2004), we scaled option net buying pressure measure by the total trading volume across all options in the class on that day.

To separate directional-motivated option trading effects from volatility-motivated option trading effects, we used the modified NBP measures proposed by Chen and Wang (2016). The modified NBP measures allowed us to distinguish between informed trading on the direction of the underlying asset price and volatility-based informed trading. Following Chen and Wang (2016), the directional-motivated demands for the kth-moneyness category of call and put options, respectively, are measured as follows:

$$NBPD^k_{C,t} = \frac{NBP^k_{C,t} - NBP^k_{P,t}}{2} \quad (1)$$

$$NBPD^{k}_{P,t} = \frac{NBP^{k}_{P,t} - NBP^{k}_{C,t}}{2}. \tag{2}$$

where $NBPD_c$ ($NBPD_p$) is the difference between the NBPs of calls (puts) and puts (calls) options divided by 2 categorized by moneyness k over the time interval t, and $k \in$ {DOTM, OTM, ATM}. Similarly, the volatility-motivated demand for the kth-moneyness category option is measured as follows:

$$NBPV^{k}_{t} = \frac{NBP^{k}_{C,t} + NBP^{k}_{P,t}}{2} \tag{3}$$

Following Jin et al. (2012), we measured the event window from days −1 to +1 relative to the announcement day. Option informed trading measures are computed during the base-, pre-, and post-event windows associated with days −50 to −11 prior to earnings announcements, days −10 to −1 prior to the event, and days +1 to +5 days after the event, respectively. Call- and put-option net-buying pressures that are directional during the base-, pre-, and post-event window are denoted as *NBPD_CALL_BASE*, *NBPD_PUT_BASE*, *NBPD_CAL_PRE*, *NBPD_PUT_PRE*, *NBPD_CALL_POST*, and *NBPD_PUT_POST* respectively. Call- and put-option net-buying pressures' volatility during the base-, pre-, and post-event window is denoted as *NBPV_BASE*, *NBPV_PRE*, and *NBPV_POST*, respectively.

Cumulative abnormal returns (CAR) for the ith stock are computed as follows:

$$CAR_{it} = \sum(r_{i,t} - r_{m,t}), \tag{4}$$

where the CRSP value-weighted market return, $r_{m,t}$, was obtained from Kenneth French's website. Cumulative abnormal stock returns during the event (post-event) window on days -1 to +1 (+6 to +90) are denoted as *XRET* (*XRET_POST*).

Similar to Jin et al. (2012), to examine whether options traders have private information on the expected change in stock prices prior to the earnings announcement date or have better processing skills of publicly disclosed information in the post-event period, we employed the following regression specifications:

$$\begin{aligned} XRET(-1,+1)_i \;&= Intercept + \alpha PRE_{NBPD^{k}_{C,i}} + \beta PRE_{NBPD^{k}_{P,i}} + \delta BASE_{NBPD^{k}_{C,i}} \\ &+\gamma BASE_{NBPD^{k}_{P,i}} + \mu PRE_{SVOLi} + \theta SURP_i + \vartheta SIZE_i + \rho MB_i \\ &+\varepsilon_i \end{aligned} \tag{5a}$$

$$\begin{aligned} XRETPOST(+6,+90)_i \;&= Intercept + \alpha POST_{NBPD^{k}_{C,i}} + \beta POST_{NBPD^{k}_{P,i}} \\ &+\delta BASE_{NBPD^{k}_{C,i}} + \gamma BASE_{NBPD^{k}_{P,i}} + XRET_i + \mu POST_{SVOLi} \\ &+\theta SURP_i + \vartheta SIZE_i + \rho MB_i + \varepsilon_i \end{aligned} \tag{5b}$$

where *PRE_SVOL* (*POST_SVOL*) is the logarithm of the volume of stocks traded during the pre-(post-)event window. If the option trading measures contain information relevant to expected changes in stock prices, then the estimated coefficients on *PRE_NBPDc* and *POST_NBPDc* should be positively related to *XRET* and *XRETPOST*, respectively. SURP is earnings surprise. Control variables are size (natural logarithm of market capitalization) and M/B (market to book ratio) of sample firms.

Following Chen and Wang (2016) and Gharghori et al. (2017), we also examined whether option trading measures of volatility-motivated informed trading are related to stock return volatility based on the following regression specifications:

$$STDEVSHORT(-1,+1)_i = Intercept + \alpha PRE\_NBPV^{k}_{i} + \beta BASE\_NBPV^{k}_{i} + \varepsilon_i \tag{6a}$$

$$STDEVLONG(+6,+90)_i = Intercept + \alpha POST\_NBPV^{k}_{i} + \beta BASE\_NBPV^{k}_{i} + \varepsilon_i \tag{6b}$$

where *STDEVSHORT (STDEVLONG)* is the standard deviation of the daily market-adjusted returns in the event window period −1 to +1 (+6 to +90) days. If option volatility trading measures contain information relevant to expected volatility changes of stocks, then the estimated coefficients on *PRE_NBPV* and *POST_NBPV* should be positively related to *STDEVSHORT* and *STDEVLONG*, respectively.

3. Data

We obtained equity options and stock-related information from the OptionMetrics database from 1 January 2005 to 30 April 2016. The database provides daily bid and ask quotes; open interest; volume; implied volatility; and Greeks, such as delta, gamma, vega, and theta, for call and put options listed on all option exchanges for underlying US equities. From this database, we collected the underlying stock-related data for daily stock bids and ask quotes, closing prices, total returns, trading volume, and outstanding shares. We collected quarterly earnings announcement dates from 2005 to 2016 from the Research Insight database. The CRSP market index data were obtained from Kenneth French's website. We merged data from these three databases based on whether firms that announce quarterly earnings during the sample period are optionable.

We selected options (calls and puts) with maturity from 10 to 60 days (Cremers and Weinbaum 2010; Jin et al. 2012). We observed many observations with zero open interest and zero volume in the data. Therefore, to address thin trading issues, we removed options with zero open interest and zero volume from the sample. Net buying pressures for both call and put options were then calculated based on the available option volume data. We excluded observations with zero net buying pressures.

Panel A of Table 1 shows the daily average of the number of option contracts during the pre-, base-, and post-windows. The number of contracts shows that call options, except for DOTM call options in the pre-window period, are traded more often than put options in all three window periods. The numbers of daily option contracts traded in the post window are the largest across option moneyness. On average, daily transaction volumes increase in the pre-window period and further in the post-window period. The increase in the number of daily contracts traded in the pre-window period may indicate that informed option investors trade during these periods. The numbers of net purchases displayed in Panel B of Table 1 are, on average, negative (except for ATM call and put options in the base- and pre-window periods and ATM put options in the post-event period), which suggests that these contracts are seller motivated. Panel C of Table 1 illustrates the net buying pressure of call and put options across the three windows. As can be seen, investors generally have net buying positions in call options in the base- and pre-windows, except DOTM calls. However, in the post-window, investors have selling positions in call options. Somehow, similar positions are observed for puts; however, the selling level in the post-window for puts in contrast to calls is high in the categories of OTM and ATM options.

Table 1. Daily average of number of option contracts and net buying pressure.

	BASE		PRE		POST	
	CALL	PUT	CALL	PUT	CALL	PUT
Panel A. Number of Contracts						
DOTM	3,355,564	3,136,258	3,467,501	3,641,739	7,012,477	6,098,234
OTM	9,445,297	7,006,597	12,213,275	9,500,400	17,278,566	12,767,542
ATM	4,208,167	2,382,670	5,899,880	3,241,505	7,223,099	4,348,246
Panel B. Net Purchases of Contracts						
DOTM	−1,214,728	−1,115,330	−1,021,626	−1,195,070	−3,681,233	−2,626,351
OTM	−1,327,765	−969,565	−895,825	−1,318,548	−4,210,721	−2,346,881
ATM	340,380	141,210	541,999	138,619	−305,685	264,422
Panel C. Net Buying Pressure						
DOTM	−9052	−1492	−6443	−783	−14,955	−3991
OTM	12,883	14,785	22,587	19,485	−952	−20,464
ATM	7093	−1948	20,287	11,138	−2727	−15,980

This table reports daily average number of option contracts, net purchase of option contracts, and net buying pressure during the base, pre, and post windows for deep-out-the-money, out-of-the-money and at-the-money call and put option contracts. Base-, pre-, and post-event windows associated with days −50 to −11 prior to earnings announcement day, days −10 to −1 prior to the event, and days +1 to +5 days after the event, respectively. The net buying pressure is calculated as the difference between the number of buyer-motivated contracts, multiplied by the absolute value of option delta.

4. Results

4.1. Directional-Motivated Options Trading

Table 2 reports the relation between informed option trading measures during the pre-event window period and the cumulative announcement return during the event window (Equation (5a)).[2] Each panel reports directional informed trading tests for each option of moneyness. As prior studies suggest that option investors may trade differently depending on the quality of an announcement (Chen and Wang 2016), we examined the effect of pre-event informed option trading measures on good or bad announcements. We define bad or good news if the cumulative abnormal return during the event window is negative or positive, respectively. We found that the coefficients on NBPD_CALL_PRE and NBPD_PUT_PRE for DOTM and OTM options, as reported in Panels A and B, are not positively related to the cumulative announcement returns. The negative coefficient of NBPD_CALL_PRE for good news in the OTM option, however, is not consistent with the informed trading hypothesis. One possible explanation for this conflicting sign is that it could be related to hedging purposes due to the moneyness of the options. Overall, these results suggest that informed option investors do not use DOTM or OTM options to trade on the expected changes on the underlying stocks' prices.

Panel C of Table 2 examines the relation between the cumulative abnormal announcement period returns and net buying pressures for ATM options. We found that, for good announcements, the net buying pressures of call options are positively related to announcement-period returns. This result suggests that informed trading occurs in ATM options during the pre-announcement period of good announcements. The (in)significant relationship between announcement return and net buying pressures of ATM options may suggest that option investors are more (less) likely to trade if the impending earnings announcement is good (bad) news. These results are consistent with the direction-learning hypothesis and findings of Kang and Park (2008) that informed investors use options to trade on the direction of the underlying. Moreover, our finding that option investors are informed on the good earnings announcement is consistent with the evidence of Whalen and Collver (2004).

Table 2. Relationship between option NBPDs of calls (puts) and event returns.

Panel A. DOTM	All	Bad News	Good News
NBPD_CALL_PRE	0.018	0.031	−0.027
	(0.31)	(0.18)	(0.12)
NBPD_PUT_PRE	0.017	0.008	−0.001
	(0.36)	(0.67)	(0.98)
NBPD_CALL_BASE	0.012	0.029 ***	−0.013 *
	(0.12)	(0.00)	(0.08)
NBPD_PUT_BASE	−0.002	−0.003	0.005
	(0.85)	(0.71)	(0.56)
Intercept	−0.029 ***	−0.206 ***	0.197 ***
	(0.00)	(0.00)	(0.00)
Observations	6134	3067	3067
Adjusted R^2	0.0024	0.1222	0.1022
Panel B. OTM			
NBPD_CALL_PRE	0.000	−0.002	−0.009 **
	(0.96)	(0.62)	(0.05)
NBPD_PUT_PRE	0.005	−0.001	0.004
	(0.37)	(0.84)	(0.46)
NBPD_CALL_BASE	−0.003	0.000	−0.009 ***
	(0.20)	(0.96)	(0.00)
NBPD_PUT_BASE	−0.002	0.000	−0.004
	(0.55)	(0.91)	(0.17)
Intercept	−0.020 ***	−0.185 ***	0.187 ***
	(0.00)	(0.00)	(0.00)
Observations	9543	4775	4768
Adjusted R^2	0.0031	0.0975	0.1052

Table 2. *Cont.*

Panel C. ATM			
NBPD_CALL_PRE	0.002	−0.001	0.008 **
	(0.56)	(0.84)	(0.03)
NBPD_PUT_PRE	0.006	0.001	0.002
	(0.17)	(0.75)	(0.68)
NBPD_CALL_BASE	0.000	−0.001	0.000
	(0.93)	(0.64)	(0.97)
NBPD_PUT_BASE	−0.001	−0.003	−0.001
	(0.70)	(0.32)	(0.76)
Intercept	−0.041 ***	−0.188 ***	0.173 ***
	(0.00)	(0.00)	(0.00)
Observations	4146	2136	2010
Adjusted R^2	0.0072	0.0976	0.0923

This table reports results for the effect of net buying pressure of call and put options on the event excess returns for different moneyness categories. The dependent variable is the cumulative abnormal stock returns (−1, +1). *PRE_NBPD* and *BASE_NBPD* are the option net buying pressure directional measures for days −10 to −2 and days −50 to −11, respectively. To conserve space, control variables are not reported. Bad news is for negative announcement returns; good news is for positive announcement returns; *p*-values are in parentheses; *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

4.2. Volatility-Motivated Options Trading

Figlewski and Frommherz (2017) and Gharghori et al. (2017) argue that transactions in the options market may be related to expected changes in the volatility of underlying asset values. To test this conjecture, following Chen and Wang (2016), we employed a measure of informed option trading based on options transactions—namely net buying pressure volatility (NBPV). This measure reflects option trading information related to the volatility (rather than the direction) of underlying asset values. Table 3 reports net buying pressures of volatility trading prior to announcement dates. The coefficient on NBPV_PRE is negative and significant for OTM options, suggesting that volatility-based trading by using OTM options prior to the event is probably hedging motivated. These results are also consistent with the findings of Kang and Park (2008), who do not find informed trading on the volatility instead they find option trading on the direction of the underlying.

Table 3. Relationship between option NBPVs and event returns volatility.

Panel A. DOTM	**All**	**Bad News**	**Good News**
NBPV_PRE	−0.009	−0.009	−0.008
	(0.10)	(0.24)	(0.27)
NBPV_BASE	−0.007 ***	−0.011 ***	−0.003
	(0.00)	(0.00)	(0.35)
Intercept	0.027 ***	0.028 ***	0.026 ***
	(0.00)	(0.00)	(0.00)
Observations	6234	3118	3116
Adjusted R^2	0.0018	0.0034	0.0001
Panel B. OTM			
NBPV_PRE	−0.003 **	−0.001	−0.005 ***
	(0.03)	(0.71)	(0.01)
NBPV_BASE	−0.006 ***	−0.005 ***	−0.007 ***
	(0.00)	(0.00)	(0.00)
Intercept	0.032 ***	0.032 ***	0.032 ***
	(0.00)	(0.00)	(0.00)
Observations	9684	4849	4835
Adjusted R^2	0.0066	0.0044	0.0091

Table 3. *Cont.*

Panel C. ATM			
NBPV_PRE	0.001	0.000	0.002
	(0.43)	(0.79)	(0.15)
NBPV_BASE	−0.002 **	−0.002	−0.001
	(0.04)	(0.13)	(0.14)
Intercept	0.032 ***	0.033 ***	0.032 ***
	(0.00)	(0.00)	(0.00)
Observations	4227	2177	2050
Adjusted R^2	0.0007	0.0002	0.001

This table reports results for the effect of net buying pressure of call and put options on the event excess returns volatility for different moneyness categories. The dependent variable is the volatility of abnormal stock returns during the event window period (−1,+1). *PRE_NBPV* and *BASE_NBPV* are the option net buying pressure volatility measures for days −10 to −2, and −50 to −11, respectively. Bad news is for negative announcement returns; good news is for positive announcement returns; *p*-values are in parentheses; ** and *** denote statistical significance at 5% and 1% levels, respectively

4.3. Post-Event Options Trading

Prior studies suggest that option investors may have superior skills compared to other investors to process publicly disclosed information. To test this conjecture, in the spirit of Jin et al. (2012), we examined whether net buying pressures measured in the post-event period (+1, +5) are positively related to abnormal stock returns in the post-event period (+6, +90). Panels A and B of Table 4 show, that for good earnings announcements, net buying pressures of DOTM and OTM call and put options in the post-event period (+1, +5) are significantly related to post-event cumulative abnormal stock returns measured during days +6 to +90 relative to the announcement dates. This significant relation, however, is absent for ATM options. Thus, we infer that option investors have information-processing skills with respect to information from earnings announcements and trade by using DOTM and OTM options on this information.

Table 4. Relationship between option NBPDs of calls (puts) and post-event returns.

Panel A. DOTM	**All**	**Bad News**	**Good News**
NBPD_CALL_POST	0.152 ***	0.144 ***	0.172 ***
	(0.00)	(0.00)	(0.00)
NBPD_PUT_POST	−0.114 ***	−0.091 ***	−0.132 ***
	(0.00)	(0.00)	(0.00)
NBPD_CALL_BASE	−0.001	−0.002	0.002
	(0.89)	(0.86)	(0.79)
NBPD_PUT_BASE	−0.003	0.015	−0.021 ***
	(0.70)	(0.14)	(0.01)
Intercept	−0.002	−0.042 ***	0.001
	(0.76)	(0.00)	(0.86)
Observations	6139	3067	3072
Adjusted R^2	0.3865	0.2109	0.3828
Panel B. OTM			
NBPD_CALL_POST	0.009	0.003	0.006
	(0.91)	(0.98)	(0.38)
NBPD_PUT_POST	−0.016	−0.008	−0.032 ***
	(0.86)	(0.96)	(0.00)
NBPD_CALL_BASE	−0.078 ***	−0.157 ***	0.002
	(0.01)	(0.01)	(0.45)
NBPD_PUT_BASE	−0.014	−0.027	0.000
	(0.66)	(0.67)	(0.94)
Intercept	−0.183 ***	−0.467 ***	−0.002
	(0.00)	(0.00)	(0.80)
Observations	9518	4761	4757
Adjusted R^2	0.0034	0.0019	0.2567

Table 4. *Cont.*

Panel C. ATM			
NBPD_CALL_POST	−0.004	−0.014	0.005
	(0.97)	(0.95)	(0.46)
NBPD_PUT_POST	−0.020	−0.039	−0.018 **
	(0.87)	(0.87)	(0.04)
NBPD_CALL_BASE	0.003	0.002	−0.001
	(0.95)	(0.98)	(0.63)
NBPD_PUT_BASE	−0.033	−0.051	−0.005 *
	(0.47)	(0.56)	(0.08)
Intercept	−0.350 ***	−0.879 ***	0.014
	(0.01)	(0.00)	(0.17)
Observations	4066	2102	1964
Adjusted R^2	0.0015	0.0029	0.2206

This table reports results for the effect of net buying pressure of call and put options on the event excess returns for different moneyness categories. The dependent variable is the post-event cumulative abnormal stock returns (+6,+90). *POST_NBPD* and *BASE_NBPD* are the option net buying pressure directional measures for days +1 to +5, and −50 to −11, respectively. To conserve space, control variables are not reported. Bad news is for negative announcement returns; good news is for positive announcement returns; *p*-values are in parentheses; *, ** and *** denote statistical significance at 10%, 5% and 1% levels, respectively.

Table 5 shows the regression results for Equation (6b) relating stock return volatility to post-event option net buying pressure. The coefficient on NBPV_POST is statistically insignificant for each category of option moneyness. Thus, we did not find evidence that options investors trade on expected stock returns volatility in the post-event period.

Table 5. Relationship between option NBPVs and post-event returns volatility.

Panel A. DOTM	All	Bad News	Good News
NBPV_POST	−0.007	−0.004	−0.008
	(0.19)	(0.61)	(0.32)
NBPV_BASE	−0.009 ***	−0.013 ***	−0.005 **
	(0.00)	(0.00)	(0.05)
Intercept	0.018 ***	0.019 ***	0.017 ***
	(0.00)	(0.00)	(0.00)
Observation	6208	3103	3105
Adjusted R^2	0.0044	0.008	0.0011
Panel B. OTM			
NBPV_POST	0.000	−0.005	0.003
	(0.99)	(0.93)	(0.14)
NBPV_BASE	0.017	0.038 *	−0.005 ***
	(0.12)	(0.07)	(0.00)
Intercept	0.028 ***	0.035 ***	0.021 ***
	(0.00)	(0.00)	(0.00)
Observations	9642	4825	4817
Adjusted R^2	0	0.0003	0.0082
Panel C. ATM			
NBPV_POST	0.006	0.012	0.002
	(0.88)	(0.89)	(0.19)
NBPV_BASE	0.005	0.011	−0.001
	(0.75)	(0.72)	(0.26)
Intercept	0.035 ***	0.048 **	0.021 ***
	(0.00)	(0.04)	(0.00)
Observations	4121	2128	1993
Adjusted R^2	−0.0005	−0.0009	0.0006

In this table *POST_NBPV* and *BASE_NBPV* are the option net buying pressure volatility measures for days +1 to +5 and −50 to −11, respectively. Bad news is for negative announcement returns; good news is for positive announcement returns; *p*-values are in parentheses; *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

5. Conclusions

Our empirical results suggest that option investors have private information on the expected direction of the underlying stocks prices prior to good earnings announcements, and they trade by using ATM options to exploit their private information. This is probably because ATM options have high liquidity and lower transaction costs compared to the other option, moneyness. Further results suggest that option investors have the processing ability of information from publicly disclosed announcements in terms of predicting the direction of stock returns during the post-event window period. In the post-event period, however, these investors do not use ATM options, but trade by using OTM and DOTM options. We found limited evidence that option investors trade on the expected volatility of the underlying stocks' prices prior to and after the announcements, and the results suggest that these transactions could be related to hedging purposes. Overall, our empirical evidence suggests that informed option traders' benefit from their private information related to both the expected direction and volatility of underlying asset values.

Author Contributions: Both authors equally contributed to the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. Chen and Wang (2016) document evidence supporting both directional and volatility trading on stock index option in the Taiwanese option market prior to 2011.
2. To conserve space, we do not report the results for the other variables in the regression model.

References

Atilgan, Yigit. 2014. Volatility spreads and earnings announcement returns. *Journal of Banking and Finance* 38: 205–15. [CrossRef]

Bollen, Nicolas P. B., and Robert E. Whaley. 2004. Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59: 711–53. [CrossRef]

Chakravarty, Sugato, Huseyin Gulen, and Stewart Mayhew. 2004. Informed trading in stock and options markets. *Journal of Finance* 59: 1235–57. [CrossRef]

Chan, Kohan, Li Ge, and Tse-Chun Lin. 2015. Informational content of option trading on acquirer announcement returns. *Journal of Financial and Quantitative Analysis* 50: 1057–82. [CrossRef]

Chen, Chao-Chun, and Shih-Hua Wang. 2016. Net buying pressure and option informed trading. *Journal of Futures Markets* 37: 238–59. [CrossRef]

Cremers, Martijn, and David Weinbaum. 2010. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis* 45: 335–67. [CrossRef]

Figlewski, Stephen, and Anja Frommherz. 2017. Volatility Leadership among Index Options. *The Journal of Derivatives* 25: 43–60. [CrossRef]

Gharghori, Philip, Edwin D. Maberly, and Annette Nguyen. 2017. Informed trading around stock split announcements: Evidence from the option market. *Journal of Quantitative & Financial Analysis* 52: 705–35.

Jin, Wen, Joshua Livnat, and Yuan Zhang. 2012. Options prices leading equity prices: Do options traders have an information advantage? *Journal of Accounting Research* 50: 401–31. [CrossRef]

Kang, Jangkoo, and Hyoung-Jin Park. 2008. The information content of net buying pressure: Evidence from the KOSPI 200 index option market. *Journal of Financial Markets* 11: 36–56. [CrossRef]

Kim, Oliver, and Robert E. Verrecchia. 1991. Market reaction to anticipated announcements. *Journal of Financial Economics* 30: 273–309. [CrossRef]

Kim, Oliver, and Robert E. Verrecchia. 1994. Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics* 17: 41–67. [CrossRef]

Patell, James M., and Mark A. Wolfson. 1981. The ex ante and ex post price effects of quarterly earnings announcements reflected in option and stock prices. *Journal of Accounting Research* 19: 434–58. [CrossRef]

Whalen, Dennis J., and Charles D. Collver. 2004. Informed Trading Around Earnings Announcements: Another Look. *Financial Review* 39: 409–34. [CrossRef]

Journal of
Risk and Financial Management

Communication

Outliers and Time-Varying Jumps in the Cryptocurrency Markets

Anupam Dutta [1,*] and Elie Bouri [2]

[1] School of Accounting and Finance, University of Vaasa, 65101 Vaasa, Finland
[2] School of Business, Lebanese American University, Beirut P.O. Box 13-5053, Lebanon; elie.elbouri@lau.edu.lb
* Correspondence: adutta@uwasa.fi

Abstract: We examine the presence of outliers and time-varying jumps in the returns of four major cryptocurrencies (Bitcoin, Ethereum, Ripple, Dogecoin, Litecoin), and a broad cryptocurrency index (CCI30). The results indicate that only Bitcoin returns are contaminated with outliers. Time-varying jumps are present in Bitcoin, Litecoin, Ripple, and the cryptocurrency index. Notably, the presence of jumps in Bitcoin is significant after correcting for outliers. The main findings point to a price instability in some major cryptocurrencies and thereby the importance of accounting for large shocks and time-varying jumps in modelling volatility in the debatable cryptocurrency markets.

Keywords: Bitcoin; cryptocurrencies; outliers; GARCH-jump; time-varying jumps

Citation: Dutta, Anupam, and Elie Bouri. 2022. Outliers and Time-Varying Jumps in the Cryptocurrency Markets. *Journal of Risk and Financial Management* 15: 128. https://doi.org/10.3390/jrfm15030128

Academic Editor: James W. Kolari

Received: 5 December 2021
Accepted: 28 February 2022
Published: 8 March 2022

1. Introduction

Cryptocurrencies are decentralised payment systems involving technological innovation called blockchain. They have attracted much attention on the financial scene as a digital asset class, capable of offering very high returns and decent diversification benefits when combined with conventional assets (Bouri et al. 2020). Several studies have focused on Bitcoin and other major cryptocurrencies in terms of price discovery (Corbet et al. 2019; Chen et al. 2020), herding (Yousaf et al. 2021), bubble formation (Bouri et al. 2019; Chaim and Laurini 2019), interconnectedness (Ji et al. 2019), market efficiency (López-Martín et al. 2021; Noda 2021), and safe-haven ability (Bouri et al. 2020; Das et al. 2020; Dutta et al. 2020; Hatemi-J. et al. 2020). Notably, cryptocurrencies are characterised by extreme return volatility that has been the subject of volatility modelling (Chu et al. 2017; Katsiampa 2017; Tiwari et al. 2019; Walther et al. 2019; Mostafa et al. 2021), especially using GARCH processes that are capable of parameterising higher order dependence and time-evolution of conditional volatility. The largest cryptocurrency, Bitcoin, is known for extreme return volatility[1] and large abrupt price variations in the form of jumps (Chaim and Laurini 2018). Furthermore, Bitcoin and other major cryptocurrencies tend to jump with geopolitical uncertainty (Bouri et al. 2020). However, there is no empirical evidence of the presence of outliers in leading cryptocurrencies[2] and the scarce academic literature available considers jump behaviour in the Bitcoin market only, overlooking the time-varying nature of jumps. Interestingly, large cryptocurrencies such as Ethereum, Ripple, Litecoin, and Dogecoin[3] have attracted significant attention from institutional investors and business communities. Furthermore, their return volatility tends to exceed that of the largest cryptocurrency, Bitcoin (see Table 1), which makes them relevant candidates for the analysis of outliers and time-varying jumps.

In this study, we extend the limited understanding of whether outliers are present in various cryptocurrencies and whether cryptocurrencies are characterised by time-varying jumps. To do this, we detect the presence of outliers and then apply GARCH-jump models capable of uncovering evidence of time-varying jumps in the daily return series.

Our paper is related to a growing strand of literature on the volatility of Bitcoin and other major cryptocurrencies (e.g., Salisu and Ogbonna 2021; Shahzad et al. 2021) during the COVID-19 outbreak, when uncertainty in the global economy and financial markets spiked and the prices of global equity indices tumbled. Notably, Bitcoin and other major cryptocurrencies experienced large increases in their prices from the second quarter of 2020 until most of 2021, driven by an accentuated trend towards digitalisation and acceptance of Bitcoin as a means of payment by large corporations (e.g., Tesla) as well as possibilities of central banks and emerging economies to adopt cryptocurrencies (Cunha et al. 2021).

Table 1. Descriptive statistics of daily returns.

	Mean	Min	Max	Standard Deviation	Skewness	Kurtosis	PP Test (p-Value)
Bitcoin	0.2964	−46.473	22.5119	3.9934	−0.81385	11.7148	0.00
Bitcoin (outlier-free)	0.2300	−31.190	19.7621	2.9552	−0.0006	7.5791	0.00
Ethereum	0.3726	−55.0714	41.2405	6.1632	0.000903	7.6691	0.00
Ripple	0.2131	−61.638	102.7463	7.0183	2.053332	33.74518	0.00
Dogecoin	0.3243	−51.4934	151.6211	7.7355	4.287955	76.4462	0.00
Litecoin	0.1675	−44.9012	51.0348	5.6424	0.33125	11.8109	0.00
CCI30	0.2457	−48.4483	19.5679	4.4000	−1.31042	11.4226	0.00

Notes: This table reports the main descriptive statistics for the return series of major cryptocurrencies for the period 8 August 2015–23 September 2021. Cryptocurrency index (CCI30). Phillips–Perron (PP).

Our contributions are on two fronts. Firstly, we identify potential outliers occurring in various cryptocurrencies, adding to prior studies that analyse the volatility dynamics of cryptocurrencies using GARCH-type models without correcting for possible outliers (e.g., Katsiampa 2017; Chu et al. 2017; Tiwari et al. 2019; Mostafa et al. 2021). Outliers are generally present in financial variables and can lead to serious distortion of model specifications, parameter estimation, and volatility forecasting (Grané and Veiga 2010; Carnero et al. 2012), which makes the detection/removal of outliers an important step in modelling volatility and in making risk-management inferences. This is very relevant to cryptocurrencies that are highly subject to price slippage that might induce so-called outliers. Secondly, we test the presence of time-varying jumps in leading cryptocurrencies[4] that are generally characterised by extreme volatility that can be associated with specific events such as forks, hacks, and thefts. Our examination adds to prior studies that focus on Bitcoin only and argues that the existence of jumps can substantially impact the structure of losses and gains related to Bitcoin (e.g., Chaim and Laurini 2018). This is crucial given that jumps represent an important element of an asset's risk and are an input into option pricing models, and thereby can help enhance the accuracy of model prediction.

The rest of the paper is structured in three sections. Section 2 describes the dataset and methods used to detect outliers and model the time-varying jumps. Section 3 presents and discusses the empirical results. Section 4 concludes.

2. Data and Methods

2.1. Data

We collected the daily prices of five leading cryptocurrencies (Bitcoin, Ethereum, Ripple, Dogecoin, and Litecoin) against USD, from cryptomarketcap.com (accessed on 18 November 2021). We also collected price data on a broad cryptocurrency index (CCI30) from https://cci30.com (accessed on 18 November 2021).

Notably, those five cryptocurrencies were selected from the largest 20 cryptocurrencies not only because they represent 65% of the market capitalisation of all cryptocurrencies but also because they have the longest common data sample period that starts from 8 August 2015. Accordingly, our sample period is 8 August 2015–23 September 2021, yielding 2239 daily price observations.

Given that our methods require stationary data, we used log return series and the summary statistics of these return series (Table 1) to exhibit evidence of stationarity as indicated by the Phillips–Perron (PP) test.

2.2. Outlier Detection Method

We follow Ané et al. (2008) in detecting the presence of outliers. Let R_t be the log return on the cryptocurrency on day t, which follows an AR(2)-GARCH(1,1) model:[5]

$$R_t = b_0 + b_1 R_{t-1} + b_1 R_{t-2} + \varepsilon_t \quad (1)$$

$$\sigma_t^2 = a_0 + a_1 \varepsilon_{t-1}^2 + a_2 \sigma_{t-1}^2 \quad (2)$$

where $\varepsilon_t = \sigma_t z_t$ which follows Student's t distribution. I_{t-1} refers to the filtration of information at time $t-1$.

R_{t+1} is considered an outlier if it does not belong to the following interval:

$$R_{t+1} \in \left[R_{t,t+1} \pm F\left(1 - \frac{\alpha}{2}\right) \sigma_{t,t+1} \right]$$

where, $R_{t,t+1}$ is the one-step ahead return forecast given by:

$$R_{t,t+1} = E(R_{t+1}/I_t) = b_0 + b_1 R_t + b_2 R_{t-1}$$

and $\sigma_{t,t+1}^2$ denotes the one-step ahead variance forecast defined as:

$$\sigma_{t,t+1}^2 = \text{var}(R_{t+1}/I_t) = a_0 + (a_1 + a_2)\sigma_t^2$$

Furthermore, $F\left(1 - \frac{\alpha}{2}\right) = P(z_t \leq 1 - \alpha/2)$ is a fractile of the assumed conditional distribution.

The above detection procedure is rolled over until the end of the sample period. Notably, the detection procedure is robust to any model misspecifications (Ané et al. 2008).

Note that a number of recent studies have used this method to detect outliers in different financial markets. Dutta (2018a), for instance, shows that outliers play a crucial role in modelling the volatility of the EU emissions market. Another study by Dutta (2018b) finds similar results for the precious metals market.

2.3. The GARCH-Jump Process

Following the model of Chan and Maheu (2002) and its recent use by Liu et al. (2021) and Li et al. (2021), the GARCH-jump specification is:

$$R_t = \pi + \sum_{i=1}^{n} \mu_i R_{t-i} + \epsilon_t \quad (3)$$

where R_t is the log return of the cryptocurrency at time t, and ϵ_t denotes the error term at time t. ϵ_t has two components:

$$\epsilon_t = \epsilon_{1t} + \epsilon_{2t} \quad (4)$$

where ϵ_{1t} is defined as:

$$\epsilon_{1t} = \sqrt{h_t} z_t, \; z_t \sim \text{Student's } t$$

$$h_t = \omega + \alpha \epsilon_{1t-1}^2 + \beta h_{t-1} \quad (5)$$

and ϵ_{2t} is a jump innovation that consists of abnormal price movements with $E(\epsilon_{2t}|L_{t-1}) = 0$, where L_{t-1} designates the information set. Now, ϵ_{2t} is defined as the discrepancy between the jump component and the expected total jump size between t–1 and t:

$$\epsilon_{2t} = \sum_{l=1}^{n_t} U_{tl} - \theta \lambda_t \quad (6)$$

where U_{tl} denotes the jump size, which is assumed to be normally distributed with mean θ and variance d^2, $\sum_{l=1}^{n_t} U_{tl}$ is the jump component, and n_t indicates the number of jumps. n_t follows a Poisson variable with an autoregressive conditional jump intensity as:

$$\lambda_t = \lambda_0 + \rho \lambda_{t-1} + \gamma \xi_{t-1} \quad (7)$$

where λ_t is the time-varying conditional jump intensity parameter, λ_0 refers to a constant jump intensity, and ξ_{t-1} indicates the intensity residual with $\lambda_t > 0$, $\lambda_0 > 0$, $\rho > 0$ and $\gamma > 0$.

The log-likelihood function is:

$$L(\Omega) = \sum_{t=1}^{T} \log f(R_t | I_{t-1}; \Omega) \tag{8}$$

where $\Omega = (\pi, \mu_i, \omega, \alpha, \beta, \theta, d, \lambda_0, \rho, \gamma)$.

3. Empirical Results

3.1. Outliers

The findings from the outlier detection process suggest that extreme observations occur only in the Bitcoin return series.[6] Overall, we have found 16 outliers during the sample period. We also document that these outliers are mainly present after the soar. It is worth noting that such outliers could arise due to different significant events or news including wars, political conflicts, cyberattacks, and economic downturns. Based on these findings, we consider both the original return series and the outlier-free return series. Table 1 shows that the standard deviation of Bitcoin returns is reduced by almost 26% after correcting for outliers, while outlier correction substantially increases the mean return of Bitcoin. This result suggests that Bitcoin returns are contaminated by more negative return outliers than positive ones. Additionally, outlier correction reduces the kurtosis and skewness for Bitcoin. Interestingly, the skewness converges to zero.

3.2. Time-Varying Jumps

The results of GARCH-Jump model are shown in Table 2. The GARCH parameters are statistically significant, and the sum of α and β indicates a high degree of volatility persistence. The jump intensity parameters (λ_0, ρ, γ) are statistically significant for Bitcoin and Litecoin, implying time-variability in the jump intensity and evidencing large abrupt price variations. Taking Bitcoin as an example, the parameter ρ (0.9741) being high and significant indicates that the time-varying jump intensity is persistent. The γ parameter, which measures the sensitivity of λ_t to past shock, ξ_{t-1}, is 0.3832, suggesting that a unit increase in ξ_{t-1} results in a dampened effect (0.3832) on the next period's jump intensity.

Overall, the jump intensity parameters satisfy the constraints that $\lambda_0 > 0$, $\rho > 0$ and $\gamma > 0$, implying that the GARCH-jump model is a proper choice for describing volatility dynamics and jump behaviour in the cryptocurrency markets. Additionally, the positive values of ρ and γ for Bitcoin, Litecoin, Dogecoin, Ethereum, and the CCI30 index indicate that the current jump intensity (λ_t) is influenced by the most recent jump intensity (λ_{t-1}) and the intensity residuals (ξ_{t-1}). The high values of ρ and γ, especially for Bitcoin and CCI30, suggest a high degree of persistence in the jump intensity. For Ripple returns, only the parameter of time-varying jump intensity is significant.

The results involving Bitcoin are generally in line with Chaim and Laurini (2018). The findings of the outlier-corrected data for Bitcoin show that jumps still exist after taking into account the presence of outliers. The likelihood ratio test suggests that the GARCH-jump model using Bitcoin outlier-free data outperforms the one using Bitcoin original data (i.e., Bitcoin data not corrected for outliers).

These findings suggest cryptocurrencies are not only characterised by time-varying volatility, but also by extreme price movements, which exceed the current respective market volatility. Such jump behavior points towards an instable condition in the market and hence the information on cryptocurrency prices could mislead the investment decisions (Dutta 2018b). Our analysis is, therefore, important for investors in making proper asset-allocation decisions.

It is also noteworthy that time dependent jumps may provide early signals of significant downturns in cryptocurrency markets. Earlier studies (Chan and Maheu 2002; Maheu and McCurdy 2004) also document that the conditional expected number of jumps in different asset classes tends to increase and that the information on such time-varying jumps could be used in predicting future market crashes. We thus conclude that the jump dynamics in cryptocurrency returns could capture the adverse impact of negative news or events (e.g., COVID-19 pandemic) on their price levels.

Table 2. Estimates of GARCH-jump model.

	Bitcoin	Bitcoin (Outlier-Free)	Litecoin	Ripple	Dogecoin	Ethereum	CCI30
π	0.0841 ***	0.0783 ***	−0.1179 **	−0.0051	0.2144 **	0.1159	0.0651 *
μ_1	0.0056	−0.0329	−0.0987 *	−0.0899	−0.1853 *	−0.0661 *	−0.5323 ***
μ_2	0.0062	0.0547	−0.1123 **	−0.1164 **	0.2188		
ω	0.0107 *	0.0606	0.0831 ***	0.1241 **	0.0844	0.0700 *	0.0441 **
α	0.1072 ***	0.1068 **	0.1553 ***	0.1455 ***	0.0981 **	0.1126 **	0.0676 ***
β	0.7739 ***	0.7455 ***	0.7249 ***	0.5666 ***	0.7252 ***	0.5165 ***	0.7865 ***
θ	−0.0831	−0.0961	0.4438 ***	0.0961	0.1176	−0.0762	−2.1729 ***
d^2	2.0400 ***	−0.9976 ***	3.8976 ***	2.8903 ***	2.1439 ***	1.6754 ***	−3.5208 ***
λ_0	0.0699 ***	0.0502 **	0.0986 ***	0.0346	0.0334	0.1345 ***	0.0014
ρ	0.9658 ***	0.9189 ***	0.7242 ***	0.9054 ***	0.7119 **	0.8764 ***	0.9958 ***
γ	0.3939 ***	0.2956 ***	0.3001 ***	0.1679	0.3973 **	0.4138 ***	0.3489 ***
Log-likelihood	−3857.14	−3201.57	−723.76	−998.61	−941.54	−788.18	−1922.98

Notes: This table shows the estimated coefficients of the GARCH-jump model, as described in Section 2.3. π and μ are parameters depicting the conditional mean (see Equation (3)). ω, α, and β are parameters depicting the conditional variance (see Equation (5)). λ_0, ρ, and γ are parameters describing the time-varying jump intensity (see Equation (7)). θ and d^2 are the mean and variance of the jump size, respectively (see Equation (6)). ***, ** and * indicate statistically significant results at 1%, 5% and 10% levels, respectively.

4. Conclusions

In this paper, we have extended the limited understanding on the presence of outliers and time-varying jumps in the cryptocurrency markets. The main results show that outliers exist only in Bitcoin returns, suggesting the importance of accounting for them, and Bitcoin returns are characterised by time-varying jumps after correcting for outliers. Litecoin is also characterised by time-varying jumps. The findings complement previous studies (Katsiampa 2017; Chu et al. 2017; Chaim and Laurini 2018) and point to the presence of abrupt price variations in some cryptocurrencies, suggesting potential suitability of including jumps when pricing options on Bitcoin and Litecoin. Given recent evidence on the importance of jumps for portfolio management, future studies could consider dynamic portfolio allocation and risk management inferences in the cryptocurrency markets with time-varying jump risk (Zhou et al. 2019).

Author Contributions: Conceptualization and methodology, A.D. and E.B.; formal analysis, A.D.; data curation, A.D. and E.B.; writing—original draft preparation, A.D. and E.B.; writing—review and editing, A.D. and E.B.; visualization, A.D.; project administration, E.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are openly available in cryptomarketcap.com and https://cci30.com (accessed on 20 November 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. Bitcoin price skyrocketed for most of 2016–2017, then crashed for most of 2018, and then experienced large up and down swings.
2. Thies and Molnár (2018) focus on the Bitcoin market. Using a Bayesian change point model, they show evidence of structural breaks in the first and second moments of the return distribution.
3. We pay a special attention to Dogeeoin due the influence of Elon Musk's tweets on the price dynamic of Dogecoin from early 2021 and therefore the possible change in the characteristics of Dogecoin after the soar of its price from that date.
4. Cryptocurrencies can be very prone to jumps due to the presence of hacks and forks.

[5] For Ethereum and CCI30 index, the AR(1)-GARCH(1,1) process appears to be the best fitted model based on the AIC and BIC values.

[6] Quite similar findings are reported by Thies and Molnár (2018) who use a Bayesian change point model and report evidence of structural breaks in the Bitcoin market.

References

Ané, Thierry, Loredana Ureche-Rangau, Jean-Benoît Gambet, and Julien Bouverot. 2008. Robust outlier detection for Asia-Pacific stock index returns. *Journal of International Financial Markets, Institutions & Money* 18: 326–43.

Bouri, Elie, Syed Jawad Hussain Shahzad, and David Roubaud. 2019. Co-explosivity in the cryptocurrency market. *Finance Research Letters* 29: 178–83. [CrossRef]

Bouri, Elie, Rangan Gupta, and Xuan Vinh Vo. 2020. Jumps in geopolitical risk and the cryptocurrency market: The singularity of Bitcoin. *Defence and Peace Economics*, 1–12. [CrossRef]

Bouri, Elie, Syed Jawad Hussain Shahzad, David Roubaud, Ladislav Kristoufek, and Brian Lucey. 2020. Bitcoin, gold, and commodities as safe havens for stocks: New insight through wavelet analysis. *The Quarterly Review of Economics and Finance* 77: 156–64. [CrossRef]

Carnero, M. Angeles, Daniel Peña, and Esther Ruiz. 2012. Estimating GARCH volatility in the presence of outliers. *Economics Letters* 114: 86–90. [CrossRef]

Chaim, Pedro, and Marcio P. Laurini. 2018. Volatility and return jumps in Bitcoin. *Economics Letters* 114: 158–63. [CrossRef]

Chaim, Pedro, and Marcio P. Laurini. 2019. Is Bitcoin a bubble? *Physica A: Statistical Mechanics and Its Applications* 517: 222–32. [CrossRef]

Chan, Wing H., and John M. Maheu. 2002. Conditional jump dynamics in stock market returns. *Journal of Business & Economic Statistics* 20: 377–89.

Chen, Zheshi, Chunhong Li, and Wenjun Sun. 2020. Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics* 365: 112395. [CrossRef]

Chu, Jeffrey, Stephen Chan, Saralees Nadarajah, and Joerg Osterrieder. 2017. GARCH modelling of cryptocurrencies. *Journal of Risk and Financial Management* 10: 17. [CrossRef]

Corbet, Shaen, Douglas J. Cumming, Brian M. Lucey, Maurice Peat, and Samuel Vigne. 2019. Investigating the Dynamics between Price Volatility, Price Discovery, and Criminality in Cryptocurrency Markets. Price Discovery, and Criminality in Cryptocurrency Markets. Available online: http://dx.doi.org/10.2139/ssrn.3384707 (accessed on 28 November 2021).

Cunha, Paulo Rupino, Paulo Melo, and Helder Sebastião. 2021. From Bitcoin to Central Bank Digital Currencies: Making Sense of the Digital Money Revolution. *Future Internet* 13: 165. [CrossRef]

Das, Debojyoti, Corlise Le Roux, Rabin K. Jana, and Anupam Dutta. 2020. Does Bitcoin hedge crude oil implied volatility and structural shocks? A comparison with gold, commodity and the US Dollar. *Finance Research Letters* 36: 101335. [CrossRef]

Dutta, Anupam. 2018a. Modeling and Forecasting the Volatility of Carbon Emission Market: The Role of Outliers, Time-varying Jumps and Oil Price Risk. *Journal of Cleaner Production* 172: 2773–81. [CrossRef]

Dutta, Anupam. 2018b. Impacts of Oil Volatility Shocks on Metal Markets: A Research Note. *Resources Policy* 55: 9–19. [CrossRef]

Dutta, Anupam, Debojyoti Das, R.K. Jana, and Xuan Vinh Vo. 2020. COVID-19 and oil market crash: Revisiting the safe haven property of gold and Bitcoin. *Resources Policy* 69: 101816. [CrossRef]

Grané, Aurea, and Helena Veiga. 2010. Wavelet-based detection of outliers in financial time series. *Computational Statistics & Data Analysis* 54: 2580–93.

Hatemi-J., Abdulnasser, Mohamed A. Hajji, Elie Bouri, and Rangan Gupta. 2020. The Benefits of Diversification between Bitcoin, Bonds, Equities and the US Dollar: A Matter of Portfolio Construction. *Asia-Pacific Journal of Operational Research 2040024*. [CrossRef]

Ji, Qiang, Elie Bouri, Chi Keung Marco Lau, and David Roubaud. 2019. Dynamic connectedness and integration in cryptocurrency markets. *International Review of Financial Analysis* 63: 257–72. [CrossRef]

Katsiampa, Paraskevi. 2017. Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters* 158: 3–6. [CrossRef]

Li, Shaoyu, Kaixuan Ning, and Teng Zhang. 2021. Sentiment-aware jump forecasting. *Knowledge-Based Systems* 228: 107292. [CrossRef]

Liu, Feng, Chuanguo Zhang, and Mengying Tang. 2021. The impacts of oil price shocks and jumps on China's nonferrous metal markets. *Resources Policy* 73: 102228. [CrossRef]

López-Martín, Carmen, Sonia Benito Muela, and Raquel Arguedas. 2021. Efficiency in cryptocurrency markets: New evidence. *Eurasian Economic Review* 11: 403–31. [CrossRef]

Maheu, John M., and Thomas H. McCurdy. 2004. News Arrival, Jump Dynamics and Volatility Components for Individual Stock Returns. *Journal of Finance* 59: 755–93. [CrossRef]

Mostafa, Fahad, Pritam Saha, Mohammad Rafiqul Islam, and Nguyet Nguyen. 2021. GJR-GARCH Volatility Modeling under NIG and ANN for Predicting Top Cryptocurrencies. *Journal of Risk and Financial Management* 14: 421. [CrossRef]

Noda, Akihiko. 2021. On the evolution of cryptocurrency market efficiency. *Applied Economics Letters* 28: 433–39. [CrossRef]

Salisu, Afees A., and Ahamuefula E. Ogbonna. 2021. The return volatility of cryptocurrencies during the COVID-19 pandemic: Assessing the news effect. *Global Finance Journal*, 100641. [CrossRef]

Shahzad, Syed Jawad Hussain, Elie Bouri, Sang Hoon Kang, and Tareq Saeed. 2021. Regime specific spillover across cryptocurrencies and the role of COVID-19. *Financial Innovation* 7: 1–24. [CrossRef]

Thies, Sven, and Peter Molnár. 2018. Bayesian change point analysis of Bitcoin returns. *Finance Research Letters* 27: 223–27. [CrossRef]

Tiwari, Aviral Kumar, Satish Kumar, and Rajesh Pathak. 2019. Modelling the dynamics of Bitcoin and Litecoin: GARCH versus stochastic volatility models. *Applied Economics* 51: 4073–82. [CrossRef]

Walther, Thomas, Tony Klein, and Elie Bouri. 2019. Exogenous Drivers of Bitcoin and Cryptocurrency Volatility: A Mixed Data Sampling Approach to Forecasting. *Journal of International Financial Markets, Institutions & Money* 63: 101133.

Yousaf, Imran, Shoaib Ali, Elie Bouri, and Anupam Dutta. 2021. Herding on fundamental/non-fundamental information during the COVID-19 outbreak and cyber-attacks: Evidence from the cryptocurrency market. *SAGE Open* 11: 21582440211029911. [CrossRef]

Zhou, Chunyang, Chongfeng Wu, and Yudong Wang. 2019. Dynamic portfolio allocation with time-varying jump risk. *Journal of Empirical Finance* 50: 113–24. [CrossRef]

Journal of *Risk and Financial Management*

MDPI

Article

Multifactor Market Indexes

Wei Liu [1] **and James W. Kolari** [2,*]

[1] USAA Bank, San Antonio, TX 78288, USA; yuna0831@gmail.com
[2] Department of Finance, Mays Business School, Texas AM University, College Station, TX 77843, USA
* Correspondence: j-kolari@tamu.edu

Abstract: This paper combines the CRSP market index with multiple factors to create a single multifactor market index. Empirical tests of different multifactor market indexes indicate that: (1) Sharpe ratios substantially increase and GRS test statistics decrease as multifactors are incrementally added to the CRSP index; and (2) the resultant multifactor market indexes are significantly priced in cross-sectional tests of associated beta loadings with *t*-values exceeding 3.0 in most cases.

Keywords: market index; market factor; multifactors; efficient portfolios

1. Introduction

Cochrane (2011) has humorously alluded to the growing list of multifactors in asset pricing as a "factor zoo". With so many contenders as factors, what factors should be used in an asset pricing model? Harvey et al. (2016) investigated over 300 proposed multifactors in asset pricing models in an effort to discern significant factors versus false discoveries. In addition, Chordia et al. (2020) examined false factor discoveries in asset pricing by studying 2 million trading strategies with real data. Both studies recommended that *t*-statistics associated with factor loadings in Fama and MacBeth (1973) cross-sectional regression tests should exceed 3 to avoid false factor discoveries. This higher statistical hurdle substantially reduces the number of acceptable factors.

Given a smaller set of factors, a multifactor mall problem still remains. For example, assuming 20 valid factors, a large number of models are conceivable. Which combinations of factors and models should be used? In this paper, we propose a solution to this problem. Specifically, we aggregate numerous factors into a single market index. We hypothesize that the CRSP market index (i.e., market portfolio proxy) can be combined with popular multifactors to create more efficient aggregate market indexes dubbed multifactor market indexes. The rationale for these new aggregate indexes is straightforward. If the CRSP index and investable multifactors provide a well-specified asset pricing model for U.S. stock returns, then combining them into a single mean-variance efficient portfolio should be possible. Assuming that a multifactor market index is a linear combination that is efficient, it should be significantly priced as a single factor in a market beta pricing model. Hence, a large number of investable multifactors can be incorporated into a single market index to develop low-dimensional and more parsimonious models. While numerous authors have proven in theory that a combination of factor-mimicking portfolios is minimum variance efficient (e.g., Huberman and Kandel (1987); Shanken (1987); Fama (1996); Shanken and Weinstein (2006); Kan and Zhou (2008), and others), no previous studies to the authors' knowledge have verified this result in empirical tests. Filling this gap in the literature, we incrementally combine popular factors with the CRSP index to show that the resultant aggregate market indexes are increasingly efficient and priced in the cross section of average stock returns.

To construct efficient multifactor market indexes, we utilize portfolio weights for the CRSP index and each multifactor based on their relative Sharpe ratios. In this regard, we incrementally add five prominent multifactors, including size, value profit, investment,

Citation: Liu, Wei, and James W. Kolari. 2022. Multifactor Market Indexes. *Journal of Risk and Financial Management* 15: 155. https://doi.org/10.3390/jrfm15040155

Academic Editor: Robert Brooks

Received: 24 February 2022
Accepted: 14 March 2022
Published: 30 March 2022

and momentum, to the value-weighted CRSP index. As expected, Sharpe ratios gradually increase as more multifactors are incorporated into the CRSP index. For example, in the sample period July 1963 to December 2016, the CRSP index has a Sharpe ratio of 0.12 compared to 0.34 for a market index combining the CRSP index with the aforementioned five multifactors. GRS time-series regression tests indicate that market indexes become gradually more efficient as multifactors are added. Importantly, cross-sectional Fama and MacBeth regression tests show that multifactor market indexes become increasingly significant as multifactors are incorporated into the CRSP index. We obtain *t*-values associated with multifactor market index beta loadings greater than 3.0 for numerous test asset portfolios, which exceeds the recommended threshold for statistical significance. Hence, multifactors improve approximations of efficient market portfolio returns in the sense that multifactor market indexes become significantly priced in the cross-section of average stock returns.

Further analyses construct an aggregate industry index using our relative Sharpe ratio weighting procedure. It is well known that industry portfolios are not priced by the CRSP index and popular multifactors. By contrast, we find that our new industry index is significantly priced using industry portfolios as test assets. We subsequently combine industry factors and popular multifactors to form a more general multifactor market index that garners high *t*-values typically ranging from 3 to 6 for a variety of test assets, in addition to insignificant mispricing errors and significant pricing of industry portfolios.

Regarding model selection, we find that the momentum factor continues to be significantly priced in cross-sectional tests even when incorporated in multimarket indexes. Hence, we interpret this evidence to mean that momentum is a possible strong factor. Of course, as more multifactors are incorporated into multifactor market indexes to boost their efficiency, the hurdle for strong factors is raised. Future studies may well find that momentum is not a strong factor in the case of more efficient market multifactor indexes. More generally, it is likely that few strong factors exist, such that low-dimensional models are possible.

We conclude that combining multifactors with the CRSP index enables the formation of single market indexes that are efficient. Unlike many studies that reject market beta using the CRSP market index and other stock market indexes, multifactor market indexes lend support for the notion that asset returns are a linear function of a general market index. An important implication is that asset pricing models incorporating a market factor can benefit from multifactor market factors. Currently, the CRSP index continues to be used as the market proxy in most asset pricing models even though it is not priced in the cross section of stock returns. Multifactor market indexes that are significantly priced can potentially lead to more parsimonious and robust asset pricing models. Since the market factor is commonly used in many areas of corporate, investment, and institutional finance, widespread applications of multifactor market indexes are possible in future studies. Another implication is that similar efficient aggregate indexes can be constructed in the real world for investment purposes, thereby benefiting many investors including those saving for retirement.

Section 2 provides background discussion of multifactor market indexes. Section 3 combines the CRSP index with popular multifactors to form a variety of multifactor market indexes based on U.S. stock returns. Descriptive statistics, times-series tests, and cross-sectional tests of multifactor market indexes are provided, in addition to results for an aggregate industry factor and discussion of momentum as a possible strong factor. Section 4 concludes.

2. Multifactor Market Indexes

The Capital Asset Pricing Model (CAPM) of Treynor (1961, 1962), Sharpe (1964), Lintner (1965), Mossin (1966), and Black (1972) is based on a mean-variance efficient market portfolio computed as the value-weighted return on all marketable assets. Roll (1977), and others have shown that this portfolio is a minimum variance portfolio if and only if beta associated with the market factor is priced for all assets in this portfolio. Unfortunately, Fama and French (1992, 1993, 1995, 1996a, 1996b) found that CAPM beta loadings associated with the market portfolio proxied by the CRSP stock index timates were not significantly priced in the cross section of average stock returns. In view of this failure, they concluded that CAPM beta was dead and proposed a three-factor model that augments the CRSP market factor with largely orthogonal size and value factors defined as zero-investment portfolios with long and short positions. The success of this multifactor innovation triggered a plethora of related studies.

Due to the CAPM's failure as discussed in the previous section, multifactor models have arisen with theoretical support from Ross' (1976) arbitrage pricing theory (APT) and Merton's (1973) intertemporal capital asset pricing model (ICAPM). Carhart (1997) added a zero-investment momentum factor to the three-factor model to study mutual fund performance. Subsequently, Chen and Zhang (2010) advanced another three-factor model by replacing size and value with profit and investment factors. A related paper was later published by Hou et al. (2014), which proposed a q-factor model with four factors (viz., market, size, investment, and return on equity) grounded in neoclassical investment q-theory. Similar to the q-factor model, Fama and French (2015) added profit and investment factors to their three-factor model to create a five-factor model. In addition, Stambaugh and Yuan (2017) proposed a four-factor model including the market and size factors plus two mispricing factors (viz., management and performance).[1] Hou et al. (2018) added a growth factor to the q-model. Barillas and Shanken (2018) formed a six-factor model by including value and momentum factors. Additionally, Fama and French (2018) added momentum to their five-factor model to form a six-factor model. Subsequently, Fama and French (2020) proposed cross-section factors developed from Fama–MacBeth regressions as well as conditional models with time-varying factor loadings. Many other models containing different factors have been proposed in the asset pricing literature.[2] This proliferation of factors has resulted in a *model mall problem*. Which model should an academic researcher or professional investment manager use? Is there a way to condense well-accepted factors into a more parsimonious model?

Using the notation and discussion in Ferson (1995, 2019), multifactor models of the expected return on the ith asset take the familiar cross-sectional form:

$$E_t(R_{i,t+1}) = \lambda_{0t} + \sum_{k=1}^{K} b_{ikt}\lambda_{kt}, \text{for all } i, \quad (1)$$

where $b_{i1t}, \ldots, b_{iKt}$ are time t conditional betas for asset i related to K risk factors, and λ_{kt} are market-wide risk premiums for $k = 1, \ldots, K$ risk factors equal to the incremental expected return per unit type-j beta. The intercept λ_{0t} is the riskless return or expected zero-beta rate conditionally uncorrelated with the K risk factor loadings if no riskless asset exists. The conditional betas are estimated from the time-series factor model:

$$R_{i,t+1} = a_{it} + \sum_{k=1}^{K} b_{ijt}F_{k,t+1}, \text{for all } i, \quad (2)$$

where $E_t(\mu_{i,t+1}F_{k,t+1}) = E_t(\mu_{i,t+1}) = 0$ for all i and k. The model is well-specified when the factor portfolios form the tangency portfolio (i.e., $a_{it} = 0$). As proven by many authors, Equation (2) implies that a combination of K factor-mimicking portfolios is the minimum variance efficient (see Grinblatt and Titman (1987); Huberman et al. (1987); Jobson and Korkie (1982); Gibbons et al. (1989); Kan and Zhou (2008), and Ferson and Siegel (2009), among others).

Extending this literature, MacKinlay (1993) argued that, if a linear combination of factor portfolios cannot identify the efficient tangency portfolio, there exists an optimal orthogonal portfolio of N assets, which when combined with K factor portfolios, forms the tangency portfolio. This unique portfolio is orthogonal to the factor portfolios.[3] Using this approach, he showed that a bound on the Sharpe ratio exists if the deviation from the single factor model (e.g., CAPM) can be accomplished by a common component of the residual variance. In other words, it is possible that the tangency portfolio can be located by identifying a complete (but limited) set of orthogonal risk factors. This tangency portfolio has the maximum squared Sharpe measure among all portfolios.

In this paper we employ multifactors to develop more efficient aggregate indexes of the stock market. The value-weighted CRSP stock market index is used to proxy the theoretical orthogonal optimal portfolio. We assume that all market information is known and investors seek the most efficient portfolio based on available information. Given the long-only portfolio excess return denoted as R_L, we add $k = 1, \ldots, K$ orthogonal factor portfolios to form new more efficient portfolios with higher Sharpe ratios. Factor portfolios are investable zero-investment portfolios.[4] We define the corresponding Sharpe ratios as

$$\begin{aligned} S_L &= \frac{\mu_L}{\sigma_L} \\ S_k &= \frac{\mu_k}{\sigma_k}, \; k = 1, \ldots, K, \end{aligned} \tag{3}$$

where μ is the excess return, and σ is the volatility of the portfolio. Combining the long portfolio with the first zero-investment (long/short) factor portfolio F_1, the aggregate index return R_I is

$$R_I = R_L + x_1 F_1, \tag{4}$$

with variance equal to $\sigma_L^2 + x_1^2\sigma_1^2$ and Sharpe ratio

$$S_I = \frac{\mu_L + x_1\mu_1}{\sqrt{\sigma_L^2 + x_1^2\sigma_1^2}}. \tag{5}$$

The first order condition $\partial S_I / \partial x_1$ gives

$$x_1 = \frac{S_1}{S_L}\frac{\sigma_L}{\sigma_1} = \frac{S_1^2}{S_L^2}\frac{\mu_L}{\mu_1}. \tag{6}$$

Now the Sharpe ratio for market index I can be written as

$$S_I = \frac{\mu_L + (\frac{S_1^2}{S_L^2}\frac{\mu_L}{\mu_1})\mu_1}{\sqrt{\sigma_L^2 + (\frac{S_1}{S_L}\frac{\sigma_L}{\sigma_1})^2\sigma_1^2}} = \frac{(1 + \frac{S_1^2}{S_L^2})\mu_L}{\sqrt{1 + \frac{S_1^2}{S_L^2}}\sigma_L} = \sqrt{1 + \frac{S_1^2}{S_L^2}}S_L > S_L \tag{7}$$

Thus, this new market index portfolio is more efficient than the long-only portfolio index. Upon continuing this process by incrementally adding more zero-investment factors to candidate portfolio P, the market index portfolio's efficiency is increased. As recognized by MacKinlay (1993), the number of true risk factors will be limited even for increasing numbers of assets in the market. Of course, if a candidate market index portfolio is the tangency portfolio, it is not possible to increase the Sharpe ratio by combining it with another zero-investment factor.[5]

It is worthwhile noting that Equation (7) can alternatively be written as

$$S_I = \sqrt{S_L^2 + S_I^2}. \tag{8}$$

After iteratively adding $k = 1, \ldots, K$ factors, the optimized Sharpe ratio is

$$S_F = \sqrt{\sum_{k=1}^{K} S_k^2} = \sqrt{\mu' \Sigma^{-1} \mu}, \quad (9)$$

which is the ex post tangency portfolio defined by Gibbons et al. (1989). Note that this approach does not provide the weight x_i in Equation (6) to form efficient portfolio indexes.

In the next section we conduct empirical tests of these concepts. First, long/short portfolios based on zero-investment portfolios (or factors) are added to the CRSP index to determine whether more efficient market indexes can be constructed. The market indexes should become increasingly efficient as zero-investment portfolios are added based on size, value, profit, capital investment, and momentum. We also form a market index that incorporates both the aforementioned multifactors plus zero-investment portfolios that are long industry portfolios and short Treasury bills. Second, we perform time-series and cross-sectional tests of multifactor market indexes to determine if they are significantly priced and therefore efficient.

3. Empirical Tests

Here we report the empirical results for multifactor market indexes based on the following popular risk factors: market (CRSP index), size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). Further analyses incorporate zero-investment industry portfolios as well. Monthly returns for the value-weighted CRSP market index, zero-investment multifactors, industry portfolios, and Treasury bills are downloaded from Kenneth French's website.[6] Using the weighting procedure in the previous section, we construct multifactor market indexes that combine the CRSP index with five multifactors.

Because factors are not uncorrelated, we orthogonalize them as follows. Denoting the current market index return as $R(I_K)$ (e.g., CRSP index) and the multifactor return (e.g., size factor) to be added to this market index as F_{k+1}, we regress F_{k+1} on the excess market return $R(I_{K,t}) - R_{f,t}$ over t sample period months:

$$F_{k+1,t} = \alpha_{k+1} + \beta_{k+1}[R(I_{K,t}) - R_{f,t}] + \epsilon_{k+1,t}. \quad (10)$$

The orthogonalized multifactor is $F_{k+1,t}^{new} = F_{k+1,t} - \beta_{k+1}[R(I_{K,t}) - R_{f,t}]$. The portfolio I_{K+1} return combining the previous portfolio I_K return plus the multifactor portfolio return is:

$$R(I_{K+1,t}) = R(I_{K,t}) + x_{k+1} F_{k+1,t}^{new}. \quad (11)$$

After substituting $F_{k+1,t}^{new}$ and rearranging terms, we have:

$$R(I_{K+1,t}) = R_{f,t} + (1 - x_{k+1}\beta_k)[R(I_{K,t}) - R_{f,t}] + x_{k+1} F_{k+1,t}. \quad (12)$$

This equation decomposes the new multifactor market index return $R(I_{K+1,t})$ into the riskless rate plus weighted premiums for the previous index excess return $R(I_{K,t}) - R_{f,t}$ and newly-added zero-investment factor return $F_{k+1,t}$. Notice that the estimated β_{k+1} coefficient in Equation (10) affects the relative weights in the construction of return series $R(I_{K+1,t})$. As noted by an anonymous referee, errors in estimated coefficients will affect these weights and, in turn, multifactor market index returns. This potential bias is beyond the scope of the present research and is therefore left for future research.

To determine the weight x_k for the new factor in this multifactor market index, we apply Equation (6). These steps are repeated to incrementally add the five popular long/short factors F_{k+1} ($k + 1 = 2, \ldots, 6$) to each successive market index $I_{K+1,t}$ in month t to create five new market indexes with respective monthly return series $R(I_{K+1,t})$ (i.e., $R(I_2, t), \ldots, R(I_6, t)$). The return series $R(I_{K,t}) = R(I_1, t)$ is the CRSP index.

3.1. Descriptive Statistics

Based on the sample period from July 1963 to December 2016, descriptive statistics in Table 1 are provided for the CRSP index (denoted I_1), five multifactors (i.e., SMB, HML, RMW, CMA, and MOM), and multifactor market indexes combining the CRSP index with various multifactors (denoted I_2 to I_6). The familiar CRSP index (I_1) has mean excess returns of 0.51 percent per month, standard deviation of 4.42 percent, and Sharpe ratio of 0.12. With the exception of momentum, the descriptive statistics for the multifactors are not too different from one another. Momentum (MOM) has a noticeably higher mean return of 0.66 percent per month compared to the excess return on the CRSP index, whereas the other multifactors have mean returns less than the excess return on the CRSP index in the range of 0.24 percent to 0.37 percent.

Table 1. Descriptive statistics for the CRSP index, five multifactors, and market indexes combining the CRSP index with various mult-factors and industry portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides descriptive statistics for monthly excess returns (over the Treasury bill rate) on the value-weighted CRSP index, zero-investment portfolio returns of five popular multifactors, and excess returns of seven market indexes combining CRSP index excess returns with these multifactors in addition to industry excess returns. We downloaded the following multifactors and industry returns from Kenneth French's website: size (SMB), value (HML), profit (RMW), capital investment (CMA), momentum (MOM), and 30 industries. Multifactor market indexes are formed using the following steps. First, the size (SMB) factor monthly returns are regressed on CRSP index (I_1) excess returns over the Treasury bill rate. The residual term from this regression is utilized as the orthogonalized factor. Second, this orthogonalized size factor is added to the CRSP index using Equation (11) to compute the return for the new multifactor market index I_2 as $R(I_2) = R(I_1) + x_1$SMB. Third, value (HML) factor returns are regressed on the new $R(I_1) + x_1$SMB portfolio excess returns to obtain the orthogonalized value factor. This residual value factor is added to the CRSP + SMB portfolio to get market index I_3 = CRSP + x_1SMB + x_2HML. Fourth, the last step is repeated to sequentially create market index I_4 = CRSP + x_1SMB + x_2HML + x_3RMW, market index I_5 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA, and market index I_6 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA + x_5MOM. We drop the CMA multifactor to form market index I_7 = CRSP + x_1SMB + x_2HML + x_3RMW + x_5MOM. Lastly, 30 industry excess returns are added to create market index I_8. For comparison purposes, mean excess returns, standard deviations of returns, and Sharpe ratios for multifactors and different market indexes are computed.

Portfolios	Mean	Std. Dev.	Sharpe Ratio
I_1 = CRSP	0.51	4.42	0.12
SMB	0.27	3.04	0.09
HML	0.37	2.81	0.13
RMW	0.24	2.23	0.11
CMA	0.31	2.01	0.15
MOM	0.66	4.22	0.16
I_2 = CRSP + SMB	0.64	4.94	0.13
I_3 = CRSP + SMB + HML	1.71	8.10	0.21
I_4 = CRSP + SMB + HML + RMW	2.66	10.09	0.26
I_5 = CRSP + SMB + HML + RMW + CMA	2.98	10.69	0.28
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	4.49	13.13	0.34
I_7 = CRSP + SMB + HML + RMW + MOM	4.22	12.72	0.33
I_8 = CRSP + SMB + HML + RMW + CMA + MOM + 30 Industry Factors	6.19	15.41	0.40

Referring to the market index results, the index portfolio denoted I_2 = CRSP + SMB in Table 1 has mean excess return, volatility, and Sharpe ratio characteristics similar to momentum. As more multifactors are added to the CRSP index, new market indexes have noticeably higher mean excess returns, volatility, and, more importantly, Sharpe ratios compared to the CRSP index. For example, for the CRSP index plus all five multifactors (denoted I_6 = CRSP + SMB + HML + RMW + CMA + MOM), the mean excess return per month jumps to 4.49 percent, standard deviation of returns to 13.13 percent, and Sharpe ratio to 0.34. The latter Sharpe ratio is almost three times that of the CRSP index. Hence, even though this market index has considerably higher total risk compared to the CRSP index, its excess return per unit risk is much higher than this commonly-used market index. We computed a variety of multifactor market indexes with different combinations of multifactors and find that dropping the CMA multifactor from the market index (i.e., I_7 = CRSP + SMB + HML + RMW + MOM) has little effect on the descriptive statistics of the market index (e.g., Tables 1 and 2 show that the Sharpe ratio only decreases from 0.34 to 0.33).

Table 2. Descriptive statistics for scaled portfolio indexes combining the CRSP index with multifactors: July 1963–December 2016. This table repeats Table 1 by scaling mean monthly portfolio index returns to contain no leverage. As discussed in the text, to orthogonalize a zero-investment factor (e.g., F_1), we regress $F_{1,t}$'s returns on the CRSP index $R(I_{1,t})$ over t sample period months as follows: $F_{1,t} = \alpha_1 + \beta_1 R(I_{1,t}) + \epsilon_{1,t}$. The orthogonalized multifactor is $F_{1,t}^{new} = F_{1,t} - \beta_1 R(I_{1,t})$. Using notation from Equation (11), the more efficient portfolio I_2's return is computed as: $R(I_{2,t}) = R(I_{1,t}) + x_1 F_{1,t}^{new}$. We apply Equation (6) to determine the weight x_1 for the factor return $F_{1,t}^{new}$ in the new market index I_2. Finally, we divide the mean return $R(I_2)$ over the sample period by the term $1 + x_1(1 - \beta_1)$. Each portfolio index is deleveraged by means of this process.

Portfolios	Mean	Std. Dev.	Sharpe Ratio
I_1 = CRSP	0.51	4.42	0.12
SMB	0.27	3.04	0.09
HML	0.37	2.81	0.13
RMW	0.24	2.23	0.11
CMA	0.31	2.01	0.15
MOM	0.66	4.22	0.16
I_2 = CRSP + SMB	0.40	3.07	0.13
I_3 = CRSP + SMB + HML	0.38	1.82	0.21
I_4 = CRSP + SMB + HML + RMW	0.33	1.27	0.26
I_5 = CRSP + SMB + HML + RMW + CMA	0.33	1.18	0.28
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.38	1.11	0.34
I_7 = CRSP + SMB + HML + RMW + MOM	0.39	1.18	0.33
I_8 = CRSP + SMB + HML + RMW + CMA + MOM + 30 Industry Factors	0.40	0.98	0.40

In constructing aggregate market indexes combining the CRSP index with the SMB, HML, RMW, CMA, and MOM multifactors, the weights for factors denoted x_k in Equation (6) for multifactor market indexes are as follows:[7]

I_2 = 0.533 CRSP + 0.467 SMB;

I_3 = 0.313 CRSP + 0.223 SMB + 0.464 HML;

I_4 = 0.231 CRSP + 0.175 SMB + 0.248 HML + 0.346 RMW;

I_5 = 0.232 CRSP + 0.175 SMB + 0.073 HML + 0.203 RMW + 0.317 CMA;

I_6 = 0.197 CRSP + 0.120 SMB + 0.123 HML + 0.159 RMW + 0.203 CMA + 0.198 MOM;

I_7 = 0.202 CRSP + 0.130 SMB + 0.202 HML + 0.236 RMW + 0.230 MOM.

In most of the indexes, SMB gets a relatively lower weight and therefore contributes less to increasing the Sharpe ratio of indexes than other factors. No individual factor appears to dominate the other factors in terms of relative weight. In general, all of the factors are important in forming aggregate indexes.

Our multifactor market indexes have noticeably higher total risk than the CRSP index due to increasing leverage in these portfolios. By deleveraging these indexes, a more accurate assessment of how multifactors affect both mean excess returns and their standard deviation can be obtained. In this respect, Sharpe ratios adjust mean excess returns for total risk but do not reveal these component effects. To adjust for leverage, we rescale the market index portfolios to a zero leverage level. As an example, consider portfolio index

I_2 combining the CRSP index with the size factor denoted F_1. We initially orthogonalize size factor returns F_1. To simplify the derivation and avoid abusing notations, we substitute $R(I_{1,t})$ for $R(I_1, t) - R_{f,t} = R_{m,t} - R_{f,t}$ in Equation (10). Now the orthogonalized multifactor is $F_{1,t}^{new} = F_{1,t} - \beta_1 R(I_{1,t})$. Portfolio index I_2's risk premium is computed as: $R(I_{2,t}) = R(I_{1,t}) + x_1 F_{1,t}^{new} = R(I_{1,t}) + x_1[F_{1,t} - \beta_1 R(I_{1,t})]$. To deleverage the index I_2's return, we divide the mean return $R(I_2)$ over the sample period by the term $1 + x_1(1 - \beta_1)$.[8] This process is repeated for portfolio indexes' returns $R(I_3)$ to $R(I_7)$. In this way, each portfolio index is deleveraged.

Portfolio index returns with no leverage are shown in Table 2. As in Table 1, the highest Sharpe ratios are attained by indexes I_6, I_7, and I_8 at 0.34, 0.33, and 0.40, respectively, which are approximately three times the CRSP index at 0.12. Notice that most of this gain in efficiency is due to decreasing the standard derviation of returns. Indexes I_2 to I_8 have lower mean returns than the CRSP index but substantially lower standard deviations of returns. Portfolios I_5 to I_8 have less than one-third of the standard deviation of CRSP index returns. Thus, adding multifactors and industry factors to the CRSP index provides sizable diversification benefits in the form of lower total risk. As multifactors and industry factors are added to the market index, new multifactor market indexes gradually become more efficient with higher Sharpe ratios due to diversification gains.

3.2. Time-Series Tests of Multifactor Market Indexes

Stambaugh (1982) created alternative market indexes combining common stocks with bonds, real estate, and consumer durables and found very high correlations between their time-series returns. Not surprisingly, empirical tests of the CAPM were not sensitive to the composition of these market indexes. Consistent with these findings, Black (1995) commented that " ... all candidates for the U.S. market portfolio are highly correlated... ", including U.S. domestic and world market indexes, equal- and value-weighted portfolios, and human capital and real estate portfolios of traded assets.[9] For this reason, he believed that the problem of selecting an appropriate market index was not severe (even though it tends to flatten the line between expected return and beta). Related work by Jagannathan and Wang (1996) augmented the value-weighted market index with a proxy for human capital to more comprehensively measure the return on aggregate wealth. Unlike Stambaugh, rather than combining returns to stock market capital and human capital to form a single market index, they treated them as two different market factors in a CAPM market model framework. Subsequent cross-sectional tests indicated that human capital was significantly priced but not the value-weighted market index.[10]

In Table 3 we report the correlation coefficients between the time-series monthly returns for different multifactors and market indexes. The multifactors themselves tend to have relatively low correlation coefficients, with the exception of HML and CMA at 0.69. As a multifactor is added to a market index, their correlation naturally increases (e.g., the CRSP index and SMB have a correlation of 0.28 compared to the correlation of market index CRSP + SMB and SMB at 0.68). In addition, as multifactors are progressively added to create new market indexes, the correlation of the CRSP index with market indexes decreases. Strikingly, the correlation between the CRSP index and market index I_6 containing the five multifactors is only 0.06. Multifactor market indexes are not only more efficient than the CRSP index but not highly correlated with this index (and other commonly-used market indexes) due to including the net effects from other risk factors with relatively low correlations with the CRSP index.

Gibbons et al. (1989) developed a time-series regression test of the CAPM. The GRS statistic tests estimated whether the α_is in the CAPM market model (1) for $i = 1, \ldots, A$ test assets jointly equal zero. Alternatively, given a set of test assets, a riskfree rate, and a market index, the results can be interpreted as a test of whether the market index is a mean-variance efficient portfolio (see Fama 2017). Using monthly returns for Fama and French's 25 size and book-to-market (value) sorted test asset portfolios downloaded from Kenneth French's website for the sample period July 1963 to December 2016, we estimated market model (1) with different market indexes. The GRS test statistics in Table 4 indicate that none of the market indexes is a mean-variance efficient portfolio (i.e., all F-values are statistically significant at the 1 percent level).[11] However, as multifactors are added to create new market indexes, GRS statistics gradually decrease from 4.77 for the CRSP index to 2.77 for market index I_7 = CRSP + SMB + HML + RMW + MOM. We infer that, for these test assets, market indexes become gradually more efficient as multifactors are added. In this regard, more multifactors are needed to achieve an insignificant GRS statistic. Given the large number of factors proposed by researchers as mentioned in the introduction, this possibility is plausible.

Table 3. Correlation coefficients between the monthly returns of multifactors and market indexes: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides correlation coefficients between different market indexes as well as multifactor (zero-investment portfolio) returns. Monthly returns for the value-weighted CRSP index and popular multifactors are downloaded from Kenneth French's website. The multifactors are: size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). The market indexes are: I_1 = CRSP index, I_2 = CRSP + SMB, I_3 = CRSP + SMB + HML, I_4 = CRSP + SMB + HML + RMW, I_5 = CRSP + SMB + HML + RMW + CMA, I_6 = CRSP + SMB + HML + RMW + CMA + MOM, I_7 = CRSP + SMB + HML + RMW + MOM, and I_8 = CRSP + SMB + HML + RMW + CMA + MOM + 30 Industry Factors. The text and Table 1 discuss the process for forming these market indexes that comprise different combinations of CRSP index excess returns, multifactors, and industry excess returns.

	CRSP	**SMB**	**HML**	**RMW**	**CMA**	**MOM**	I_2	I_3	I_4	I_5	I_6	I_7	I_8
I_1	1.00	0.28	−0.26	−0.23	−0.38	−0.13	0.90	0.51	0.33	0.15	0.06	0.21	0.27
SMB		1.00	−0.08	−0.35	−0.10	−0.02	0.68	0.48	0.23	0.16	0.13	0.19	0.18
HML			1.00	0.07	0.69	−0.19	−0.24	0.63	0.63	0.78	0.59	0.45	0.12
RMW				1.00	−0.04	0.11	−0.34	−0.21	0.41	0.34	0.37	0.43	0.26
CMA					1.00	−0.01	−0.34	0.29	0.25	0.55	0.49	0.21	0.38
MOM						1.00	−0.11	−0.24	−0.16	−0.14	0.46	0.47	0.50
I_2							1.00	0.61	0.36	0.19	0.11	0.25	0.29
I_3								1.00	0.80	0.79	0.57	0.56	0.33
I_4									1.00	0.94	0.75	0.79	0.47
I_5										1.00	0.81	0.75	0.53
I_6											1.00	0.96	0.77
I_7												1.00	0.73
I_8													1.00

Table 4. GRS tests using 25 size-value sorted test asset portfolios. This table reports GRS statistic results proposed by Gibbons et al. (1989), which tests whether all estimated intercepts (αs) for the test assets jointly equal zero. Our test assets are the monthly returns for the 25 size-value portfolios downloaded from Kenneth French's data website for the sample period July 1963 to December 2016. The GRS test is conducted for the CAPM market model estimated with the following market indexes: I_1 = CRSP index, I_2 = CRSP + SMB, I_3 = CRSP + SMB + HML, I_4 = CRSP + SMB + HML + RMW, I_5 = CRSP + SMB + HML + RMW + CMA, I_6 = CRSP + SMB + HML + RMW + CMA + MOM, I_7 = CRSP + SMB + HML + RMW + MOM, and I_8 = CRSP + SMB + HML + RMW + CMA + MOM + 30 Industry Factors. The text and Table 1 discuss the process for forming these market indexes that represent different combinations of CRSP index excess returns, multifactors, and industry excess returns. The GRS statistic follows an $F(25, 642 - 25 - 1)$ distribution. Greater GRS test values indicate larger absolute values of estimated αs.

	I_1 = **CRSP**	I_2	I_3	I_4	I_5	I_6	I_7	I_8
F-value	4.78	4.70	3.85	3.38	3.55	2.86	2.77	2.81

3.3. Cross-Sectional Fama-MacBeth Tests of Multifactor Market Indexes

An important test of the efficiency of market indexes augmented with multifactors is whether they are significantly priced in the cross section of average stock returns. For this purpose, using monthly excess returns for 25 size-value sorted test asset portfolios downloaded from French's website, we conduct Fama and MacBeth (1973) tests of market beta associated with different market indexes incorporating popular multifactors. Tests incorporating industry factors are provided in the forthcoming Section 3.6.

We begin by estimating time-series regressions for the full sample period using the excess returns for each of the 25 portfolios and 1 of the market indexes. Monthly index returns are scaled to contain no leverage (see Table 2). As such, the estimated index premiums can be compared to the historical market premium (i.e., 0.51 percent per month in Table 1). Rolling monthly cross-sectional regressions are estimated with excess returns in month t and the estimated full sample betas for all sample period months $t = 1, \ldots, T$. From these regressions we estimate the market price of beta risk, or $\hat{\lambda}_{Mt}$, which are averaged over $t = 1, \ldots, T$ sample months to obtain $\hat{\lambda}_M$. This period-by-period regression approach has the advantage that the t-statistic associated with $\hat{\lambda}_M$ takes into account the covariance of regression residuals and the independent variables without requiring estimates of the covariances (see Fama 2017). The analyses are repeated for each of the portfolio market indexes.

Results for the cross-sectional tests are reported in Table 5. Consistent with earlier studies, the CRSP index (I_1) is not significantly priced, as the market price of risk $\hat{\lambda}_M = -0.43$ ($t = -1.06$). Additionally, index I_2 = CRSP + SMB produces a market proxy that is not priced with $\hat{\lambda}_M$ = 0.00 ($t = 0.00$). However, upon adding both SMB and HML to the CRSP index to form index I_3, we obtain $\hat{\lambda}_M$ = 0.34 ($t = 3.80$), which is economically meaningful and statistically significant at less than a one percent level. Further adding RMA, CMA, and MOM multifactors to the CRSP index to form index I_6 yields similar results. Notice that the strongest finding is for multifactor market index I_7 = CSRP + SMB + HML + RMW + MOM (excluding CMA), which yields $\hat{\lambda}_M$ = 0.38 ($t = 4.04$). In addition, the estimated mispricing term $\hat{\alpha}$ is insignificantly different from zero for the following market indexes: I_3 = CRSP + SMB + HML, I_4 = CRSP + SMB + HML + RMW, and I_7 = CRSP + SMB + HML + RMW + MOM. Finally, adjusted R^2 values jump from 9 percent for the CRSP index to over 50 percent for the combined CRSP plus multifactor portfolios, with the exception of the CRSP + SMB portfolio.[12] For market index I_7 = CRSP + SMB + HML + RMW + MOM, the adjusted R^2 value is 0.71, which indicates a fairly strong goodness-of-fit.

Importantly, the t-values associated with a number of multifactor market indexes well exceed recent standards for the significance of an asset pricing factor. As mentioned earlier, Harvey et al. (2016) and Chordia et al. (2020) have recommended that acceptable factors should exceed a t-statistic threshold of 3.0 or more. We infer that market indexes combining the CRSP index and multifactor portfolios provide market portfolio proxies that are priced in the cross-section of average stock returns and therefore are relatively more efficient portfolios than the CRSP index.

3.4. Cross-Sectional Fama-MacBeth Tests of Multifactors

Here we compare the cross-sectional test results for the CAPM, three-factor, and five-factor models, in addition to the five-factor model augmented with the momentum factor. The results in Table 6 show that the five-factor model plus momentum has insignificant mispricing (i.e., $\hat{\alpha}$ = 0) and a higher estimated adjusted R^2 value (i.e., 83 percent) than the other models. Momentum has a t-value of 4.71, which exceeds the t-values of other factors. Across different models, the SMB, HML, and RMW factors are consistently significant and exceed 3 in some instances, CMA is not priced, and the multifactor models substantially boost the estimated adjusted R^2 values relative tthe CAPM. Note that our multifactor market index I_7 (see Table 5) outperforms the three- and five-factor models, with similar R^2 value but insignificant $\hat{\alpha}$ mispricing, and performs almost as well as the five-factor model plus momentum.

Table 5. Cross-sectional asset pricing tests of market beta for different market indexes using 25 size-value sorted test asset portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of the value-weighted CRSP index as well as market indexes combining the CRSP index with popular multifactors. Downloaded from Kenneth French's website, the multifactors are: size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). Multifactor market indexes are formed using the following steps. First, the size (SMB) factor monthly returns are regressed on CRSP index (I_1) excess returns over the Treasury bill rate. The residual term from this regression is utilized as the orthogonalized factor. Second, this orthogonalized size factor is added to the CRSP index using Equation (11) to compute the return for the new multifactor market index I_2 as $R(I_2) = R(I_1) + x_1$SMB. Third, value (HML) factor returns are regressed on the new $R(I_1) + x_1$SMB portfolio excess returns to obtain the orthogonalized value factor. This residual value factor is added to the CRSP + SMB portfolio to get market index I_3 = CRSP + x_1SMB + x_2HML. Fourth, the last step is repeated to sequentially create market index I_4 = CRSP + x_1SMB + x_2HML + x_3RMW, market index I_5 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA, and market index I_6 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA + x_5MOM. In addition, we drop the CMA multifactor to form market index I_7 = CRSP + x_1SMB + x_2HML + x_3RMW + x_5MOM. Monthly returns for 25 size-value test assets are downloaded from French's data website (i.e., the value firm characteristic corresponds to the book-to-market ratio). A time-series regression using monthly excess returns is run for the full sample period to estimate CAPM betas for each of the 25 tests assets using the CRSP index. Time-series regression analyses are repeated for the other market indexes. Following the procedure in the text and Table 2, monthly market index returns are scaled to contain no leverage. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with excess returns and full sample betas for all sample months. The resultant monthly series of estimated market prices of beta risk are averaged over all sample months to estimate $\hat{\lambda}_M$ (and associated t-statistics are in parentheses).

Indexes	$\hat{\alpha}$	$\hat{\lambda}_M$	**Adj. R^2**
I_1 = CRSP	1.20	−0.43	0.09
	(3.19)	(−1.06)	
I_2 = CRSP + SMB	0.73	0.00	0.00
	(2.970	(0.00)	
I_3 = CRSP + SMB + HML	0.08	0.34	0.54
	(0.34)	(3.80)	
I_4 = CRSP + SMB + HML + RMW	0.22	0.28	0.61
	(0.86)	(3.70)	
I_5 = CRSP + SMB + HML + RMW + CMA	0.48	0.21	0.56
	(2.06)	(3.27)	
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.54	0.26	0.61
	(2.44)	(3.43)	
I_7 = CRSP + SMB + HML + RMW + MOM	0.24	0.38	0.71
	(0.95)	(4.04)	

We next consider whether a multifactor is priced when the market index contains the multifactor portfolio. First, we create residual returns for the 25 size-value portfolios (denoted $R_{p,t}^{res}$) as follows:

$$R_{p,t}^{res} = R_{p,t} - R_{f,t} - \beta_p[R(I_{K,t}) - R_{f,t}], \tag{13}$$

where $R_{p,t} - R_{f,t}$ is the excess portfolio return in month t, and $R(I_{K,t}) - R_{f,t}$ is the excess market index return as defined in Equation (11), and β_p is the estimated beta coefficient for portfolio $p = 1, \ldots, 25$. We employ market indexes I_1 to I_7 defined earlier in this section. Second, using monthly residual returns and multifactor returns for the full sample period, we run time-series regressions of residual returns on the multifactor returns for SMB, HML, RMW, CMA, and MOM for each of the 25 portfolios. Third, and last, using monthly average portfolios' residual returns, monthly rolling cross-sectional regressions are run as before to estimate the market prices of risk for the multifactors.

Table 6. Cross-sectional asset pricing tests of the CRSP market factor and five popular multifactors: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of the market factor plus five popular multifactors. Downloaded from Kenneth French's website, the factors are: CRSP market index (M), size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). Test assets are the Fama–French 25 size-value portfolios downloaded from French's data website. For the full sample period, a time-series regression is run using monthly excess returns for test assets as the dependent variable and the CRSP value-weighted market index excess return plus different multifactor returns as the independent variables. Using estimated factor loadings for each of the 25 test asset portfolios, cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with monthly excess returns and full sample betas for all sample months. Resultant monthly series of estimated prices of beta risk for the kth factor ($k = 1, \ldots, 6$) are averaged over all sample months to estimate $\hat{\lambda}_k$ (and associated t-statistics are in parentheses).

Index	$\hat{\alpha}$	$\hat{\lambda}_M$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{RMW}$	$\hat{\lambda}_{CMA}$	$\hat{\lambda}_{MOM}$	Adj. R^2
CAPM	1.20	−0.43						0.09
	(3.19)	(−1.06)						
Three-factor	1.27	−0.73	0.22	0.40				0.67
	(4.83)	(−2.31)	(1.78)	(3.50)				
Five-factor	1.02	−0.53	0.30	0.36	0.48	−0.02		0.74
	(3.56)	(−1.58)	(2.47)	(3.15)	(2.82)	(−0.10)		
Five-factor + MOM	0.28	0.26	0.33	0.39	0.61	−0.14	2.94	0.83
	(0.82)	(0.69)	(2.68)	(3.47)	(3.44)	(−0.80)	(4.71)	

The empirical results in Table 7 indicate that, with the exception of momentum (MOM), multifactors are not normally priced in the cross-section when the market index contains the respective multifactors. For example, for market indexes I_1 and I_2 comprised of the CRSP index and CRSP + SML, respectively, HML is significantly priced with $\hat{\lambda}_{HML} = 0.50$ ($t = 4.91$) and 0.61 ($t = 4.93$). However, for market index I_3 (viz., CRSP + SMB + HML) containing the HML portfolio, the multifactor HML is not priced with $\hat{\lambda}_{HML} = -0.25$ ($t = -1.13$). RMW is priced using market indexes I_1, I_2, I_3, and I_6 but not for I_4 and I_5 containing the RMW portfolio.[13] Unlike the other multifactors, MOM consistently remains significantly priced across all market indexes, even when the MOM portfolio is included in market indexes I_6 and I_7. For this reason, we will refer to MOM as a possible strong factor in forthcoming analyses of two-factor models comprised of multifactor market indexes and the momentum factor. From these results we infer that, with the exception of MOM, multifactors are not generally priced when they are incorporated in the market index. Lastly, as multifactors are added to the market index, the explanatory power of the multifactors diminishes (e.g., from 90 percent for the CRSP market index I_1 to only 42 percent for the multifactor market index I_7 containing SMB, HML, RMW, and MOM).[14]

Table 7. Cross-sectional asset pricing tests of multifactors using the excess residual returns of 25 size-value sorted test assets with respect to market index excess returns: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of five popular multifactors. Downloaded from Kenneth French's website, the multifactors are: size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). Test assets are monthly excess returns (denoted $R_{i,t} - R_{f,t}$) for Fama–French 25 size-value portfolios downloaded from French's data website. The residuals for test assets are computed as $R^{res}_{i,t} = R_{i,t} - R_{f,t} - \hat{\beta}_i R(I_{K,t})$, $i = 1, \ldots, 25$ where $R(I_{K,t})$ is the monthly excess return for one of seven different market indexes ($K = 1, \ldots, 7$). The text and Table 1 discuss the process for forming market indexes that represent different combinations of the CRSP index and multifactors. We denote these market indexes as follows: I_1 = CRSP index, I_2 = CRSP + SMB, I_3 = CRSP + SMB + HML, I_4 = CRSP + SMB + HML + RMW, I_5 = CRSP + SMB + HML + RMW + CMA, and I_6 = CRSP + SMB + HML + RMW + CMA + MOM, in addition to I_7 = CRSP + SMB + HML + RMW + MOM. For the full sample period, a time-series regression is run using monthly residual returns as the dependent variable and multifactor returns as the independent variables to estimate multifactor betas for each of the 25 test asset portfolios. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with monthly residual returns and full sample betas for all sample months. The resultant monthly series of estimated multifactor prices of beta risk for the kth factor ($k = 1, \ldots, 5$) are averaged over all sample months to estimate $\hat{\lambda}_k$ (and associated t-statistics are in parentheses).

Index	$\hat{\alpha}$	$\hat{\lambda}_{SMB}$	$\hat{\lambda}_{HML}$	$\hat{\lambda}_{RMW}$	$\hat{\lambda}_{CMA}$	$\hat{\lambda}_{MOM}$	Adj. R^2
I_1	0.03	0.11	0.59	0.76	0.08	3.42	0.90
	(1.21)	(0.83)	(4.91)	(4.30)	(0.38)	(6.01)	
I_2	0.09	−0.33	0.61	0.87	0.09	3.34	0.91
	(5.39)	(−1.59)	(4.93)	(4.88)	(0.40)	(5.98)	
I_3	0.15	−0.20	−0.25	0.79	−0.35	3.50	0.72
	(2.81)	(−1.02)	(−1.13)	(4.37)	(−2.17)	(6.11)	
I_4	0.63	0.22	0.17	0.38	−0.12	2.44	0.66
	(3.60)	(1.31)	(0.57)	(1.70)	(−0.86)	(4.18)	
I_5	0.24	0.14	−0.47	0.32	−0.58	3.23	0.80
	(1.39)	(0.97)	(−1.45)	(1.38)	(−2.74)	(5.48)	
I_6	0.83	0.35	0.57	0.62	−0.01	1.51	0.70
	(3.63)	(2.40)	(1.69)	(2.63)	(−0.03)	(2.63)	
I_7	0.90	0.37	0.53	0.58	0.03	1.25	0.42
	(3.94)	(2.27)	(2.10)	(2.37)	(0.17)	(2.34)	

3.5. Robustness Tests with Different Test Asset Portfolios

As a robustness check, we repeat the cross-sectional tests in Table 8 using 25 size-value sorted portfolios for a variety of different test asset portfolios available on French's data website. The following test assets are used: 25 value-investment portfolios, 25 profit-investment portfolios, 25 size-investment portfolios, 25 size-profit portfolios, 32 size-value-investment portfolios, 32 size-value-profit portfolios, 32 size-profit-investment portfolios, and 30 industry portfolios. In general, the results for these portfolios corroborate our findings in Table 5. The t-values associated with market indexes I_4 to I_7 have ranges as follows: 2.32 to 2.52 for 25 value-investment portfolios in Panel A; 3.34 to 3.61 for 25 profit-investment portfolios in Panel B; 4.16 to 4.54 for 25 size-investment portfolios in panel C; 3.85 to 3.89 for 25 size-profit portfolios in Panel D; 3.33 to 3.61 for 32 size-value-investment portfolios in Panel E; 4.49 to 4.91 for 32 size-value-profit portfolios in Panel F; 5.85 to 6.53 for 32 size-profit-investment portfolios in Panel G; and –0.01 to 0.59 for 30 industry portfolios in Panel H. Highlighting the results for the size-profit-investment portfolios in Panel G, the results are as follows: $\hat{\lambda}_M$ = 0.45 ($t = 5.85$) for I_4, $\hat{\lambda}_M$ = 0.37 ($t = 6.20$) for I_5, $\hat{\lambda}_M$ = 0.42 ($t = 6.53$) for I_6, and $\hat{\lambda}_M$ = 0.51 ($t = 6.36$) for I_7. Mispricing terms α are insignificant for these multifactor market indexes with the exception of I_6. Adjusted R^2 values range from 0.67 to 0.76. Together, these results corroborate our earlier findings in Table 5 using size-value test asset portfolios. In addition, the results are comparable to the the three-factor, five-factor, and five-factor plus momentum models in Table 6.

Table 8. Robustness checks for cross-sectional asset pricing tests of market beta for different market indexes using a variety of test asset portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of the value-weighted CRSP index as well as market indexes combining the CRSP index with popular multifactors. Downloaded from Kenneth French's website, the multifactors are: size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). Multifactor market indexes are formed using the following steps. First, the size (SMB) factor monthly returns are regressed on CRSP index (I_1) excess returns over the Treasury bill rate. The residual term from this regression is utilized as the orthogonalized factor. Second, this orthogonalized size factor is added to the CRSP index using Equation (11) to compute the return for the new multifactor market index I_2 as $R(I_2) = R(I_1) + x_1$SMB. Third, value (HML) factor returns are regressed on the new $R(I_1) + x_1$SMB portfolio excess returns to obtain the orthogonalized value factor. This residual value factor is added to the CRSP + SMB portfolio to get market index I_3 = CRSP + x_1SMB + x_2HML. Fourth, the last step is repeated to sequentially create market index I_4 = CRSP + x_1SMB + x_2HML + x_3RMW, market index I_5 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA, and market index I_6 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA + x_5MOM. In addition, we drop the CMA multifactor to form market index I_7 = CRSP + x_1SMB + x_2HML + x_3RMW + x_5MOM. Following the procedure in the text and Table 2, monthly market index returns are scaled to contain no leverage. Monthly returns for a variety of test assets are downloaded from French's data website: 25 value-investment portfolios, 25 profit-investment portfolios, 25 size-investment portfolios, 25 size-profit portfolios, 32 size-value-investment portfolios, 32 size-value-profit portfolios, 32 size-profit-investment portfolios, and 30 industry portfolios (see French's website for details of these portfolios). A time-series regression is run using monthly excess returns for the CRSP index for the full sample period to estimate CAPM betas for each of the tests assets. Time-series regression analyses are repeated for the other market indexes. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with excess returns and full sample betas for all sample months. The resultant monthly series of estimated market prices of beta risk are averaged over all sample months to estimate $\hat{\lambda}_M$ (and associated *t*-statistics are in parentheses).

Panel A: 25 Value-Investment Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	0.77	2.93	−0.11	−0.34	−0.04
I_2 = CRSP + SMB	0.49	2.41	0.13	0.75	−0.01
I_3 = CRSP + SMB + HML	0.32	1.60	0.21	2.40	0.51
I_4 = CRSP + SMB + HML + RMW	0.33	1.60	0.19	2.32	0.48
I_5 = CRSP + SMB + HML + RMW + CMA	0.45	2.40	0.16	2.47	0.54
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.49	2.60	0.21	2.52	0.53
I_7 = CRSP + SMB + HML + RMW + MOM	0.30	1.38	0.29	2.39	0.50
Panel B: 25 Profit-Investment Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	1.29	4.91	−0.69	−2.23	0.25
I_2 = CRSP + SMB	1.05	5.13	−0.35	−1.95	0.19
I_3 = CRSP + SMB + HML	0.45	2.10	0.09	0.74	−0.02
I_4 = CRSP + SMB + HML + RMW	0.00	0.01	0.43	3.61	0.52
I_5 = CRSP + SMB + HML + RMW + CMA	0.36	1.70	0.26	3.36	0.54
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.46	2.32	0.29	3.34	0.59
I_7 = CRSP + SMB + HML + RMW + MOM	0.17	0.65	0.44	3.40	0.56

Table 8. *Cont.*

Panel C: 25 Size-Investment Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	1.04	3.83	−0.28	−0.84	0.01
I_2 = CRSP + SMB	0.62	2.94	0.07	0.43	−0.03
I_3 = CRSP + SMB + HML	0.20	0.94	0.30	2.32	0.39
I_4 = CRSP + SMB + HML + RMW	0.07	0.31	0.41	4.16	0.60
I_5 = CRSP + SMB + HML + RMW + CMA	0.42	1.90	0.30	4.52	0.62
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.53	2.43	0.34	4.54	0.64
I_7 = CRSP + SMB + HML + RMW + MOM	0.16	0.70	0.48	4.35	0.63
Panel D: 25 Size-Profit Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	0.47	1.41	0.21	0.55	−0.02
I_2 = CRSP + SMB	0.31	1.37	0.24	1.40	0.15
I_3 = CRSP + SMB + HML	0.18	0.83	0.29	2.12	0.39
I_4 = CRSP + SMB + HML + RMW	0.19	0.91	0.30	3.86	0.85
I_5 = CRSP + SMB + HML + RMW + CMA	0.39	1.87	0.30	3.85	0.75
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.52	2.54	0.31	3.85	0.75
I_7 = CRSP + SMB + HML + RMW + MOM	0.32	1.57	0.32	3.89	0.84
Panel E: 32 Size-Value-Investment Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	0.51	1.53	0.22	0.56	−0.01
I_2 = CRSP + SMB	0.33	1.49	0.27	1.63	0.19
I_3 = CRSP + SMB + HML	0.22	1.11	0.29	3.06	0.58
I_4 = CRSP + SMB + HML + RMW	0.29	1.39	0.25	3.33	0.46
I_5 = CRSP + SMB + HML + RMW + CMA	0.48	2.40	0.20	3.40	0.44
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.52	2.64	0.27	3.61	0.48
I_7 = CRSP + SMB + HML + RMW + MOM	0.26	1.27	0.36	3.56	0.54
Panel F: 32 Size-Value-Profit Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	1.17	3.54	−0.42	−1.15	0.02
I_2 = CRSP + SMB	0.60	2.71	0.08	0.48	−0.02
I_3 = CRSP + SMB + HML	−0.01	−0.05	0.40	4.23	0.56
I_4 = CRSP + SMB + HML + RMW	0.13	0.58	0.32	4.79	0.73
I_5 = CRSP + SMB + HML + RMW + CMA	0.40	1.87	0.26	4.49	0.66
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.49	2.38	0.31	4.54	0.67
I_7 = CRSP + SMB + HML + RMW + MOM	0.21	0.92	0.40	4.91	0.75
Panel G: 32 Size-Profit-Investment Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	1.04	3.62	−0.33	−0.94	0.00
I_2 = CRSP + SMB	0.57	2.78	0.08	0.50	−0.02
I_3 = CRSP + SMB + HML	0.15	0.75	0.32	2.59	0.26
I_4 = CRSP + SMB + HML + RMW	−0.03	−0.12	0.45	5.85	0.67
I_5 = CRSP + SMB + HML + RMW + CMA	0.30	1.45	0.37	6.20	0.69
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.43	2.16	0.42	6.53	0.75
I_7 = CRSP + SMB + HML + RMW + MOM	0.10	0.50	0.51	6.36	0.76
Panel H: 30 Industry Portfolios					
Indexes	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	**Adj. R^2**
I_1 = CRSP	0.67	2.99	−0.06	−0.20	−0.03
I_2 = CRSP + SMB	0.66	3.64	−0.03	−0.18	−0.03
I_3 = CRSP + SMB + HML	0.71	4.24	−0.06	−0.52	0.01
I_4 = CRSP + SMB + HML + RMW	0.60	3.42	0.01	0.11	−0.03
I_5 = CRSP + SMB + HML + RMW + CMA	0.61	3.31	0.00	−0.01	−0.04
I_6 = CRSP + SMB + HML + RMW + CMA + MOM	0.60	3.04	0.03	0.35	−0.02
I_7 = CRSP + SMB + HML + RMW + MOM	0.55	2.88	0.05	0.59	−0.01

Not surprisingly, in Panel H's results for industry portfolios, market indexes I_4 to I_7 are not significantly priced. It is well-known that common market indexes and popular multi-factors are not priced using exogenous industry portfolios. Hence, a shortfall in the asset priciing literature is the absence of an aggregate industry factor.

3.6. Index Construction with Industry Factors

Our findings above indicate that efficient market indexes can be constructed from the CRSP index in combination with popular multifactors. These market indexes do a good job of pricing widely-used test assets with characteristics similar to the multifactors, but are not useful in pricing industry portfolios. Unfortunately, there is no single industry factor available.

Industry returns are different from other common risk factors based on firm characteristics (e.g., size, book-to-market, profit, and capital investment) or stock characteristics (e.g., winner and loser stock returns). A problem in forming an aggregate industry index is that different industries tend to be independent of one another.[15] A shared common risk factor across industries (other than the market factor) is not directly observed. We propose a possible solution to this problem. Using the 30 industry test assets, we initially form a pure long/short portfolio that is long a particular industry and short an equal quantity of Treasury bills. We then combine these 30 industry factors to construct a single industry index based on our multifactor market index methods using relative Sharpe ratio weights in Section 2. In addition, we combine the CRSP index, multifactors, and industry factors to construct a new multifactor market index.

Regarding our earlier cross-sectional tests, we found that the results for a multifactor market index were independent of the order with which multifactors were added one-by-one to the obtained more efficient index. This robustness holds as long as the multifactors themselves are priced factors. Each time we add a new factor to the base market index, we only add in the net part which is not included in the base market index obtained in previous step. It is important to note that, if the new factor has large noise, it will not contribute to increasing the efficiency of the market index. Its noise component will be treated as the net part to add to the base market index and, subsequently, will remain embedded in the market index through the iterative steps of adding different multifactors to the index. To mitigate this potential noise, we add industry factors with relatively smaller volatility prior to higher volatility industry factors. Following this ordered procedure, we compute a single industry index denoted IND.

Cross-sectional asset pricing tests of industry index IND are provided in Table 9. As shown there, when using the 30 industry portfolios as the test assets, this industry index is positively priced, i.e., $\hat{\lambda}_{IND} = 0.82$ ($t = 2.50$).[16] For other test assets based on firm characteristics, however, IND is normally significantly priced but negatively so and not priced for profit-investment and size-profit portfolios.

For comparison purposes, we also construct a simple industry index denoted $SIND$, which is defined as an equal-weighted portfolio of the 30 industry factors. The cross-sectional test results in Table 10 show that $SIND$ is not priced in the 30 industry portfolios and most other test assets, with the exceptions of being negatively priced for value-investment and profit-investment portfolios.

As a last step, we add the five popular multifactors plus 30 industry portfolios to the CRSP index using relative Sharpe ratio weights as defined in Section 2. Factors are added in the order discussed above, i.e., smaller volatility factors before other factors. The obtained multifactor market index denoted I_8 yields the cross-sectional test results in Table 11. All test assets, including industry portfolios (i.e., significant at the 5 percent level), are priced by this aggregate market index. Most of the t-values for different test assets exceed the recommended 3.0 threshold. For size-profit-investment portfolios the t-value reaches a high of 5.9, which is extraordinary in view of the very low and insignificant t-values associated with the market factor in almost all published asset pricing studies. Except for the value-investment portfolios, R^2 values range from 41 percent to 76 percent, which implies relatively high goodness-of-fit for a single factor model. Interestingly, all of the estimated mispricing terms $\hat{\alpha}$ using I_8 as the market factor are insignificant. These mispricing results bode well for the single factor CAPM. We infer that multifactor market index I_8 provides a superior market index for researchers to employ in asset pricing studies.

Table 9. Cross-sectional asset pricing tests of a single industry index incorporating 30 industry factors using relative Sharpe ratio weights for a variety test asset portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of a single industry index incorporating 30 industry factors denoted IND. Each industry factor is defined as a pure long/short portfolio that is long a particular industry and short an equal quantity of Treasury bills. The single industry index IND is formed by combining the 30 industry factors based on the relative Sharpe ratio weighted methods defined in Section 2. Monthly returns for a variety of test assets are downloaded from French's data website: 30 industry portfolios, 25 size-value portfolios, 25 value-investment portfolios, 25 profit-investment portfolios, 25 size-investment portfolios, 25 size-profit portfolios, 32 size-value-investment portfolios, 32 size-value-profit portfolios, and 32 size-profit-investment portfolios (see French's website for details of these portfolios). A time-series regression using monthly excess returns is run for the full sample period to estimate industry betas for each of the tests assets with respect to industry index IND. Following the procedure in the text and Table 2, monthly industry index returns are scaled to contain no leverage. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with excess returns and full sample industry betas for all sample months. The resultant monthly series of estimated market prices of industry beta risk are averaged over all sample months to estimate $\hat{\lambda}_{IND}$ (and associated t-statistics are in parentheses).

Test Assets	$\hat{\alpha}$	$\hat{\lambda}_{IND}$	Adj. R^2
30 industry portfolios	0.06	0.82	0.65
	(0.19)	(2.50)	
25 size-value portfolios	2.08	−2.23	0.29
	(5.26)	(−3.71)	
25 value-investment portfolios	1.63	−1.54	0.32
	(3.56)	(−2.28)	
25 profit-investment portfolios	0.47	0.19	−0.04
	(1.24)	(0.32)	
25 size-investment portfolios	2.47	−2.85	0.57
	(5.46)	(−4.27)	
25 size-profit portfolios	−0.15	1.39	0.12
	(0.416)	(2.74)	
32 size-value-investment portfolios	2.45	−2.79	0.49
	(5.11)	(−4.11)	
32 size-value-profit portfolios	1.92	−1.94	0.07
	(3.98)	(−2.50)	
32 size-profit-investment portfolios	1.53	−1.37	0.04
	(4.43)	(−2.69)	

3.7. Is Momentum a Strong Factor?

Because momentum was significantly priced after removing the excess returns of different multifactor market indexes from the 25 size-value test assets (see Table 7), we designated it as a possible strong factor. As mentioned earlier, strong factors supplement multifactor market indexes to improve model specification.

In Table 12 we report the cross-sectional tests for two-factor models comprised of multifactor market indexes I_5, I_6, I_7, and I_8 augmented with the momentum factor (MOM). Note that I_6, I_7, and I_8 contain the momentum multifactor. Referring to Table 12, we find that momentum loadings are significantly priced in the following tests: Panel A for the 25-size-value test asset portfolios in combination with I_5, I_6, and I_7 but not I_8; Panel D for 25 size-investment portfolios in combination with I_8; Panel F for 32 size-value-investment portfolios in combination with I_5, I_6, and I_7 but not I_8; and Panel H for 32 size-profit-investment portfolios in combination with all four multifactor market indexes.

Table 10. Cross-sectional asset pricing tests of a simple industry index incorporating equal-weighted industry factors for a variety test asset portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of a simple industry index denoted as $SIND$. Each industry factor is defined as a pure long/short portfolio that is long a particular industry and short an equal quantity of Treasury bills. The single industry index $SIND$ is formed by combining the 30 industry factors into an equal-weighted portfolio. Monthly returns for a variety of test assets are downloaded from French's data website: 25 value-investment portfolios, 25 profit-investment portfolios, 25 size-investment portfolios, 25 size-profit portfolios, 32 size-value-investment portfolios, 32 size-value-profit portfolios, 32 size-profit-investment portfolios, and 30 industry portfolios (see French's website for details of these portfolios). A time-series regression using monthly excess returns is run for the full sample period to estimate industry betas for each of the tests assets with respect to market index $SIND$. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with excess returns and full sample industry betas for all sample months. The resultant monthly series of estimated market prices of industry beta risk are averaged over all sample months to estimate $\hat{\lambda}_{SIND}$ (and associated t-statistics are in parentheses).

Test Assets	$\hat{\alpha}$	$\hat{\lambda}_{SIND}$	Adj. R^2
30 industry portfolios	0.64	−0.03	−0.03
	(2.98)	(−0.09)	
25 size-value portfolios	0.89	−0.16	−0.03
	(2.27)	(−0.34)	
25 value-investment portfolios	0.51	0.16	−0.03
	(3.56)	(−2.28)	
25 profit-investment portfolios	1.21	−0.67	0.16
	(4.45)	(−1.89)	
25 size-investment portfolios	0.81	−0.08	−0.04
	(2.78)	(−0.20)	
25 size-profit portfolios	0.06	0.62	0.14
	(0.17)	(1.52)	
32 size-value-investment portfolios	0.23	0.51	0.08
	(0.66)	(1.22)	
32 size-value-profit portfolios	0.63	0.08	−0.03
	(1.83)	(0.20)	
32 size-profit-investment portfolios	0.67	0.02	−0.03
	(2.23)	(0.05)	

We interpret these results to suggest that momentum is a possible strong factor for all four multifactor market indexes tested, as it continues to be significantly priced in some test assets even if included in respective multifactor market indexes. Nonetheless, as the efficiency of multifactor market indexes increases via the addition of more feasible factors (e.g., based on cross-sectional t-values), the hurdle for strong factors such as momentum will increase. It is conceivable that future research using more efficient multifactor market indexes will eliminate momentum as a possible strong factor.

3.8. Discussion

Multifactor market indexes have both academic and practical applications. In academic studies, as discussed in Section 2, many different asset pricing models are popular in the literature nowadays. Certainly more models with innovative factors will be proposed in coming years. Which factors should be used by researchers? Our multifactor market index approach enables researchers to reduce this problem to a manageable set of aggregate indexes and strong factors. In this way, not only can all discovered factors be incorporated into multifactor aggregate indexes, but parsimonious models can be specified for broad usage in academic research. For practitioners, multifactor market indexes represent investable strategies to construct efficient portfolios for investment purposes. Hence, portfolio managers can utilize significant factors in academic studies to boost their returns per unit risk for clients. In turn, Markowitz's (1959) mean-variance portfolio theory can be applied to create well diversified, efficient portfolios.

Table 11. Cross-sectional asset pricing tests of a multifactor market index incorporating five popular multifactors and 30 industry factors for a variety test asset portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of a multifactor market index denoted I_8 incorporating five popular multifactors and 30 industry factors. Each industry factor is defined as a pure long/short portfolio that is long a particular industry and short an equal quantity of Treasury bills. Index I_8 is formed by adding five popular multifactors (viz., size, value, profit, investment, and momentum) plus 30 industry factors to the CRSP index based on the relative Sharpe ratio weighted methods defined in Section 2. Monthly returns for a variety of test assets are downloaded from French's data website: 30 industry portfolios, 25 value-investment portfolios, 25 profit-investment portfolios, 25 size-investment portfolios, 25 size-profit portfolios, 32 size-value-investment portfolios, 32 size-value-profit portfolios, and 32 size-profit-investment portfolios (see French's website for details of these portfolios). A time-series regression using monthly excess returns is run for the full sample period to estimate CAPM betas for each of the tests assets with respect to market index I_8. Following the procedure in the text and Table 2, monthly market index returns are scaled to contain no leverage. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with excess returns and full sample betas for all sample months. The resultant monthly series of estimated market prices of beta risk are averaged over all sample months to estimate $\hat{\lambda}_M$ (and associated *t*-statistics are in parentheses).

Test Assets	$\hat{\alpha}$	$\hat{\lambda}_M$	Adj. R^2
30 industry portfolio	0.33	0.20	0.41
	(1.27)	(2.01)	
25 size-value portfolios	−0.20	0.59	0.63
	(−0.82)	(4.43)	
25 value-investment portfolios	0.36	0.20	0.12
	(1.49)	(2.11)	
25 profit-investment portfolios	0.06	0.39	0.76
	(0.25)	(4.15)	
25 size-investment portfolios	−0.02	0.47	0.65
	(−0.08)	(4.02)	
25 size-profit portfolios	0.12	0.38	0.79
	(0.56)	(3.51)	
32 size-value-investment portfolios	−0.00	0.46	0.62
	(−0.02)	(3.82)	
32 size-value-profit portfolios	−0.17	0.59	0.68
	(−0.71)	(5.03)	
32 size-profit-investment portfolios	−0.07	0.49	0.71
	(−0.34)	(5.91)	

Table 12. Cross-sectional asset pricing tests of a two-factor model containing different multifactor market indexes augmented with the momentum factor for a variety of test asset portfolios: July 1963–December 2016. Based on CRSP stock return data in the sample period July 1963 to December 2016, this table provides Fama and MacBeth (1973) cross-sectional tests of two-factor models comprised of different multifactor market indexes augmented with the momentum factor. Downloaded from Kenneth French's website, the multifactors are: size (SMB), value (HML), profit (RMW), capital investment (CMA), and momentum (MOM). Multifactor market indexes are formed using the following steps. First, the size (SMB) factor monthly returns are regressed on CRSP index (I_1) excess returns over the Treasury bill rate. The residual term from this regression is utilized as the orthogonalized factor. Second, this orthogonalized size factor is added to the CRSP index using Equation (11) to compute the return for the new multifactor market index I_2 as $R(I_2) = R(I_1) + x_1$SMB. Third, value (HML) factor returns are regressed on the new $R(I_1) + x_1$SMB portfolio excess returns to obtain the orthogonalized value factor. This residual value factor is added to the CRSP + SMB portfolio to get market index I_3 = CRSP + x_1SMB + x_2HML. Fourth, the last step is repeated to sequentially create market index I_4 = CRSP + x_1SMB + x_2HML + x_3RMW, market index I_5 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA, and market index I_6 = CRSP + x_1SMB + x_2HML + x_3RMW + x_4CMA + x_5MOM. We drop the CMA multifactor to form market index I_7 = CRSP + x_1SMB + x_2HML + x_3RMW + x_5MOM. In addition, we form market index I_8 = CRSP + x_1SMB + x_2HML + x_3RMW + x_5MOM + (x_6 to x_{35}) × (1 to 30) industry factors, which are defined as the industry index return minus the Treasury bill rate. Following the procedure in the text and Table 2, monthly market index returns are scaled to contain no leverage. Monthly returns for a variety of test assets are downloaded from French's data website: 25 value-investment portfolios, 25 profit-investment portfolios, 25 size-investment portfolios, 25 size-profit portfolios, 32 size-value-investment portfolios, 32 size-value-profit portfolios, 32 size-profit-investment portfolios, and 30 industry portfolios (see French's website for details of these portfolios). A time-series regression is run for the full sample period to estimate multifactor market and momentum factors' betas for each of the tests assets. Time-series regression analyses are repeated for the I_5, I_6, I_7, and I_8 multifactor market indexes. Cross-sectional tests are conducted by estimating monthly rolling cross-sectional regressions with excess returns and full sample betas for all sample months. The resultant monthly series of estimated market prices of beta risk are averaged over all sample months to estimate $\hat{\lambda}_M$ (and associated t-statistics are in parentheses).

Panel A: 25 Size-Value Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.74	3.43	0.25	3.82	1.87	3.42	0.65
I_6, MOM	0.74	3.43	0.50	4.50	1.87	3.42	0.65
I_7, MOM	0.37	1.43	0.73	5.43	2.47	3.93	0.79
I_8, MOM	−0.27	−1.02	0.48	2.65	0.55	0.66	0.64
Panel B: 25 Value-Investment Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.46	2.65	0.17	2.36	0.06	0.10	0.52
I_6, MOM	0.46	2.65	0.15	1.20	0.06	0.10	0.52
I_7, MOM	0.30	1.36	0.27	1.48	0.43	0.64	0.48
I_8, MOM	0.18	0.73	0.08	0.73	−0.57	−1.10	0.44
Panel C: 25 Profit-Investment Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.47	2.46	0.25	3.27	0.53	1.02	0.57
I_6, MOM	0.47	2.46	0.29	2.74	0.53	1.02	0.57
I_7, MOM	0.09	0.36	0.42	2.99	0.45	0.86	0.55
I_8, MOM	0.00	0.01	0.37	3.47	0.59	1.11	0.76

Table 12. *Cont.*

Panel D: 25 Size-Investment Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.51	2.94	0.29	4.49	0.49	0.74	0.62
I_6, MOM	0.51	2.94	0.32	2.76	0.49	0.74	0.62
I_7, MOM	0.16	0.76	0.48	3.82	0.82	1.43	0.62
I_8, MOM	0.03	0.14	0.57	4.02	1.66	2.65	0.65
Panel E: 25 Size-Profit Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.44	2.50	0.29	3.65	0.11	0.19	0.75
I_6, MOM	0.44	2.50	0.26	2.67	0.11	0.19	0.75
I_7, MOM	0.25	1.35	0.27	2.70	0.16	0.27	0.84
I_8, MOM	−0.03	−0.14	0.30	3.15	0.10	0.17	0.81
Panel F: 32 Size-Value-Investment Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.60	3.25	0.26	4.18	1.27	2.56	0.49
I_6, MOM	0.60	3.25	0.41	4.05	1.27	2.56	0.49
I_7, MOM	0.30	1.45	0.68	4.08	2.15	3.29	0.60
I_8, MOM	−0.14	−0.57	0.33	2.82	0.02	0.04	0.70
Panel G: 32 Size-Value-Profit Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.46	2.26	0.26	4.56	0.31	0.53	0.66
I_6, MOM	0.46	2.27	0.27	2.56	0.31	0.53	0.66
I_7, MOM	0.20	0.85	0.38	3.42	0.57	1.04	0.74
I_8, MOM	−0.34	−1.26	0.35	2.91	−0.26	−0.44	0.77
Panel H: 32 Size-Profit-Investment Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.54	3.02	0.34	5.54	1.36	2.85	0.76
I_6, MOM	0.54	3.02	0.50	6.20	1.36	2.85	0.76
I_7, MOM	0.21	1.09	0.59	6.96	1.51	3.30	0.76
I_8, MOM	−0.04	−0.18	0.51	7.01	1.24	2.81	0.71
Panel I: 30 Industry Portfolios							
Factors	$\hat{\alpha}$	***t*-Value**	$\hat{\lambda}_M$	***t*-Value**	$\hat{\lambda}_{MOM}$	***t*-Value**	**Adj. R^2**
I_5, MOM	0.65	3.79	0.02	0.28	0.32	0.65	−0.01
I_6, MOM	0.65	3.79	0.06	0.61	0.32	0.65	−0.01
I_7, MOM	0.60	3.36	0.11	0.96	0.43	0.87	0.03
I_8, MOM	0.30	1.34	0.20	1.98	0.37	0.84	0.39

4. Conclusions

This paper sought to mitigate the model mall problem associated with the growing list of factors in asset pricing. To do this, we proposed the construction of efficient multifactor market indexes that combine the CRSP index with popular investable multifactors. To demonstrate this approach, size, value, profit, capital investment, and momentum multifactors were sequentially added to the CRSP index. In the sample period July 1963 to December 2016, as multifactors were added to the CRSP index, multimarket market indexes became increasingly less correlated with the CRSP index and more efficient than the CRSP index. Importantly, market betas associated with most of these multifactor market indexes were significantly priced in cross-sectional asset pricing tests with economically meaningful market prices of risk. With the exception of industry portfolio test assets, the *t*-values associated with multifactor market index betas generally exceeded the recommended 3.0 threshold for the significance of an asset pricing factor. In addition, we found that multifactors were less likely to be significantly priced in cross-sectional tests when the market index contained the respective multifactors. One exception was the momentum factor, which was significantly priced even when contained in multifactor market indexes.

We inferred that momentum may be a strong factor in the sense that it can serve to augment multifactor market indexes in models comprised of two or more factors.

Further analyses demonstrated how a single industry index can be constructed by combining numerous industry indexes. This new aggregate industry index was significantly priced among industry portfolios, for which the CRSP index and multifactors are typically not priced. When popular multifactors and industry factors were combined with the CRSP index into a multifactor market index, the resultant market factor was significantly priced across a wide variety of test assets including industry portfolios. Exemplary of our findings, a very high t-value of 5.91 was achieved for size-value-investment portfolios. In tests of this market index across different test assets, the goodness-of-fit was relatively high and comparable to traditional multifactor models, and no mispricing was detected.

When multifactor market indexes were augmented with the momentum factor in a two-factor model, momentum loadings were significant in a number of different test assets, but not all test assets. We conclude that momentum is a possible strong factor that is priced at times even when contained in the two-factor model's multifactor market index. Future research is needed to determine if, as multifactor market indexes become more efficient with the inclusion of more feasible factors, momentum remains a priced factor.

Based on our empirical results, we conclude that: (1) multifactor market indexes are more efficient than the CRSP market index, and (2) market beta is significantly priced for these multivariate market index proxies. Hence, we infer that multifactors are not separate risk factors but rather help jointly in combination with a general stock market index to better proxy efficient market portfolios. An important implication of our findings is that asset pricing models incorporating a market factor (including unconditional and conditional empirical models) would benefit from employing multifactor market indexes rather than the CRSP index or other general market indexes. More parsimonious low-dimensional models can be developed that incorporate multifactor market indexes to substantially reduce factor and model selection problems. That is, rather than creating an increasing number of asset pricing models featuring different multifactors, most of the multifactors can be productively utilized to form more efficient multifactor market indexes. Most investable factors will likely be absorbed into a multifactor market index. Noninvestable factors, such as market volatility, macroeconomic state variables, etc., cannot be combined into an investable multifactor market index and, therefore, could be prospects as strong factors in a low-dimensional model. More generally, by combining many proven factors into a single aggregate market index, it is possible that better asset pricing models can be constructed that capture the collective significance of numerous risk dimensions associated with a wide array of factors.

Future research is recommended on efficient multifactor market indexes to reduce the problems of factor and model selection in various asset pricing applications, such as event studies of corporate and other major events, investment performance studies of mutual funds and hedge funds, and other studies that rely upon asset pricing models in their empirical analyses. Importantly, a practical implication is that investors can readily form multifactor market indexes by combining (for example) a tradeable S&P 500 index with ETFs related to size, value, momentum, and other multifactors. These new aggregate indexes have the potential to outperform the S&P 500 index, a common benchmark used to evaluate institutional investors. Investors seeking portfolios with higher returns per unit risk would benefit from efficient ETFs and other multifactor index products, especially those saving for retirement.

Author Contributions: W.L. worked on the empirical results; J.W.K. wrote the text. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data is available from Kenneth French's website.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1 See also Lettau and Pelger (2020), who utilized Principal Component Analysis (PCA) methods to reduce the number of potential factors to a parsimonious set of latent factors. They found that five factors with economic content help to explain the cross-section and time-series of returns.

2 For example, the three moment CAPM (Rubinstein (1973) and Kraus and Litzenberger (1976)), consumption CAPM (Breeden 1979), conditional CAPM (Jagannathan and Wang (1996) and Ferson and Harvey (1999)), liquidity-based models (Pastor and Stambaugh (2003), Acharya and Pedersen (2005), and Li et al. (2019)), intertemporal CAPM (ICAPM) (Petkova 2006), interest-rate-based models Campbell (1996), cross-factor models Fama and French (2020), among others.

3 As such, the intercept equals zero, and the beta factor loading matrix is unchanged.

4 Numerous papers have found that various long-only market indexes are not mean-variance efficient portfolios (e.g., see Gibbons et al. (1989); Gibbons (1982); Jobson and Korkie (1982); Kandel (1984); Shanken (1985, 1986); Kandel and Stambaugh (1987a, 1987b); Gibbons et al. (1989); Haugen and Baker (1991); MacKinlay and Richardson (1991); Zhou (1993); Brière et al. (2013), and others). According to Brennan and Lo (2010), it is virtually impossible for long-only market indexes to be efficient portfolios. Supporting this proposition, many studies have found that short positions are needed to achieve efficiency (e.g., see Pulley (1981); Levy (1983); Kallberg and Ziemba (1983); Kroll et al. (1984); Green and Hollifield (1992); Jagannathan and Ma (2003); Brennan and Lo (2010); Levy and Ritov (2010), and others). More generally, Kothari et al. (1995) have argued that the equity portfolio most highly correlated with the market portfolio is efficient.

5 Also, if the expected return of the zero-cost factor portfolio is zero, the Sharpe ratio will not be increased by adding it to a candidate market index. Note that any zero-cost asset that is uncorrelated with the tangency portfolio has an expected return equal to the riskless rate such that its excess return is zero.

6 As shown there, the definitions of the multifactors are as follows: size (SMB) is the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios; value (HML) is the average return on the two value portfolios minus the average return on the two growth portfolios; profit (RMW) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios; capital investment (CMA) is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios; and momentum (MOM) is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios.

7 These weights are averages using different orders of entry in forming indexes based on a rotation of factors. For example, I_3 can be formed by starting with CRSP, adding SMB, and then adding HML. Alternatively, we can start with SMB, add HML, and then CRSP. Finally, we could add them in the order HML, SMB, and CRSP. We average weights for each factor across these rotational combinations of order entry. While the order in which portfolios is combined changes their relative weights, the Sharpe ratios and other performance metrics of respective aggregate indexes as well as forthcoming cross-sectional tests were little changed.

8 Rearranging terms, we have: $R(I_{2,t}) = R(I_{1,t}) + x_1 F_{1,t}^{new} = R(I_{1,t}) + x_1[F_{1,t} - \beta_1 R(I_{1,t})]$. Since both $R(I_{1,t})$ and $F_{1,t}$ are deleveraged portfolios, we can deleverage $R(I_{2,t})$ by dividing its return by $1 + x_1(1 - \beta_1)$.

9 See also Roll (1977, p. 130), who also observed that most proxies for the market portfolio are very highly correlated.

10 They also included a time-varying market factor to capture beta instability over time (i.e., the BAA minus AAA bond yield spread), which was found to be significantly priced. Hence, they concluded that, even though the static CAPM assuming constant beta over time is not supported, the conditional CAPM allowing betas and expected returns to vary over time is supported.

11 The GRS test statistic has a noncentral F distribution with degrees of freedom N (25 portfolios) and $T - N - 1$ (T 654 months). As the noncentrality parameter increases, the probability of rejecting a false null hypothesis tends to increase. According to tests in Gibbons et al. (1989, pp. 1130–38), the power of our tests should be sufficient to detect deviations from the efficiency of the index.

12 Following standard practice, adjusted R^2 values are estimated by regressing the average excess returns for test asset portfolios in the full sample period on their full sample beta estimates.

13 By contrast, CMA is positive but not significantly priced for market indexes I_1 and I_2 based on the CRSP index and CRSP + SML market index, respectively, but it is significantly priced with market indexes I_3 and I_5 containing the CMA portfolio. These unexpected results for CMA are difficult to interpret due to being negatively (rather than positively) priced when significant.

14 Multifactor market index I_7 does not eliminate the significance of the SMB, HML, RMW, and MOM factors even though these factors are contained in this index. However, the markedly lower R^2 value of 42 percent for this market index indicates that these factors' residual explanatory power is substantially diminished by market index I_7.

15 Early work on industry as an asset pricing factor by King (1966) and Meyers (1973) found that, based on principal components analyses of U.S. stock returns, most components could not be identified with specific industry groups. Components were only weakly associated with industry classifications.

16 Unreported in Table 9, we also tested a two-factor model with the excess return on the CRSP index (I_1) and industry factor (IND). Using 30 industry portfolios as test assets, both factors are significantly priced with t-values of 2.12 and 2.98, respectively. The correlation between these two factors was relatively high at 0.60.

References

Acharya, Viral V., and Lasse H. Pedersen. 2005. Asset pricing with liquidity risk. *Journal of Financial Economics* 77: 375–410. [CrossRef]

Barillas, Francisco, and Jay Shanken. 2018. Comparing asset pricing models. *Journal of Finance* 73: 715–54. [CrossRef]

Black, Fischer. 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45: 444–54. [CrossRef]

Black, Fischer. 1995. Estimating expected return. *Financial Analysts Journal* 49: 168–71. [CrossRef]

Breeden, Douglas T. 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics* 7: 265–96. [CrossRef]

Brennan, Douglas T., and Andrew W. Lo. 2010. Impossible frontiers. *Management Science* 56: 905–23. [CrossRef]

Brière, Marie, Bastien Drut, Valerie Mignon, Kim Oosterlink, and Ariane Szafarz. 2013. Is the market portfolio efficient? A new test to of mean-variance efficiency when all assets are risky. *Finance* 34: 7–41. [CrossRef]

Campbell, John Y. 1996. Understanding risk and return. *Journal of Political Economy* 104: 298–345. [CrossRef]

Carhart, Mark M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52: 57–82. [CrossRef]

Chen, Long, and Lu Zhang. 2010. A better three-factor model that explains more anomalies. *Journal of Finance* 65: 563–94.

Chordia, Tarun, Amit Goyal, and Alessio Saretto. 2020. Anomalies and false rejections. *Review of Financial Studies* 33: 2134–79. [CrossRef]

Cochrane, John H. 2011. Presidential address: Discount rates. *Journal of Finance* 56: 1047–108. [CrossRef]

Fama, Eugene F. 1996. Multifactor portfolio efficiency and multifactor asset pricing. *Journal of Financial and Quantitative Analysis* 31: 441–65. [CrossRef]

Fama, Eugene F. 2017. *Cross-Section versus Time-Series Tests of Asset Pricing Models*. Working Paper. Chicago: University of Chicago.

Fama, Eugene F., and James D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81: 607–36. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47: 427–65. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 1993. The cross-section of expected returns. *Journal of Financial Economics* 33: 3–56. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 1995. Size and book-to-market factors in earnings and returns. *Journal of Finance* 50: 131–56. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 1996a. The CAPM is wanted, dead or alive. *Journal of Finance* 51: 1947–58. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 1996b. Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51: 55–84. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116: 1–22. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2018. Choosing factors. *Journal of Financial Economics* 128: 234–52. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2020. Comparing cross-section and time-series factor models. *Review of Financial Studies* 33: 1892–926. [CrossRef]

Ferson, Wayne E. 1995. Theory and empirical testing of asset pricing models. In *Handbooks in Operations Research and Management Science*. Edited by Robert A. Jarrow, Vojislav Maksimovic and William T. Ziemba. Amsterdam: Elsevier, pp. 145–200.

Ferson, Wayne E. 2019. *Empirical Asset Pricing Models and Methods*. Cambridge: MIT Press.

Ferson, Wayne, and Campbell R. Harvey. 1999. Conditioning variables and the cross-section of stock returns. *Journal of Finance* 54: 1325–60. [CrossRef]

Ferson, Wayne E., and Andrew F. Siegel. 2009. Testing portfolio efficiency with conditioning informaiton. *Review of Financial Studies* 22: 2735–58. [CrossRef]

Gibbons, Michael R. 1982. Multivariate tests of financial models: A new approach. *Journal of Financial Economics* 10: 3–27. [CrossRef]

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken. 1989. A test of the efficiency of a given portfolio. *Econometrica* 57: 1121–52. [CrossRef]

Green, Richard C., and Burton Hollifield. 1992. When will mean-variance efficient portfolios be well diversified? *Journal of Finance* 47: 1785–809. [CrossRef]

Grinblatt, Mark, and Sheridan Titman. 1987. The relation between mean-variance efficiency and arbitrage pricing. *Journal of Business* 60: 97–112. [CrossRef]

Harvey, Campbell R., Yang Liu, and Heqing Zhu. 2016. . . . and the cross-section of expected returns. *Review of Financial Studies* 29: 5–68. [CrossRef]

Haugen, Robert A., and Nardin L. Baker. 1991. The efficient market inefficiency of capitalization-weighted stock portfolios. *Journal of Portfolio Managment* 17: 35–40. [CrossRef]

Hou, Kewei, Chen Xue, and Lu Zhang. 2014. Digesting anomalies: An investment approach. *Review of Financial Studies* 28: 650–705. [CrossRef]

Hou, Kewei, Haitao Mo, Chen Xue, and Lu Zhang. 2018. q^5. Working Paper. Columbus: Ohio State University.

Huberman, Gur, and Shmuel Kandel. 1987. Mean-variance spanning. *Journal of Finance* 42: 873–88. [CrossRef]

Huberman, Gur, Shmuel Kandel, and Roger Stambaugh. 1987. Mimicking portfolios and exact arbitrage pricing. *Journal of Finance* 42: 1–9. [CrossRef]

Jagannathan, Ravi, and Tongshu Ma. 2003. Risk reduction in large portfolios: A role for portfolio weight constraints. *Journal of Finance* 58: 1651–84. [CrossRef]

Jagannathan, Ravi, and Zhenyu Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance* 51: 3–53. [CrossRef]
Jobson, John D., and Bob Korkie. 1982. Potential performance and tests of portfolio efficiency. *Journal of Financial Economics* 10: 433–65. [CrossRef]
Kallberg, Jerry G., and William T. Ziemba. 1983. Comparison of alternative utility functions in portfolio selection problems. *Management Science* 29: 1257–76. [CrossRef]
Kan, Raymond, and Guofu Zhou. 2008. Tests of mean-variance spanning. *Annals of Economics and Finance* 13: 145–93.
Kandel, Shmuel. 1984. On the exclusion of assets from tests of the mean variance efficiency of the market portfolio. *Journal of Finance* 39: 63–75. [CrossRef]
Kandel, Shmuel, and Roger F. Stambaugh. 1987a. A mean variance framework for tests of asset pricing models. *Review of Financial Studies* 2: 125–56. [CrossRef]
Kandel, Shmuel, and Roger F. Stambaugh. 1987b. On correlations and inferences about mean/variance efficiency. *Journal of Financial Economics* 18: 61–90. [CrossRef]
King, Benjamin F. 1966. Market and industry factors in stock price behavior. *Journal of Business* 39: 139–90. [CrossRef]
Kothari, Sagar P., Jay Shanken, and Richard G. Sloan. 1995. Another look at the cross-section of expected stock returns. *Journal of Finance* 50: 185–224. [CrossRef]
Kraus, Alan, and Robert H. Litzenberger. 1976. Skewness preference and the valuation of risk assets. *Journal of Finance* 31: 1085–100.
Kroll, Yoram, Haim Levy, and Harry Markowitz. 1984. Mean-variance versus direct utility maximization. *Journal of Finance* 39: 47–61. [CrossRef]
Lettau, Martin, and Marcus Pelger. 2020. Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies* 33: 2274–325. [CrossRef]
Levy, Haim. 1983. The capital asset pricing model: Theory and empiricism. *Economic Journal* 93: 145–65. [CrossRef]
Levy, Moshe, and Yaacov Ritov. 2010. Mean-variance efficient portfolios with many assets: 50% short. *Quantitative Finance* 10: 1461–71. [CrossRef]
Li, Hongtao, Robert Novy-Marx, and Mikhail Velikov. 2019. Liquidity risk and asset pricing. *Critical Finance Review* 8: 223–55. [CrossRef]
Lintner, John. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* 47: 13–37. [CrossRef]
MacKinlay, A. Craig. 1993. *Multifactor Models Do Not Explain Deviations from the CAPM*. Working Paper. Philadelphia: University of Pennsylvania.
MacKinlay, A. Craig, and Matthew P. Richardson. 1991. Using generalized method of moments to test mean/variance efficiency. *Journal of Finance* 46: 511–27. [CrossRef]
Markowitz, Harry M. 1959. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.
Merton, Robert C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41: 867–87. [CrossRef]
Meyers, Stewart L. 1973. A re-examination of market and industry factors in stock price behavior. *Journal of Finance* 28: 695–705. [CrossRef]
Mossin, Jan. 1966. Equilibrium in a capital asset market. *Econometrica* 34: 768–83. [CrossRef]
Pastor, Lubos, and Robert F. Stambaugh. 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111: 642–85. [CrossRef]
Petkova, Ralitsa. 2006. Do the Fama-French factors proxy for innovations in predictive variables? *Journal of Finance* 61: 581–612. [CrossRef]
Pulley, Larry B. 1981. General mean-variance approximation to expected utility for short holding periods. *Journal of Financial and Quantitative Analysis* 16: 361–73. [CrossRef]
Roll, Richard. 1977. A critique of the asset pricing theory's tests, part I: On past and potential future testability of the theory. *Journal of Financial Economics* 4: 129–76. [CrossRef]
Ross, Stephen A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13: 341–60. [CrossRef]
Rubinstein, Mark E. 1973. The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis* 8: 61–69. [CrossRef]
Shanken, Jay. 1985. Multivariate tests of the zero-beta CAPM. *Journal of Financial Economics* 14: 327–48. [CrossRef]
Shanken, Jay. 1986. On the exclusion of assets from tests of the mean variance efficieny of the market. *Journal of Finance* 41: 331–37. [CrossRef]
Shanken, Jay. 1987. Multivariate proxies and asset pricing relations: Living with the Roll critique. *Journal of Financial Economics* 18: 91–110. [CrossRef]
Shanken, Jay, and Mark I. Weinstein. 2006. Economic forces and the stock market revisited. *Journal of Empirical Finance* 13: 129–44. [CrossRef]
Sharpe, William F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* 19: 425–42.
Stambaugh, Roger F. 1982. On the exclusion of assets from tests of the two-parameter model. *Journal of Financial Economics* 10: 237–68. [CrossRef]

Stambaugh, Roger F., and Yu Yuan. 2017. Mispricing factors. *Review of Financial Studies* 30: 1270–315. [CrossRef]
Treynor, Jack L. 1961. *Market Value, Time, and Risk*. Rochester: SSRN, August 8.
Treynor, Jack L. 1962. Toward a theory of market value of risky assets. *Unpublished manuscript*.
Zhou, Guofu. 1993. Asset pricing tests under alternative distributions. *Journal of Finance* 48: 1927–42. [CrossRef]

Journal of
Risk and Financial Management

Article

A Singular Stochastic Control Approach for Optimal Pairs Trading with Proportional Transaction Costs

Haipeng Xing

Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794, USA; haipeng.xing@stonybrook.edu; Tel.: +1-631-632-1892

Abstract: Optimal trading strategies for pairs trading have been studied by models that try to find either optimal shares of stocks by assuming no transaction costs or optimal timing of trading fixed numbers of shares of stocks with transaction costs. To find optimal strategies that determine optimally both trade times and number of shares in a pairs trading process, we use a singular stochastic control approach to study an optimal pairs trading problem with proportional transaction costs. Assuming a cointegrated relationship for a pair of stock log-prices, we consider a portfolio optimization problem that involves dynamic trading strategies with proportional transaction costs. We show that the value function of the control problem is the unique viscosity solution of a nonlinear quasi-variational inequality, which is equivalent to a free boundary problem for the singular stochastic control value function. We then develop a discrete time dynamic programming algorithm to compute the transaction regions, and show the convergence of the discretization scheme. We illustrate our approach with numerical examples and discuss the impact of different parameters on transaction regions. We study the out-of-sample performance in an empirical study that consists of six pairs of U.S. stocks selected from different industry sectors, and demonstrate the efficiency of the optimal strategy.

Keywords: free-boundary problem; pairs trading; stochastic control; trading strategies; transaction costs; transaction regions

Citation: Xing, Haipeng. 2022. A Singular Stochastic Control Approach for Optimal Pairs Trading with Proportional Transaction Costs. *Journal of Risk and Financial Management* 15: 147. https://doi.org/10.3390/jrfm15040147

Academic Editor: David Allen

Received: 14 February 2022
Accepted: 12 March 2022
Published: 23 March 2022

1. Introduction

Pairs trading is one of proprietary statistical arbitrage tools used by many hedge funds and investment banks. It is a short-term trading strategy that first identifies two stocks whose prices are associated in a long-run equilibrium and then trades on temporary deviations of stock prices from the equilibrium. Though pairs trading is a simple market neutral strategy, it has been used and discussed extensively by industrial practitioners in the last several decades; see detailed discussion in Vidyamurthy (2004), Whistler (2004), Ehrman (2006), Lai and Xing (2008), and references therein.

Besides its wide practice in financial industry, pairs trading also draws much attention from academic researchers. For instance, Gatev et al. (2006) examined the risk and returns of pairs trading using daily data collected from the U.S. equity market and concluded that the strategy in general produces profit higher than transaction costs. To investigate the pairs trading strategy analytically, Elliott et al. (2005) modeled the spread of returns as a mean-reverting process and proposed a trading strategy based on the model. This motivates subsequent researchers to formulate pairs trading rules as stochastic control problems for an Ornstein–Uhlenbeck (OU) process and a correlated stock price process. In particular, Mudchanatongsuk et al. (2008) assumed the log-relationship between a pair of stock prices follows a mean-reverting process, and considered a self-financing portfolio strategy that only allows positions that were long in one stock and short in the other with equal dollar amounts. They then formulated a portfolio optimization based stochastic control problem and obtained the optimal solution to this control problem in closed form via the corresponding Hamilton–Jacobi–Bellman (HJB) equation. Relaxing the equal dollar constraint, Tourin and Yan (2013) extended Mudchanatongsuk et al. (2008)'s approach and study pairs trading strategies with arbitrary amounts in each stock without any transaction costs.

Instead of deriving optimal weights of stocks in pairs trading, another line of study on pairs trading strategies fixes the number of traded shares for each stock during the entire trading process and considers only the optimal timing of trades in the presence of fixed or proportional transaction costs. Specifically, Leung and Li (2015) studies the optimal timing to open or close the position subject to fixed transaction costs and the effect of stop-loss level under the OU process by constructing the value function directly. Zhang and Zhang (2008), Song and Zhang (2013), and Ngo and Pham (2016) studied the optimal pairs trading rule that is based on optimal switching among two (buy and sell) or three (buy, sell, and flat) regimes with a fixed commission cost for each transaction, and solve the problem by finding viscosity solutions to the associated HJB equations (quasi-variational inequalities). Lei and Xu (2015) studied the optimal pairs trading rule of entering and exiting the asset market in finite horizon with proportional transaction cost for two convergent assets. Note that, although transaction costs are considered in these strategies, since the number of traded shares of stocks are fixed during the entire trading period, these strategies are still far from traders' practical experience in reality.

The above study on optimal pairs trading focuses either on optimal trading shares without transaction costs or optimal trading times with fixed trading shares in the presence of transaction costs. To relax the assumption of fixed trading shares in the latter study, this paper uses a singular stochastic control approach to study the joint effect of optimal trading shares and optimal trading times in pairs trading process with proportional transaction costs. For convenience, we assume the same diffusion and Urnstein–Uhlenbeck processes for one stock and its spread with the other stock as those in Mudchanatongsuk et al. (2008). However, different from Mudchanatongsuk et al. (2008) who used a trading rule which requires to short one stock and long the other in equal dollar amounts, we consider a delta-neutral rule under which the ratio of traded shares for two stocks is fixed and this fixed ratio is determined by the cointegration relationship of two stocks. Hence, when the number of shares of one stock is determined, based on the rule of delta neutral, the number of shares for the other stock is also determined. Besides the weight of shares need to be optimally chosen, we also assume a proportional transaction cost for each trade and hence the optimal times of trading also needs to be decided.

With the above assumptions, we solve the optimal pair trading problem by the singular stochastic control approach in Davis et al. (1993). As the overall transaction cost based on the above assumption depends on both trading times and the numbers of shares in each trade, we compute the terminal utility of wealth over a fixed horizon and formulate the problem of choosing optimal trading times and the number of shares as a singular stochastic control problem. We derive the Hamilton–Jacobi–Bellman equations for this problem, and show that the value function of the problem is the unique viscosity solution of a quasi-variational inequality. We further argue that the quasi-variational inequality is equivalent to a free boundary problem so that the state space consisting of one stock price and its spread with the other stock can be naturally divided into three transaction regions: long the first stock and short the second, short the first and long the second, and no transaction. The implied transaction regions can help us determine not only optimal times of each transaction, but also the optimal number of shares in each transaction. To compute the boundaries of these transaction regions, we develop a numerical algorithm that is based on discrete time dynamic programming to solve the equation for the negative exponential utility function, and show that the numerical solution converges to the unique continuous-time solution of the problem.

To demonstrate the advantage of joint consideration of optimal shares and optimal trading times in pair trading, we carry out both simulation and empirical studies. Specifically, we study the time-varying transaction regions (or trading boundaries) for a specific set of model parameters, and investigate the impact of variations of model parameters on transaction regions and performance of the optimal strategy. For comparison purposes, we also consider a benchmark strategy based on the deviation of the spread from its long-term mean and is popular among practitioners. In both simulation studies and real data analysis, we show that the optimal trading strategy performs better than the benchmark strategy.

The rest of the paper is organized as follows. Section 2 first formulates the model and then derive the Hamilton–Jacobi–Bellman equations associated with the singular stochastic control problems. It shows the existence and uniqueness of the viscosity solution for the variational inequalities, which are equivalent to the portfolio optimization problem, and reduces the problem into a free boundary problem. Section 2 also considers the optimal trading problem with exponential utility functions. In Section 3, we discretize the free boundary problem and propose a discrete time dynamic programming algorithm. We also demonstrate that the solution of the discretized problem converges to the viscosity solution of the variational inequalities. Sections 4 and 5 provide simulation and empirical studies of the model and the optimal trading strategy, and compare its performance with a benchmark trading strategy. Some concluding remarks are given in Section 6.

2. A Pairs Trading Problem with Proportional Transaction Costs

2.1. Model Specification

Consider a pair of two stocks P and Q, and let $p(t)$ and $q(t)$ denote their prices at time t, respectively. We assume that the price of stock P follows a geometric Brownian motion,

$$dp(t) = \mu p(t)dt + \sigma p(t)dB(t), \tag{1}$$

where μ and σ are the drift and the volatility of stock P, and $B(t)$ is a standard Brownian motion defined on a filtered probability space and specified later. Denote $x(t)$ the difference of the logarithms of the two stock prices, i.e.,

$$x(t) = \log q(t) - \log p(t) = \log(q(t)/p(t)). \tag{2}$$

We assume that the spread follows an Ornstein–Uhlenbeck process

$$dx(t) = \kappa(\theta - x(t))dt + \nu dW(t), \tag{3}$$

where $\kappa > 0$ is the speed of mean reversion, and θ is the long-term equilibrium level to which the spread reverts. We assume that $(B(t), W(t))$ is a two-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, and the instantaneous correlation coefficient between $B(t)$ and $W(t)$ is ρ, i.e.,

$$E[dW(t)dB(t)] = \rho dt. \tag{4}$$

The above assumptions are same as those in Mudchanatongsuk et al. (2008). With these assumptions, we can express the dynamics of $q(t)$ as

$$dq(t) = \left[\mu + \kappa(\theta - x(t)) + \frac{1}{2}\nu^2 + \rho\sigma\nu\right]q(t)dt + \sigma q(t)dB(t) + \nu q(t)dW(t). \tag{5}$$

In the presence of proportional transaction costs, the investor pays $0 < \zeta_p, \zeta_q < 1$ and $0 < \eta_p, \eta_q < 1$ of the dollar value transacted on purchase and sale of the underlying stocks P and Q. Denote $L_p(t)$ and $M_p(t)$ two nondecreasing and non-anticipating processes and represent the cumulative number of shares of stock P bought or sold, respectively, within the time interval $[0, t]$, $0 \leq t \leq T$. Let $y_p(t)$ be the number of shares held in stock P, i.e., $y_p(t) = L_p(t) - M_p(t)$, and similarly, we define $L_q(t)$, $M_q(t)$, and $y_q(t) = L_q(t) - M_q(t)$ for stock Q. Denote $g(t)$ the dollar value of the investment in bond which pays a fixed risk-free rate of r. Then, the investor's position in two stocks and the bond is driven by

$$dy_p(t) = dL_p(t) - dM_p(t), \qquad dy_q(t) = dL_q(t) - dM_q(t) \tag{6}$$

and

$$dg(t) = rg(t)dt + b_p p(t)dM_p(t) - a_q q(t)dL_q(t) + b_q q(t)dM_q(t) - a_p p(t)dL_p(t), \quad (7)$$

where $a_i = 1 + \zeta_i$ and $b_i = 1 - \eta_i$ for $i = p, q$.

We then need to choose a rule to determine the number of shares of stocks P and Q bought or sold at time t. Note that, Mudchanatongsuk et al. (2008) assumed no transaction cost and considered the strategy that always shorts one stock and longs the other in equal dollar amount, i.e., $p(t)dL_p(t) + q(t)dM_q(t) = 0$ or $p(t)dM_p(t) + q(t)dL_q(t) = 0$ at time t. Lei and Xu (2015) and Ngo and Pham (2016) considered a delta-neutral strategy that always long one share of a stock and short one share of the other stock, i.e., $dy_p(t) = -dy_q(t) = 1$ or $dy_p(t) = -dy_q(t) = -1$ at time t. Here, we also consider a delta-neutral strategy that requires the total of positive and negative delta of two assets is zero, hence it suggests that the number of shares of stock P bought (or sold) at time t are same as the number of shares of stock Q sold (or bought), i.e.,

$$dL_p(t) = dM_q(t), \qquad dM_p(t) = dL_q(t). \quad (8)$$

Equation (8) implies that

$$dy_q(t) = -dy_p(t)$$

at any time t. Comparing to Lei and Xu (2015) and Ngo and Pham (2016), we remove the constraint $dy_p(t) = -dy_q(t) = 1$ or -1 and allow $y_p(t) = -y_q(t)$ to be a control variable. Using Equations (5) and (8), the dynamics of $g(t)$ in Equation (7) can be simplified as

$$dg(t) = rg(t)dt - \left(a_p - b_q e^{x(t)}\right)p(t)dL_p(t) + \left(b_p - a_q e^{x(t)}\right)p(t)dM_p(t). \quad (9)$$

The process $(L_p(t), M_p(t))$ together with our delta-neutral strategy provides us an admissible trading strategy. For convenience, we denote $\mathcal{T}(g_0)$ the set of admissible trading strategies that an investor starts at time zero with amount g_0 of the investment in bond and zero holdings in two stocks (i.e., $y_p(0) = y_q(0) = 0$), which indicates that the numbers of shares held in stocks P and Q at time t are $y_p(t)$ and $-y_p(t)$, respectively. For notational convenience, we omit the subscript of $y_p(t)$ and denote $y_p(t)$ as $y(t)$ in our discussion. Then, Equations (1), (3), (6) and (9) compose the market model in the time interval $[0, T]$, which describes a stochastic process of $(p(t), x(t), y_p(t), g(t))$ in $\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

Denote the terminal value of the pairs trading portfolio by $J(x(T), p(T), y(T))$. Note that, under our assumption, $y(T)$ indicates that the investor's positions in stocks P and Q are $y(T)$ and $-y(T)$, respectively, then the liquidated value of the portfolio is

$$J(p(T), x(T), y(T)) = A_+(p(T), x(T))y(T)\mathbb{1}_{\{y(T)\geq 0\}} + A_-(p(T), x(T))y(T)\mathbb{1}_{\{y(T)<0\}}, \quad (10)$$

where

$$A_+(p, x) = (b_p - a_q e^x)p, \qquad A_-(p, x) = (a_p - b_q e^x)p.$$

Furthermore, if the investment in bond at terminal time T is $g(T)$, the terminal wealth of the investor is given by $g(T) + J(p(T), x(T), y(T))$. Suppose that the investor's utility $U : \mathbb{R} \longrightarrow \mathbb{R}$ is a concave and increasing function with $U(0) = 0$. We assume that the investor's goal is to maximize the expected utility of terminal wealth under the market model (1), (3), (6) and (9),

$$\begin{aligned} V(t, p, x, y, g) = \sup_{(L_p(t), M_p(t)) \in \mathcal{T}(g_0)} & E\Big\{U(g(T) + J(p(T), x(T), y(T))|p(t) = p, \\ & x(t) = x, y_t = y, g(t) = g\Big\}. \end{aligned} \quad (11)$$

Furthermore, given trading strategies (L_p, M_p), the total trading cost incurred over $[t, T]$ can be expressed as

$$C(L_p, M_p; t, T) = \int_t^T e^{r(T-u)} A_-(p(u), x(u)) dL_p(u) - \int_t^T e^{r(T-u)} A_+(p(u), x(u)) dM_p(u) \\ - J(p(T), x(T), y(T)). \tag{12}$$

and the total profit over $[t, T]$ is $-C(L_p, M_p; t, T)$.

2.2. The Hamilton-Jacobi-Bellman Equations and Free Boundary Problems

We now derive the Hamilton–Jacobi–Bellman (HJB) equations, associated with the stochastic control problems, for the utility maximization problem (11). Consider a class of trading strategies such that $L_p(t)$ and $M_p(t)$ are absolutely continuous processes, given by

$$L_p(t) = \int_0^t l(u) du, \qquad M_p(t) = \int_0^t m(u) du,$$

where $l(u)$ and $m(u)$ are positive and uniformly bounded by $\xi < \infty$. Then, (1), (3), (6) and (9) provides us a system of stochastic differential equations with controlled drift, and the Bellman equation for a value function denoted by V^ξ is

$$\mathcal{L}_{1,o} V^\xi + \sup_{0 \le l_t, m_t \le \xi} \left\{ \left[\mathcal{L}_{1,b} V^\xi \right] l_t - \left[\mathcal{L}_{1,s} V^\xi \right] m_t \right\} = 0,$$

for $(t, p, X, y, g) \in [0, T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, in which the operators $\mathcal{L}$, $\mathcal{B}$, and $\mathcal{S}$ are defined as

$$\mathcal{L}_{1,o} := \frac{\partial}{\partial t} + \kappa(\theta - x) \frac{\partial}{\partial x} + \mu p \frac{\partial}{\partial p} + r g \frac{\partial}{\partial g} + \frac{1}{2} \nu^2 \frac{\partial^2}{\partial x^2} + \rho \nu \sigma p \frac{\partial^2}{\partial p \partial x} + \frac{1}{2} \sigma^2 p^2 \frac{\partial^2}{\partial p^2},$$

$$\mathcal{L}_{1,b} := \frac{\partial}{\partial y} - \left(a_p - b_q e^{x(t)}\right) p(t) \frac{\partial}{\partial g},$$

$$\mathcal{L}_{1,s} := \frac{\partial}{\partial y} - \left(b_p - a_q e^{x(t)}\right) p(t) \frac{\partial}{\partial g}.$$

The optimal trading strategy is then determined by considering the following three possible cases:

(i) buying stock P and sell stock Q at the same rate $l(t) = \xi$ (i.e., $m(t) = 0$) when

$$\mathcal{L}_{1,b} V^\xi \ge 0, \qquad \mathcal{L}_{1,s} V^\xi > 0; \tag{13}$$

(ii) selling stock P and buy stock Q at rate $m(t) = \xi$ (i.e., $l(t) = 0$) when

$$\mathcal{L}_{1,b} V^\xi < 0, \qquad \mathcal{L}_{1,s} V^\xi \le 0; \tag{14}$$

(iii) doing nothing (i.e., $l(t) = m(t) = 0$) when

$$\mathcal{L}_{1,b} V^\xi \le 0, \qquad \mathcal{L}_{1,s} V^\xi \ge 0. \tag{15}$$

Note that the case $\mathcal{L}_{1,b} V^\xi > 0$ and $\mathcal{L}_{1,s} V^\xi < 0$ can not occur, as all value functions are increasing functions of g.

The above argument shows that the optimization problem (11) is a free boundary problem in which the optimal trading strategy is defined by the inequalities (i), (ii), and (iii) for a given value function. Besides, the state space $[0, T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ is partitioned into *buy*, *sell*, and *no-transaction* regions for stock P, which are characterized by inequalities (13), (14) and (15), respectively. For sufficiently large ξ, the state space remains divided into a *buy region* $\mathcal{B}$, a *sell region* $\mathcal{S}$, and a *no-transaction region* $\mathcal{N}$ for stock P, which are correspondingly the *sell region*, the *buy region*, and the *no transaction region* for stock Q due to Equation (8). Obviously, the buy and sell regions for stock P are disjoint, as it is not optimal to buy and sell the same stock at the same time. We denote the boundaries between the no-transaction region $\mathcal{N}$ and the buy and sell regions $\mathcal{B}$ and $\mathcal{S}$ as $\partial\mathcal{B}$ and $\partial\mathcal{S}$, respectively.

Let $\xi \to \infty$, the class of admissible trading strategies becomes $\mathcal{T}(g_0)$. We can guess that the state space is still divided into three regions, a region of buying P and selling Q, a region of selling P and buying Q, and a no-transaction region. Then, the optimal trading strategy requires an immediate move to the boundaries of buy or sell regions, if the state is in the buy region $\mathcal{B}$ or the sell region $\mathcal{S}$. Actually, we can obtain equations that each of the value functions should satisfy as follows.

(i) In region $\mathcal{B}$ of buying P and selling Q, the value function remains constant along the path of the state, dictated by the optimal trading strategy, and therefore, for $\delta y \geq 0$

$$V(t,p,x,y,g) = V(t,p,x,y+\delta y, g-(a_p - b_q e^x)p\delta y), \tag{16}$$

where δy is the number of shares of stock P bought and stock Q sold by the investor. δy can be any positive value up to the number required to take the state to $\partial\mathcal{B}$, so letting $\delta y \to 0$ in (16) yields

$$\mathcal{L}_{1,b}V = 0. \tag{17}$$

(ii) Similarly, in region $\mathcal{S}$ of selling P and buying Q, the value function obeys the following equation for $\delta y \geq 0$

$$V(t,p,x,y,g) = V(t,p,x,y-\delta y, g+(b_p - a_q e^x)p\delta y), \tag{18}$$

where δy is the number of shares of stock P sold and stock Q bought by the investor. δy can be any positive value up to the number required to take the state to $\partial\mathcal{S}$, so letting $\delta y \to 0$ in (18) yields

$$\mathcal{L}_{1,s}V = 0. \tag{19}$$

(iii) In the no-transaction region, the value function obeys the same set of equations obtained for the class of absolutely continuous trading strategies, and therefore the value function is given by

$$\mathcal{L}_{1,o}V = 0, \tag{20}$$

and the pair of inequalities, shown above in (15), also hold. Note that, due to the continuity of the value function, if it is known in the no-transaction region, it can be determined in both the buy and sell regions by (17) and (19), respectively.

In the buy region $\mathcal{B}$, $\mathcal{L}_{1,s}V < 0$, and, in the sell region $\mathcal{S}$, $\mathcal{L}_{1,b}V > 0$. Additionally, from the two pairs of inequalities (13) and (14), we may conjecture that $\mathcal{L}_{1,o}V$ in (20) is negative in both the buy region $\mathcal{B}$ and the sell region $\mathcal{S}$. Therefore, the above set of equations can be summarized as the following fully nonlinear partially differential equations (PDE):

$$\min\Big\{-\mathcal{L}_{1,b}V, \mathcal{L}_{1,s}V, -\mathcal{L}_{1,o}V\Big\} = 0 \tag{21}$$

for $(t,p,X,y,g) \in [0,T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. Note that the above discussion also yields the following free boundary problem for the singular stochastic control value function:

$$\begin{cases} \mathcal{L}_{1,b}V &= 0 & \text{in } \mathcal{B} \\ \mathcal{L}_{1,s}V &= 0 & \text{in } \mathcal{S} \\ \mathcal{L}_{1,o}V &= 0 & \text{in } \mathcal{N} \\ V(T,p,x,y,g) &= U(g+J(p,x,y)). & \end{cases} \tag{22}$$

We next show that the value function given by (11) is a constrained viscosity solution of the variational inequality (21) on $[0,T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, and it is the unique bounded constrained viscosity solution of (21). The proof is given in the Appendix A.

Theorem 1. *The value function $V(t,p,x,y,g)$ is a constrained viscosity solution of* (21) *on* $[0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$.

Theorem 2. *Let u be a bounded upper semicontinuous viscosity subsolution of* (21)*, and v a bounded from below lower semicontinuous viscosity supersolution of* (21)*, such that $u(T,\mathbf{x})\le v(T,\mathbf{x})$ for all $\mathbf{x}\in\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$. Then $u\le v$ on $[0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$.*

2.3. Optimal Trading with Exponential Utility Functions

We next assume that the investor has the negative exponential utility function

$$U(z)=1-\exp(-\gamma z), \tag{23}$$

where γ is the constant absolute risk aversion (CARA) parameter such that $-U''(z)/U'(z)=\gamma$. For Equation (21), this utility function can reduce much of computational effort and is easy to interpret. Note that for the utility function (23), the definition of the value function (11) can be expressed as

$$V(t,p,x,y,g)=1-\exp\Big(-\gamma g e^{r(T-t)}\Big)H(t,p,x,y), \tag{24}$$

where $H(t,p,x,y)$ is a convex nonincreasing continuous function in y and defined by

$$\begin{aligned} H(t,p,x,y) &= \inf_{Lp(t),M_p(t)\in\mathcal{T}(g_0)} E\Big\{\exp[-\gamma J(p(T),x(T),y(T)]\,|\,p(t)=p,x(t)=x,y(t)=y\Big\} \\ &= 1-V(t,p,x,y,0). \end{aligned}$$

Plug (24) into (21), and define the following operators for $H(t,p,x,y)$ on $[0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}$,

$$\mathcal{L}_{2,o}H=\frac{\partial H}{\partial t}+\kappa(\theta-x)\frac{\partial H}{\partial x}+\mu p\frac{\partial H}{\partial p}+\frac{1}{2}\nu^2\frac{\partial^2 H}{\partial x^2}+\rho\nu\sigma p\frac{\partial^2 H}{\partial p\partial x}+\frac{1}{2}\sigma^2p^2\frac{\partial^2 H}{\partial p^2},$$

$$\mathcal{L}_{2,b}H=\frac{\partial H}{\partial y}+\gamma e^{r(T-t)}A_-(p,x)H,$$

$$\mathcal{L}_{2,s}H=\frac{\partial H}{\partial y}+\gamma e^{r(T-t)}A_+(p,x)H.$$

Then (21) is transformed into the following PDE for $H(t,p,x,y)$

$$\min\Big\{\mathcal{L}_{2,b}H,-\mathcal{L}_{2,s}H,\mathcal{L}_{2,o}H\Big\}=0 \tag{25}$$

with the following boundary conditions

$$H(T,p,x,y)=\exp\big\{-\gamma J(p,x,y)\big\}.$$

Correspondingly, the free boundary problem (22) becomes

$$\left\{\begin{array}{rcll} \mathcal{L}_{2,o}H &=& 0 & y\in[Y_b(t,p,x),Y_s(t,p,x)] \\ \mathcal{L}_{2,b}H &=& 0 & y\le Y_b(t,p,x) \\ \mathcal{L}_{2,s}H &=& 0 & y\ge Y_s(t,p,x) \\ H(T,p,x,y) &=& \exp\big\{-\gamma J(p,x,y)\big\}. & \end{array}\right. \tag{26}$$

in which $Y_b(t,p,x)$ and $Y_s(t,p,x)$ are the buy and sell boundaries for stock P, respectively. Note that the function $H(t,p,x,y)$ is evaluated in the four-dimensional space $[0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$. Furthermore, this suggests that while (t,u_t,w_t) is inside the no-transaction region, the dynamics of $h(t,u,w,y)$ are driven by two-dimensional standard Brownian motions

$\{z_t, t \geq 0\}$ and $\{w_t, t \geq 0\}$ with correlation ρ. In the buy and sell regions, it follows from (26) that

$$H(t,p,x,y) = \exp\{-\gamma e^{r(T-t)} A_-(p,x)[y - Y_b(t,p,x)]\} H(t,p,x,Y_b(t,p,x)), \quad y \leq Y_b(t,p,x),$$
$$H(t,p,x,y) = \exp\{-\gamma e^{r(T-t)} A_+(p,x)[y - Y_s(t,p,x)]\} H(t,p,x,Y_s(t,p,x)), \quad y \geq Y_s(t,p,x).$$

3. Discretization and a Numerical Algorithm

The solution of the PDE (21) or (25) can be obtained by turning the stochastic differential Equations (1), (3), (6) and (9) into Markov chains and then applying the discrete time dynamic programming algorithm. The discrete state is $\mathbb{X} = (\chi, \mathbb{p}, \mathbb{x}, \vartheta, \mathbb{g})$, whose elements denote time, price of stock P, spread, number of shares of stock P, and amount in the bank in a discrete space. The value function, denoted by $\mathbb{V}$, are given a value at the final time by using the boundary conditions for the continuous value functions over the discrete subspace $(\mathbb{p}, \mathbb{x}, \vartheta, \mathbb{g})$, and then they are estimated by proceeding backward in time by using the discrete time algorithm. As in the continuous time case, this algorithm is the same for both value functions and is derived below for a value function denoted by $\mathbb{V}^\delta(\chi, \mathbb{p}, \mathbb{x}, \vartheta, \mathbb{g})$, where ρ is a discretization parameter, which depends on the discrete time interval t_δ. If t_δ and the resolution of the ϑ-axis ϑ_δ are sent to zero, then the above discrete value function converges to a viscosity subsolution and a viscosity supersolution of the PDE (21). Therefore, all the discrete value functions converge to their continuous counterparts; this is due to the uniqueness of the viscosity solution.

Consider an evenly spaced partition of the time interval $[0, T]$: $\chi = \{\delta, 2\delta, \ldots, n\delta\}$, where $\delta = T/n$, and two evenly spaced partitions of the space intervals $\mathbb{z} = \{0, \pm\sqrt{\delta}, \pm 2\sqrt{\delta}, \ldots, \}$ and $\mathbb{w} = \{0, \pm\sqrt{\delta}, \pm 2\sqrt{\delta}, \ldots, \}$. The grid $\mathbb{p}$ is defined by $\mathbb{z}$ via the following transformation,

$$\mathbb{p}_i = \exp\Big((\mu - \frac{1}{2}\sigma^2)T + \mathbb{z}_i \sigma \sqrt{T}\Big). \tag{27}$$

Note that the SDE (3) implies that the asymptotic distribution of $X(t)$ is Normal $(\theta, \nu^2/(2\kappa))$, we define grid $\mathbb{x}$ by

$$\mathbb{x}_j = \theta + \frac{\nu}{\sqrt{2\kappa}} \mathbb{w}_j. \tag{28}$$

Denote $\chi_i = i\delta$ for $i = 1, \ldots, n-1$. The dynamics (1) and (3) of $P(t)$ and $X(t)$ implies the following transition density for $(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i))$,

$$\begin{pmatrix} \mathbb{p}(\chi_{i+1}) \\ \mathbb{x}_{\chi_{i+1}} \end{pmatrix} \Bigg| \begin{pmatrix} \mathbb{p}(\chi_i) \\ \mathbb{x}_{\chi_i} \end{pmatrix} \sim N\left(\begin{pmatrix} \log \mathbb{p}(\chi_i) + (\mu - \frac{1}{2}\sigma^2)\delta \\ (1 - \delta\kappa)\mathbb{x}(\chi_i) + \delta\kappa\theta \end{pmatrix}, \begin{pmatrix} \delta\sigma^2, \delta\rho\sigma\nu \\ \delta\rho\sigma\nu, \delta\nu^2 \end{pmatrix} \right). \tag{29}$$

We also note that the discrete time equation for the amount in the bank $\mathbb{g}(\chi)$ is

$$\mathbb{g}(\chi_{i+1}) = \mathbb{g}(\chi_i) \exp(r\delta).$$

Given the grid defined above, the discrete time dynamic programming principle is invoked, and the following discretization scheme is proposed for PDE (21):

$$\begin{aligned} \mathbb{V}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta, \mathbb{g}(\chi_i)) = \max\Big\{ & \\ & \mathbb{V}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta + \xi, \mathbb{g}(\chi_i) - (a_p - b_q e^{\mathbb{x}(\chi_i)})\mathbb{p}(\chi_i)\xi), \\ & \mathbb{V}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta - \xi, \mathbb{g}(\chi_i) + (b_p - a_q e^{\mathbb{x}(\chi_i)})\mathbb{p}(\chi_i)\xi), \\ & E\{\mathbb{V}^\delta(\chi_{i+1}, \mathbb{p}(\chi_{i+1}), \mathbb{x}(\chi_{i+1}), \vartheta, \mathbb{g}(\chi_{i+1}))\} \Big\}. \end{aligned} \tag{30}$$

where $\xi > 0$ is a real constant and $i = 0, \ldots, n-1$. This scheme is based on the principle that the investor's policy is the choice of the optimum transaction. We next show that, as the discretization parameter $\delta \to 0$, the solution $\mathbb{V}^\delta$ of (30) converges to the value function V, or, equivalently, to the unique constrained viscosity solution of (21).

Theorem 3. *The solution $\mathbb{V}^\delta$ of (30) converges locally uniformly as $\delta \to 0$ to the unique continuous constrained viscosity solution of (21).*

For the exponential utility function $U(z) = 1 - \exp(-\gamma z)$, the value function V can be expressed as (24), and its discretization scheme is given by

$$\mathbb{V}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta, \mathbb{g}(\chi_i)) = 1 - \exp\Big(- \gamma \mathbb{g}(\chi_i) e^{r(T-\chi_i)}\Big)\mathbb{H}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta).$$

Then, the discretization scheme (30) can be reduced to

$$\begin{aligned}\mathbb{H}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta) = \min\Big\{ & F_b(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \xi) \cdot \mathbb{H}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta + \xi), \\ & F_s(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \xi) \cdot \mathbb{H}^\delta(\chi_i, \mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \vartheta - \xi),\ E\{\mathbb{H}^\delta(\chi_{i+1}, \mathbb{p}(\chi_{i+1}), \mathbb{x}(\chi_{i+1}), \vartheta)\}\Big\}.\end{aligned} \tag{31}$$

where

$$F_b(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \xi) = \exp\{\gamma \xi A_-(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i)) e^{r(T-\chi_i)}\},$$
$$F_s(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i), \xi) = \exp\{ - \gamma \xi A_+(\mathbb{p}(\chi_i), \mathbb{x}(\chi_i)) e^{r(T-\chi_i)}\}.$$

4. Simulation Studies

4.1. Buy and Sell Regions

We use the numerical algorithm proposed in Section 2 to studies the buy and sell boundaries of the pairs trading strategy. Our study focuses on two aspects of the problem. The first is the property of buy and sell boundaries (or no transaction regions) for a given set of model parameters, and the other is the impact of different model parameters on the shape of buy and sell boundaries. Without loss of the generality, we assume the time horizon $T = 1$ and $p(0) = 1$ in all our simulation studies.

We first consider a baseline scenario. The parameter values in the baseline scenario are $\mu = 0.2, \sigma = 0.4, \theta = 0.1, \kappa = 1, \nu = 0.15, \rho = 0.5, r = 0.01, \gamma = 5$ and $\zeta_p = \zeta_q = \xi_p = \xi_q = 0.0005$. For convenience, we label the setting of the baseline parameter values as Scenario 1 or (S1). We discretize the state space (t, p, x, y, g) and use the developed Markov chain approximation to solve the discretized optimization problem. Figure 1 shows the buy and sell surfaces of (S1) at time $t = 0.05, 0.35, 0.65$, and 0.95. To better read the figure, we also show in Figures 2 and 3 the buy and sell boundaries of (S1) at prices $p = 0.845, 1.095, 1.400, 2.108$, and $x = 0.023, 0.092, 0.157, 0.266$, respectively. These points are chosen such that they correspond to the 24%, 48%, 72%, and 96% quantiles of the distribution of $p(T)$ and asymptotic distribution of $x(t)$, respectively. We find the following from these figures. First, at a given time and a given price level, the no transaction region becomes narrower when the spread gets larger, and the no transaction region moves from the negative to the positive when the spread turns from the negative to the positive. For example, at $t = 0.05$ and $p(t) = 0.845$, the no transaction region changes from $[-9.4, -8.0]$ at $x(t) = 0.023$ to $[-4.6, -3.4]$ at $x(t) = 0.092$, $[-0.7, 0.2]$ at $x(t) = 0.157$, and $[3.2, 3.7]$ at $x(t) = 0.266$. Second, at a given time and a given spread level, the no transaction region becomes narrower when the price $p(t)$ gets larger, and the no transaction region moves up when the price becomes larger. For instance, at $t = 0.05$ and $x(t) = 0.023$, the no transaction region changes from $[-9.4, -8.0]$ at $p(t) = 0.845$ to $[-6.8, -5.6]$ at $p(t) = 1.095$, $[-4.9, -3.9]$ at $p(t) = 1.400$, and $[-2.7, -2.0]$ at $p(t) = 2.108$. Note that the movement of the no transaction region with respect to price change but with a fixed spread level is relatively smaller than that with respect to spread change but with a fixed price level. Third, when time ellipses from 0 to 1, the no transaction region moves upward. For instance, at the fixed price-spread level $(p(t), x(t)) = (1.095, 0.092)$, the no transaction intervals at

$t = 0.05, 0.35, 0.65$ and 0.95 are $[-2.6, -1.6]$, $[-2.1, -1.2]$, $[-1.5, -0.7]$, and $[-0.8, -0.2]$, respectively.

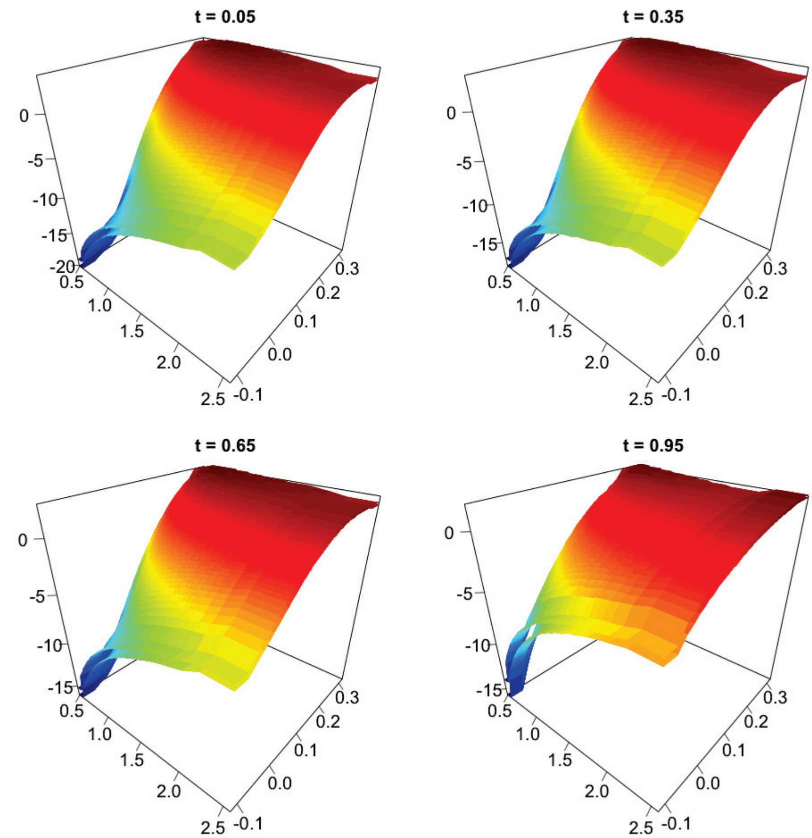

Figure 1. Buy and sell boundaries of the baseline scenario (S1) at different times.

We then discuss the impact of different parameter values on the buy and sell boundaries (or no transaction regions). Besides the parameter values in (S1), we now consider other 18 sets of parameter values, labeled as Scenarios 2–19. In each of Scenarios 2–19, all parameters values are same as those in (S1) except one parameter is changed as the specification; see Table 1 that summarizes parameter values in all 19 scenarios. For example, Scenario 2 uses parameter values $\mu = 0.1$ and assume all other parameters $\sigma, \theta, \kappa, \nu, \rho, r, \gamma$ and $\zeta_p = \zeta_q = \xi_p = \xi_q$ have same values as those in (S1). We discretize the state space (t, p, x, y, g), and use the developed Markov chain approximation to solve the discretized optimization problem for Scenarios 2–19.

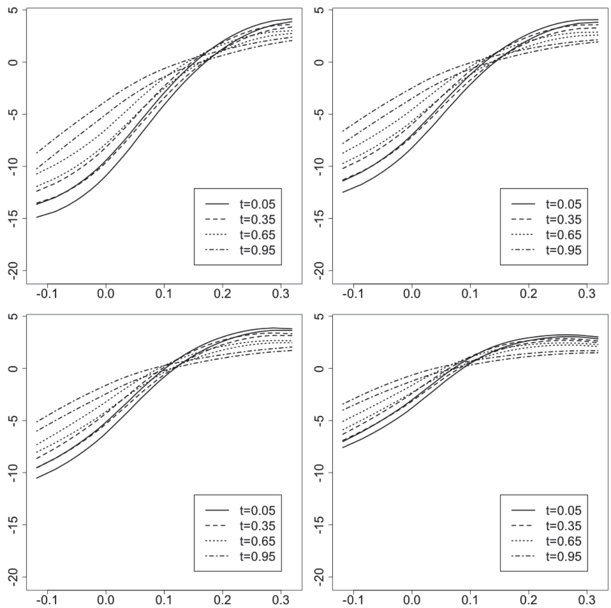

Figure 2. Buy and sell boundaries of at prices $P_t = 0.845$ (**top left**), 1.095 (**top right**), 1.400 (**bottom left**), and 2.108 (**bottom right**) and different times.

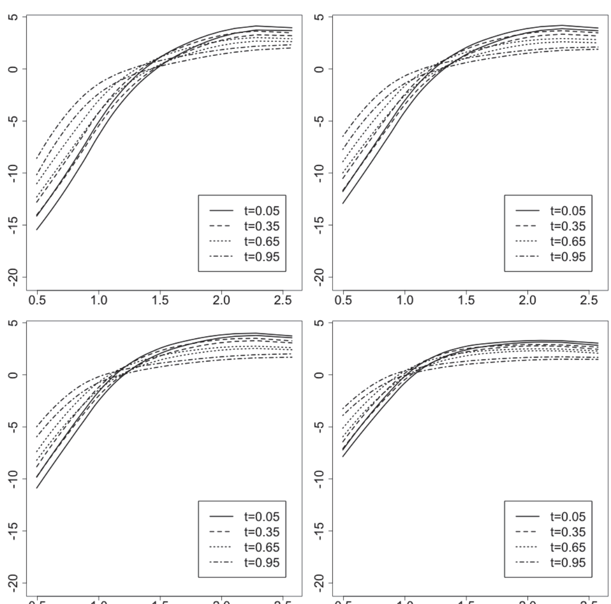

Figure 3. Buy and sell boundaries of at spread $X_t = 0.023$ (**top left**), 0.092 (**top right**), 0.157 (**bottom left**), and 0.266 (**bottom right**) and different times.

To compare the buy and sell boundaries (or no transaction regions) among different scenarios, we plot the buy and sell boundaries over time at four fixed points $(p^{(1)}, x^{(1)}) = (0.9, 0.09)$, $(p^{(2)}, x^{(2)}) = (0.9, 0.12)$, $(p^{(3)}, x^{(3)}) = (1.5, 0.09)$, and $(p^{(4)}, x^{(4)}) = (1.5, 0.12)$, respectively. Figures 4–12 demonstrate variations of the buy and sell boundaries over time for different values of $\mu, \sigma, \theta, \kappa, \nu, \rho, r, \gamma, \zeta_p (= \zeta_q = \xi_p = \xi_q)$, respectively. In each figure, we plot the buy and sell boundaries for $(p^{(i)}, x^{(i)}), i = 1, 2, 3, 4$ on the top left, top right, bottom left, and bottom right, respectively, we also use the solid (dashed, dotted) lines to represent the baseline value (the smaller value, the larger value) of the parameter under comparison. Figure 4 suggests that when μ increases, the buy and sell boundaries move downward at all four points. Figure 5 indicates that when σ increases, the buy and sell boundaries move upward at $(p^{(1)}, x^{(1)})$ and $(p^{(2)}, x^{(2)})$, but move downward at $(p^{(3)}, x^{(3)})$ and $(p^{(4)}, x^{(4)})$. Figure 6 shows that, when θ increases, the buy and sell boundaries move downward at all four points. Figure 7 indicates that, when κ increases, the buy and sell boundaries move downward, and the magnitude of such movement is larger at $(p^{(1)}, x^{(1)})$ than the other three points. Figure 8 shows that, when ν increases, the buy and sell boundaries move upward at $(p^{(i)}, x^{(i)}), i = 1, 2, 3$, but move downward at $(p^{(4)}, x^{(4)})$. Figure 9 suggests that, when the correlation ρ changes from the negative to the positive, the buy and sell boundaries move downwards at $(p^{(1)}, x^{(1)})$ and $(p^{(2)}, x^{(2)})$, but move upward at $(p^{(3)}, x^{(3)})$ and $(p^{(4)}, x^{(4)})$. Figure 10 indicates that variations of interest rate r have little impact on the buy and sell boundaries. Figure 11 shows that, when the risk aversion parameter γ increases, the buy and sell boundaries move upward at $(p^{(i)}, x^{(i)}), i = 1, 2, 3$, but move downward at $(p^{(4)}, x^{(4)})$. Figure 12 suggests that, when the transaction cost increases, the center of the no transaction region seems to not change, but the region gets wider.

Table 1. Parameter values of different scenarios.

(S1) $\mu = 0.2, \sigma = 0.4, \theta = 0.1, \kappa = 1, \nu = 0.15, \rho = 0.5,$ $r = 0.01, \gamma = 5$ and $\zeta_p = \zeta_q = \xi_p = \xi_q = 0.0005$.		
(S2) $\mu = 0.1$	(S8) $\kappa = 0.8$	(S14) $r = 0.005$
(S3) $\mu = 0.3$	(S9) $\kappa = 1.2$	(S15) $r = 0.03$
(S4) $\sigma = 0.2$	(S10) $\nu = 0.1$	(S16) $\gamma = 3$
(S5) $\sigma = 0.6$	(S11) $\nu = 0.2$	(S17) $\gamma = 8$
(S6) $\theta = -0.05$	(S12) $\rho = -0.2$	(S18) $\zeta_p = \zeta_q = \xi_p = \xi_q = 0.0001$
(S7) $\theta = 0.3$	(S13) $\rho = 0.6$	(S19) $\zeta_p = \zeta_q = \xi_p = \xi_q = 0.0010$

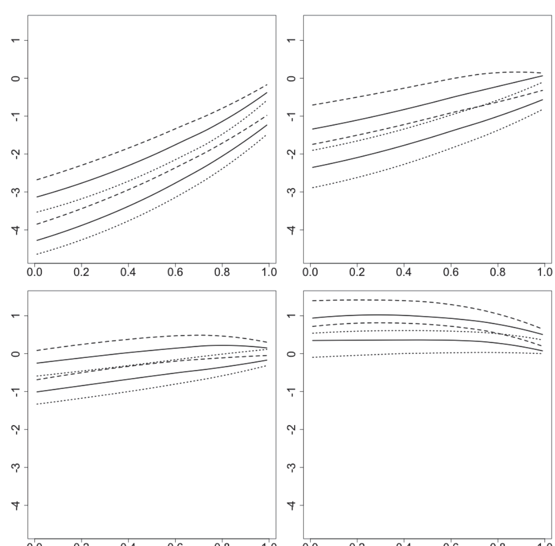

Figure 4. Buy and sell boundaries of at fixed prices $(p^{(i)}, x^{(i)}), i = 1, 2, 3, 4$ for $\mu = 0.1$ (dashed), 0.2 (solid), 0.3 (dotted).

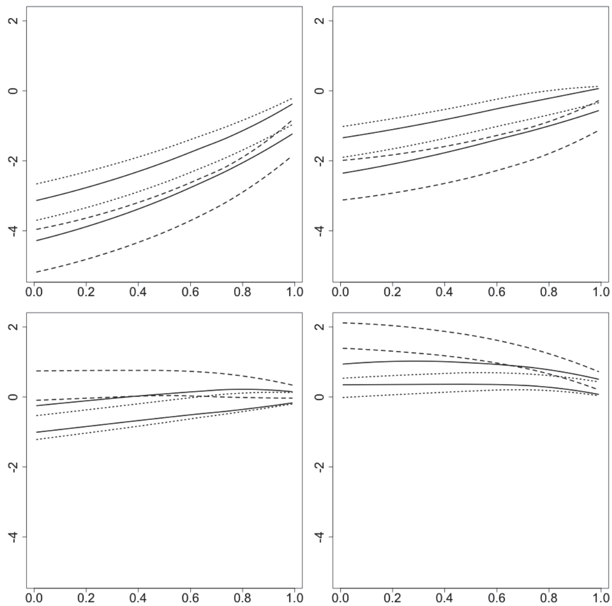

Figure 5. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\sigma = 0.2$ (dashed), 0.4 (solid), 0.6 (dotted).

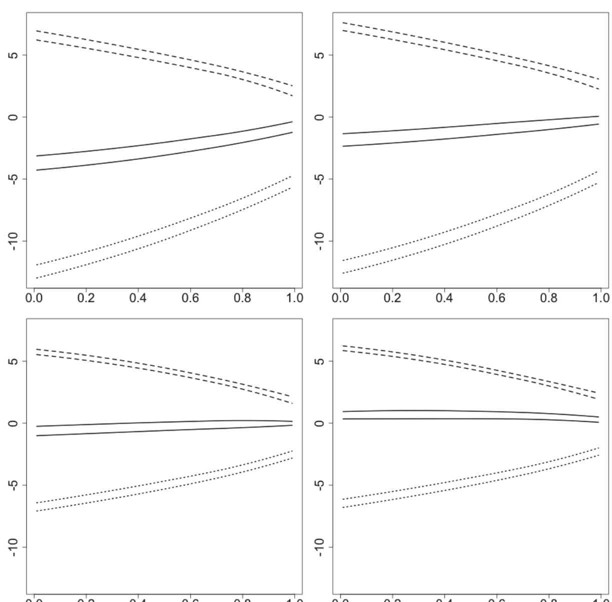

Figure 6. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\theta = -0.05$ (dashed), 0.1 (solid), 0.3 (dotted).

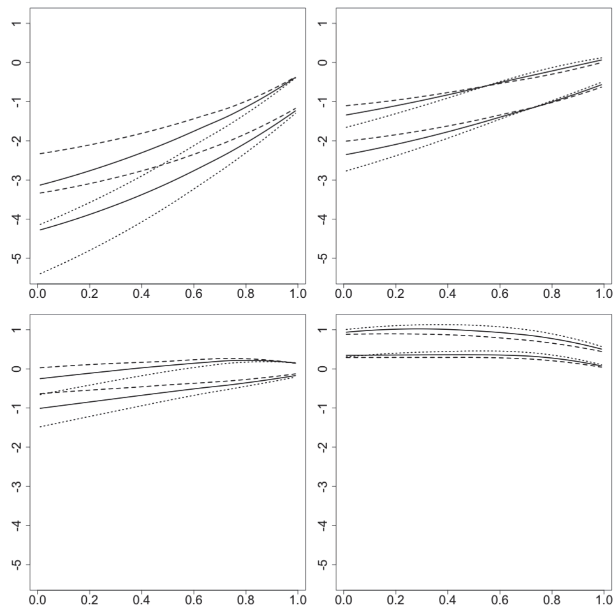

Figure 7. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\kappa = 0.8$ (dashed), 1 (solid), and 1.2 (dotted).

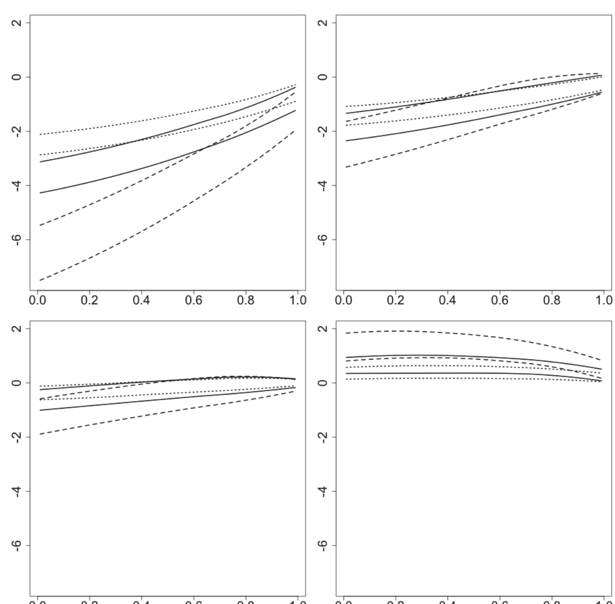

Figure 8. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\nu = 0.1$ (dashed), 0.15 (solid), and 0.2 (dotted).

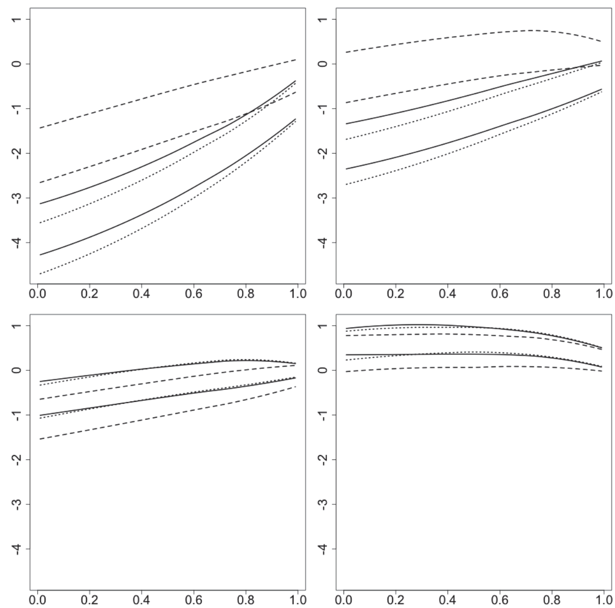

Figure 9. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\rho = -0.2$ (dashed), 0.5 (solid), and 0.6 (dotted).

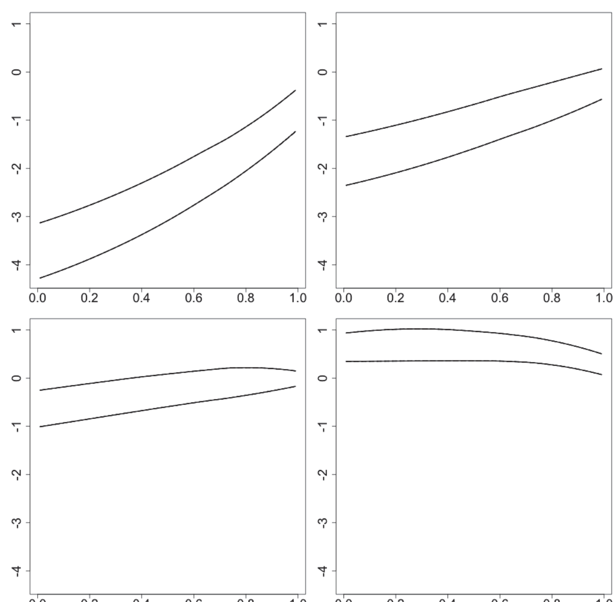

Figure 10. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $r = 0.005$ (dashed), 0.01 (solid), and 0.03 (dotted).

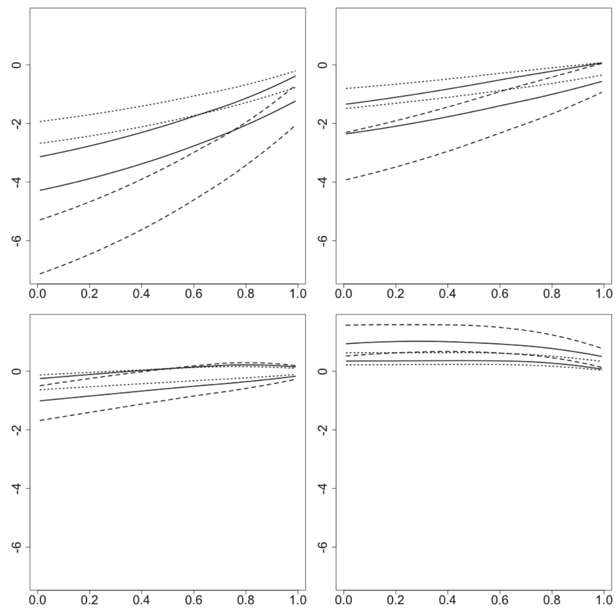

Figure 11. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\gamma = 3$ (dashed), 5 (solid), and 8 (dotted).

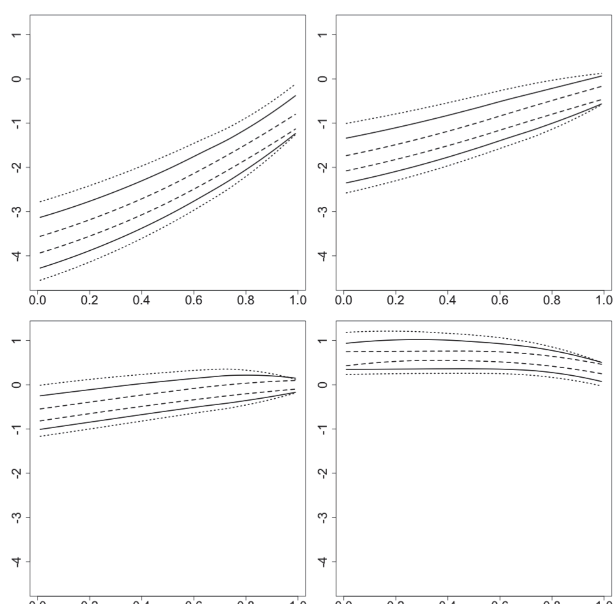

Figure 12. Buy and sell boundaries of at fixed price $(p^{(i)}, x^{(i)})$, $i = 1, 2, 3, 4$ for $\zeta_p = \zeta_q = \xi_p = \xi_q =$ 0.0001 (dashed), 0.0005 (solid), and 0.0010 (dotted).

4.2. Performance of the Strategy

We also perform simulation studies to investigate the performance of the optimal trading strategy. For comparison purpose, we also consider a benchmark strategy that is analogous to the relative-value arbitrage strategy used in Gatev et al. (2006) and based on standard deviation of the spread. Specifically, the strategy opens a position when the spread exceeds twice the standard deviation of the spread process, and closes the position when either price converges or the maturity is reached. As the benchmark strategy doesn't specify the number of shares of stocks that should be bought or sold, we assume that the number of shares of stocks traded each time is one.

We simulate the price process p_t and the spread process x_t to compare the performance of the benchmark strategy and our strategy in scenarios (S1)–(S19). Assume that $T = 1$, and we discretize the time interval $(0, 1]$ as $\{0.01, 0.02, \ldots, 0.99, 1\}$, so that we have 100 trading periods. For each scenario, we simulate 1000 paths of $\{(p_t, x_t)|t = 0, 0.01, \ldots, 0.99, 1, p_0 = 1\}$, and for each simulated path (p_t, x_t), we implement the benchmark strategy and the optimal strategy at $t = 0.01, 0.02, \ldots, 0.99$ and close the position at $T = 1$. Let $i = b, o$ represent the benchmark and the optimal strategies, respectively. For each realized trading strategies, denote $N^{(i)}$ as the number of trades (i.e., buy and sell) among the 100 trading periods and $PL^{(i)} = -C^{(i)}(L_p, M_p; 0, 1)$ the total profit made during the trading process. Note that the benchmark strategy trades only one share of stock each time while the number of shares of stocks in the optimal strategy are "optimally" chosen based on the buy and sell regions, we define $PS^{(i)}$ as the the average profit (or loss) generated from the maximum number of shares of stocks during the trading process. That is, $PS^{(i)} := -C^{(i)}(L_p, M_p; 0, 1) / \max_t |Y_t^{(i)}|$, where $Y_t^{(i)}$ is the number of shares of stock P at $t = 0.01, 0.02, \ldots, 0.99$.

Table 2 summarizes the mean and standard error of $N^{(i)}$, $PL^{(i)}$, and $PS^{(i)}$ ($i = o, b$) for 1000 paths in each scenario. We note that the total numbers of trades $N^{(o)}$ in the optimal strategy range from 45.736 to 55.821 for (S1)–(S17), and increases (or decreases) significantly when the transaction costs decreases (or increases) in (S18) and (S19). In comparison to this, the total numbers of trades $N^{(b)}$ in the benchmark strategy are much smaller, essentially, between 1 and 2. This suggests the benchmark strategy is much more conservative than the optimal strategy. For the realized profit over the trading period, $PL^{(o)}$ is much larger than $PL^{(b)}$ as the optimal strategy can choose to buy or sell the "optimal" number of shares of stock pairs, while the benchmark strategy only buy or sell one share of stock pair. $PS^{(o)}$ and $PS^{(b)}$ remove the impact of number of shares of traded stocks, and provide the average earning per traded stock, and we notice that $PS^{(o)}$ is still significantly higher than $PS^{(b)}$.

Table 2. Performance of strategies.

	$N^{(o)}$	$PL^{(o)}$	$PS^{(o)}$	$N^{(b)}$	$PL^{(b)}$	$PS^{(b)}$
(S1)	52.289 (0.247)	0.349 (0.019)	0.048 (0.004)	1.094 (0.084)	0.005 (0.002)	0.005 (0.002)
(S2)	53.218 (0.241)	0.389 (0.020)	0.051 (0.004)	1.094 (0.084)	0.006 (0.002)	0.006 (0.002)
(S3)	51.348 (0.253)	0.318 (0.019)	0.046 (0.004)	1.094 (0.084)	0.004 (0.002)	0.004 (0.002)
(S4)	52.999 (0.208)	0.378 (0.019)	0.054 (0.003)	1.094 (0.084)	0.007 (0.002)	0.007 (0.002)
(S5)	51.896 (0.275)	0.326 (0.019)	0.040 (0.005)	1.094 (0.084)	0.003 (0.004)	0.003 (0.004)
(S6)	49.299 (0.235)	0.357 (0.020)	0.032 (0.003)	1.094 (0.084)	0.003 (0.002)	0.003 (0.002)
(S7)	54.233 (0.262)	0.344 (0.019)	0.064 (0.005)	1.094 (0.084)	0.008 (0.003)	0.008 (0.003)
(S8)	55.821 (0.304)	0.359 (0.021)	0.062 (0.006)	1.094 (0.084)	0.005 (0.003)	0.005 (0.003)
(S9)	45.736 (0.254)	0.266 (0.016)	0.046 (0.003)	1.094 (0.084)	0.005 (0.002)	0.005 (0.002)
(S10)	46.347 (0.292)	0.228 (0.016)	0.042 (0.005)	1.052 (0.083)	0.004 (0.002)	0.004 (0.002)
(S11)	57.689 (0.212)	0.489 (0.022)	0.053 (0.003)	1.206 (0.084)	0.007 (0.002)	0.007 (0.002)
(S12)	46.774 (0.248)	0.325 (0.015)	0.065 (0.003)	1.140 (0.086)	0.008 (0.001)	0.008 (0.001)
(S13)	53.516 (0.245)	0.361 (0.020)	0.045 (0.004)	1.140 (0.087)	0.006 (0.002)	0.006 (0.002)
(S14)	54.027 (0.232)	0.579 (0.032)	0.048 (0.004)	1.094 (0.084)	0.005 (0.002)	0.005 (0.002)
(S15)	50.031 (0.266)	0.219 (0.012)	0.049 (0.004)	1.094 (0.084)	0.005 (0.002)	0.005 (0.002)
(S16)	52.300 (0.247)	0.347 (0.019)	0.048 (0.004)	1.094 (0.084)	0.005 (0.002)	0.005 (0.002)
(S17)	52.261 (0.247)	0.357 (0.019)	0.050 (0.004)	1.094 (0.084)	0.006 (0.002)	0.006 (0.002)
(S18)	73.801 (0.286)	0.339 (0.019)	0.045 (0.004)	1.094 (0.084)	0.006 (0.002)	0.006 (0.002)
(S19)	42.996 (0.222)	0.339 (0.019)	0.049 (0.004)	1.094 (0.084)	0.004 (0.002)	0.004 (0.002)

5. Real Data Studies

We test our model with real market data in this section. We present the sample and explain our methodology first, and then show the results and discussion.

A key step of implementing pairs trading strategy is to select two stocks for pairs trading. Gatev et al. (2006) illustrate how this can be done by using stock price data. An alternative to this approach is to use fundamentals analysis to select two stocks that have almost the same risk factor exposures; see Vidyamurthy (2004). In this study, we consider a hybrid of these two approaches. Specifically, we restrict two stocks P and Q to belong to the same industry sector. Table 3 lists six pairs of stocks selected from four different sectors. For each pair of stocks P and Q, we compute the spread by regressing log price of stock Q on the log price of stock P, and the fitted values of the regression is considered as the "transformed" price of P. Figure 13 shows six pairs of the original prices of Q and transformed prices of P over time.

Table 3. Six pairs of stocks selected from different industries.

Sector	Stock Q	Stock P
Consumer goods	Apple Inc. (AAPL)	Procter & Gamble Co. (PG)
Consumer goods	Coca-Cola Co. (KO)	PepsiCo, Inc. (PEP)
Technology	Alphabet Inc Class A (GOOGL)	Microsoft Corporation (MSFT)
Technology	AT&T Inc. (T)	Verizon Communications Inc. (VZ)
Industrial goods	Boeing Corporation (BA)	General Electric Company (GE)
Financial	Goldman Sachs Group Inc. (GS)	JPMorgan Chase & Co. (JPM)

Figure 13. Original (solid) and transformed (dashed) prices of six pairs of stocks.

We then apply the optimal strategy and the benchmark strategy in Section 4.2 to test the out-of-the-sample performance. Specifically, we use the past three years of the historical data of each pair to estimate the model parameter, and run a unit-root test to conclude if the spread x_t is a stationary process. If x_t is not stationary, we do not implement any strategies. Otherwise, we implement both the optimal strategy and the benchmark strategy. Note that the optimal strategy can optimally choose the number of shares of stocks in each trade, while we still trade one unit of stock in the benchmark strategy. Table 4 shows the number of trades $N^{(i)}$, the accumulated profit (in U.S. dollars) at maturity $PL^{(i)}$, and the average profit per traded share $PS^{(i)}$ over two testing periods, for $i = o$ (the optimal strategy) and $i = b$ (the benchmark strategy). Table 4 suggests that the benchmark strategy is much more conservative than the optimal strategy. Besides, the average profits per traded share $PS^{(o)}$ of the optimal strategy are much larger than that of the benchmark strategy except for the stock pair (KO, PEP).

Table 4. Performance of strategies.

Pairs	Year	$N^{(o)}$	$PL^{(o)}$	$PS^{(o)}$	$N^{(b)}$	$PL^{(b)}$	$PS^{(b)}$
(AAPL, PG)	2014	58	8.56	2.173	0	0	0
	2015	70	25.439	3.91	0	0	0
(BA, GE)	2014	97	27.866	1.292	0	0	0
	2015	165	168.543	1.982	20	0.455	0.455
(T, VZ)	2014	127	77.908	2.158	2	0.603	0.603
	2015	131	115.587	2.883	0	0	0
(GOOGL, MSFT)	2015	103	94.271	6.734	8	1.623	1.623
	2016	135	65.957	6.296	0	0	0
(GS, JPM)	2015	100	7.654	0.195	6	−2.54	−2.54
	2016	200	94.542	2.375	8	−1.66	−1.66
(KO, PEP)	2015	142	37.51	0.675	22	10.154	10.154
	2016	165	217.878	4.059	4	5.983	5.983

6. Concluding Remark

The problem of optimal pairs trading has been studied by many academic researchers and financial practitioners. Existing models and methods try to find either the optimal shares of stocks by assuming no transaction costs, or the optimal timing of trading fixed number of shares of stocks with transaction costs. In contrast to these analysis, this paper studies the joint effect of optimal shares and optimal trading times in pairs trading process with proportional transaction costs. Under the assumption that the investor's aim is to maximize the expected utility of terminal wealth, the optimal pair trading problem can be written as a singular stochastic control problem and solved by the approach in Davis et al. (1993). We then demonstrate the advantage of joint consideration of optimal shares and optimal trading times in pair trading via simulation and empirical studies.

The following issues may need further investigation to make this study more practical. First, our approach can be easily extended for nonexponential utility functions. In such a case, the optimization problem involves five (instead of four) variables, and the numerical algorithm in our paper needs to be modified to adapt for five variables. Second, our approach can be extended to solve the optimal co-integration trading, which involves n stocks with m co-integration relationship. Third, many empirical studies suggest that stock price processes can be better approximated by incorporating jumps. Using the framework and algorithms developed in Xing et al. (2017), the method developed here can be extended to the case that price processes follow geometric jump-diffusion processes. In such a case, the value function of the corresponding variational inequalities involve integro-differential equations, which can be solved by extending our numerical algorithm.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Proof of Theorems

Proof of Theorem 1. In our case, the state $\mathbf{X}$ is $(s, \mathbf{x})$, where $\mathbf{x} = (p, x, y, g)$. Let $\mathbf{X}_0 = (s_0, p_0, x_0, y_0, G_0)$, it follows that there exists an optimal trading strategy, dictated by the pair of processes $(L_p^*(t), M_p^*(t)$, where $\mathbf{X}_0^*(t) = (t, p_0^*(t), x_0^*(t), y_0^*(t), g_0^*(t))$ is the optimal trajectory, with $\mathbf{X}_0^*(s_0) = \mathbf{X}_0$.

(i) First, we prove that V is a viscosity subsolution of (21) on $[0, T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R})$. For this, we must show that, for all smooth functions $\phi(\mathbf{X})$, such that $V(\mathbf{X}) - \phi(\mathbf{X})$ has a local maximum at $\mathbf{X}_0$, the following inequality holds:

$$\min\Big\{ -\mathcal{B}\phi(\mathbf{X}_0), \mathcal{S}\phi(\mathbf{X}_0), -\mathcal{L}\phi(\mathbf{X}_0)\Big\} \le 0. \tag{A1}$$

Without loss of generality, we assume that $V(\mathbf{X}_0) = \phi(\mathbf{X}_0)$ and $V \le \phi$ on $[0, T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. We argue by contradiction: if the arguments inside the operator of (A1) satisfy $-\mathcal{B}\phi(\mathbf{X}_0) > 0$ and $\mathcal{S}\phi(\mathbf{X}_0) > 0$, then there exists $\theta > 0$, such that $-\mathcal{L}\phi(\mathbf{X}_0) > \theta$. From the fact that ϕ is smooth, the above inequalities become $-\mathcal{B}\phi(\mathbf{X}) > 0$, $\mathcal{S}\phi(\mathbf{X}) > 0$, and $-\mathcal{L}\phi(\mathbf{X}) > \theta$, where $\mathbf{X} = (t, p, x, y, g) \in \mathscr{B}(\mathbf{X}_0)$, a neighborhood of $\mathbf{X}_0$. In Lemma 1, it is shown that $\mathbf{X}_0^*(t)$ has no jumps, P-a.s., at $\mathbf{X}_0 = \mathbf{X}_0^*(s_0)$. Hence, $\tau(\omega)$, defined by

$$\tau(\omega) = \inf\{t \in (s_0, T] : \mathbf{X}_0^*(t) \notin \mathscr{B}(\mathbf{X}_0)\},$$

is positive P-a.s., and therefore the integral along $\mathbf{X}_0^*(t)$

$$\begin{aligned} -\theta \int_{s_0}^{\tau} dt >& E\int_{s_0}^{\tau} \mathcal{B}\phi(\mathbf{X}_0^*(t))dL^*(t) - E\int_{s_0}^{\tau} \mathcal{S}\phi(\mathbf{X}_0^*(t))dM^*(t) + E\int_{s_0}^{\tau} \mathcal{L}\phi(\mathbf{X}_0^*(t))dt \\ =& E\{I_1\} - E\{I_2\} + E\{I_3\}, \end{aligned} \tag{A2}$$

where $(L^*(t), M^*(t))$ is the optimal trading strategy at $\mathbf{X}_0$. Applying Itô's formula to $\phi(\mathbf{X})$, where the state dynamics are given by (1)–(6), we get

$$E\{\phi(\mathbf{X}_0^*(\tau))\} = \phi(\mathbf{X}_0) + E\{I_1\} - E\{I_2\} + E\{I_3\}. \tag{A3}$$

Since $V(\mathbf{X}) \le \phi(\mathbf{X})$, for all $\mathbf{X} \in \mathcal{B}(\mathbf{X}_0)$, and $V(\mathbf{X}_0) = \phi(\mathbf{X}_0)$, (A2) and (A3) yield

$$E\{V(\mathbf{X}_0^*(\tau))\} \le V(\mathbf{X}_0) + E\{I_1\} - E\{I_2\} + E\{I_3\} < V(\mathbf{X}_0) - \theta\int_{s_0}^{\tau} dt,$$

which violates the dynamic programming principle, together with the optimality of $(L^*(t), M^*(t))$. Therefore, at least one of the arguments inside the minimum operator of (A1) is nonpositive, and hence the value function is a viscosity subsolution of (21).

(ii) In the second part of the proof, we show that V is a viscosity supersolution of (21). For this, we must show that, for all smooth functions $\phi(\mathbf{X})$, such that $V(\mathbf{X}) - \phi(\mathbf{X})$ has a local minimum at $\mathbf{X}_0$, the following inequality holds:

$$\min\Big\{ -\mathcal{B}\phi(\mathbf{X}_0), \mathcal{S}\phi(\mathbf{X}_0), -\mathcal{L}\phi(\mathbf{X}_0)\Big\} \ge 0, \tag{A4}$$

where, without loss of generality, $V(\mathbf{X}_0) = \phi(\mathbf{X}_0)$ and $V(\mathbf{X}) \geq \phi(\mathbf{X})$ on $[0, T] \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. In this case, we prove that each argument of the minimum operator of (A4) is non-negative.

Consider the trading strategy $L(t) = L_0 > 0$, $s_0 \leq t \leq T$, and $M(t) = 0$, $s_0 \leq t \leq T$. By the dynamic programming principle,

$$V(s_0, p_0, x_0, y_0, g_0) \geq V(s_0, p_0, x_0, y_0 + L_0, g - (a_p - b_q e^{X_0}) p_0 L_0).$$

This inequality holds for $\phi(s, p, x, y, g)$ as well, and, by taking the left-hand side to the right-hand side, dividing by L_0, and sending $L_0 \to 0$, we get $\mathcal{B}\phi(\mathbf{X}_0) \leq 0$. Similarly, by using the trading strategy $L(t) = 0, s_0 \leq t \leq T$, and $M(t) = M_0 > 0$, $s_0 \leq t \leq T$, the second argument inside the minimum operator is found to be non-negative.

Finally, consider the case where no trading is applied. By the dynamic programming principle

$$E\{V(\mathbf{X}_0^d(t))\} \leq V(s_0, p_0, x_0, y_0, g_0), \tag{A5}$$

where $\mathbf{X}_0^d(t)$ is the state trajectory of starting at s_0, when $M(t) = L(t) = 0$, $s_0 \leq t \leq T$, given by (1)–(6) as

$$\mathbf{X}_0^d(t) = (t, p(t), x(t), y_0, g(t))$$

and $\mathbf{X}_0^d(t) \in \mathscr{B}(\mathbf{X}_0)$. Therefore, by applying Itô's rule on $\phi(s, X, B, y, G)$, inequality (A5) yields

$$E\left\{ \int_{s_0}^{t} \mathcal{L}\phi(\mathbf{X}_0^d(\xi)) d\xi \right\} \leq 0,$$

and, by letting $t \to s_0$, the third argument inside the minimum operator is found to be non-negative. This complete the proof. □

Lemma A1. *Assume that $-\mathcal{B}\phi(\mathbf{X}_0) > 0$, and denote the event that the optimal trajectory $\mathbf{X}_0^*(t)$ has a jump of size ϵ, along the direction $(0, 0, 0, 1, -(a_p - b_q e^{x_0}) p_0)$ by $A(\omega)$. Assume that the state (after the jump) is $(s_0, p_0, x_0, y_0 + \epsilon, -(a_p - b_q e^{x_0}) B_0 \epsilon) \in \mathscr{B}(\mathbf{X}_0)$. Then,*

$$\Big(\mathcal{B}\phi(\mathbf{X}_0)\Big) P(A) \geq 0, \tag{A6}$$

therefore $P(A) = 0$. Similarly, if $\mathcal{S}\phi(\mathbf{X}_0) > 0$, then the optimal trajectory has no jumps along the direction $(0, 0, 0, -1, (b_p - a_q e^{x_0}) p_0)$, P-a.s. at $\mathbf{x}_0$.

Proof. By the principle of dynamic programming,

$$\begin{aligned} &V(s_0, p_0, x_0, y_0, g_0) = E\{V(s_0, p_0, x_0, y_0 + \epsilon, -(a_p - b_q e^{x_0}) B_0 \epsilon)\} \\ &= \int_{A(\omega)} V(s_0, p_0, x_0, y_0 + \epsilon, -(a_p - b_q e^{x_0}) B_0 \epsilon) dP + \int_{A(\omega)} V(s_0, p_0, x_0, y_0, g_0) dP, \end{aligned}$$

and therefore

$$\int_{A(\omega)} \Big[\phi(s_0, p_0, x_0, y_0 + \epsilon, -(a_p - b_q e^{x_0}) B_0 \epsilon) - \phi(s_0, p_0, x_0, y_0, g_0)\Big] dP \geq 0,$$

since $V(\mathbf{X}) \leq \phi(\mathbf{X})$ for all $\mathbf{X} \in \mathcal{B}(\mathbf{X}_0)$ and $V(\mathbf{X}_0) = \phi(\mathbf{X}_0)$. Therefore,

$$\limsup_{\epsilon \to 0} \left\{ \int_{A(\omega)} \frac{\phi(s_0, p_0, x_0, y_0 + \epsilon, -(a_p - b_q e^{x_0}) p_0 \epsilon) - \phi(s_0, p_0, x_0, y_0, g_0)}{\epsilon} dP \right\} \geq 0,$$

and, by Fatou's lemma,

$$\int_{A(\omega)} \limsup_{\epsilon \to 0} \left\{ \frac{\phi(s_0, p_0, x_0, y_0 + \epsilon, -(a_p - b_q e^{x_0}) B_0 \epsilon) - \phi(s_0, p_0, x_0, y_0, G_0)}{\epsilon} \right\} dP \geq 0,$$

which implies (A6). □

Proof of Theorem 3. Let

$$V^{\delta}(t,p,x,y,g) = \begin{cases} \mathbb{V}^{\delta}(\chi,\mathbb{p},\mathbb{x},y,\mathbb{g}) & \text{if } t \in [\chi,\chi+\delta), y \in [\nu,\nu+\kappa\delta), \\ Z(\mathbb{p},\mathbb{x},y,\mathbb{g}) & \text{if } t = T \end{cases}$$

and

$$\underline{V}(\mathbf{X}) = \lim_{\mathbf{Y}\to\mathbf{X}} \inf_{\delta\to 0}\{\mathbb{V}^{\delta}(\mathbf{Y})\} \quad \text{and} \quad \overline{V}(\mathbf{X}) = \lim_{\mathbf{Y}\to\mathbf{X}} \sup_{\delta\to 0}\{\mathbb{V}^{\delta}(\mathbf{Y})\}, \tag{A7}$$

where $\mathbf{X} = (t,p,x,y,g)$. We will show that $\underline{V}(\mathbf{X})$ and $\overline{V}(\mathbf{X})$ are a viscosity supersolution and a viscosity subsolution of (21), respectively. Combining this with the uniqueness of the viscosity solution of (21) yields $\underline{V}(\mathbf{X}) \geq \overline{V}(\mathbf{X})$ on $[0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$. The opposite inequality is true by the definition of $\underline{V}(\mathbf{X})$ and $\overline{V}(\mathbf{X})$, and therefore

$$\underline{V}(\mathbf{X}) = \overline{V}(\mathbf{X}) = V(\mathbf{X}),$$

which, together with (A7), also implies the local uniform convergence of $\mathbb{V}^{\delta}$ to V.

Note that we only prove that $\underline{V}$ is a viscosity supersolution of (21), as the arguments for $\overline{V}$ is identical. Let $\mathbf{X}_0$ be a local minimum of $\underline{V}-\phi$ on $[0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$, for $\phi \in C^{1,2}([0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R})$. Without loss of generality, we may assume that $\mathbf{X}_0$ is a strict local minimum, that $\underline{V}(\mathbf{X}_0) = \phi(\mathbf{X}_0)$, and that $\phi \leq -2\times\sup_{\delta}\{||\mathbb{V}^{\delta}||_{\infty}\}$ outside the vall $\mathscr{B}(\mathbf{X}_0,R)$, $R>0$, where $\underline{V}(\mathbf{X})-\phi(\mathbf{X}) \geq 0$.

Then, there exist sequences $\delta_n \in \mathbb{R}^+$ and $\mathbf{Y}_n \in [0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}$, such that

$$\delta_n \to 0, \mathbf{Y}_n \to \mathbf{X}_0, \mathbb{V}^{\delta_n}(\mathbf{Y}_n) \to \underline{V}(\mathbf{X}_0), \mathbf{Y}_n \text{ if a global minimum point of } \mathbb{V}_j^{\delta_n} - \phi.$$

Let $h_n = \mathbb{V}^{\delta_n} - \phi$; then

$$h_n \to 0 \text{ and } \mathbb{V}_j^{\delta_n}(\mathbf{X}) \geq \phi(\mathbf{X}) + h_n(\mathbf{X}) \quad \text{for any } \mathbf{X} \in [0,T]\times\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}. \tag{A8}$$

To show that $\underline{V}$ is a viscosity supersolution of (21), it suffices to show that

$$\min\Big\{-\mathcal{B}\phi(\mathbf{X}_0), \mathcal{S}\phi(\mathbf{X}_0), -\mathcal{L}\phi(\mathbf{X}_0)\Big\} \geq 0. \tag{A9}$$

Let $\mathbb{Y}_n = (s_i,\mathbb{p}_n,\mathbb{x}_n,y_n,\mathbb{g}_n)$, where $s_i \in [\chi_i,\chi_i+\delta_n)$ and $y_{\delta_n} \in [\vartheta_n,\vartheta_n+\kappa\delta_n)$. Denote $\mathbb{Y}_n^{(0)} = (\chi_n,\mathbb{p}_n,\mathbb{x}_n,y_n,\mathbb{g}_n)$,

$$\mathbb{Y}_n^{(1)} = \big(\chi_n,\mathbb{p}_n,\mathbb{x}_n,\vartheta_n+\kappa\delta_n,\mathbb{g}_n-(a_p-b_q e^{\mathbb{x}_n})\mathbb{p}_n\kappa\delta_n\big),$$

$$\mathbb{Y}_n^{(2)} = \big(\chi_n,\mathbb{p}_n,\mathbb{x}_n,\vartheta_n-\kappa\delta_n,\mathbb{g}_n+(b_p-a_q e^{\mathbb{x}_n})\mathbb{p}_n\kappa\delta_n\big).$$

Then,

$$\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) = \max\Big\{\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(1)}), \mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(2)}), E\{\mathbb{V}^{\delta_n}(\mathbb{Y}_{n+1}^{(0)})\}\Big\}.$$

Now, we look at the following three cases.

Case 1. It holds that $\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) = \mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(1)})$. Then (A8) implies that

$$\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) \geq \phi(\mathbb{Y}_n^{(1)}) + \mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) - \phi(\mathbb{Y}_n^{(0)}),$$

and therefore

$$0 \geq \liminf_n \left\{ \frac{\phi(\mathbb{Y}_n^{(1)}) - \phi(\mathbb{Y}_n^{(0)})}{\delta_n} \right\} \geq \lim_{\delta \to 0} \inf \left\{ \frac{\phi(\mathbb{Y}_0^{(1)}) - \phi(\mathbb{Y}_0^{(0)})}{\delta} \right\}$$
$$= \frac{\partial \phi(\mathbf{X}_0)}{\partial y} - (a_p - e^{x_0(t)}) p_0(t) \frac{\partial \phi(\mathbf{x}_0)}{\partial g}.$$

Case 2. It holds that $\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) = \mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(2)})$. Arguing similarly to case 1, we get

$$0 \geq -\left(\frac{\partial \phi(\mathbf{X}_0)}{\partial y} - (b_p - a_q e^{x_0(t)}) p_0(t) \frac{\partial \phi(\mathbf{X}_0)}{\partial g} \right).$$

Case 3. It holds that $\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) = E\{\mathbb{V}^{\delta_n}(\mathbb{Y}_{n+1}^{(0)})\}$. Then (A8) implies that

$$\mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) \geq E\{\phi(\mathbb{Y}_{n+1}^{(0)})\} + \mathbb{V}^{\delta_n}(\mathbb{Y}_n^{(0)}) - \phi(\mathbb{Y}_{n+1}^{(0)}),$$

and therefore

$$0 \geq \liminf_n \left\{ \frac{\phi(\mathbb{Y}_{n+1}^{(0)}) - \phi(\mathbb{Y}_n^{(0)})}{\delta_n} \right\} \geq \lim_{\delta \to 0} \inf \left\{ \frac{\phi(\mathbb{Y}_1^{(0)}) - \phi(\mathbb{Y}_0^{(0)})}{\delta} \right\} = \mathcal{L}\phi(\mathbf{X}_0).$$

Combining the results in cases 1–3 yields (A9), and the proof is complete. □

References

Davis, Mark H. A., Vassilios G. Panas, and Thaleia Zariphopoulou. 1993. European option pricing with transaction costs. *SIAM Journal on Control and Optimization* 31: 470–93. [CrossRef]

Ehrman, Douglas S. 2006. *The Handbook of Pairs Trading: Strategies Using Equities, Options, and Futures*. Hoboken: John Wiley and Sons.

Elliott, Robert J., John Van Der Hoek, and William P. Malcolm. 2005. Pairs Trading. *Quantitative Finance* 5: 271–76. [CrossRef]

Gatev, Evan, William N. Goetzmann, and K. Geert Rouwenhorst. 2006. Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies* 19: 797–827. [CrossRef]

Lai, Tze Leung, and Haipeng Xing. 2008. *Statistical Models and Methods for Financial Markets*. New York: Springer.

Lei, Yaoting, and Jing Xu. 2015. Costly arbitrage through pairs trading. *Journal of Economic Dynamics and Control* 56: 1–19. [CrossRef]

Leung, Tim, and Xin Li. 2015. Optimal mean reversion trading with transaction costs and stop-loss exit. *International Journal of Theoretical and Applied Finance* 18: 1550020. [CrossRef]

Mudchanatongsuk, Supakorn, James A. Primbs, and Wilfred Wong. 2008. Optimal pairs trading: A stochastic approach. Paper presented at 2008 American Control Conference, Seattle, WA, USA, June 11–13.

Ngo, Minh-Man, and Huyen Pham. 2016. Optimal switching for the pairs trading rule: A viscosity solutions approach. *Journal of Mathematical Analysis and Applications* 441: 403–25. [CrossRef]

Song, Qingshuo, and Qing Zhang. 2013. An optimal pairs-trading rule. *Automatica* 49: 3007–14. [CrossRef]

Tourin, Agnès, and Raphael Yan. 2013. Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control* 37: 1972–81. [CrossRef]

Vidyamurthy, Ganapathy. 2004. *Pairs Trading—Quantitative Methods and Analysis*. New York: John Wiley and Sons.

Whistler, Mark. 2004. *Trading Pairs—Capturing Profits and Hedging Risk with Statistical Arbitrage Strategies*. New York: John Wiley and Sons.

Xing, Haipeng, Yang Yu, and Tiong Wee Lim. 2017. European option pricing under geometric levy processes with proportional transaction costs. *Journal of Computational Finance* 21: 101–27.

Zhang, Hanqin, and Qing Zhang. 2008. Trading a mean-reverting asset: Buy low and sell high. *Automatica* 44: 1511–18. [CrossRef]

Journal of
Risk and Financial Management

MDPI

Article

A Procedure to Set Prices and Select Inventory in Thinly Traded Markets Using Data from eBay

Xinbo Hu and Paul J. Zak *

Center for Neuroeconomics Studies, Claremont Graduate University, 160 E. 10th St., Claremont, CA 91711-6165, USA; hellen.xbh@gmail.com
* Correspondence: paul@neuroeconomicstudies.org

Abstract: Prices respond to equate supply and demand. However, price-setting in low-volume or "thin" markets is a challenge as is determining which items to carry. We present an algorithm that takes into account a store's fixed costs, the cost of goods sold, prices, and listing duration to determine the portfolio of items to maximize profits. Prices can then be assigned as a mark-up over cost. The usefulness of this approach is demonstrated by applying it to a store on eBay in which the seller needs to meet a profit threshold. The findings identify how sellers of unusual items can effectively determine which items to list and how to set price to reach profit goals.

Keywords: pricing; market volume; portfolio profitability; Poisson model

JEL Classification: D4 (Market Structure and Pricing); L1 (Market Structure; Firm Strategy; and Market Performance)

Citation: Hu, Xinbo, and Paul J. Zak. 2022. A Procedure to Set Prices and Select Inventory in Thinly Traded Markets Using Data from eBay. *Journal of Risk and Financial Management* 15: 297. https://doi.org/10.3390/jrfm15070297

Academic Editors: James W. Kolari and Thanasis Stengos

Received: 15 November 2021
Accepted: 4 July 2022
Published: 5 July 2022

1. Introduction

A thin market is characterized by few buyers and sellers with few transactions (Armstrong 2006). Artwork, antique collectables, and real estate are often traded in thin markets (Knight 2002). In contrast, a thick market has high trading volume in which prices equate supply and demand (Acemoglu 2007). Understanding how to set prices is especially important when trading in markets with unusual items in which few are sold. (Khezr 2015).

Prices in thin markets tend to be volatile because suppliers often guess when choosing what to charge (Coslor 2016). Lacking the robust pricing rules used in thick markets (Hanson 2003), prices in thin markets use algorithms (Tesauro and Kephart 2002; Yu et al. 2011). Most algorithms focus on fixed costs (Deng and Yano 2006) or macroeconomic variables such as GDP, employment rates, interest rates, or price indices (Cirman et al. 2015). Profits can also be volatile in thin markets with sales of very few items determining gain or loss. Sellers of items in thin markets face the challenge of building portfolios of items and setting prices to reach profit goals (Severini 2017). The choice of inventory to accumulate and sell in thin markets typically fails to account for the duration of shelf time and the timing of cash flows that substantially influences small businesses' survival (Elmaghraby and Keskinocak 2003). In addition, the advent of e-commerce has increased the need to understand, and potentially improve, how prices are set in low volume long-tail markets (Kendall and Tsui 2011).

The present paper presents an algorithmic approach that sets prices for a portfolio of thinly traded items subject to a cash flow constraint. The algorithm is applied to data from a seller of unique items on eBay to demonstrate how to implement the methodology.

2. Method

The nonsmoothness of sales in thin markets can be modelled as a Poisson distribution. An event is the sale of a single item in the store. This contrasts with typical sales models in which transactions include multiple items purchased at the same time. The probability of a given number of items selling in a month is given by,

$$\text{Pois}(\lambda) = \frac{\lambda^k e^{-\lambda}}{k!},$$

where k is the number of items sold and λ is the expected frequency of sales. The probability of at least k items being sold ("ProbSold") is 1 minus the cumulative distribution function (CDF) of the distribution.

Sales frequency is affected by the quantity of goods available to sell, the portfolio of goods that vary in their demand, and prices for each type of good. One can solve for the types of goods to list by modifying Guadix et al. (2011) to include a fixed-cost threshold that must be met every month and segmenting the listed items into one of four categories: (1) high profit, sold fast (HF); (2) high profit, sold slow (HS); (3) low profit, sold fast (LF); (4) low profit, sold slow (LS).

3. Results

We obtained data from a vintage auto parts store on eBay to demonstrate the above method. The store owner obtains parts from junkyards, refurbishes them, and lists them for sale. The customers are individual car owners and repair shops. The uniqueness of listed items means that there are few competitors and therefore uncertainty about market-clearing prices. The data are a sample of 319 items sold over three months of approximately 5000 that the seller has listed. The data include part type, shelf duration, listed price, and item cost. These data were used to calculate gross profit (sale price − item cost). The store owner also shared his estimated monthly fixed costs, but was unable to supply the costs of the time spent sourcing parts, cleaning and refurbishing them, researching prices, and packing and mailing items. The algorithm below is easily modified to reflect these additional costs.

3.1. Listed Items and Duration

The average revenue per sale was USD 19.62 (SD = USD 73.73) and average cost of goods sold was USD 8.56 (SD = USD 7.28). This produced average gross profit of USD 111.06 (SD = USD 68.37). The average listing time prior to sale was 14.1 months (SD = 11.89). The sample of 319 items was used to reflect the population of 5000 items used in the analysis.

We parameterized the Poisson distribution using a month as the event period and the event as the number of items sold per month. The parameter λ is the expected frequency of monthly sales. The dataset shows this to be 68.65. The probability of at least k items being sold in a month is $1 - e^{-68.65} \sum_{i=0}^{\lfloor k \rfloor} \frac{68.65^i}{i!}$ as shown in Figure 1.

The value λ = 68.85 for items sold per month maximizes expected gross profit (the product of expected items sold at an average gross profit per item of USD 111.06), earning the owner USD 5798. The value of lambda is scalable with listing quantity, showing that listing more items generates more revenue and profit (Table 1).

Table 1. The maximum expected gross profit increases with increased listings.

Listed Quantity	319	600	900	1200	1500	1800	2100
Lambda (λ)	68.65	129.14	193.70	258.27	322.84	387.41	451.97
Maximum expected gross profit	USD 5798	USD 11,633	USD 18,045	USD 24,559	USD 31,142	USD 37,775	USD 44,444

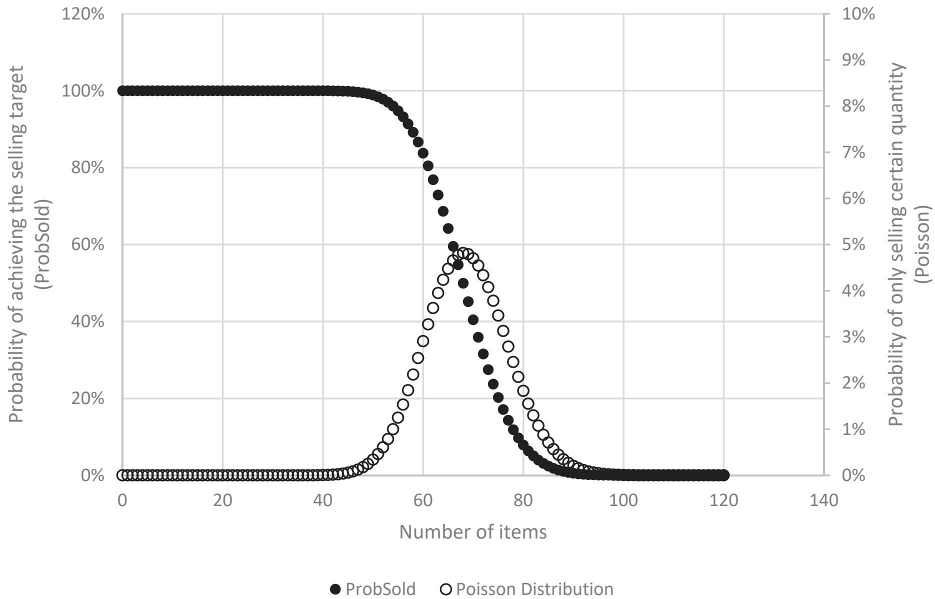

Figure 1. The Poisson distribution for λ = 68.65 showing the probability of selling at least k items in a month.

3.2. Determining the Distribution of Listed Items

The store owner faces a binding revenue constraint that, prior to this analysis, he sought to reach by guessing the types of items to list each month. This experience-based assessment can be complemented by a data-driven algorithm. We used the store's monthly fixed costs, listing price, historical listing time, and cost of each item to develop an algorithm to select items to meet a profit goal. We wrote code that iterates the number of items in each class, extending a previous approach (Guadix et al. 2011). The procedure, written in SAS v 9.4 software, first calculates the cost per item and assigns a portion of the store's fixed costs (USD 2000) to each item in the store. The assigned fixed cost per item is multiplied by the number of months each item remains in the store unsold. Then, the assigned fixed cost per item is compared to the item's gross profit. The item is kept in the portfolio if the gross profit is greater than its cost. This process continues until the portfolio contains only items that generate positive profits. The code converged after 10 iterations, producing a portfolio of 81 items. The flow chart in Figure 2 depicts how the algorithm builds the portfolio. Appendix A shows the iterative process in detail and includes the SAS code.

The optimized portfolio of items varies considerably from the seller's current inventory. High profit items H (=HF + HS) make up 71.60% of the optimized mix compared to 50.47% in the seller's current store. Low profit items L (=LF + LS) are 28.40% of the new mix, compared to 49.53% previously. Nearly all of the high profit items sell fast in the optimized portfolio and there are no low profit slow selling (LS) items. The proportion of each product class in the optimal mix is shown in Table 2.

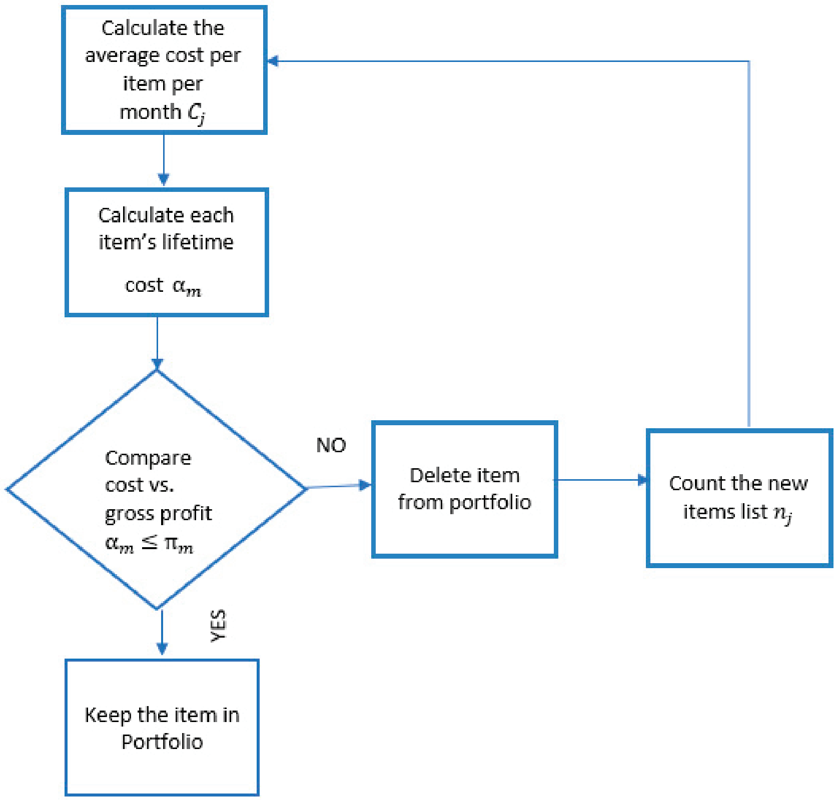

Figure 2. Flow chart of portfolio screening process.

Table 2. The distribution of the optimal portfolio of items. Gross profit and cost are the contribution to overall profit and the cost of goods sold. Weighted profit is the percent of gross profit multiplied by average gross profit (USD 140.57).

Median Split Category	Gross Profit	Item Cost	Weighted Profit (per Item)
High profit, sold fast (HF)	70.37%	50.62%	USD 104.85
High profit, sold slow (HS)	1.23%	1.23%	USD 22.43
Low profit, sold fast (LF)	28.40%	48.15%	USD 5.93
Low profit, sold slow (LS)	0.00%	0.00%	USD 0.00

The optimized portfolio has an average cost of items sold of USD 9.11 (SD = USD 7.41) and an average gross profit per item of USD 140.57 (SD = USD 67.18). The average profit in the optimal portfolio is 26.6% higher than in the seller's present portfolio (t-test, $p = 0.0003$). The new portfolio has an average listing duration of 2.81 months (SD = 1.99) compared to 14.1 months for current listings (Figure 3); this is an 80% reduction in turnover duration ($p = 0.0000$). The total expected monthly profit for the optimized portfolio is USD 4052.02 (USD 140.57*81/2.81) that is 61% higher than original monthly profit of USD 2512.63 (USD 111.06*319/14.1). Further, the algorithm we developed reduced the standard deviation of duration by 83.3%. Every month, approximately 81 new items would be added to the store's inventory (Table 2).

In the optimized portfolio, compared to the original set of items, the duration of parts for cars from the 1990s was significantly lower (p = 0.048) as was the duration of items for cars from Japan (p = 0.0172).

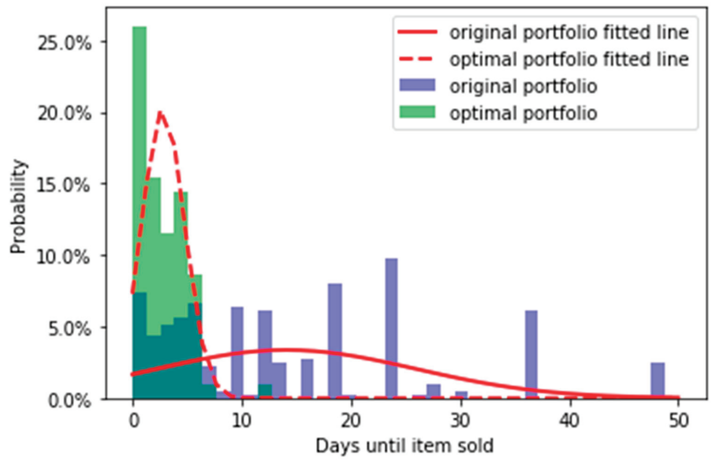

Figure 3. The histogram of duration for original and optimal portfolio.

3.3. Duration and Prices

One way to determine prices is a linear markup over cost. Estimating a linear regression using the original set of listed items (Figure 4), the average markup over cost is 668%. The estimated pricing equation is

$$Item\ Price = 53.877 + 7.6822 \times Item\ Cost.$$

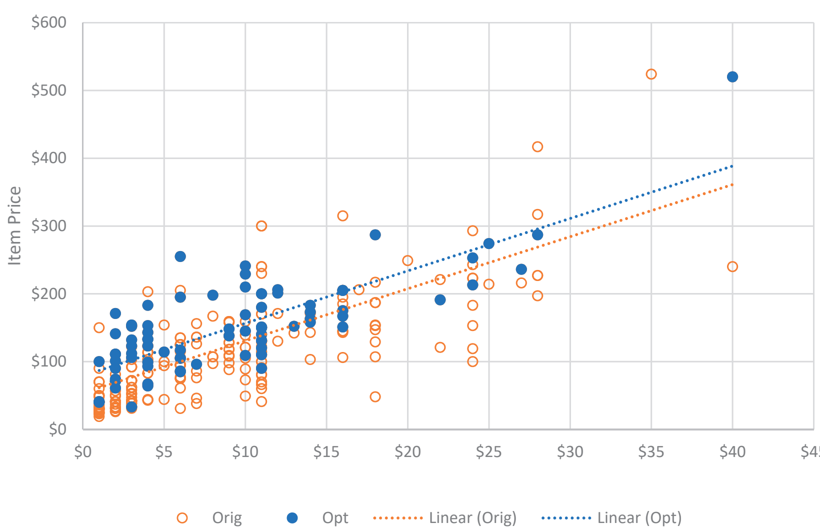

Figure 4. The relationship between cost and price for the original (Orig) and optimal (Opt) portfolio.

In the optimized portfolio, the average markup over cost is 673%. The pricing equation is

$$Item\ Price = 79.23 + 7.7322 \times Item\ Cost.$$

Using the number of items sold per month (69) and the optimized distribution of item types (Table 2), the optimized listings include 49 HF and 20 LF items. Assuming the markup of each group follows the same ratio, the markups are 821% for HF and 335% for LF goods.

The correlation between price and duration in the optimized portfolio was 0.56, substantially higher than the 0.04 correlation in the original portfolio (z = 4.67, p = 0.0000). The higher correlation was due to the selection algorithm.

4. Discussion

The growth of eBay and other online stores has been phenomenal. In 2019, USD 22 billion of goods were sold on eBay from 1.3 billion listings and 182 million users (Lin 2020). In the US, 40% of items sold on eBay were used rather than new. Setting a price of items with high volume is easy—one simply searches eBay to establish the average and range of prices. However, sellers in thin markets have a paucity of information on prices and must also choose a selection of items to source and sell. Our contribution provides online and physical retailers in thin markets with an algorithmic approach to stocking their stores.

The algorithm developed in this paper offers sellers of items in thinly traded markets a methodology to improve the distribution of items sold, while at the same time increasing the likelihood that profit thresholds are met. The algorithm can select the type of items to list by calculating the profitability of item categories and is relatively straightforward to implement (Appendix A). The approach we have developed can easily be generalized to include additional costs, including storage, overhead, administrative, utilities, and taxes.

The optimal portfolio did not sacrifice the seller's high per item markup. Indeed, the markup that maximizes profits was shown to be higher than that currently used. Prices in thin markets are typically only weakly correlated with demand (Sudhir 2001). An additional insight from our approach is that the seller should avoid listing LS (low profit, sold slow) items. The algorithm thus provides a screen that reduces the sourcing, cleaning, and listing of LS items, eliminating an entire class of inventory to manage. The cutoff for LS items can be varied from the median split used here by assessing the profit contribution of such items when they sell. An exception to this rule could be rare or antique items that might sell slowly but could drive visitors to a store as a form of advertising (Barney and Hesterly 2014).

Author Contributions: P.J.Z. conceived of the project, obtained the data, designed the analysis and wrote the paper. X.H. wrote the code, ran simulations, estimated the model and wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data contain personally identifiable information and must be kept confidential.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Iterative procedure to build portfolios of HS, LS, HF, HS items.

1. Let n_j be the number of items kept in the portfolio and let j be the number of iterations (j = 0, 1, 2, ...).
   a. when j = 0, $n_0 = 319$.
2. Let C be the total cost of items based on current cycle's portfolio.
   a. $C_j = \cos t / n_j$

3. Let α be the cost of an item, T be the duration of the listing, m be the index of items kept in the jth cycle. The cost to source an item is assumed to increase in proportion to duration based on data from the original portfolio,
   a. $\alpha_m = C_j * T_m$
4. Compare each items' profit π_m vs. cost α_m such that if $\pi_m > \alpha_m$, keep the item in the portfolio. Otherwise, delete the item from the portfolio.
5. Recalculate n_j
6. Repeat this process until there is no $\pi_m < \alpha_m$ to obtain the final optimal portfolio.

SAS code

```
proc import out=my_Ebay
datafile='C:\SAS\Ebay'
dbms=xlsx replace;
sheet="Orig data"my_Ebay
run;

/*Add column of 1/T=orig d*/
data my_Ebay;
set my_Ebay;
orig_d=1/Dur__Months;
run;
sheet="Orig data";
run;

proc contents data=my_Ebay;
run;

/*Add column of 1/T=orig_d*/
data my_Ebay;
set my_Ebay;
orig_d=1/Dur__Months;
k=_N_;
fact_k=fact(k);
run;

/*calculate the total orig d*/
proc sql;
select sum(orig_d)/count(*) into: final_d
from my_Ebay;
quit;
%put &final_d.;

/*test count*/
proc sql;
select count(*) from my_Ebay;
quit;

/*calculate p*/
data my_Ebay;
set my_Ebay;
p=exp(-&final_d.)*(&final_d.*k)/fact_k;
run;
```

References

Acemoglu, Daron. 2007. Equilibrium bias of technology. *Econometrica* 75: 1371–409. [CrossRef]

Armstrong, Mark. 2006. Competition in two-sided markets. *The RAND Journal of Economics* 37: 668–91. [CrossRef]

Barney, Jay, and William S. Hesterly. 2014. *Strategic Management and Competitive Advantage: Concepts and Cases*, 5th ed. Englewood Cliffs: Prentice Hall, pp. 283–301.

Cirman, Andreja, Marko Pahor, and Miroslav Verbic. 2015. Determinants of time on the market in a thin real estate market. *Inzinerine Ekonomika-Engineering Economics* 26: 4–11. [CrossRef]

Coslor, Erica. 2016. Transparency in an opaque market: Evaluative frictions between "thick" valuation and "thin" price data in the art market. *Accounting, Organizations and Society* 50: 13–26. [CrossRef]

Deng, Shiming, and Candace A. Yano. 2006. Joint production and pricing decisions with setup costs and capacity constraints. *Management Science* 52: 741–56. [CrossRef]

Elmaghraby, Wedad, and Pınar Keskinocak. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. *Management Science* 49: 1287–309. [CrossRef]

Guadix, Jose, Luis Onieva, Jesus Munuzuri, and Pablo Cortés. 2011. An overview of revenue management in service industries: An application to car parks. *The Service Industries Journal* 31: 91–105. [CrossRef]

Hanson, Robin. 2003. Combinatorial information market design. *Information Systems Frontiers* 5: 107–19. [CrossRef]

Kendall, Todd D., and Kevin Tsui. 2011. The economics of the long tail. *The BE Journal of Economic Analysis & Policy* 11: 76. [CrossRef]

Khezr, Peyman. 2015. Time on the market and price change: The case of Sydney housing market. *Applied Economics* 47: 485–98. [CrossRef]

Knight, John R. 2002. Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics* 30: 213–37. [CrossRef]

Lin, Ying. 2020. 10 eBay Statistics You Need to Know in 2020. *Oberlo.com*. Available online: https://www.oberlo.com/blog/ebay-statistics (accessed on 15 October 2021).

Severini, Thomas A. 2017. *Introduction to Statistical Methods for Financial Models*. Boca Raton: Chapman and Hall/CRC.

Sudhir, Karunakaran. 2001. Competitive pricing behavior in the auto market: A structural analysis. *Marketing Science* 20: 42–60. [CrossRef]

Tesauro, Gerald, and Jeffrey O. Kephart. 2002. Pricing in agent economies using multi-agent Q-learning. *Autonomous Agents and Multi-Agent Systems* 5: 289–304. [CrossRef]

Yu, Rongshan, Wenxian Yang, and Susanto Rahardja. 2011. Optimal real-time price based on a statistical demand elasticity model of electricity. Paper presented at the 2011 IEEE First International Workshop on Smart Grid Modeling and Simulation (SGMS), Brussels, Belgium, October 17; pp. 90–95.

Journal of
Risk and Financial Management

Article

Dynamic Conditional Bias-Adjusted Carry Cost Rate Futures Hedge Ratios

Dean Leistikow, Yi Tang * and Wei Zhang

Gabelli School of Business, Fordham University, 140 West 62nd Street, New York, NY 10023, USA; leistikow@fordham.edu (D.L.); wzhang167@fordham.edu (W.Z.)
* Correspondence: ytang@fordham.edu; Tel.: +1-646-312-8292

Abstract: This paper proposes new dynamic conditional futures hedge ratios and compares their hedging performances along with those of common benchmark hedge ratios across three broad asset classes. Three of the hedge ratios are based on the upward-biased carry cost rate hedge ratio, where each is augmented in a different bias-mitigating way. The carry cost rate hedge ratio augmented with the dynamic conditional correlation between spot and futures price changes generally: (1) provides the highest hedging effectiveness and (2) has a statistically significantly higher hedging effectiveness than the other hedge ratios across assets, sub-periods, and rolling window sizes.

Keywords: carry cost rate; hedge ratio; conditional hedge ratio; bias adjustments

JEL Classification: G01; G11; G12; G17

Citation: Leistikow, Dean, Yi Tang, and Wei Zhang. 2022. Dynamic Conditional Bias-Adjusted Carry Cost Rate Futures Hedge Ratios. *Journal of Risk and Financial Management* 15: 12. https://doi.org/10.3390/jrfm15010012

Academic Editor: Shigeyuki Hamori

Received: 23 November 2021
Accepted: 29 December 2021
Published: 3 January 2022

1. Introduction

The goal of futures hedging is to reduce the firm's risk and increase its value. A key consideration for a futures hedger is the ratio of futures assets to short over the number of spot assets long, i.e., the futures hedge ratio. Not surprisingly, a large literature has evolved regarding the best hedge ratio estimation technique, i.e., that which is expected to reduce the hedger's risk the most. Wang et al. (2015) empirically tested many hedge ratio estimation methods and found that they fail to beat the naïve hedge ratio.

This paper examines dynamic conditional futures hedge ratios. It focuses on: (1) the economics-based carry-cost rate hedge ratio introduced in Leistikow et al. (2020) and employed in Leistikow and Chen (2019), (2) a hedge ratio based on the Engle (2002) statistics-based dynamic conditional correlation model, and (3) three bias-adjusted versions of the carry-cost rate hedge ratio, one of which incorporates the Engle (2002) dynamic conditional correlation model. It compares the hedge ratios' hedging performances across three broad asset classes to those of each other and those of two common benchmark hedge ratios. The carry-cost rate hedge ratio when augmented with the dynamic conditional correlation between the spot and futures price changes generally (1) provides the highest hedging effectiveness and (2) has a statistically significantly higher hedging effectiveness than either that of the common benchmark hedge ratios or the other approaches across: assets, sub-periods, and rolling window sizes. Moreover, it explains why the carry-cost rate hedge ratio, when augmented with the dynamic conditional correlation between the spot and futures price changes, performs better than the naïve hedge ratio.

Section 2 of the paper discusses the hedge ratios studied. Section 3 specifies the data employed. Section 4 presents statistics regarding the hedge ratios, where the hedge ratios are calculated using a 1008 trading day (i.e., 4 years) rolling window. Alternative rolling window sizes (of 2 and 6 years) are analyzed and found to yield similar results; these results are available in Appendices that will be provided upon request from the authors. Section 5 shows the hedge ratios' hedging performances and relative hedging performances. Section 6 concludes the paper.

2. Risk-Minimizing Hedge Ratio Estimation

For each unit of the spot asset, the hedge ratio (h) for the one-unit futures contract that minimizes the future period's risk is:

$$h = \frac{\rho \sigma_s}{\sigma_F}, \tag{1}$$

where: σ_s is the standard deviation of the spot price change (ΔS)[1], σ_F is the standard deviation of the futures price change (ΔF) and ρ is the correlation between ΔS and ΔF. In this section, we discuss alternative empirical hedge ratio measures.

2.1. The Traditional Hedge Ratio

The traditional way to estimate the risk-minimizing hedge ratio on day d is to run the ordinary least square regression:

$$\Delta S_d = a + h_t \Delta F_d + \epsilon_d, \tag{2}$$

where the slope coefficient of ΔF, h_t, is the "traditional" hedge ratio. In this study, we estimate Equation (2) on day d using a rolling-window out-of-sample procedure, where ΔS_d and ΔF_d are daily spot and future price changes over the past 1008 trading days (i.e., day d-1007 to day d), respectively.[2] The 1008-trading days, i.e., 4-year, rolling window is updated daily. Moreover, we use the same 1008-trading days rolling window when estimating alternative models (discussed below) for the hedge ratio estimated on d to ensure that they are compared on an equal basis. This "traditional" hedge ratio is discussed in mainstream textbooks, e.g., Hull (2018) and has traditionally been used as the benchmark for alternative hedge ratios that have been proposed over the last 40 years.

2.2. The "Dynamic Conditional" Hedge Ratio

Equation (2) is often considered a static model in that the historical data are assumed to be equally informative about future realizations. In this subsection, we relax this assumption by estimating an autoregressive-based model that allows more recent observations to be weighted more heavily. Our rationale is that if the dynamics of the underlying variable are indeed time-varying, an appropriately chosen dynamic conditional model is likely to outperform the static model. Engle (2002) proposes a new class of multivariate models called dynamic conditional correlation (DCC) models that are used to predict the time-varying correlations between two financial assets. These models have the flexibility of univariate GARCH models coupled with parsimonious parametric models for the correlations.[3] Following Engle (2002), we estimate the conditional covariance between ΔS and ΔF:

$$\Delta S_d = \alpha_0^{\Delta S} + \varepsilon_{\Delta S,d}, \tag{3}$$

$$\Delta F_d = \alpha_0^{\Delta F} + \varepsilon_{\Delta F,d}, \tag{4}$$

$$E_d\left[\varepsilon^2_{\Delta S,d+1}\right] \equiv \sigma^2_{\Delta S,d+1} = \beta_0^{\Delta S} + \beta_1^{\Delta S}\varepsilon^2_{\Delta S,d} + \beta_2^{\Delta S}\sigma^2_{\Delta S,d}, \tag{5}$$

$$E_d\left[\varepsilon^2_{\Delta F,d+1}\right] \equiv \sigma^2_{\Delta F,d+1} = \beta_0^{\Delta F} + \beta_1^{\Delta F}\varepsilon^2_{\Delta F,d} + \beta_2^{\Delta F}\sigma^2_{\Delta F,d}, \tag{6}$$

$$E_d[\varepsilon_{\Delta S,d+1}\varepsilon_{\Delta F,d+1}] \equiv \sigma_{\Delta S\Delta F,d+1} = \rho_{\Delta S\Delta F,d+1} \cdot \sigma_{\Delta S,d+1} \cdot \sigma_{\Delta F,d+1}, \tag{7}$$

$$\rho_{\Delta S\Delta F,d+1} = \frac{q_{\Delta S\Delta F,d+1}}{\sqrt{q_{\Delta S\Delta S,d+1} \cdot q_{\Delta F\Delta F,d+1}}}, \tag{8}$$

$$q_{\Delta S\Delta F,d+1} = \overline{\rho}_{\Delta S\Delta F} + a_1 \cdot (\varepsilon_{\Delta S,d} \cdot \varepsilon_{\Delta F,d} - \overline{\rho}_{\Delta S\Delta F}) + a_2 \cdot (q_{\Delta S\Delta F,d} - \overline{\rho}_{\Delta S\Delta F}) \tag{9}$$

In the above system of equations, ΔS_d and ΔF_d denote the daily spot and futures price changes over the past 1008 trading days up to day d, respectively, and $E_d[.]$ denotes the expectation operator conditional on day d information. $\sigma^2_{\Delta S,d+1}$ is the day-d expected conditional variance of daily spot price changes (ΔS), $\sigma^2_{\Delta F,d+1}$ is the day-d expected conditional variance of daily futures price changes (ΔF), and $\sigma^2_{\Delta S\Delta F,d+1}$ is the day-d expected conditional covariance between ΔS and ΔF. $\rho_{\Delta S\Delta F,d+1} = q_{\Delta S\Delta F,d+1}/\sqrt{q_{\Delta S\Delta S,d+1} \cdot q_{\Delta F\Delta F,d+1}}$ is the day-d expected conditional correlation between ΔS and ΔF; $\overline{\rho}_{\Delta S\Delta F}$ is the unconditional correlation between ΔS and ΔF. To ease the parameter convergence, we follow Bali and Engle (2010) and Bali et al. (2017) and use correlation targeting, assuming that the time-varying correlation mean reverts to the sample correlation, $\overline{\rho}_{\Delta S\Delta F}$. For each day d + 1, the DCC hedge ratio is defined as the ratio of Equations (6) and (7): $h_{dcc} = \sigma_{\Delta S\Delta F,d+1}/\sigma^2_{\Delta F,d+1}$. Therefore, h_{dcc} is purely out-of-sample, estimated using the information available up to the formation of the hedged portfolio (i.e., day d).

2.3. The Simple Carry-Cost-Rate Hedge Ratio

Unfortunately, the traditional hedge ratio ignores the economic connection between the spot and futures prices. Leistikow et al. (2020) derived a hedge ratio that incorporates the carry-cost hypothesis that links spot and futures prices as follows:

$$S(1+c)^T = F, \tag{10}$$

where c is the annualized spot asset's carry-cost rate (hereinafter CCR) to the futures' maturity,[4] T is the years to the futures contract's maturity, and S and F are the asset's contemporaneous spot and futures prices, respectively.

From the carry-cost hypothesis, it follows that $\frac{\partial S}{\partial F} = (1+c)^{-T}$; thus, $\frac{\sigma_{\Delta S}}{\sigma_{\Delta F}} = (1+c)^{-T}$. Assuming ΔS and ΔF are perfectly correlated, the day d instantaneous hedge ratio (h_{ccr}) is derived as[5]:

$$h_{ccr} = (1+c)^{-T} \tag{11}$$

Therefore, h_{ccr} is calculated quickly and simply, since it only requires knowledge of c and T on day d.

2.4. The Augmented Carry-Cost-Rate Hedge Ratios

The simple CCR-based hedge ratio discussed in the previous subsection is easier to implement and has a firmer economic foundation than the statistics-based approaches discussed earlier. However, h_{ccr} is biased upward because the correlation between ΔS and ΔF is almost certainly less than its assumed level of one.[6] In this subsection, we introduce several ways to augment h_{ccr} to correct its upward bias.

First, we augment h_{ccr} with the dynamic conditional correlation between ΔS and ΔF. It was estimated as part of the dynamic conditional hedge ratio discussed earlier in Section 2.2. We calculate this hedge ratio, denoted h_{ccr_dcc}, on day d as:

$$h_{ccr\_dcc} = h_{ccr} \cdot \rho_{dcc}, \tag{12}$$

where ρ_{dcc} comes from Equation (8) based on daily data over the period from d-1007 to day d.

Second, we augmented h_{ccr} with the correlation (ρ) between daily ΔS and ΔF over the period from d-1007 to day d. We calculated this hedge ratio, denoted h_{ccr_corr}, on day d as:

$$h_{ccr\_corr} = h_{ccr} \cdot \rho \tag{13}$$

Finally, we use the Leistikow et al. (2020) BAM bias-adjusted h_{ccr} on day d, denoted h_{ccr_bam}:

$$h_{ccr\_bam} = h_{ccr} \cdot BAM \tag{14}$$

where BAM is the bias-adjustment multiplier, defined as the average ratio of the average daily h_t and average h_{ccr} for each futures contract employed in the hedges over the 1008 trading days between day d-1007 to day d. In this way, each of the three bias-adjustment factors:

ρ_{dcc}, ρ, and BAM, proxy h_{ccr}'s bias based on the same amount of data, 1008 trading days for now.

2.5. The "Naive" Hedge Ratio

Despite the widespread acceptance of h_t and the development of more complex statistical variants of it, Wang et al. (2015) find that h_t and its more complex statistical variants do not reduce risk better than does the "naive" hedge ratio of 1 (h_1). Therefore, we use h_1 as well as the regression-based hedge ratio (h_t) as benchmarks.

3. Data

We study the S&P500, Japanese Yen (JY), and gold since they represent a broad cross section of asset classes, their carry-cost rates are known and objectively calculable, and their spot and futures markets trade actively. All data are downloaded via a Bloomberg terminal, which are updated daily during our sample period.

CME e-mini S&P500 futures contracts mature on the 3rd Friday of their maturity month and have been the most liquid equity futures since about 2000 when they overtook the regular S&P500 futures. We study their March, June, September, and December contract maturities because they are the most liquid contract maturities and are very liquid (except when they are far from and very near to their maturities). These futures contract maturity months are alternatingly used as the hedging instrument where the hedging instrument is the near maturity futures contract until it is one week to its maturity at which point it switches to the second near maturity futures contract. The JY futures contracts mature two business days before the 3rd Wednesday in the maturity month. For liquidity reasons, the near futures maturity contract is used as the hedging instrument for the IMM Japanese Yen, until the 1st Friday of the maturity month. For COMEX gold, we use the liquid February, April, June, August, October, and December futures contract maturities and roll the hedging instrument from the near to the second near futures maturity two weeks earlier than for the JY due to gold's earlier liquidity shift from the near to the second near futures due to gold futures' delivery options. The gold futures mature the 3rd last business day of the maturity month.

The S&P500 data begins with the Mar '98, while the JY data begins with the Mar '97, and the gold data begins with the Feb '91 futures contract as the hedging instrument. The Jun '20 futures is the last futures used as the hedging instrument for all the assets.

The carry-cost rate is the US risk-free short-term interest rate[7] minus the S&P500 dividend yield, the excess of the US over the JY risk-free short-term interest rate, and the US risk-free short-term interest rate, for the S&P500, JY and gold, respectively.

4. Statistics for the Hedge Ratios and Their Differences from the Benchmarks

In this section, we provide statistics for the various daily profit hedge ratios (and their differences from the benchmark hedge ratios) discussed in Section 2. After the first 4 years (given the 4 years rolling window), for each subsequent S&P500 and JY futures contract an average hedge ratio is calculated based on the 252/4 trading days that the contract is the hedging instrument since their contracts follow a quarterly rotation. The corresponding gold average hedge ratio is calculated based on the 252/6 trading days, since gold futures contracts follow a bimonthly rotation. These futures' average hedge ratios represent the underlying data for the Table 1 statistics. Table 1A reports the descriptive statistics for the average hedge ratios for the S&P500 in Panel A, JY in Panel B, and gold in Panel C.[8] In each panel, the first column presents the results of the traditional regression hedge ratio (h_t), which serves as Table 1A's benchmark hedge ratio. Columns 2–6 represent the descriptive statistics for the differences between h_t and the: dynamic conditional hedge ratio (h_{dcc}), the simple carry-cost-rate hedge ratio (h_{ccr}), and, finally, the carry-cost-rate hedge ratio augmented with (i) the dynamic conditional correlation estimated from the DCC model (h_{ccr_dcc}), (ii) the simple correlation between the daily spot and futures price changes (h_{ccr_corr}), and (iii) the bias-adjustment multiplier (h_{ccr_bam}), respectively.

Column (1) of Table 1A shows that the mean traditional hedge ratios (h_t) for the S&P500, JY, and gold are between 0.8975 and 0.9568. Their standard deviations are very small, ranging from 0.0151 to 0.0469. Thus, as expected, their daily hedge ratios are highly persistent since they are estimated using 1008 trading-days rolling windows.

Table 1. The Benchmark Hedge Ratios and Differences from them.

A. The Benchmark Hedge Ratio and Differences from it When h_t is the Benchmark.						
Panel A. S&P500						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9568	0.0073	0.0441	0.0133	0.0153	−0.0012
std	0.0151	0.0323	0.0156	0.0164	0.014	0.0144
count	74	74	74	74	74	58
t-stat		1.9382	24.3608	6.9605	9.3919	−0.614
signif lev		10%	1%	1%	1%	
Panel B. JY						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9554	−0.01	0.0423	0	−0.0065	−0.0036
std	0.018	0.0389	0.0177	0.0232	0.0208	0.0162
count	78	78	78	78	78	62
t-stat		−2.2673	21.1487	0.0174	−2.757	−1.7318
signif lev		5%	1%		1%	10%
Panel C. Gold						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.8975	0.0112	0.0982	0.0233	0.0098	0.0058
std	0.0469	0.0582	0.0483	0.0387	0.0189	0.0287
count	153	153	153	153	153	129
t-stat		2.3805	25.1193	7.4559	6.3829	2.3166
signif lev		5%	1%	1%	1%	5%
Panel D. HR Differences Aggregated across Assets						
		$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean		0.0002	0.0715	0.0123	0.0084	0.0015
std		0.0429	0.0451	0.0238	0.0215	0.0156
count		222	222	222	222	174
t-stat		0.0775	23.619	7.6897	5.8467	1.2684
signif lev			1%	1%	1%	

The count is the number of futures contracts used as hedge instruments, or, alternatively, the number of hedges since each hedge has a single hedging instrument. For each futures, the "average hedge ratio" is determined from the ≈252/4 dynamic daily hedge ratios calculated for a quarterly maturing futures contract (e.g., as for the S&P500 and JY, whereas for gold there are ≈252/6 dynamic daily hedge ratios since the futures maturities are bimonthly). Since the BAM is calculated from 4 years of futures contracts, the h_{ccr_bam} count is 4 × 4 (6 × 4) for the quarterly futures contract maturities of the S&P500 and JY (bimonthly maturities of gold).

The "aggregated across assets" count is equal-weighted. It uses the most recent differences for each asset for roughly the same period as the e−mini S&P500 (since its data series is the shortest). Since the gold futures data started earlier and there are 50% more gold futures maturities/years, a little less than the 1st half of the gold futures contract results do not enter into the aggregation across assets.

Table 1. *Cont.*

B. The Benchmark Hedge Ratio and Differences from it When h_1 is the Benchmark						
Panel A. S&P500						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	1	−0.0359	0.0009	−0.0299	−0.0279	−0.0396
std	0	0.0321	0.0024	0.0058	0.0054	0.0106
count	74	74	74	74	74	58
t-stat		−9.6302	3.3772	−44.7285	−44.1444	−28.4181
signif lev		1%	1%	1%	1%	1%
Panel B. JY						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	1	−0.0546	−0.0023	−0.0446	−0.0511	−0.0431
std	0	0.0391	0.0025	0.0156	0.0132	0.0123
count	78	78	78	78	78	62
t-stat		−12.3317	−8.2224	−25.2679	−34.2633	−27.5771
signif lev		1%	1%	1%	1%	1%
Panel C. Gold						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	1	−0.0913	−0.0043	−0.0791	−0.0927	−0.0913
std	0	0.0671	0.0039	0.0385	0.0465	0.0362
count	153	153	153	153	153	129
t-stat		−16.8137	−13.8224	−25.4034	−24.6593	−28.6446
signif lev		1%	1%	1%	1%	1%
Panel D. HR Differences Aggregated across Assets						
		$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean		−0.0721	−0.0008	−0.06	−0.0639	−0.0694
std		0.059	0.0024	0.0368	0.039	0.0422
count		222	222	222	222	174
t-stat		−18.2127	−4.9113	−24.2818	−24.4085	−21.6855
signif lev		1%	1%	1%	1%	1%

The count is the number of futures contracts used as hedge instruments, or, alternatively, the number of hedges since each hedge has a single hedging instrument. For each futures, the "average hedge ratio" is determined from the ≈252/4 dynamic daily hedge ratios calculated for a quarterly maturing futures contract (e.g., as for the S&P500 and JY, whereas for gold there are ≈252/6 dynamic daily hedge ratios since the futures maturities are bimonthly). Since the BAM is calculated from 4 years of futures contracts, the h_{ccr_bam} count is 4 × 4 (6 × 4) for the quarterly futures contract maturities of the S&P500 and JY (bimonthly maturities of gold).

The "aggregated across assets" count is equal-weighted. It uses the most recent differences for each asset for roughly the same period as the e−mini S&P500 (since its data series is the shortest). Since the gold futures data started earlier and there are 50% more gold futures maturities/years, a little less than the 1st half of the gold futures contract results do not enter into the aggregation across assets.

Columns 2–6 of Table 1A show that the mean differences between the other hedge ratios and h_t are small (<0.0233 in absolute value across the assets); except for the upward-biased h_{ccr}; its maximum mean difference is 0.0982 (for gold). Therefore, the bias adjustments to the CCR hedge ratio largely mitigate the simple CCR hedge ratio's upward bias. Similar to h_{ccr}, both h_{ccr_dcc}, and h_{ccr_corr} are larger than h_t and the difference is significant at the 1% level; however, this does not hold for the JY. The paper's statistical significance tests are two-tailed tests, except for the one-tailed $h_{ccr} - h_t$ test in that h_{ccr} is upward biased relative to h_t.

Table 1A, Panel D has equal-weighted, roughly contemporaneous, aggregated across asset results that end with the assets' June 2020 futures contracts. To generate these results, since the S&P500 data/results start later than the other assets' (starting with the Mar 2002 futures), all of its results are included but only the contemporaneous JY (starting with the Mar 2002 futures) and roughly contemporaneous gold results are included in the aggregation (starting with the April 2008 futures). The gold results included in the equal-weighted aggregation begin later, since gold has 50% more futures maturities/years than the others and therefore gets the same number of futures contracts in 2/3 the time. The aggregated across assets mean $h_{ccr} - h_t$ was above 0.07, while the means for the other differences were not more than 0.0123 in absolute value. Hedge ratios h_{ccr}, h_{ccr_dcc}, and h_{ccr_corr} are all larger than h_t at the 1% significance level.

Table 1B below is the same as Table 1A, except that Table 1B's benchmark hedge ratio is h_1. The other hedge ratios' means, except for the S&P500 h_{ccr}, are significantly less than 1 at the 1% confidence level for all assets. For the Panel D, aggregations across assets, while h_{ccr}'s mean is barely below 1, the other hedge ratios' means are between 0.06 and 0.0721 below 1. All of the hedge ratio means are below h_1 at the 1% significance level. The fact that, for this study period and these assets, the upward biased h_{ccr} is significantly below 1 portends poor h_1 hedging performance; this conjecture is strongly supported by the results in the next section.

5. The Benchmark Hedge Ratios' Hedging Performances and Hedging Performance Differences from Those of the Benchmark Hedge Ratios

Next, we calculated each hedge ratio's out-of-sample hedge effectiveness (HE) and compared it to those of the two benchmark hedge ratios. The HE is defined as the percentage profit variance reduction for each contract over our sample period:

$$HE = 1 - \frac{Var(\text{hedged profits})}{Var(\text{unhedged profits})} \tag{15}$$

$$\text{Hedged profit} = \Delta S - h \cdot \Delta F \tag{16}$$

$$\text{Unhedged profit} = \Delta S \tag{17}$$

where $Var(.)$ denotes variance, ΔS and ΔF are, respectively, the daily spot and futures price changes on day d + 1, and h denotes a hedge ratio on day d calculated from one of the hedge ratios described in Section 2. Therefore, the HE of a hedge ratio increases with the risk reduced.

Table 2A reports the hedge effectiveness results, where h_t's HE is the benchmark. Column 1 is h_t's HE while the remaining columns represent the hedge effectiveness differences in between those for the various hedge ratios and for h_t.

Table 2A, column (1) shows that h_t's average HE for the S&P500, JY, and gold are between 0.8464 and 0.9388. Their corresponding standard deviations are small, ranging from 0.0378 to 0.1477. These results indicate that h_t's out-of-sample HE is highly persistent.

The HE for h_t is higher than that for either h_{dcc} or h_{ccr}. The latter result suggests that the benefit of h_{ccr} through recognizing the economic link between the spot and the future prices is less than the cost of its upward bias. The finding that h_{dcc} does not outperform h_t is consistent with previous studies that show that complicated time-series hedge ratios do not yield superior HE performance.[9] Therefore, modeling the dynamic conditional hedge ratio based on spot and futures price changes alone is insufficient to improve hedge effectiveness.

Columns 4–6 of Table 2A compare the augmented h_{ccr}s' HEs with $h_{t's}$ HE where the h_{ccr} is augmented with the conditional correlation (h_{ccr_dcc}), the unconditional correlation (h_{ccr_corr}), and the bias-adjustment multiplier (h_{ccr_bam}), respectively. Column (4) shows that the HE for h_{ccr_dcc} is higher than that for h_t for each asset, where the difference in HE is generally statistically significant at the 5% level. When the HE differences are aggregated across assets, the HE for h_{ccr_dcc} is statistically significantly higher than that for h_t at the 1% significant level. Columns (5) and (6) show that the HEs for h_{ccr_corr} and h_{ccr_bam} are not typically significantly different from that for h_t. Therefore, the h_{ccr_dcc} result suggests that incorporating the economic relation between the spot and futures prices and properly modeling their time-varying conditional correlations jointly produce the best hedging performance.

Table 2. The Benchmark HRs' Hedge Effectiveness and Differences from them.

A. The Benchmark HR's Hedge Effectiveness and Differences from it When h_t is the Benchmark						
Panel A. S&P500						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9388	−0.0014	−0.0011	0.0035	0.0003	0.0001
std	0.0378	0.0215	0.005	0.0137	0.0022	0.002
count	74	74	74	74	74	58
t-stat		−0.5447	−1.9024	2.1813	1.1068	0.2169
signif lev			10%	5%		
Panel B. JY						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9107	−0.0023	−0.0018	0.0003	0.0008	0.0001
std	0.0479	0.0098	0.0051	0.0028	0.003	0.0022
count	78	78	78	78	78	62
t-stat		−2.0244	−3.1242	1.0111	2.2843	0.5062
signif lev		5%	1%		5%	
Panel C. Gold						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.8464	−0.0184	−0.0115	0.0055	−0.0006	−0.0006
std	0.1477	0.1955	0.026	0.0259	0.0046	0.0071
count	153	153	153	153	153	129
t-stat		−1.1639	−5.4726	2.6309	−1.5674	−0.9196
signif lev			1%	1%		
Panel D. HE Differences Aggregated across Assets						
		$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean		−0.0023	−0.0085	0.0032	0	0.0001
std		0.0189	0.0197	0.0124	0.0041	0.0022
count		222	222	222	222	174
t-stat		−1.8015	−6.4151	3.8774	−0.1053	0.4399
signif lev		10%	1%	1%		
B. The Benchmark HR's Hedge Effectiveness and Differences from it When h_1 is the Benchmark						
Panel A. S&P500						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.9378	−0.0004	−0.0002	0.0044	0.0012	0.0015
std	0.0406	0.0228	0.0005	0.0167	0.003	0.0047
count	74	74	74	74	74	58
t-stat		−0.1566	−2.6257	2.2837	3.5766	2.3963
signif lev			5%	5%	1%	5%

Table 2. *Cont.*

Panel B. JY						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.9087	−0.0003	0.0002	0.0023	0.0028	0.0019
std	0.0492	0.0126	0.0004	0.0057	0.0071	0.0055
count	78	78	78	78	78	62
t-stat		−0.1944	4.2624	3.554	3.4394	2.7357
signif lev			1%	1%	1%	1%
Panel C. Gold						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.8343	−0.0063	0.0006	0.0176	0.0115	0.014
std	0.1629	0.1958	0.0011	0.0373	0.0251	0.0219
count	153	153	153	153	153	129
t-stat		−0.3963	7.3019	5.8403	5.692	7.2647
signif lev			1%	1%	1%	1%
Panel D. HE Differences Aggregated across Assets						
		$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean		0.0063	0.0001	0.0118	0.0086	0.0096
std		0.0296	0.0005	0.0249	0.0184	0.0185
count		222	222	222	222	174
t-stat		3.1852	4.0726	7.0864	6.9614	6.8561
signif lev		1%	1%	1%	1%	1%

Table 2B reports the same results; however, h_1's HE is the benchmark. The HE for h_1 is statistically significantly lower (generally at the 1% level) than that for all the other hedge ratios except for h_{dcc} and, for the S&P500, h_{ccr}. Our finding differs from that of Wang et al. (2015) and neither should be interpreted as a general result. The low HE for h_1 was anticipated earlier when we noted that the upward biased h_{ccr} was significantly less than 1. Our low HE for h_1 is not a general result since the carry-cost rate varies across assets, currency denominations, and time; thus, in the very improbable case that the upward biased h_{ccr} exceeded 1 by its bias, the risk minimizing hedge ratio would be h_1.

Table 3 repeats the HE analyses, except that it analyzes the first and second halves of our sample period separately to see if the HE and HE difference results are stable across sub-periods. The only HE difference from the benchmark's HE that is nearly always statistically significantly positive across both benchmarks, both halves, and all 3 assets is that for h_{ccr_dcc} and the JY in the 1st half when h_t is the benchmark is the main exception. The HEs for the other two augmented h_{ccr}s are statistically significantly higher than those for h_1 for 8 of the 12 individual asset results. However, their HEs are not generally higher for those of h_t. Table 3, Panel D, shows for the aggregations across assets that the HE for h_{ccr_dcc} is statistically significantly higher than that for either benchmark in both sub-periods at the 1% level. The HEs for the other two augmented h_{ccr}s are also statistically significantly higher than those for h_1 in both halves at the 1% level.

Table 3. Stability of the Benchmark HRs' HEs and Differences from them.

Panel A. S&P500						
1st half: HE and HE differences when the benchmark is h_t						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9427	−0.0007	0.0003	0.0013	0.0008	0.0001
std	0.0248	0.0045	0.0036	0.0028	0.0017	0.0022
count	37	37	37	37	37	21
t-stat		−0.9852	0.493	2.8216	3.0666	0.2467
signif lev				1%	1%	
2nd half: HE and HE differences when the benchmark is h_t						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9348	−0.002	−0.0025	0.0057	−0.0003	0
std	0.0475	0.0303	0.0058	0.019	0.0026	0.0019
count	37	37	37	37	37	37
t-stat		−0.3998	−2.6145	1.8057	−0.6361	0.0712
signif lev			5%	10%		
1st half: HE and HE differences when the benchmark is h_1						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.9429	−0.001	0.0001	0.0011	0.0006	0.0003
std	0.0256	0.0042	0.0003	0.0039	0.0026	0.0037
count	37	37	37	37	37	21
t-stat		−1.4123	1.0029	1.6317	1.4325	0.3594
signif lev						
2nd half: HE and HE differences when the benchmark is h_1						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.9327	0.0001	−0.0004	0.0078	0.0019	0.0022
std	0.0513	0.0322	0.0005	0.0229	0.0032	0.0051
count	37	37	37	37	37	37
t-stat		0.0259	−3.9637	2.0694	3.5117	2.5698
signif lev			1%	5%	1%	5%
Panel B. JY						
1st half: HE and HE differences when the benchmark is h_t						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9072	−0.0019	−0.002	0	0.0008	−0.0002
std	0.0553	0.0117	0.0059	0.0034	0.0036	0.0013
count	39	39	39	39	39	23
t-stat		−0.9886	−2.1761	−0.065	1.3773	−0.8266
signif lev			5%			

Table 3. *Cont.*

2nd half: HE and HE differences when the benchmark is h_t						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9142	−0.0026	−0.0015	0.0007	0.0008	0.0004
std	0.0397	0.0076	0.0041	0.002	0.0023	0.0026
count	39	39	39	39	39	39
t-stat		−2.1798	−2.3064	2.0882	2.0626	0.8715
signif lev		5%	5%	5%	5%	
1st half: HE and HE differences when the benchmark is h_1						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.9049	0.0004	0.0002	0.0022	0.0031	0.0017
std	0.0574	0.0148	0.0005	0.0067	0.0085	0.0066
count	39	39	39	39	39	23
t-stat		0.1703	2.804	2.0646	2.2411	1.248
signif lev			1%	5%	5%	
2nd half: HE and HE differences when the benchmark is h_1						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.9126	−0.001	0.0002	0.0024	0.0024	0.002
std	0.0398	0.01	0.0003	0.0046	0.0053	0.0049
count	39	39	39	39	39	39
t-stat		−0.601	3.7447	3.2522	2.8669	2.6087
signif lev			1%	1%	1%	5%
Panel C. Gold						
1st half: HE and HE differences when the benchmark is h_t						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.9013	−0.0333	−0.0012	0.0053	−0.0001	−0.0007
std	0.155	0.2766	0.018	0.0334	0.0029	0.0051
count	76	76	76	76	76	52
t-stat		−1.051	−0.5814	1.3761	−0.2459	−0.9525
signif lev						
2nd half: HE and HE differences when the benchmark is h_t						
	h_t	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$
mean	0.7922	−0.0036	−0.0216	0.0058	−0.0011	−0.0005
std	0.118	0.0226	0.0286	0.0156	0.0057	0.0082
count	77	77	77	77	77	77
t-stat		−1.4157	−6.6423	3.2315	−1.6364	−0.5428
signif lev			1%	1%		
1st half: HE and HE differences when the benchmark is h_1						
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.8993	−0.0313	0.0009	0.0073	0.002	0.0028
std	0.1648	0.2734	0.0014	0.0385	0.0198	0.0108

Table 3. *Cont.*

			1st half: HE and HE differences when the benchmark is h_1			
count	76	76	76	76	76	52
t-stat		−0.9975	5.4563	1.6599	0.8745	1.8718
signif lev			1%			10%
			2nd half: HE and HE differences when the benchmark is h_1			
	h_1	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$
mean	0.7701	0.0184	0.0004	0.0278	0.021	0.0215
std	0.1337	0.0411	0.0006	0.0334	0.0263	0.0242
count	77	77	77	77	77	77
t-stat		3.9315	6.0318	7.3037	6.9998	7.8227
signif lev		1%	1%	1%	1%	1%
			Panel D. HE Differences Aggregated across Assets:			
			1st half: HE differences when the benchmark is h_t			
	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$	
Mean	−0.0045	−0.0066	0.0011	0.0005	0.0002	
std	0.016	0.0189	0.0062	0.0036	0.002	
count	111	111	111	111	63	
t-stat	−2.9776	−3.7015	1.9475	1.3209	0.7654	
signif lev	1%	1%	10%			
			2nd half: HE differences when the benchmark is h_t			
	$h_{dcc} - h_t$	$h_{ccr} - h_t$	$h_{ccr\_dcc} - h_t$	$h_{ccr\_corr} - h_t$	$h_{ccr\_bam} - h_t$	
mean	−0.0001	−0.0103	0.0053	−0.0005	0	
std	0.0213	0.0204	0.0162	0.0044	0.0024	
count	111	111	111	111	111	
t-stat	−0.028	−5.3351	3.4564	−1.2058	0.0331	
signif lev		1%	1%			
			1st half: HE differences when the benchmark is h_1			
	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$	
mean	0.0023	0.0001	0.0079	0.0072	0.0082	
std	0.0255	0.0003	0.0198	0.0193	0.0165	
count	111	111	111	111	63	
t-stat	0.931	4.322	4.2188	3.9366	3.9247	
signif lev		1%	1%	1%	1%	
			2nd half: HE differences when the benchmark is h_1			
	$h_{dcc} - h_1$	$h_{ccr} - h_1$	$h_{ccr\_dcc} - h_1$	$h_{ccr\_corr} - h_1$	$h_{ccr\_bam} - h_1$	
mean	0.0104	0.0001	0.0158	0.01	0.0105	
std	0.0329	0.0006	0.0287	0.0174	0.0196	
count	111	111	111	111	111	
t-stat	3.3385	2.4052	5.7886	6.0397	5.6234	
signif lev	1%	5%	1%	1%	1%	

Table 4 directly tests the HE differences across the various bias-adjusted h_{ccr} approaches. The HE is statistically significantly higher for h_{ccr_dcc} than it is for the other bias-adjustment approaches for the individual assets (other than for the JY) and for the aggregation across assets.

Table 4. Hedge Effectiveness Differences across h_{ccr} Bias-Adjustment Methods.

	Panel A. S&P500		
	$h_{ccr\_dcc} - h_{ccr\_corr}$	$h_{ccr\_dcc} - h_{ccr\_bam}$	$h_{ccr\_corr} - h_{ccr\_bam}$
mean	0.0032	0.0038	−0.0001
std	0.0152	0.0164	0.0018
count	74	58	58
t-stat	1.8087	1.7788	−0.4114
signif lev	10%	10%	
	Panel B. JY		
	$h_{ccr\_dcc} - h_{ccr\_corr}$	$h_{ccr\_dcc} - h_{ccr\_bam}$	$h_{ccr\_corr} - h_{ccr\_bam}$
mean	−0.0005	0	0.0008
std	0.0037	0.0028	0.0035
count	78	62	62
t-stat	−1.1004	0.1176	1.6953
signif lev			10%
	Panel C. Gold		
	$h_{ccr\_dcc} - h_{ccr\_corr}$	$h_{ccr\_dcc} - h_{ccr\_bam}$	$h_{ccr\_corr} - h_{ccr\_bam}$
mean	0.0061	0.0064	0.0001
std	0.0261	0.0254	0.0087
count	153	129	129
t-stat	2.8874	2.8625	0.162
signif lev	1%	1%	
	Panel D. HE Differences Aggregated across Assets		
	$h_{ccr\_dcc} - h_{ccr\_corr}$	$h_{ccr\_dcc} - h_{ccr\_bam}$	$h_{ccr\_corr} - h_{ccr\_bam}$
mean	0.0033	0.0036	−0.0004
std	0.0137	0.0141	0.0039
count	222	174	174
t-stat	3.5345	3.3777	−1.5283
signif lev	1%	1%	

Table 5 repeats the analyses of Table 4 except that it analyzes the first and second halves separately to see if the HE difference results are stable across halves. Though the sample sizes in each half are not large, the HE is generally higher for h_{ccr_dcc} than for the other bias-adjustment approaches across assets and halves (the JY is an exception). While the HE aggregation across assets for the 1st half is not statistically significant, it is significant at the 1% level for the 2nd half.

Table 5. Stability of the HE Differences across h_{ccr} Bias-Adjustment Methods.

	$h_{ccr\_dcc} - h_{ccr\_corr}$	$h_{ccr\_dcc} - h_{ccr\_bam}$	$h_{ccr\_corr} - h_{ccr\_bam}$
Panel A. S&P500 1st Half			
mean	0.0004	0.0007	0.0002
std	0.0025	0.0033	0.0016
count	37	21	21
t-stat	1.0708	0.9055	0.7225
signif lev			
S&P500 2nd Half			
mean	0.0059	0.0056	−0.0003
std	0.0211	0.0203	0.0019
count	37	37	37
t-stat	1.7108	1.6902	−0.9254
signif lev	10%	10%	
Panel B. JY 1st Half			
mean	−0.0008	−0.0004	0.0014
std	0.0049	0.0039	0.0053
count	39	23	23
t-stat	−1.0677	−0.5421	1.2426
signif lev			
JY 2nd Half			
mean	−0.0001	0.0003	0.0004
std	0.0018	0.0019	0.0019
count	39	39	39
t-stat	−0.2797	1.0794	1.3287
signif lev			10%
Panel C. Gold 1st Half			
mean	0.0053	0.0066	0.0011
std	0.033	0.0346	0.0039
count	76	52	52
t-stat	1.4141	1.3841	2.1034
signif lev	10%	10%	5%
Gold 2nd Half			
mean	0.0068	0.0063	−0.0006
std	0.0169	0.0169	0.0107
count	77	77	77
t-stat	3.5346	3.2453	−0.4598
signif lev	1%	1%	

Table 5. *Cont.*

Panel D. Aggregated across Assets 1st Half			
	$h_{ccr\_dcc} - h_{ccr\_corr}$	$h_{ccr\_dcc} - h_{ccr\_bam}$	$h_{ccr\_corr} - h_{ccr\_bam}$
mean	0.0007	0.0006	−0.0003
std	0.007	0.0071	0.0048
count	111	63	63
t-stat	1.0398	0.7008	−0.5409
signif lev			
Aggregated across Assets 2nd Half			
mean	0.0058	0.0053	−0.0005
std	0.0177	0.0166	0.0032
count	111	111	111
t-stat	3.4493	3.3615	−1.6791
signif lev	1%	1%	10%

6. Conclusions

The goal of futures hedging is to reduce the firm's risk and increase its value. A key consideration for a futures hedger is the ratio of futures assets to short over the number of spot assets long, i.e., the futures hedge ratio.

This paper proposes new dynamic conditional futures hedge ratios. It studies: (1) the economics-based carry-cost rate hedge ratio introduced in Leistikow et al. (2020) and employed in Leistikow and Chen (2019), (2) a hedge ratio based on the Engle (2002) statistics-based dynamic conditional correlation model, and (3) three bias-adjusted versions of the carry-cost rate hedge ratio, where one uses the Engle (2002) dynamic conditional correlation model to adjust the bias. The hedge ratios' values and hedging performances are compared to those of two common benchmark hedge ratios (the traditional and "naive" hedge ratios) across three broad asset classes, two sub-periods, and three rolling window sizes.

The newly proposed economics-based carry-cost rate hedge ratio augmented with the Engle (2002) dynamic conditional correlation between the spot and futures prices nearly always (1) provides the highest hedging effectiveness and (2) has a statistically significantly higher hedging effectiveness than that of the other hedge ratio approaches across assets, sub-periods, and rolling window sizes.

All of the bias-adjusted carry-cost rate hedge ratios provide statistically significantly superior hedging performance across assets, sub-periods, and rolling window sizes to that of the "naive" hedge ratio of 1, h_1, advocated by Wang et al. (2015). However, this should not be interpreted as a general result. The poor hedging performance of h_1 is expected when the upward biased carry-cost rate hedge ratio, h_{ccr}, is significantly less than 1 as it was for these assets and time periods. Nevertheless, the carry-cost rate varies across assets, currency denominations, and time; thus, in the rare instances that the upward biased h_{ccr} is above 1 by its bias, the risk minimizing hedge ratio would also be h_1.

Unlike h_1, the statistics-based traditional hedge ratio, though inefficient compared to h_{ccr}, is correlated with the changing carry-cost rate. Of the three augmented carry-cost rate hedge ratios, only the one augmented with the Engle (2002) dynamic conditional correlation between the spot and futures prices nearly always provided higher hedging performance than the traditional hedge ratio.

Overall, our findings suggest that an effective hedge-ratio model needs to consider both the economic motivation of the model parameters and the dynamic nature of the correlations between a futures contract and its underlying asset. On the other hand, we acknowledge that our time-series models are estimated using historical data, which may be susceptible to backward bias. In future research, more fruitful work can be done by incorporating forward-looking data (e.g., options data) as part of the information set into multivariate time-series models.

Author Contributions: Conceptualization, methodology, and writing: D.L. and Y.T., Data-gathering, analysis, and proofreading: W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data came from Bloomberg.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. Many attribute this approach to Ederington (1979). Ferguson and Leistikow (1999) show that the ΔS term should be reduced by the period's carry-cost. Henceforth, we define ΔS as the carry-cost adjusted spot price change. Ferguson and Leistikow (1999) further shows that the carry-cost-adjusted spot price change hedge ratio calculation method is similar to, but theoretically superior to, the ECM hedge ratio advanced and studied by Kroner and Sultan (1993) and Ghosh and Clayton (1996).
2. The 1008 trading days, i.e., 4 years, rolling window size is arbitrary. For a robustness check, we replicate our analyses using 2- and 6-years rolling windows. Our results remain intact. The results are available upon request.
3. This modeling has been shown to produce superior empirical performance in a variety of situations. See, e.g., Engle (2002), Chiang et al. (2007), and Baur and Lucey (2010).
4. For a general carry-cost rate discussion, see Brennan (1958).
5. In discrete time hedging, T in Equation (11) is replaced by the hedge lift date's years to maturity, i.e., $T - t$.
6. For a more complete discussion of the advantages and disadvantages of the carry cost rate based hedge ratios, see Leistikow et al. (2020).
7. The goal was to get a long (at least back to 12 July 1990) series for the short term (ideally overnight for the daily data and weekly for the weekly data) US nearly riskless interest rate. Bloomberg's US 1-week Repo rate data begins on 23 July 1998 and has several gaps. Their overnight repo rate data begins at the same time as the 1-week repo rate data but is missing for about 100 more dates. Surprisingly, the 1-week rate average was about 2.5 basis points less than the overnight rate, but still this seems a minor difference. From these 2 series, we created a merged 1-week/1-day repo series: it is the 1-week repo rate unless it is unavailable, then it is the overnight repo rate -2.5 basis points. Bloomberg also has data on two 1-month interest rate series (repo rate and Libor) that go back farther than the above discussed (preferred, but unavailable) shorter (1-week/1-day) interest rates. The 1-month repo rate was about 20 basis points lower than the Libor on average. From these two 1-month rates, we created a 1 month merged rate series—it was the repo rate when available and Libor-20 basis points, otherwise. Finally, we created an overall series from the two merged series we just created (i.e., the 1-week/1-night series and the 1-month series). Since the merged 1 month series averaged 0.5 basis points less than the merged 1week/1night series, it is the merged 1 week/day repo series unless it is unavailable, in which case it is the 1 month series +0.5 basis points. This final merged series is the short term nearly riskless US interest rate used in CCR calculations for the various assets.
8. For the S&P500 and JY, which have quarterly maturities, we lose the first 16 contracts to construct the first 1008-days rolling window. The gold contracts have bimonthly maturities, so we lose the first 24 contracts to construct the first 1008-days rolling window. Therefore, in this section, the results of the S&P500, JY, and gold are based on 74, 78, and 153 contracts, respectively.
9. There are more statistics-based hedge ratios such as the GARCH method and its variants (e.g., Sarno and Valente (2000), Shaffer and Demaskey (2005), Alizadeh et al. (2008), Lee et al. (2009), and Wang et al. (2015). Alexander and Barbosa (2007) find the GARCH model hedging performance to be inferior to that for h_t, while Lien (2009) finds it is inferior except in small samples under special conditions.

References

Alexander, Carol, and Andreza Barbosa. 2007. Effectiveness of Minimum-Variance Hedging. *Journal of Portfolio Management* 33: 46–59. [CrossRef]

Alizadeh, Amir, Nikos Nomikos, and Panos Pouliasis. 2008. A Markov Regime Switching Approach for Hedging Energy Commodities. *Journal of Banking & Finance* 32: 1970–83.

Bali, Turan G., and Robert F. Engle. 2010. The Intertemporal Capital Asset Pricing Model with Dynamic Conditional Correlations. *Journal of Monetary Economics* 57: 377–90. [CrossRef]

Bali, Turan G., Robert F. Engle, and Yi Tang. 2017. Dynamic Conditional Beta Is Alive and Well in the Cross Section of Daily Stock Returns. *Management Science* 63: 3531–997. [CrossRef]

Baur, Dirk, and Brian Lucey. 2010. Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *Financial Review* 45: 217–29. [CrossRef]
Brennan, M. 1958. The Supply of Storage. *American Economic Review* 48: 50–71.
Chiang, Thomas, Bang Jeon, and Huimin Li. 2007. Dynamic Correlation Analysis of Financial Contagion: Evidence from Asian Markets. *Journal of International Money and Finance* 26: 1206–28. [CrossRef]
Ederington, Louis. 1979. The Hedging Performance of the New Futures Markets. *Journal of Finance* 34: 157–70. [CrossRef]
Engle, Robert F. 2002. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics* 20: 339–50. [CrossRef]
Ferguson, Robert, and Dean Leistikow. 1999. Futures Hedge Profit Measurement, Error-Correction Model vs. Regression Approach Hedge Ratios, and Data Error Effects. *Financial Management* 28: 118–25. [CrossRef]
Ghosh, Asim, and Ronnie Clayton. 1996. Hedging with International Stock Index Futures: An Intertemporal Error Correction Model. *Journal of Financial Research* 19: 477–91. [CrossRef]
Hull, John. 2018. *Options, Futures and Other Derivatives*, 10th ed. London: Pearson Education Inc.
Kroner, Kenneth, and Jahangir Sultan. 1993. Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures. *Journal of Financial and Quantitative Analysis* 28: 535–51. [CrossRef]
Lee, Cheng-Few, Kehluh Wang, and Yan Long Chen. 2009. Hedging and Optimal Hedge Ratios for International Index Futures Markets. *Review of Pacific Basin Financial Markets and Policies* 12: 593–610. [CrossRef]
Leistikow, Dean, Ren-Raw Chen, and Yuewu Xu. 2020. *Spot Asset Carry Cost Rates and Futures Hedge Ratios*. Working Paper. Bronx: Gabelli School of Business, Fordham University, forthcoming in *Review of Quantitative Finance and Accounting*.
Leistikow, Dean, and Ren-Raw Chen. 2019. Carry Cost Rate Regimes and Futures Hedge Ratio Variation. *Journal of Risk and Financial Management* 12: 78. [CrossRef]
Lien, Donald. 2009. Note on the Hedging Effectiveness of GARCH Models. *International Review of Economics and Finance* 18: 110–12. [CrossRef]
Sarno, Lucio, and Giorgio Valente. 2000. The Cost of Carry Model and Regime Shifts in Stock Index Futures Markets: An Empirical Investigation. *The Journal of Futures Markets* 20: 603–24. [CrossRef]
Shaffer, David, and Andrea Demaskey. 2005. Currency Hedging Using the Mean-Gini Framework. *Review of Quantitative Finance and Accounting* 25: 125–37. [CrossRef]
Wang, Yidong, Chongfeng Wu, and Li Yang. 2015. Hedging with Futures: Does Anything Beat the Naïve Hedging Strategy? *Management Science* 61: 2870–89. [CrossRef]

Journal of
Risk and Financial Management

MDPI

Article

Forecasting Commodity Prices Using the Term Structure

Yasmeen Idilbi-Bayaa and Mahmoud Qadan *

School of Business, Faculty of Social Sciences, University of Haifa, Haifa 3498838, Israel; Yasmeen.idilbi@gmail.com
* Correspondence: mqadan@univ.haifa.ac.il

Abstract: The aim of this study is to test the ability of the yield curve on US government bonds to forecast the future evolution in the prices of commodities often used in as raw materials. We consider the monthly prices of nine commodities for more than 30 years. Our findings, confirmed by several parametric and non-parametric tests, are robust and indicate that the ability to forecast future performance changes over time. Specifically, between 1986 and the early 2000s the yield curve was quite successful in forecasting monthly changes in commodity prices, but that success diminished in the period following. One possible explanation for this outcome is the increased flow of capital into the commodity market resulting in stronger correlations with the equity markets and a breakdown of the obvious relationship between commodities and business cycle. Our findings are important for asset pricing, commodity traders and policy makers.

Keywords: forecasting; commodity market; metals; term structure; yield spread

JEL Classification: C53; E3; E4; Q02

Citation: Idilbi-Bayaa, Yasmeen, and Mahmoud Qadan. 2021. Forecasting Commodity Prices Using the Term Structure. *Journal of Risk and Financial Management* 14: 585. https://doi.org/10.3390/jrfm14120585

Academic Editor: James W. Kolari

Received: 5 November 2021
Accepted: 30 November 2021
Published: 4 December 2021

1. Introduction

The literature regards the term structure curve, which plots the yield of government bonds against their maturity, as an indicator with valuable information about the current and future states of the economy (e.g., Harvey 1989; Abdymomunov 2013; Gogas et al. 2015; McMillan 2021b). The U.S. Federal Reserve, among other policymakers and institutional market participants, has always looked at the difference between the yields on long- and short-term sovereign bonds as an indication of where the economy is heading. Thus, the forecasting ability of the yield spread has become something of a stylized fact among macroeconomists.

Many studies, detailed in the literature review section, have demonstrated the ability of yield spreads to predict future economic situations effectively. They have established that the spreads contain a great deal of information about future economic activity and are accurate predictors of economic growth. However, there have been fewer comprehensive attempts to understand the dynamic relationships between the evolution in the prices of commodities and the shape of the yield curve. This question has become particularly relevant in the wake of the unconventional monetary policy used in the last two decades, which has not been employed since the Great Depression during 1930s. Thus, our goal is to fill this gap by exploring the ability of the term structure to forecast the future evolution in the prices of commodities. To accomplish this task, we use 30 years of data about nine commodities often used as raw materials: coal, gold, silver, oil, platinum, palladium, zinc, ethanol, and natural gas. We also use various proxies for the yield spread combined from 30-year, 10-year, 2-year, 1-year and 3-month interest rates.

Our findings indicate that yield spreads are generally positively correlated with future changes in the price of commodities. Our results are robust to controlling for real economic and financial variables. Adding any or all of these potential alternative explanatory variables only marginally affects the coefficients or their statistical significance. In our analysis, we also utilize the Engle's (2002) dynamic conditional correlation procedure (DCC). The overall picture confirms the existence of time-varying correlation between the yield spread and future price movements in commodities. To determine why this result emerged, we consider sub-samples that are determined endogenously using the Bai and Perron (2003) tests. These structural break tests confirm that the positive correlation is economically and statistically significant mainly for the period prior to 2004. On the other hand, in the period following (2004–2020) it seems that the yield spread has been unable to predict changes in commodities in any significant way.

The weakening correlation between the variables of interest may be related to the massive capital inflows from individual and institutional investors into the commodity market in the early 2000s (e.g., Tang and Xiong 2012). Many studies justified these inflows to commodities by their relatively low correlation with financial markets and, accordingly, the potential diversification benefits (e.g., Gorton and Rouwenhorst 2006; Daskalaki et al. 2017). However, the considerable inclusion of commodities in investors' portfolios resulted in the financialization of commodities, yielding a strong correlation between commodity prices and equity markets (e.g., Hu et al. 2020) and a breakdown of the obvious relationship between commodities and cyclical phases of the economy.

Lastly, between 1986 and December 2020, we found eight periods during which the yield spread was negative or equal to zero. The non-parametric tests conducted to track the future evolution in commodity prices following flat or downward-sloped yield curves indicate that such situations can be a successful timing to embark on long positions in several commodities for investors planning to hold for a relatively long period of time. Recently, with the outbreak of the Coronavirus late in February 2020, the U.S. 1-year yield was 1.43%, and the 10-year was about flat (1.46%). Tracking the commodity prices in the few months following indicate significant shrink in prices. However, the prices recovered sharply after 2–4 quarters. This recent case, among the others observed, confirms that specific shapes of the yield curve may generate abnormal returns for investors (see Table A1 in the online Appendix A).

The mechanism underlying our conjecture here accords with the empirical evidence confirming a strong relationship between the business cycle and commodity prices (Labys et al. 1999; Chevallier et al. 2014). In addition, the literature has established that financial markets, including the yield curve, move more quickly than real markets (Saar and Yagil 2015). Accordingly, one should observe a causal relationship from this macro-financial predictor to the commodity market.

The paper contributes to the existing literature in several ways. First, our study sheds light on the link between yield spreads and long-term prices in the commodity market. While previous studies have focused on interest rates in level rather than the difference between long and short-term sovereign bond yields (e.g., Dai and Kang 2021), little is known about the information content of the term structure for commodities. Second, we add to the literature documenting the time-varying relationship between real economic and financial variables by examining the structural breaks in the long-term correlation between yield spreads and the commodity market (e.g., Chinn and Kucko 2015). Third, our findings are especially useful for policy makers and central banks because long-term predictions about commodity prices are essential in targeting inflation and promoting overall economic stability (e.g., Garner 1989; Orlowski 2017; Fasanya and Awodimila 2020). Finally, modeling and forecasting future innovations in commodity prices are important for both market participants and scholars. Predictions in this area play a vital role in portfolio optimization and risk management. Indeed, investors are attracted to commodities due to their inflation-hedging properties (e.g., Beckmann and Czudaj 2013; Bampinas and Panagiotidis 2015; Levine et al. 2018; Umar et al. 2019), and their possible contribution to diversifying risks (e.g., Gagnon et al. 2020).

The remainder of this study proceeds as follows. Section 2 reviews the literature. Section 3 describes the data and the construction of the key variables. Section 4 presents the methodology. Section 5 details the empirical findings and discusses the results. Section 6 checks the robustness of the findings, and Section 7 concludes.

2. Literature Review

The macroeconomic literature has established that future real economic growth is positively correlated with lagged interest spreads (e.g., Stock and Watson 1989; Estrella and Hardouvelis 1991; Plosser and Rouwenhorst 1994; McMillan 2021a). In parallel, the link between interest rates and commodities has been also investigated and can be classified into two categories. The first addresses the effect of the interest rate level on commodity prices. This line of literature has established that commodity prices increase significantly in response to reductions in real interest rates (Akram 2009; Arango et al. 2011). The second category explores the effect of shocks in the commodity market (mainly oil prices) on long-term interest rates (e.g., Ioannidis and Ka 2018). Recent studies use the Granger (1969) causality test and provide evidence that not only do interest rates drive commodity prices, but also that commodity prices drive income and interest rates (e.g., Harvey et al. 2017). Despite these extensive efforts, the examination of the ability of yield spreads to predict future innovations in commodity prices has attracted relatively less attention in the literature.

The literature points to several reasons why yield spreads forecast future real economic activity. One reason relates to the expectation theory, according to which when the yield curve flattens, market participants expect short-term interest rates to fall due to a recession. This expectation translates into a drop in long-term interest rates, as deteriorating market conditions during recessions might explain the decline in short-term rates. Indeed, economic depressions are often associated with job loss, increased uncertainty, business failures, and credit line contractions. Consequently, if people anticipate a slowdown in economic activity, there will likely be a drop in the demand for credit, which in turn leads to a decline in long-term interest rates. On the other hand, if market participants anticipate an upturn in the economy, future short-term interest rates will be expected to rise, leading to a steepening of the yield curve. Thus, while falling yield spreads preceding recessions are caused by both aforementioned factors, the decline in expectations about short-term rates is the more important one (Hamilton and Kim 2002).

The second explanation is related to the countercyclical monetary policy according to which economic expansion is accompanied by inflation. To control inflation, central banks follow a countercyclical monetary policy by raising short-term interest rates. Tight monetary policies are used to stabilize output growth and cause the yield spread to drop. This measure is aimed at reducing the anticipation of inflation to levels below the current inflation rate. Consequently, short-term interest rates rise more than long-term interest rates do, and the yield curve flattens. As real interest rates remain high, spending decreases, causing an economic slowdown. Estrella (2005) provided a theoretical model wherein the yield spread explains both output and inflation. The author showed that the predictive ability of the yield spread depends upon the reaction of the given monetary policy. By the same logic, in a recessionary economy, central banks will reduce short-term interest rates as part of a countercyclical monetary policy. Thus, a lower yield spread or a flat yield curve is a harbinger of economic downturn.

The third explanation of why the yield curve slope is a leading indicator of economic output is referred to as the inter-temporal consumption model. As per Harvey (1989), during expansionary periods people have a stable level of consumption, whereas during recessions, when income is falling, they tend to reduce their consumption. Hence, if people anticipate a decline in economic activity, they have an incentive to save in the current period by selling short-term assets and buying bonds, which will ensure a stable income during the low-income period. As a result, long-term bond prices rise, which in turn reduces their yields, and short-term bonds trade at increased rates.

Finally, there are various empirical works dealing with the relationship between the factors affecting macroeconomic fundamentals and commodity prices. Variables such as an increase in economic activity (e.g., Duarte et al. 2021), economic uncertainty (Qadan and Nama 2018), the exchange rate of the dollar (e.g., Churchill et al. 2019) and the market index (Kagraoka 2016) are capable of affecting commodity prices. Considering the ability of the spread in bond yields to anticipate future economic activity, it is very important to have some understanding of its role in providing information about the future prices of commodities.

3. Data

Our sample consists of monthly data on nine commodities—oil, silver, gold, platinum, palladium, zinc, ethanol, coal, and natural gas—obtained based on the availability of the data. These commodities are used in many industries as raw materials. Data about silver and gold come from the Chicago Mercantile Exchange (CME). Data about platinum palladium and natural gas come from the New York Mercantile Exchange (NYMEX). The data on zinc and copper come from the London Bullion Market Association (LBMA). Information about coal comes from the International Exchange (ICE), whereas the data on WTI oil are taken from the Federal Reserve Bank of St. Louis.

The largest sample period used is that for oil, gold and silver, and ranges from January 1986 to December 2020, while the smallest sample is that for coal and ranges from January 2009 to December 2020. Our starting point for each commodity is simply due to the availability of information about their prices. We use the International Monetary Fund's International Financial Statistics database for the rates for the 3-month, 1-year, 2-year, 10-year and 30-year Treasury bills. Table 1 reports the descriptive statistics of the key variables used in this study and outlines the sample period. Panel A reports the descriptive statistics of the six proxies used to capture the yield spread, while Panel B reports the descriptive statistics of the commodities employed here.

Table 1. Descriptive Statistics. Panel A—yield rates. Panel B—commodities.

Panel A

	Y_{3M}	Y_1	Y_2	Y_{10}	Y_{30}
Mean	3.173	3.455	3.761	4.845	5.343
Median	3.055	3.390	3.920	4.680	5.155
Maximum	9.140	9.570	9.680	9.520	9.610
Minimum	0.010	0.100	0.130	0.620	1.270
Std. Dev.	2.557	2.623	2.653	2.280	2.052
Skewness	0.258	0.237	0.214	0.199	0.162
Kurtosis	1.845	1.824	1.793	1.981	2.046
J-Bera	28.001	28.118	28.716	20.950	17.767
#Obs.	420	420	420	420	420
Sample Period	1986:01 to 2020:12	1986:01 to 2020:12	1986:01 to 2020:12	1986:01 to 2020:12	1986:01 to 2020:12

Panel B

	COAL	ETHNL	GOLD	NGAZ	OIL	PLDM	PLTNM	SLVR	ZINC
Mean	83.85	1.89	730.10	3.75	44.15	459.31	822.34	11.20	1879.99
Median	82.65	1.77	425.55	2.92	31.90	309.75	680.50	6.72	1891.75
Maximum	130.90	3.62	1970.80	13.92	140.97	2508.80	2180.70	48.58	4474.00
Minimum	49.95	0.82	255.00	1.17	11.13	76.35	336.40	3.56	746.75
Std. Dev.	21.12	0.50	480.20	2.24	28.99	440.48	440.90	8.26	807.30
Skewness	0.280	0.672	0.792	1.721	0.879	2.111	0.813	1.418	0.528
Kurtosis	2.142	3.121	2.151	6.416	2.671	8.299	2.611	4.823	2.765
J-Bera	6.30	14.20	56.52	360.54	55.93	797.60	48.57	198.86	13.68
#Obs.	144	187	420	368	420	417	417	420	281
Sample Period	2009:01 to 2020:12	2005:06 to 2020:12	1986:01 to 2020:12	1990:05 to 2020:12	1986:01 to 2020:12	1986:04 to 2020:12	1986:04 to 2020:12	1986:01 to 2020:12	1997:08 to 2020:12

Notes: Panel A of the table reports the descriptive statistics of the Treasury yield rates, whereas Panel B reports those of the commodity prices. Y_{3M}, Y_1, Y_2, Y_{10} and Y_{30}, are US treasury yields on 3-month, 1-year, 2-year, 10-year and 30-year bonds, respectively—all denominated in annual terms.

4. Method

The empirical economic literature defines the yield spread as the difference between the yield rates on long-term and short-term government bonds. In fact, there is no precise theory that defines how the yield spread should be calculated, and the choice of creating a proxy for the yield spread is somewhat arbitrary. Indeed, the literature provides many proxies for the yield spread including the difference between the yields on 10-year bonds and 3-month bonds (Estrella and Hardouvelis 1991), the difference between 10-year and 1-year interest rates (Stock and Watson 1989) and the difference between yields on 30-year and 3-month bonds (Duffee 1998). Given the mixed definitions of the yield spread, we utilize as broad a spectrum of bonds as possible, specifically, the differences in the yields on 10-year and 3-month Treasury bonds, 10-year and 1-year bonds, 10-year and 2-year bonds, 30-year and 3-month bonds, 30-year and 1-year bonds and 30-year and 2-year bonds.

We formulate the following model to trace the effect of the current yield spread (at time t) on the cumulative rate of change in the "h" subsequent months or quarters.

$$R_{t+h} = \beta_0^h + \beta_1^h \left(Y_{Long,t} - Y_{Short,t}\right) + B' X_t + v_{t+h} \tag{1}$$

where

$$R_{t+h} = \left(\frac{12}{h} \times 100\right) \times \left(\ln(P_{t+h}) - \ln(P_t)\right) \tag{2}$$

R_{t+h} is the rate of change in the price of the commodity in annual terms. If, for example, $h = 1$, then R_{t+1} captures the cumulative return of one period (say, a quarter) ahead. If $h = 4$, then R_{t+4} captures the cumulative returns for the coming twelve months (four quarters). The difference between the yield rates on long-term and short-term government bonds is given by $\left(Y_{Long,t} - Y_{Short,t}\right)$ and v_{t+h} is the forecast error. Given the possibility that the forecast error might be correlated, we use Newey and West's (1987) corrected covariance estimator. The estimated coefficients guarantee consistency in the presence of both heteroscedasticity and autocorrelation (HAC) of unknown form. X_t denotes a matrix of additional explanatory variables. In line with the literature, we use the U.S. dollar exchange rate (Churchill et al. 2019), the S&P500 (Kagraoka 2016), the industrial production index (Duarte et al. 2021) and the economic policy uncertainty index (Huang et al. 2021).

To depict the dynamic correlation between the current yield spread and the future price direction of commodities, we use the established multivariate concept of dynamic conditional correlation generalized autoregressive conditional heteroscedasticity (DCC GARCH). Engle (2002) developed this state-of-the-art method, which has been used extensively to quantify dynamic relationships over time. In the following, we provide a very basic description of this methodology.[1]

The dynamic conditional correlation estimator is an extension of the constant conditional correlation model suggested by Bollerslev (1990). According to Bollerslev's procedure, the correlation matrix ρ is constant. That is, $H_t = D_t \rho D_t$, where $D_t = diag\{\sqrt{h_{i,t}}\}$ and $h_{i,t}$ represents the i-th univariate (G)ARCH(p, q) process. In other words,

$$D_t = \begin{pmatrix} h_{1t} & 0 & 0 & \cdots & 0 \\ 0 & h_{2t} & 0 & \cdots & 0 \\ 0 & 0 & h_{3t} & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & & h_{nt} \end{pmatrix} \tag{3}$$

According to Engle (2002), ρ is allowed to vary in time t. Thus,

$$H_t \equiv D_t \rho_t D_t \tag{4}$$

The correlation matrix is then given by:

$$\rho_t = \begin{pmatrix} 1 & q_{12,t} & q_{13,t} & \cdots & q_{1n,t} \\ q_{21,t} & 1 & q_{23,t} & \cdots & q_{2n,t} \\ q_{31,t} & q_{32,t} & 1 & \cdots & q_{3n,t} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{n1,t} & q_{n2,t} & q_{n3,t} & \cdots & 1 \end{pmatrix} \tag{5}$$

The correlation matrix is a positive definite one because of the positive nature of H_t. Given that $Q_t = (q_{ij,t})$, then:

$$Q_t = (1 - \alpha - \beta)\overline{Q} + \alpha\eta_{t-1}\eta'_{t-1} + \beta Q_{t-1} \tag{6}$$

where $\eta_t = \varepsilon_{i,t}/\sqrt{h_{i,t}}$ are the standardized residuals from the (G)ARCH model, $\overline{Q}$ is a $n \times n$ matrix and represents the unconditional variance matrix of the standardized error terms η_t and computed as $\overline{Q} = E[\eta_t \times \eta'_t]$. α and β are non-negative scalars and satisfy the mean-revering assumption (i.e., $\alpha + \beta < 1$). Q_t is a positive definite matrix that determines the structure of dynamics and Q_t^{*-1} normalizes the elements in Q_t;

$$Q_t^{*-1} = \begin{pmatrix} \frac{1}{\sqrt{q_{11t}}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{q_{22t}}} & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{\sqrt{q_{33t}}} & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{q_{nnt}} \end{pmatrix} \tag{7}$$

In order to estimate the parameters of H_t specifically $\Phi = (\alpha, \beta)$, the following log-likelihood function is maximized:

$$L(\Phi) = -0.5\sum_{i=1}^{T}\left(n\log(2\pi) + \log(|H_t|) + y'_t H_t^{-1} y_t\right) \tag{8}$$

5. Empirical Findings

Table 2 reports the estimation results of the reduced form of Equation (1). That is, future commodity returns (R_{t+h}) are regressed against the current yield spread only. $R_{t+h} = \beta_0^h + \beta_1^h(Y_{Long,t} - Y_{Short,t}) + v_t^h$. In this table, we utilize the difference between 10-year and 3-month bond yields (Y_{10}-Y_{3M}) as a proxy for the yield spread. We also use four forecasting horizons (h; h = 1, 2, 3 and 4) where h = 1 indicates forecasting of one quarter ahead and h = 4 indicates forecasting four quarters ahead. Panel A of the table reports the estimation results with respect to the entire sample, Panel B covers 1986 to 2003, and Panel C covers 2004 to 2020.

Table 2. Estimation results of Equation (1) with the Y_{10}-Y_{3M} indicator. Panel A: entire sample. Panel B: sample period 1986–2003. Panel C: sample period 2004–2020.

Panel A												
Forecast Horizon	**Oil (1986:01–2020:12)**				**Silver (1986:01–2020:12)**				**Gold (1986:01–2020:12)**			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	10.61 (0.42)	−3.78 (0.56)	0.003	139	5.33 (0.48)	−0.37 (0.92)	0.000	139	4.89 (0.24)	0.08 (0.97)	0.000	138
2	9.03 (0.29)	−3.08 (0.46)	0.004	138	3.97 (0.45)	0.39 (0.88)	0.000	138	4.74 (0.10)	0.15 (0.92)	0.000	138
3	7.41 (0.24)	−2.30 (0.46)	0.004	137	1.85 (0.67)	1.40 (0.50)	0.003	137	4.56 * (0.06)	0.11 (0.93)	0.000	137
4	4.29 (0.41)	−0.91 (0.72)	0.001	136	−0.23 (0.95)	2.40 (0.17)	0.014	136	4.47 ** (0.04)	0.06 (0.95)	0.000	136

Table 2. *Cont.*

Panel A												
Forecast Horizon	**Platinum (1986:04–2020:12)**				**Palladium (1986:04–2020:12)**				**Zinc (1997:08–2020:12)**			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	5.23	−1.48	0.001	139	19.38 *	−6.14	0.011	139	3.18	−0.02	0.000	93
	(0.45)	(0.66)			(0.06)	(0.22)			(0.75)	(0.99)		
2	3.68	−0.72	0.000	138	15.63 **	−3.95	0.009	138	−1.03	2.62	0.005	92
	(0.44)	(0.76)			(0.04)	(0.28)			(0.89)	(0.49)		
3	1.27	0.45	0.0005	136	10.53 *	−1.16	0.001	137	−4.31	4.53	0.021	91
	(0.73)	(0.80)			(0.09)	(0.70)			(0.51)	(0.17)		
4	−0.18	1.17	0.004	136	8.18	0.24	0.000	136	−6.91	**6.04 ****	0.05	90
	(0.95)	(0.43)			(0.14)	(0.93)			(0.23)	**(0.04)**		
Forecast Horizon	**Ethanol (2005:06–2020:12)**				**Coal (2009:01–2020:12)**				**Natural Gas (1990:05–2020:12)**			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	3.52	−2.70	0.012	62	−7.04	4.08	0.006	48	4.24	−1.36	0.0003	122
	(0.86)	(0.79)			(0.66)	(0.59)			(0.79)	(0.86)		
2	1.49	−2.32	0.002	61	−10.92	5.85	0.0.23	47	2.11	−0.40	0.000	121
	(0.90)	(0.70)			(0.37)	(0.31)			(0.84)	(0.94)		
3	−1.69	−0.64	0.004	60	−15.40	7.67	0.06	46	−1.00	1.02	0.001	120
	(0.84)	(0.88)			(0.12)	(0.10)			(0.91)	(0.80)		
4	−5.01	0.61	0.001	59	−14.55	**7.14 ***	0.06	45	−3.51	2.39	0.005	119
	(0.43)	(0.85)			(0.11)	**(0.09)**			(0.61)	(0.47)		
Panel B												
Forecast Horizon	**Oil (1986:01–2003:12)**				**Silver (1986:01–2003:12)**				**Gold (1986:01–2003:12)**			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	11.73	−3.38	0.003	71	−6.16	4.15	0.019	71	−4.99	3.60	0.025	70
	(0.42)	(0.63)			(0.40)	(0.25)			(0.37)	(0.18)		
2	10.25	−2.58	0.004	71	−8.19	**5.85 ****	0.07	71	−4.51	**3.36 ****	0.06	71
	(0.29)	(0.58)			(0.13)	**(0.03)**			(0.18)	**(0.04)**		
3	10.62	−2.91	0.01	71	−7.71 *	**5.28 *****	0.09	71	−3.99	**2.74 ****	0.06	71
	(0.13)	(0.39)			(0.07)	**(0.01)**			(0.13)	**(0.03)**		
4	9.55	−2.26	0.009	71	−7.57 **	**5.20 *****	0.114	71	−3.54	**2.40 ****	0.06	71
	(0.12)	(0.44)			(0.04)	**(0.004)**			(0.13)	**(0.04)**		
Forecast Horizon	**Platinum (1986:04–2003:12)**				**Palladium (1986:04–2003:12)**				**Zinc (1997:08–2003:12)**			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	−1.01	2.90	0.009	71	11.53	−4.63	0.007	71	−17.87	9.02	0.10	25
	(0.89)	(0.44)			(0.40)	(0.49)			(0.11)	(0.12)		
2	−3.28	**4.35 ***	0.046	71	4.58	−0.001	0.000	71	−15.85 *	**9.50 ****	0.18	25
	(0.51)	**(0.07)**			(0.66)	(0.99)			(0.06)	**(0.03)**		
3	−4.65	**4.74 ****	0.085	70	0.13	2.20	0.004	71	−14.08 **	**8.39 ****	0.20	25
	(0.23)	**(0.014)**			(0.99)	(0.60)			(0.04)	**(0.02)**		
4	−4.98	**4.91 *****	0.112	71	−1.30	2.99	0.009	71	−13.17 **	**8.38 *****	0.26	25
	(0.15)	**(0.004)**			(0.87)	(0.43)			(0.03)	**(0.01)**		

Table 2. *Cont.*

Panel B

Forecast Horizon	Natural Gas (1990:05–2003:12)			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	8.41 (0.76)	1.11 (0.93)	0.0002	54
2	−0.57 (0.97)	5.51 (0.50)	0.009	54
3	−5.24 (0.69)	7.99 (0.19)	0.03	54
4	−7.17 (0.48)	**9.42** * **(0.05)**	0.07	54

Panel C

Forecast Horizon	Oil (2004:01–2020:12)				Silver (2004:01–2020:12)				Gold (2004:01–2020:12)			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	9.51 (0.67)	−4.22 (0.70)	0.002	68	17.23 (0.20)	−5.16 (0.44)	0.009	68	14.40 ** (0.02)	−3.29 (0.27)	0.019	68
2	7.77 (0.59)	−3.65 (0.60)	0.004	67	16.94 * (0.07)	−5.45 (0.23)	0.022	67	14.54 *** (0.002)	−3.29 (0.14)	0.033	67
3	3.90 (0.72)	−1.62 (0.76)	0.001	66	12.35 (0.11)	−2.87 (0.44)	0.009	66	13.92 *** (0.00)	−2.81 (0.14)	0.03	66
4	−1.66 (0.85)	0.66 (0.88)	0.0004	65	8.09 (0.20)	−0.78 (0.80)	0.001	65	13.53 *** (0.00)	−2.64 (0.12)	0.04	65
Forecast Horizon	**Platinum (2004:01–2020:12)**				**Palladium (2004:01–2020:12)**				**Zinc (2004:01–2020:12)**			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	11.84 (0.31)	−6.18 (0.29)	0.017	68	27.41 * (0.07)	−7.64 (0.31)	0.016	68	13.99 (0.29)	−4.68 (0.48)	0.008	68
2	11.20 (0.17)	−6.21 (0.12)	0.04	67	27.36 ** (0.013)	−8.18 (0.13)	0.036	67	6.80 (0.52)	−0.91 (0.86)	0.0005	67
3	7.84 (0.21)	−4.25 (0.17)	0.029	66	21.92 ** (0.02)	−4.88 (0.26)	0.02	66	0.98 (2.38)	2.38 (0.59)	0.005	66
4	5.27 (0.30)	−2.97 (0.23)	0.023	65	18.91 ** (0.02)	−2.94 (0.43)	0.01	65	−3.42 (0.67)	4.67 (0.24)	0.022	65

Forecast Horizon	Natural Gas (2004:01–2020:12)			
h	**C**	**(Y_{10}-Y_{3M})**	**R^2**	**N**
1	1.95 (0.92)	−4.07 (0.67)	0.003	68
2	5.14 (0.71)	−5.95 (0.38)	0.012	67
3	3.37 (0.76)	−5.48 (0.31)	0.016	66
4	0.49 (0.96)	−4.21 (0.32)	0.016	65

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_{3M} denotes the yield spread calculated as the difference between the yield rates on 10-year and 3-month government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

The picture that emerges indicates an insignificant positive tendency of the yield spread to forecast future changes in the commodities used. Panel B, however, presents a different picture. Except for oil and palladium, we find that, when considering 1986–2003, the yield spread has a positive effect on the future prices of the rest of the commodities. The ability of the current yield spread to predict future innovations in commodity prices is manifested in both the statically significant positive coefficients and the relatively high R^2 (for example, R^2 values are 6%, 7%, 11.2%, 11.4% and 26% for gold, natural gas, platinum, silver and zinc, respectively.). A steeper yield curve is always viewed as an indication that the growth in future output is about to rise. Thus, the positive correlation detected indicates that an increase in the slope at time t will have a positive impact on the future prices of commodities.

While Table 2 regresses the commodity returns against the yield spread only, in Table 3 we present the full estimation of Equation (1) after controlling for additional explanatory variables. The sample period considered in Table 3 is for 1986–2003. The results for the entire sample and the period after 2004 appear in Table A2 (in the online Appendix A). The overall picture is maintained as evident by the significant positive coefficients of the yield spreads even after controlling for real and financial economic variables in the period prior to 2004. The results hold true for all commodities except for oil and palladium.

Table 3. Estimation results of Equation (1) for the period 1986–2003.

Forecast Horizon	Oil (1986:01–2003:12)								Silver (1986:01–2003:12)							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	11.80	−4.97	0.48	−2.20	11.89	0.03	−0.06	71	−8.22	4.93	0.58	−10.74	−1.92	0.00	−0.04	71
	(0.76)	(−0.62)	(0.06)	(−0.11)	(0.67)	(0.26)			(−1.06)	(1.23)	(0.15)	(−1.06)	(−0.22)	(−0.07)		
2	10.61	−3.25	−9.41c	−5.24	10.78	−0.02	0.00	71	−9.31	**6.50b**	0.58	−5.78	−3.10	−0.01	0.01	71
	(1.06)	(−0.63)	(−1.82)	(−0.4)	(0.95)	(−0.19)			(−1.64)	(2.22)	(0.2)	(−0.78)	(−0.48)	(−0.27)		
3	10.22	−2.73	−5.10	−4.14	3.21	0.00	−0.03	71	**−9.37b**	**6.04a**	0.53	−9.09	−2.79	−0.02	0.06	71
	(1.39)	(−0.72)	(−1.35)	(−0.43)	(0.38)	(−0.03)			(−2.11)	(2.64)	(0.23)	(−1.57)	(−0.55)	(−0.4)		
4	8.01	−1.83	−2.85	−7.46	2.58	0.02	−0.03	71	**−8.92b**	**5.92a**	−0.02	−7.43	−2.73	−0.01	0.09	71
	(1.25)	(−0.56)	(−0.86)	(−0.9)	(0.35)	(0.44)			(−2.36)	(3.05)	(−0.01)	(−1.51)	(−0.63)	(−0.46)		
Forecast Horizon	**Gold (1986:01–2003:12)**								**Platinum (1986:04–2003:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	−7.28	3.49	4.47	−10.20	3.19	0.00	0.01	70	−2.64	2.25	−1.79	**−22.72b**	7.66	−0.10	0.06	71
	(−1.23)	(1.17)	(1.52)	(−1.39)	(0.5)	(0.1)			(−0.34)	(0.57)	(−0.45)	(−2.26)	(0.87)	(−1.57)		
2	−4.01	**4.01b**	−1.50	3.21	−5.20	−0.02	0.04	70	−5.10	**4.63c**	−1.73	**−14.33b**	2.42	−0.03	0.07	71
	(−1.14)	(2.23)	(−0.83)	(0.7)	(−1.3)	(−0.72)			(−1.01)	(1.78)	(−0.66)	(−2.17)	(0.42)	(−0.75)		
3	−3.74	**3.04b**	−0.49	0.67	−2.80	−0.02	0.02	70	−5.71	**4.17b**	0.32	**−10.25c**	5.89	−0.02	0.09	71
	(−1.35)	(2.14)	(−0.35)	(0.19)	(−0.89)	(−0.93)			(−1.42)	(2.02)	(0.16)	(−1.97)	(1.29)	(−0.49)		
4	−3.62	**2.94b**	−0.87	−0.42	−3.48	−0.02	0.03	70	−5.19	**4.32b**	−0.21	−5.51	4.59	−0.02	0.09	71
	(−1.46)	(2.31)	(−0.68)	(−0.13)	(−1.24)	(−0.84)			(−1.44)	(2.33)	(−0.11)	(−1.17)	(1.12)	(−0.7)		
Forecast Horizon	**Palladium (1986:04–2003:12)**								**Zinc (1997:08–2003:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	15.89	−10.17	**−14.00b**	−13.30	38.73	−0.18	0.13	71	**−19.18c**	5.87	8.85	−27.86	9.80	−0.02	0.07	25
	(1.2)	(−1.49)	(−2.05)	(−0.77)	(2.57)	(−1.6)			(−1.76)	(1.00)	(1.59)	(−1.15)	(0.66)	(−0.18)		
2	6.75	−4.82	−3.68	−12.06	**32.29a**	−0.08	0.06	71	**−17.44b**	7.48	5.42	−26.55	4.01	0.04	0.15	25
	(0.65)	(−0.9)	(−0.68)	(−0.89)	(2.72)	(−0.9)			(−2.13)	(1.7)	(1.3)	(−1.47)	(0.36)	(0.38)		
3	3.16	−4.11	3.39	−4.46	**37.41a**	−0.02	0.13	71	**−15.76b**	**7.29c**	3.40	−24.21	−2.48	0.10	0.21	25
	(0.38)	(−0.95)	(0.78)	(−0.41)	(3.9)	(−0.35)			(−2.4)	(2.07)	(1.02)	(−1.67)	(−0.28)	(1.26)		
4	0.88	−2.11	3.50	−4.59	**30.30a**	−0.01	0.10	71	**−15.02b**	**7.85b**	2.03	−12.48	1.41	0.09	0.22	25
	(0.11)	(−0.53)	(0.88)	(−0.46)	(3.46)	(−0.23)			(−2.59)	(2.52)	(0.69)	(−0.97)	(0.18)	(1.37)		

Table 3. *Cont.*

Forecast Horizon	Natural gas (1990:05–2003:12)							
h	C	(Y_{10}-Y_{3M})	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	4.31	−3.20	18.12	−12.32	36.33	0.04	−0.04	54
	(0.15)	(−0.23)	(1.33)	(−0.33)	(1.14)	(0.16)		
2	−6.06	4.49	4.89	−34.36	21.29	−0.03	−0.03	54
	(−0.34)	(0.52)	(0.57)	(−1.43)	(1.06)	(−0.18)		
3	−16.59	9.11	5.02	**−50.56a**	20.50	0.12	0.16	54
	(−1.32)	(1.51)	(0.83)	(−3.02)	(1.46)	(1.15)		
4	−15.38	**10.68b**	3.90	**−32.98b**	10.68	0.11	0.13	54
	(−1.51)	(2.19)	(0.8)	(−2.44)	(0.94)	(1.33)		

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_{3M} denotes the yield spread calculated as the difference between the yield rates on 10-year and 3-month government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. a, b and c denote statistical significance at the 1%, 5%, and 10% levels, respectively.

The statistically significant positive coefficients accord with the empirical literature postulating that commodity prices (specifically silver, gold, platinum, zinc and natural gas) are tightly linked with the business cycle and the state of the economy (e.g., Batten et al. 2010; Kucher and McCoskey 2017; Jahan and Serletis 2019). This economically and statistically significant relationship confirms that these metals are used extensively in various industries, making them more exposed to the expected phases in the economic cycle.

In investigating whether the relationship between yield spreads and future innovations in commodity prices is stable over time, we made two major findings. The first is that the correlation between commodity prices and yield spreads is not stable over time. This finding is evident in Engle's (2002) dynamic conditional correlations depicted in Figures 1–3, which illustrate the dynamic conditional correlation between the yield spreads and the commodity prices two, three and four quarters ahead. This finding accords with recent studies maintaining time-varying relationship between yield spread and future economic output (e.g., Kuosmanen et al. 2019; Chinn and Kucko 2015).

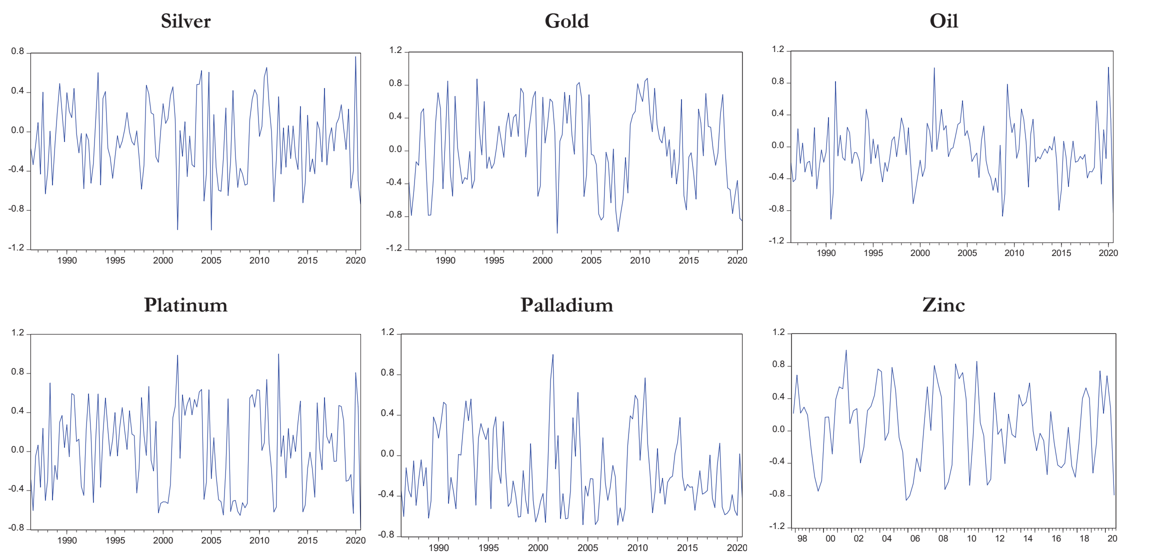

Figure 1. *Cont.*

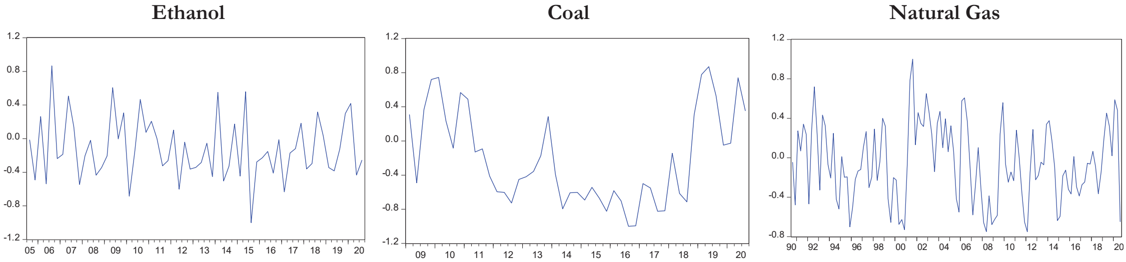

Figure 1. Dynamic correlation between the yield spread and two quarters ahead.

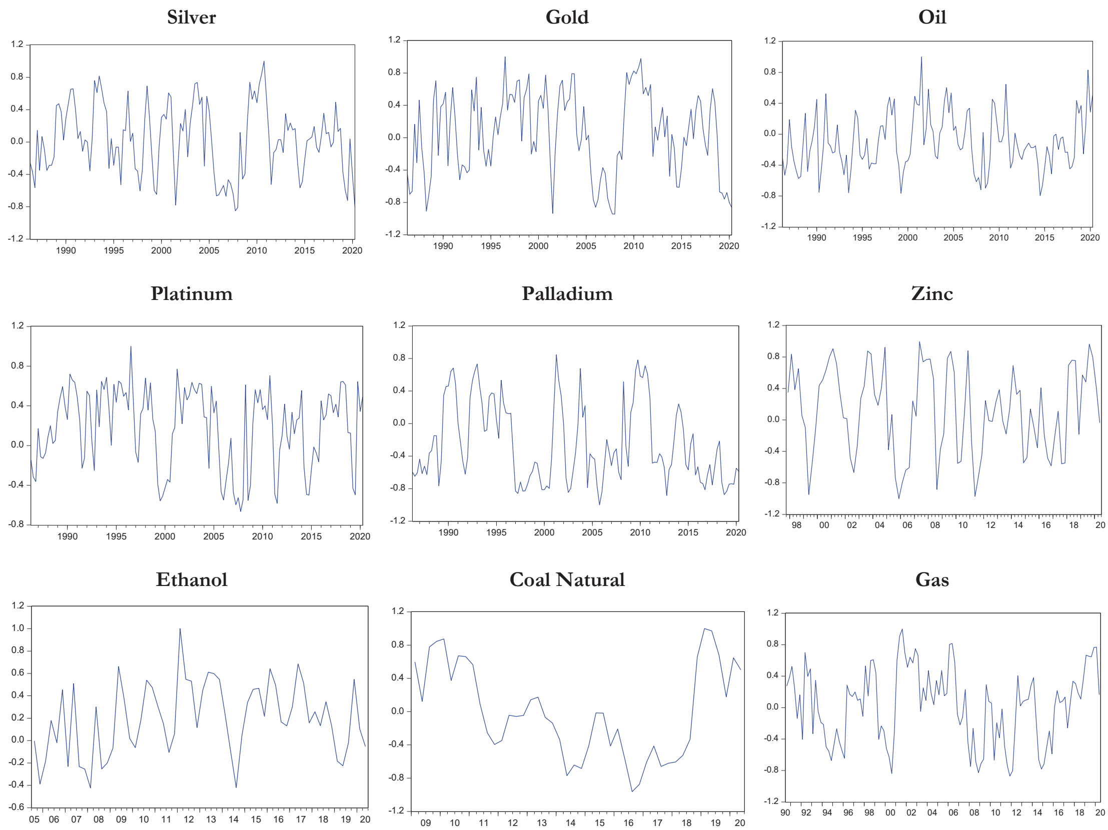

Figure 2. Dynamic correlation between the yield spread and three quarters ahead.

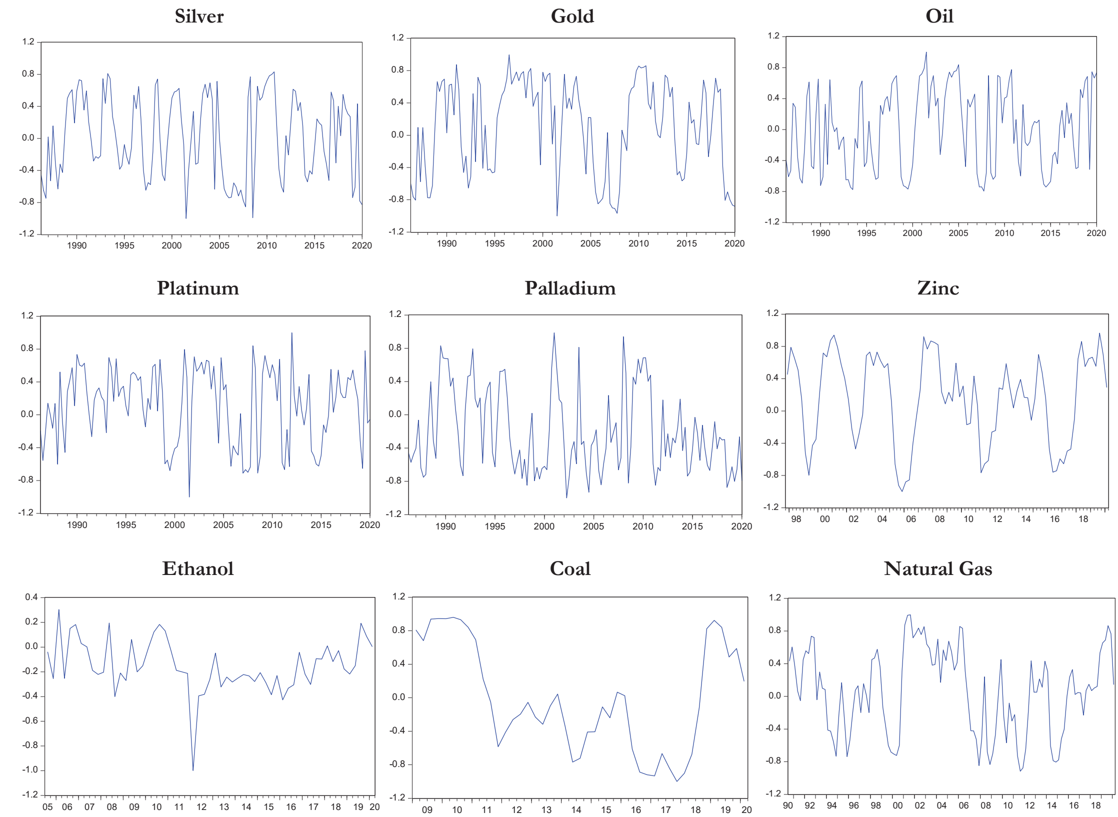

Figure 3. Dynamic correlation between the yield spread and four quarters ahead.

Second, using the Bai and Perron (2003) structural break test we find that the early 2000s is the period associated with structural breaks in the relationship between the yield curve and the prices of commodities. Table 4 presents the results of this test and indicates the dates detected as structural break points. In general, the findings of Bai and Perron (2003) test point to the 2003–3004 as the period in which there was a structural break in the relationship between yield spreads and future commodity prices. These findings accord with earlier studies that date the start of the financialization commodities to the early 2000s (Hamilton and Wu 2015; Henderson et al. 2015). In other words, this period marks the start of the increased exposure of portfolio managers, individuals and hedge funds to commodities.

Table 4. Bai and Perron (2003) multiple break-point test.

	h = 2 (Two Quarters Ahead)								
	Silver	**Oil**	**Gold**	**PLTNM**	**PLDM**	**Zinc**	**ETHNL**	**Coal**	**NatGas**
Break Point #1	2003Q2	=	2001Q2	=	=	=	=	=	=
Break Point #2	2011Q2	=	2012Q3	=	=	=	=	=	=
Break Point #3	=	=	=	=	=	=	=	=	=

Table 4. *Cont.*

h = 3 (Three Quarters Ahead)									
	Silver	**Oil**	**Gold**	**PLTNM**	**PLDM**	**Zinc**	**ETHNL**	**Coal**	**NatGas**
Break Point #1	2003Q2	=	2001Q1	=	1996Q3	=	=	2011Q1	=
Break Point #2	2011Q1	=	2012Q2	=	=	=	=	2015Q4	=
Break Point #3	=	=	=	=	=	=	=	=	=
h = 4 (Four Quarters Ahead)									
	Silver	**Oil**	**Gold**	**PLTNM**	**PLDM**	**Zinc**	**ETHNL**	**Coal**	**NatGas**
Break Point #1	2003Q1	1995Q4	2001Q2	1998Q4	1996Q2	=	=	2011Q1	2000Q3
Break Point #2	2011Q1	=	2012Q2	2010Q4	2001Q2	=	=	2015Q4	2006Q2
Break Point #3	=	=	=	=	2008Q4	=	=	=	=

Notes: We tested for any structural break using the Bai and Perron (2003) multiple break-point test. The values listed in the table are those of the break dates. The vast majority of the commodities point to the 2000s as the structural break points. "=" denotes that no significant breakpoint was detected between the future return and the current yield spread.

This finding is in line with prior studies documenting the weakening ability of the term structure to predict future economic activity. Early on, Stock and Watson (2003) and Giacomini and Rossi (2006) maintained that the yield spread's ability to forecast economic expansion has weakened since the 1980s, but its predictive ability remains strong only for recessions. Other works raise questions regarding the stability of the term spread's predictive content (e.g., Wheelock and Wohar 2009). Evgenidis et al. (2020) confirm the time-varying nature of the yield spread's predictive ability, mainly during the 2000s.

One possible factor explaining this break between commodities and the most reliable indicator of future economic activity is the financialization of commodities. For a long time, commodities were viewed as a segmented market offering significant diversification benefits in light of the low—even negative—correlation between their returns and the stock market (e.g., Bodie and Rosansky 1980; Demiralay et al. 2019). This characteristic prompted traders, financial institutions and institutional investors to consider this new asset class as a useful diversifier in their portfolios. A byproduct of this development is the acceleration in the financialization of these commodities, which in turn fueled a rapid increase in their co-movements with equity markets (e.g., Qadan et al. 2019). This evolution may explain the breakdown of the obvious relationship between commodities and the expected economic evolution.

We also test the extent to which the dynamic correlation of commodity "i" co-moves with that of commodity "j." A quick glance at Figures 1–3 shows the apparent co-movements between some of these commodities. Table 5 presents the simple correlation between the DCC values. Some of the correlation values are negative and statistically significant. For example, we detect a negative correlation between the prices of ethanol and gold, gold and natural gas, natural gas and palladium, oil, and palladium. On the other hand, the majority of the other cases are associated with statistically significant positive correlations, particularly for precious metals. For example, the correlation between the DCC values of gold and silver is 0.73, and that between gold and platinum is 0.534. Overall, this picture reveals that the conditional slope among commodities is largely connected—a clear indication of their similar reaction to the current yield spread.

Table 5. Pearson's correlation between the DCC values of the commodities.

	Coal	**Ethanol**	**Gold**	**Nat.Gas**	**Oil**	**PLDM**	**PLTNM**	**Silver**	**Zinc**
Coal	1.00								
Ethanol	**0.647 *****	1.00							
Gold	0.090	**−0.230 ***	1.00						
Nat.Gas	0.124	**0.321 ****	**−0.153 ***	1.00					
Oil	**0.692 *****	**0.519 *****	**0.167 ***	**0.627 *****	1.00				
PLDM	**0.407 *****	0.205	**0.166 ***	**−0.222 ****	**−0.141 ***	1.00			

Table 5. *Cont.*

	Coal	Ethanol	Gold	Nat.Gas	Oil	PLDM	PLTNM	Silver	Zinc
PLTNM	0.120	−0.037	**0.534** ***	**0.354** ***	**0.406** ***	**0.327** ***	1.00		
Silver	0.249	−0.071	**0.730** ***	−0.065	0.091	**0.487** ***	**0.630** ***	1.00	
Zinc	**0.493** ***	**0.277** **	−0.154	0.086	**0.196** *	**0.278** ***	0.081	0.159	1.00

Notes: ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

6. Robustness Checks

6.1. Additional Proxies for Yield

Previous studies suggested capturing the yield spread using different proxies. To develop a broader picture regarding the interaction between the yield spread and the future evolution in commodity prices, we depart from the standard yield spread used in the literature (Y_{10}-Y_{3M}) and test other proxies: the difference between 10-year and 1-year interest rates (Y_{10}-Y_1), between 10-year and 2-year yields (Y_{10}-Y_2), between 30-year and 3-months yields (Y_{30}-Y_{3M}), between 30-year and 1-year bond yields (Y_{30}-Y_1) and between 30-year and 2-year bond yields (Y_{30}-Y_2).

Tables 6–10 report the estimation results of the prediction model. Table 6 presents the estimation results given Y_{10}-Y_1 as the yield spread. In Table 7, Y_{10}-Y_2 proxies for the yield spread. In Table 8, Y_{30}-Y_{3M} proxies for the yield spread. Table 9 utilizes Y_{30}-Y_1 as the yield spread, and Table 10 utilizes Y_{30}-Y_2 to proxy for the yield spread. The overall picture is maintained as evident by the significant positive coefficients in the period prior to 2004 (Panel B in each table), but the insignificant results in the period that follows (Panels C). The regression results that include the other explanatory variables reflect very similar picture. They appear in Tables A3–A7 in the online Appendix A.

A closer glance at the results in Panel B of Table 6 confirms that the yield spread, defined as Y_{10}-Y_1, is an efficient predictor of the future prices of silver, gold, platinum and zinc. The resulting R^2 for silver ranges between 0.03 when forecasting one quarter ahead (h = 1), and 0.19 when forecasting the prices one year ahead (h = 4). We find that the regression R^2 for gold ranges between 0.04 (for h = 1) and 0.15 (for h = 4), for platinum it ranges between 0.01 (for h = 1) and 0.12 (for h = 4), and finally it ranges between 0.09 (for h = 1) and 0.26 for zinc (h = 4). This picture is essentially replicated in Tables 7–10. Moreover, Panel B of Table 10 provides strong support for these findings. The resulting R^2 for zinc ranges between 0.08 (for h = 1), and 0.34 (for h = 4). By and large, these findings confirm the premise that metal prices are positively correlated with macroeconomic activity (e.g., Fama and French 1988).

Table 6. Estimation results of Equation (1) with the Y_{10}-Y_1 indicator. Panel A: entire sample. Panel B: sample period 1986–2003. Panel C: sample period 2004–2020.

Panel A												
Forecast Horizon	**Oil (1986:01–2020:12)**				**Silver (1986:01–2020:12)**				**Gold (1986:01–2020:12)**			
h	**C**	**(Y_{10}-Y_1)**	**R^2**	**N**	**C**	**(Y_{10}-Y_1)**	**R^2**	**N**	**C**	**(Y_{10}-Y_1)**	**R^2**	**N**
1	10.51 (0.25)	−4.45 (0.38)	0.003	139	3.28 (0.61)	1.01 (0.79)	0.000	139	3.82 (0.40)	0.84 (0.72)	0.001	138
2	8.59 (0.30)	−3.40 (0.47)	0.004	138	2.60 (0.67)	1.43 (0.70)	0.002	138	3.65 (0.41)	0.94 (0.68)	0.003	138
3	7.34 (0.35)	−2.71 (0.55)	0.005	137	1.03 (0.86)	2.26 (0.52)	0.007	137	3.41 (0.42)	0.94 (0.67)	0.004	137
4	4.95 (0.51)	−1.56 (0.72)	0.002	136	−0.65 (0.90)	3.19 (0.35)	0.021	136	3.30 (0.42)	0.89 (0.67)	0.004	136

Table 6. *Cont.*

Forecast Horizon	Platinum (1986:04–2020:12)				Palladium (1986:04–2020:12)				Zinc (1997:08–2020:12)			
Panel A												
h	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N
1	3.98 (0.45)	−0.88 (0.78)	0.000	139	18.34 ** (0.04)	−6.63 (0.25)	0.01	139	3.24 (0.80)	−0.06 (0.99)	0.000	93
2	3.15 (0.54)	−0.49 (0.87)	0.0003	138	15.69 * (0.08)	−4.79 (0.39)	0.01	138	0.19 (0.99)	2.13 (0.71)	0.003	92
3	1.71 (0.73)	0.23 (0.94)	0.0001	136	12.24 (0.18)	−2.59 (0.63)	0.004	137	−2.26 (0.84)	3.74 (0.46)	0.01	91
4	0.49 (0.92)	0.94 (0.72)	0.002	136	10.16 (0.27)	−1.10 (0.83)	0.001	136	−4.33 (0.69)	5.12 (0.27)	0.03	90
Forecast Horizon	**Ethanol (2005:06–2020:12)**				**Coal (2009:01–2020:12)**				**Natural gas (1990:05–2020:12)**			
h	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N
1	2.00 (0.88)	−1.97 (0.80)	0.001	62	−4.91 (0.79)	3.21 (0.71)	0.004	48	5.07 (0.71)	−2.16 (0.76)	0.001	122
2	0.20 (0.98)	−1.70 (0.78)	0.001	61	−8.31 (0.61)	4.90 (0.52)	0.02	47	2.75 (0.82)	−0.90 (0.89)	0.0002	121
3	−1.68 (0.84)	−0.72 (0.90)	0.000	60	−12.06 (0.40)	6.51 (0.34)	0.04	46	0.96 (0.93)	−0.12 (0.98)	0.000	120
4	−4.84 (0.46)	0.56 (0.91)	0.000	59	−11.26 (0.42)	5.98 (0.37)	0.04	45	−1.10 (0.91)	1.19 (0.82)	0.001	119
Panel B												
Forecast Horizon	**Oil (1986:01–2003:12)**				**Silver (1986:01–2003:12)**				**Gold (1986:01–2003:12)**			
h	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N
1	12.42 (0.31)	−4.79 (0.43)	0.005	71	−7.50 (0.15)	**6.26 *** **(0.07)**	0.03	71	−5.59 (0.14)	**5.00 **** **(0.03)**	0.04	70
2	10.64 (0.33)	−3.56 (0.50)	0.006	71	−9.37 * (0.05)	**8.28 **** **(0.03)**	0.11	71	−5.59 * (0.07)	**5.06 ***** **(0.01)**	0.10	71
3	9.74 (0.32)	−3.03 (0.53)	0.009	71	−8.99 ** (0.03)	**7.64 **** **(0.01)**	0.14	71	−5.41 ** (0.04)	**4.52 ***** **(0.004)**	0.13	71
4	8.86 (0.34)	−2.35 (0.62)	0.007	71	−8.93 ** (0.01)	**7.60 ***** **(0.003)**	0.19	71	−5.18 ** (0.04)	**4.26 ***** **(0.003)**	0.15	71
Forecast Horizon	**Platinum (1986:04–2003:12)**				**Palladium (1986:04–2003:12)**				**Zinc (1997:08–2003:12)**			
h	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N	C	$(Y_{10}\text{-}Y_1)$	R^2	N
1	−0.64 (0.93)	3.39 (0.35)	0.01	71	14.98 (0.29)	−8.42 (0.31)	0.02	71	−16.75 (0.14)	9.49 (0.12)	0.09	25
2	−2.72 (0.68)	5.09 (0.12)	0.05	71	8.46 (0.54)	−2.89 (0.72)	0.004	71	−14.68 (0.15)	**10.01 *** **(0.09)**	0.17	25
3	−3.93 (0.52)	**5.45 *** **(0.07)**	0.09	70	4.74 (0.73)	−0.63 (0.93)	0.000	71	−13.22 (0.18)	**8.97 *** **(0.09)**	0.20	25
4	−4.48 (0.43)	**5.85 *** **(0.04)**	0.12	71	2.72 (0.84)	0.80 (0.91)	0.001	71	−12.50 (0.17)	**9.12 *** **(0.06)**	0.26	25

Table 6. *Cont.*

Panel B

Forecast Horizon	Natural Gas (1990:05–2003:12)			
h	C	$(Y_{10}-Y_1)$	R^2	N
1	10.73 (0.67)	−0.22 (0.99)	0.000	54
2	2.97 (0.89)	4.30 (0.66)	0.004	54
3	−0.41 (0.98)	6.44 (0.47)	0.02	54
4	−1.84 (0.92)	7.84 (0.30)	0.04	54

Panel C

Forecast Horizon	Oil (2004:01–2020:12)				Silver (2004:01–2020:12)				Gold (2004:01–2020:12)			
h	C	$(Y_{10}-Y_1)$	R^2	N	C	$(Y_{10}-Y_1)$	R^2	N	C	$(Y_{10}-Y_1)$	R^2	N
1	8.26 (0.60)	−3.95 (0.64)	0.002	68	15.22 (0.10)	−4.46 (0.46)	0.006	68	13.82 ** (0.01)	−3.34 (0.29)	0.02	68
2	6.04 (0.65)	−2.99 (0.69)	0.002	67	15.87 * (0.06)	−5.53 (0.34)	0.019	67	14.21 *** (0.01)	−3.52 (0.26)	0.03	67
3	4.29 (0.73)	−2.10 (0.76)	0.002	66	12.56 (0.12)	−3.41 (0.55)	0.012	66	13.88 *** (0.01)	−3.18 (0.29)	0.04	66
4	−0.09 (0.99)	−0.28 (0.97)	0.000	65	9.14 (0.25)	−1.59 (0.78)	0.004	65	13.71 *** (0.004)	−3.12 (0.26)	0.05	65
Forecast Horizon	Platinum (2004:01–2020:12)				Palladium (2004:01–2020:12)				Zinc (2004:01–2020:12)			
h	C	$(Y_{10}-Y_1)$	R^2	N	C	$(Y_{10}-Y_1)$	R^2	N	C	$(Y_{10}-Y_1)$	R^2	N
1	8.60 (0.24)	−4.83 (0.30)	0.01	68	22.90 ** (0.01)	−5.63 (0.44)	0.01	68	13.37 (0.46)	−4.91 (0.58)	0.01	68
2	9.07 (0.19)	−5.64 (0.23)	0.03	67	24.29 *** (0.004)	−7.25 (0.32)	0.02	67	7.93 (0.65)	−1.80 (0.82)	0.002	67
3	7.55 (0.25)	−4.66 (0.27)	0.03	66	21.46 ** (0.01)	−5.25 (0.43)	0.02	66	3.58 (0.83)	0.97 (0.89)	0.000	66
4	5.67 (0.37)	−3.65 (0.33)	0.03	65	19.62 ** (0.03)	−3.82 (0.54)	0.01	65	0.15 (0.99)	2.97 (0.66)	0.01	65

Forecast Horizon	Natural Gas (2004:01–2020:12)			
h	C	$(Y_{10}-Y_1)$	R^2	N
1	0.56 (0.97)	−3.69 (0.63)	0.002	68
2	2.48 (0.86)	−4.99 (0.49)	0.01	67
3	1.93 (0.88)	−5.27 (0.41)	0.01	66
4	−0.77 (0.94)	−3.97 (0.47)	0.01	65

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_1 denotes the yield spread calculated as the difference between the yield rates on 10-year and 1-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7. Estimation results of Equation (1) with the Y_{10}-Y_2 indicator. Panel A: entire sample. Panel B: sample period 1986–2003. Panel C: sample period 2004–2020.

Panel A												
Forecast Horizon	**Oil (1986:01–2020:12)**				**Silver (1986:01–2020:12)**				**Gold (1986:01–2020:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	11.27 (0.20)	−6.39 (0.30)	0.004	139	2.30 (0.70)	2.18 (0.65)	0.001	139	2.96 (0.48)	1.86 (0.50)	0.004	138
2	8.47 (0.27)	−4.25 (0.44)	0.004	138	1.81 (0.75)	2.56 (0.58)	0.004	138	2.87 (0.48)	1.91 (0.47)	0.01	138
3	7.27 (0.32)	−3.41 (0.52)	0.005	137	0.69 (0.90)	3.20 (0.47)	0.01	137	2.61 (0.51)	1.92 (0.45)	0.01	137
4	5.27 (0.45)	−2.29 (0.66)	0.003	136	−0.55 (0.91)	3.99 (0.35)	0.02	136	2.50 (0.51)	1.86 (0.45)	0.01	136
Forecast Horizon	**Platinum (1986:04–2020:12)**				**Palladium (1986:04–2020:12)**				**Zinc (1997:08–2020:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	3.51 (0.50)	−0.71 (0.85)	0.0002	139	16.64 * (0.06)	−6.94 (0.33)	0.01	139	3.79 (0.76)	−0.54 (0.94)	0.000	93
2	2.90 (0.55)	−0.40 (0.91)	0.0001	138	15.04 * (0.09)	−5.55 (0.42)	0.01	138	1.10 (0.93)	1.78 (0.78)	0.002	92
3	2.18 (0.65)	−0.13 (0.97)	0.000	136	12.77 (0.15)	−3.80 (0.57)	0.01	137	−0.98 (0.93)	3.41 (0.56)	0.01	91
4	1.29 (0.77)	0.49 (0.88)	0.0005	136	10.88 (0.21)	−2.05 (0.74)	0.003	136	−2.64 (0.80)	4.73 (0.38)	0.02	90
Forecast Horizon	**Ethanol (2005:06–2020:12)**				**Coal (2009:01–2020:12)**				**Natural gas (1990:05–2020:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	1.64 (0.91)	−1.94 (0.84)	0.0004	62	−4.20 (0.82)	3.23 (0.74)	0.003	48	6.92 (0.59)	−4.28 (0.62)	0.002	122
2	−0.35 (0.97)	−1.50 (0.84)	0.001	61	−6.71 (0.69)	4.59 (0.62)	0.01	47	5.23 (0.65)	−3.24 (0.67)	0.002	121
3	−1.33 (0.88)	−1.11 (0.87)	0.001	60	−9.94 (0.51)	6.12 (0.47)	0.03	46	3.88 (0.72)	−2.62 (0.71)	0.002	120
4	−4.08 (0.55)	0.05 (0.99)	0.000	59	−8.80 (0.56)	5.31 (0.53)	0.03	45	1.61 (0.87)	−0.77 (0.90)	0.0003	119
Panel B												
Forecast Horizon	**Oil (1986:01–2003:12)**				**Silver (1986:01–2003:12)**				**Gold (1986:01–2003:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	12.81 (0.24)	−7.33 (0.31)	0.01	71	−6.76 (0.17)	**8.23 *** **(0.07)**	0.04	71	−5.11 (0.15)	**6.71 **** **(0.02)**	0.05	70
2	10.36 (0.30)	−4.83 (0.46)	0.01	71	−8.65 * (0.06)	**11.18 **** **(0.02)**	0.13	71	−5.24 * (0.07)	**6.93 ***** **(0.004)**	0.13	71
3	8.78 (0.34)	−3.34 (0.57)	0.01	71	−8.73 ** (0.03)	**10.73 ***** **(0.01)**	0.18	71	−5.34 ** (0.03)	**6.45 ***** **(0.001)**	0.18	71
4	7.32 (0.39)	−1.74 (0.77)	0.003	71	−8.63 *** (0.01)	**10.65 ***** **(0.001)**	0.24	71	−5.25 ** (0.03)	**6.22 ***** **(0.00)**	0.21	71

Table 7. *Cont.*

Panel B												
Forecast Horizon	**Platinum (1986:04–2003:12)**				**Palladium (1986:04–2003:12)**				**Zinc (1997:08–2003:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	−0.19 (0.98)	4.41 (0.37)	0.01	71	14.57 (0.25)	−11.71 (0.27)	0.02	71	−14.39 (0.22)	10.23 (0.17)	0.08	25
2	−1.92 (0.77)	6.49 (0.15)	0.05	71	9.81 (0.44)	−5.60 (0.58)	0.01	71	−13.13 (0.19)	11.82 * (0.09)	0.18	25
3	−2.94 (0.62)	**6.78 * (0.09)**	0.09	70	6.75 (0.59)	−3.07 (0.75)	0.004	71	−12.20 (0.18)	10.99 * (0.06)	0.22	25
4	−3.43 (0.54)	**7.32 ** (0.05)**	0.13	71	4.49 (0.71)	−0.74 (0.93)	0.0003	71	−11.53 (0.16)	11.25 ** (0.02)	0.29	25
Forecast Horizon	**Natural Gas (1990:05–2003:12)**											
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**								
1	12.09 (0.59)	−1.61 (0.91)	0.0002	54								
2	7.03 (0.72)	2.16 (0.85)	0.001	54								
3	3.90 (0.83)	4.94 (0.64)	0.01	54								
4	1.86 (0.91)	7.49 (0.40)	0.02	54								
Panel C												
Forecast Horizon	**Oil (2004:01–2020:12)**				**Silver (2004:01–2020:12)**				**Gold (2004:01–2020:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	9.12 (0.58)	−5.25 (0.61)	0.002	68	14.73 (0.13)	−4.77 (0.53)	0.005	68	13.80 *** (0.01)	−3.85 (0.31)	0.02	68
2	5.56 (0.67)	−3.09 (0.71)	0.002	67	15.76 * (0.08)	−6.31 (0.38)	0.02	67	14.40 *** (0.01)	−4.22 (0.26)	0.03	67
3	4.61 (0.71)	−2.68 (0.73)	0.002	66	13.68 (0.10)	−4.82 (0.49)	0.02	66	14.39 *** (0.01)	−4.07 (0.25)	0.04	66
4	1.36 (0.91)	−1.44 (0.85)	0.001	65	10.82 (0.17)	−3.11 (0.64)	0.01	65	14.36 *** (0.00)	−4.11 (0.22)	0.06	65
Forecast Horizon	**Platinum (2004:01–2020:12)**				**Palladium (2004:01–2020:12)**				**Zinc (2004:01–2020:12)**			
h	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**	**C**	**(Y_{10}-Y_2)**	**R^2**	**N**
1	7.53 (0.31)	−4.74 (0.36)	0.01	68	21.92 ** (0.02)	−5.74 (0.50)	0.01	68	15.05 (0.41)	−6.99 (0.51)	0.01	68
2	8.12 (0.25)	−5.78 (0.26)	0.02	67	23.87 *** (0.01)	−8.05 (0.33)	0.02	67	10.20 (0.57)	−3.84 (0.69)	0.01	67
3	7.92 (0.23)	−5.67 (0.24)	0.03	66	22.88 *** (0.01)	−7.17 (0.36)	0.03	66	6.35 (0.71)	−1.01 (0.91)	0.001	66
4	6.56 (0.29)	−4.90 (0.25)	0.04	65	21.70 *** (0.01)	−6.01 (0.40)	0.03	65	3.36 (0.84)	0.99 (0.90)	0.001	65

Table 7. *Cont.*

Panel C				
Forecast Horizon	Natural Gas (2004:01–2020:12)			
h	C	(Y_{10}-Y_2)	R^2	N
1	0.76 (0.96)	−4.42 (0.63)	0.002	68
2	2.15 (0.88)	−5.52 (0.52)	0.01	67
3	2.29 (0.86)	−6.38 (0.40)	0.01	66
4	−0.56 (0.96)	−4.76 (0.47)	0.01	65

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_2 denotes the yield spread calculated as the difference between the yield rates on 10-year and 2-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 8. Estimation results of Equation (1) with the Y_{30}-Y_{3M} indicator. Panel A: entire sample. Panel B: sample period 1986–2003. Panel C: sample period 2004–2020.

Panel A												
Forecast Horizon	Oil (1986:01–2020:12)				Silver (1986:01–2020:12)				Gold (1986:01–2020:12)			
h	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N
1	10.14 (0.32)	−2.70 (0.48)	0.002	139	3.39 (0.63)	0.60 (0.84)	0.0003	139	4.43 (0.37)	0.27 (0.88)	0.0002	138
2	8.54 (0.36)	−2.16 (0.54)	0.003	138	2.55 (0.70)	0.95 (0.73)	0.001	138	4.26 (0.37)	0.33 (0.85)	0.001	138
3	7.38 (0.40)	−1.77 (0.60)	0.004	137	0.96 (0.88)	1.49 (0.57)	0.01	137	3.99 (0.38)	0.34 (0.83)	0.001	137
4	4.49 (0.59)	−0.80 (0.80)	0.001	136	−0.53 (0.93)	2.00 (0.42)	0.01	136	3.99 (0.36)	0.27 (0.87)	0.001	136
Forecast Horizon	Platinum (1986:04–2020:12)				Palladium (1986:04–2020:12)				Zinc (1997:08–2020:12)			
h	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N
1	4.83 (0.43)	−0.96 (0.70)	0.001	139	19.20 * (0.07)	−4.65 (0.29)	0.01	139	1.46 (0.92)	0.76 (0.87)	0.0004	93
2	3.92 (0.51)	−0.67 (0.78)	0.001	138	16.41 (0.11)	−3.41 (0.42)	0.01	138	−2.14 (0.87)	2.40 (0.57)	0.01	92
3	1.72 (0.76)	0.15 (0.95)	0.000	136	11.49 (0.27)	−1.33 (0.74)	0.002	137	−5.11 (0.68)	3.66 (0.33)	0.02	91
4	0.34 (0.95)	0.68 (0.73)	0.002	136	8.62 (0.40)	−0.01 (0.99)	0.000	136	−7.57 (0.52)	4.71 (0.17)	0.05	90
Forecast Horizon	Ethanol (2005:06–2020:12)				Coal (2009:01–2020:12)				Natural gas (1990:05–2020:12)			
h	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N
1	1.03 (0.95)	−0.81 (0.89)	0.0002	62	−10.26 (0.62)	4.11 (0.53)	0.01	48	4.35 (0.78)	−1.08 (0.85)	0.0002	122
2	−0.25 (0.98)	−0.87 (0.85)	0.001	61	−13.06 (0.47)	4.96 (0.39)	0.02	47	2.07 (0.88)	−0.28 (0.95)	0.000	121
3	−1.98 (0.84)	−0.33 (0.94)	0.0002	60	−17.59 (0.26)	6.28 (0.22)	0.06	46	−0.74 (0.95)	0.66 (0.89)	0.0003	120
4	−5.05 (0.51)	0.45 (0.90)	0.001	59	−16.06 (0.30)	5.65 (0.27)	0.06	45	−3.24 (0.77)	1.70 (0.68)	0.003	119

Table 8. *Cont.*

Panel B												
Forecast Horizon	**Oil (1986:01–2003:12)**				**Silver (1986:01–2003:12)**				**Gold (1986:01–2003:12)**			
h	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**
1	13.64 (0.31)	−3.74 (0.42)	0.01	71	−8.02 (0.14)	**4.36 * (0.07)**	0.03	71	−5.61 (0.19)	**3.29 * (0.05)**	0.03	70
2	10.91 (0.35)	−2.46 (0.52)	0.01	71	−10.57 ** (0.04)	**6.02 ** (0.03)**	0.11	71	−5.54 (0.11)	**3.29 ** (0.02)**	0.08	71
3	10.37 (0.32)	−2.29 (0.52)	0.01	71	−10.29 ** (0.02)	**5.64 *** (0.01)**	0.14	71	−5.37 * (0.06)	**2.95 *** (0.01)**	0.11	71
4	8.64 (0.38)	−1.44 (0.68)	0.005	71	−9.87 *** (0.01)	**5.44 *** (0.004)**	0.18	71	−4.89* (0.08)	**2.65 ** (0.02)**	0.11	71
Forecast Horizon	**Platinum (1986:04–2003:12)**				**Palladium (1986:04–2003:12)**				**Zinc (1997:08–2003:12)**			
h	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**
1	−1.58 (0.85)	2.69 (0.34)	0.01	71	13.62 (0.40)	−4.86 (0.43)	0.01	71	−18.45 (0.11)	**7.06 * (0.09)**	0.10	25
2	−4.02 (0.57)	3.98 (0.10)	0.05	71	6.92 (0.65)	−1.14 (0.85)	0.001	71	−17.37 * (0.09)	**7.92 * (0.05)**	0.20	25
3	−5.57 (0.39)	**4.38 ** (0.04)**	0.10	70	1.89 (0.90)	0.97 (0.86)	0.001	71	−16.08 * (0.09)	**7.34 ** (0.04)**	0.25	25
4	−6.05 (0.30)	**4.60 ** (0.02)**	0.14	71	−0.58 (0.97)	2.13 (0.69)	0.01	71	−15.39 * (0.08)	**7.45 ** (0.02)**	0.33	25
Forecast Horizon	**Natural gas (1990:05–2003:12)**											
h	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**								
1	10.84 (0.71)	−0.20 (0.98)	0.000	54								
2	1.57 (0.95)	3.48 (0.67)	0.01	54								
3	−3.79 (0.87)	5.79 (0.43)	0.02	54								
4	−6.18 (0.75)	7.15 (0.24)	0.06	54								

Panel C												
Forecast Horizon	**Oil (2004:01–2020:12)**				**Silver (2004:01–2020:12)**				**Gold (2004:01–2020:12)**			
h	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**	**C**	**$(Y_{30}$-$Y_{3M})$**	**R^2**	**N**
1	5.77 (0.75)	−1.43 (0.82)	0.0004	68	17.50 * (0.09)	−3.79 (0.42)	0.01	68	16.24 *** (0.01)	−3.14 (0.18)	0.03	68
2	5.29 (0.73)	−1.57 (0.78)	0.001	67	18.60 ** (0.04)	−4.66 (0.28)	0.02	67	16.74 *** (0.003)	−3.30 (0.16)	0.05	67
3	3.24 (0.82)	−0.89 (0.87)	0.001	66	15.15 * (0.07)	−3.26 (0.42)	0.02	66	16.34 *** (0.002)	−3.05 (0.16)	0.06	66
4	−1.50 (0.91)	0.41 (0.93)	0.0002	65	11.55 (0.15)	−2.02 (0.60)	0.01	65	16.08 *** (0.001)	−2.98 (0.13)	0.07	65

Table 8. *Cont.*

Panel C												
Forecast Horizon	Platinum (2004:01–2020:12)				Palladium (2004:01–2020:12)				Zinc (2004:01–2020:12)			
h	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N	C	(Y_{30}-Y_{3M})	R^2	N
1	11.93 (0.14)	−4.47 (0.21)	0.01	68	27.18 *** (0.01)	−5.38 (0.33)	0.01	68	12.91 (0.54)	−2.90 (0.67)	0.004	68
2	12.74 (0.11)	−5.13 (0.16)	0.04	67	28.83 *** (0.003)	−6.52 (0.22)	0.03	67	6.80 (0.73)	−0.66 (0.92)	0.0004	67
3	9.95 (0.16)	−3.96 (0.19)	0.04	66	24.38 *** (0.01)	−4.56 (0.34)	0.03	66	1.45 (0.94)	1.52 (0.79)	0.003	66
4	7.52 (0.27)	−3.09 (0.24)	0.04	65	21.37 ** (0.04)	−3.16 (0.47)	0.02	65	−2.77 (0.88)	3.11 (0.55)	0.02	65
Forecast Horizon	Natural gas (2004:01–2020:12)											
h	C	(Y_{30}-Y_{3M})	R^2	N								
1	−1.68 (0.92)	−1.37 (0.81)	0.000	68								
2	1.87 (0.91)	−2.89 (0.61)	0.004	67								
3	1.14 (0.93)	−3.01 (0.55)	0.01	66								
4	−1.64 (0.89)	−2.16 (0.62)	0.01	65								

Notes: The forecast horizon (h) is in quarters. Y_{30}-Y_{3M} denotes the yield spread calculated as the difference between the yield rates on 10-year and 3-month government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 9. Estimation results of Equation (1) with the Y_{30}-Y_1 indicator. Panel A: entire sample. Panel C: sample period 2004–2020.

Panel A												
Forecast Horizon	Oil (1986:01–2020:12)				Silver (1986:01–2020:12)				Gold (1986:01–2020:12)			
h	C	(Y_{30}-Y_1)	R^2	N	C	(Y_{30}-Y_1)	R^2	N	C	(Y_{30}-Y_1)	R^2	N
1	9.77 (0.28)	−2.90 (0.43)	0.002	139	1.75 (0.79)	1.55 (0.63)	0.002	139	3.58 (0.44)	0.74 (0.69)	0.001	138
2	7.95 (0.34)	−2.18 (0.52)	0.003	138	1.51 (0.81)	1.63 (0.59)	0.004	138	3.40 (0.45)	0.83 (0.65)	0.003	138
3	7.13 (0.36)	−1.90 (0.57)	0.004	137	0.45 (0.94)	1.98 (0.49)	0.01	137	3.09 (0.47)	0.86 (0.62)	0.005	137
4	4.93 (0.51)	−1.15 (0.72)	0.002	136	−0.66 (0.90)	2.37 (0.39)	0.02	136	3.07 (0.46)	0.78 (0.64)	0.01	136
Forecast Horizon	Platinum (1986:04–2020:12)				Palladium (1986:04–2020:12)				Zinc (1997:08–2020:12)			
h	C	(Y_{30}-Y_1)	R^2	N	C	(Y_{30}-Y_1)	R^2	N	C	(Y_{30}-Y_1)	R^2	N
1	3.68 (0.52)	−0.50 (0.84)	0.0002	139	17.84 * (0.07)	−4.63 (0.32)	0.01	139	1.52 (0.91)	0.81 (0.87)	0.0004	93
2	3.41 (0.53)	−0.50 (0.84)	0.0005	138	16.10 * (0.09)	−3.75 (0.40)	0.01	138	−1.001 (0.94)	2.08 (0.64)	0.005	92
3	2.09 (0.69)	−0.03 (0.99)	0.000	136	12.74 (0.18)	−2.18 (0.60)	0.01	137	−3.21 (0.80)	3.09 (0.45)	0.01	91
4	0.96 (0.84)	0.46 (0.83)	0.001	136	10.23 (0.27)	−0.85 (0.83)	0.001	136	−5.19 (0.66)	4.02 (0.28)	0.03	90

Table 9. *Cont.*

Panel A

Forecast Horizon	**Ethanol (2005:06–2020:12)**				**Coal (2009:01–2020:12)**				**Natural gas (1990:05–2020:12)**			
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**
1	−0.37 (0.98)	−0.20 (0.97)	0.000	62	−7.99 (0.70)	3.47 (0.61)	0.01	48	4.98 (0.73)	−1.52 (0.79)	0.0005	122
2	−1.43 (0.90)	−0.38 (0.94)	0.000	61	−10.36 (0.57)	4.21 (0.50)	0.02	47	2.60 (0.84)	−0.58 (0.91)	0.0001	121
3	−1.98 (0.83)	−0.36 (0.93)	0.0002	60	−14.18 (0.38)	5.35 (0.34)	0.04	46	0.99 (0.93)	−0.10 (0.98)	0.000	120
4	−4.89 (0.51)	0.41 (0.91)	0.000	59	−12.75 (0.04)	4.74 (0.39)	0.04	45	−1.07 (0.92)	0.85 (0.84)	0.001	119

Panel B

Forecast Horizon	**Oil (1986:01–2003:12)**				**Silver (1986:01–2003:12)**				**Gold (1986:01–2003:12)**			
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**
1	13.74 (0.24)	−4.59 (0.30)	0.01	71	−8.56 * (0.09)	**5.60 ** (0.04)**	0.04	71	−5.65 (0.16)	**4.02 ** (0.03)**	0.04	70
2	10.92 (0.30)	−2.99 (0.44)	0.01	71	−10.78 ** (0.03)	**7.42 ** (0.02)**	0.14	71	−6.00 * (0.07)	**4.26 *** (0.005)**	0.12	71
3	9.39 (0.32)	−2.20 (0.54)	0.01	71	−10.63 *** (0.01)	**7.04 *** (0.003)**	0.19	71	−6.14 ** (0.03)	**4.03 *** (0.001)**	0.17	71
4	7.91 (0.37)	−1.31 (0.71)	0.004	71	−10.29 *** (0.005)	**6.85 *** (0.001)**	0.24	71	−5.87 ** (0.03)	**3.79 *** (0.001)**	0.19	71

Forecast Horizon	**Platinum (1986:04–2003:12)**				**Palladium (1986:04–2003:12)**				**Zinc (1997:08–2003:12)**			
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**
1	−0.97 (0.90)	2.89 (0.35)	0.01	71	15.75 (0.27)	−7.15 (0.30)	0.02	71	−17.19 (0.15)	7.11 (0.12)	0.09	25
2	−3.10 (0.65)	4.27 (0.12)	0.05	71	9.86 (0.48)	−3.11 (0.64)	0.01	71	−16.06 (0.12)	**8.03 * (0.07)**	0.19	25
3	−4.49 (0.47)	**4.66 * (0.05)**	0.10	70	5.65 (0.68)	−1.04 (0.86)	0.001	71	−15.05 (0.11)	**7.54 ** (0.04)**	0.24	25
4	−5.10 (0.36)	**5.02 ** (0.02)**	0.15	71	2.85 (0.83)	0.56 (0.92)	0.0004	71	−14.51* (0.09)	**7.75 ** (0.02)**	0.32	25

Forecast Horizon	**Natural gas (1990:05–2003:12)**			
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**
1	12.74 (0.62)	−1.23 (0.90)	0.0002	54
2	4.83 (0.83)	2.35 (0.77)	0.002	54
3	0.74 (0.97)	4.38 (0.54)	0.01	54
4	−1.10 (0.95)	5.68 (0.34)	0.03	54

Table 9. *Cont.*

Panel C												
Forecast Horizon	**Oil (2004:01–2020:12)**				**Silver (2004:01–2020:12)**				**Gold (2004:01–2020:12)**			
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**
1	4.55 (0.79)	−0.99 (0.87)	0.0002	68	15.55 (0.13)	−3.23 (0.52)	0.01	68	15.62 *** (0.01)	−3.15 (0.23)	0.02	68
2	3.67 (0.80)	−0.97 (0.86)	0.0004	67	17.50 * (0.06)	−4.59 (0.33)	0.02	67	16.35 *** (0.003)	−3.44 (0.18)	0.05	67
3	3.55 (0.79)	−1.12 (0.83)	0.001	66	15.24 * (0.07)	−3.62 (0.41)	0.02	66	16.21 *** (0.002)	−3.29 (0.17)	0.06	66
4	−0.10 (0.99)	−0.19 (0.97)	0.000	65	12.42 (0.12)	−2.61 (0.53)	0.02	65	16.13 *** (0.001)	−3.29 (0.13)	0.08	65
Forecast Horizon	**Platinum (2004:01–2020:12)**				**Palladium (2004:01–2020:12)**				**Zinc (2004:01–2020:12)**			
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**
1	8.85 (0.27)	−3.45 (0.31)	0.01	68	22.91 ** (0.03)	−3.90 (0.48)	0.01	68	12.24 (0.55)	−2.86 (0.69)	0.004	68
2	10.67 (0.15)	−4.65 (0.18)	0.03	67	25.87 *** (0.01)	−5.76 (0.28)	0.02	67	7.79 (0.69)	−1.18 (0.86)	0.001	67
3	9.57 (0.17)	−4.17 (0.18)	0.04	66	23.83 *** (0.01)	−4.74 (0.33)	0.03	66	3.82 (0.84)	0.57 (0.93)	0.0004	66
4	7.76 (0.23)	−3.51 (0.19)	0.05	65	21.88 ** (0.03)	−3.70 (0.41)	0.02	65	0.47 (0.98)	1.92 (0.73)	0.01	65
Forecast Horizon	**Natural Gas (2004:01–2020:12)**											
h	**C**	**(Y_{30}-Y_1)**	**R^2**	**N**								
1	−3.02 (0.86)	−0.87 (0.88)	0.0002	68								
2	−0.63 (0.97)	−2.01 (0.71)	0.002	67								
3	−0.26 (0.98)	−2.66 (0.59)	0.01	66								
4	−2.82 (0.81)	−1.83 (0.68)	0.004	65								

Notes: The forecast horizon (h) is in quarters. Y_{30}-Y_1 denotes the yield spread calculated as the difference between the yield rates on 10-year and 1-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 10. Estimation results of Equation (1) with the Y_{30}-Y_2 indicator. Panel A: entire sample. Panel B: sample period 1986–2003. Panel C: sample period 2004–2020.

Panel A												
Forecast Horizon	**Oil (1986:01–2020:12)**				**Silver (1986:01–2020:12)**				**Gold (1986:01–2020:12)**			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	10.05 (0.24)	−3.63 (0.37)	0.003	139	0.99 (0.87)	2.32 (0.53)	0.003	139	2.96 (0.50)	1.28 (0.55)	0.003	138
2	7.63 (0.33)	−2.40 (0.52)	0.003	138	0.94 (0.87)	2.30 (0.51)	0.01	138	2.85 (0.49)	1.33 (0.51)	0.01	138
3	6.90 (0.34)	−2.12 (0.56)	0.004	137	0.30 (0.96)	2.45 (0.46)	0.01	137	2.52 (0.53)	1.38 (0.47)	0.01	137
4	5.09 (0.46)	−1.47 (0.68)	0.003	136	−0.39 (0.94)	2.66 (0.40)	0.02	136	2.51 (0.51)	1.29 (0.48)	0.01	136

Table 10. *Cont.*

Panel A												
Forecast Horizon	**Platinum (1986:04–2020:12)**				**Palladium (1986:04–2020:12)**				**Zinc (1997:08–2020:12)**			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	3.25 (0.56)	−0.33 (0.91)	0.000	139	16.08 * (0.09)	−4.42 (0.40)	0.01	139	1.81 (0.89)	0.75 (0.89)	0.0003	93
2	3.20 (0.54)	−0.47 (0.87)	0.0003	138	15.32 * (0.09)	−4.00 (0.42)	0.01	138	−0.29 (0.98)	1.95 (0.69)	0.003	92
3	2.49 (0.62)	−0.28 (0.92)	0.0002	136	13.01 (0.15)	−2.77 (0.55)	0.01	137	−2.13 (0.86)	2.90 (0.52)	0.01	91
4	1.64 (0.72)	0.12 (0.96)	0.0001	136	10.71 (0.23)	−1.32 (0.76)	0.002	136	−3.72 (0.75)	3.73 (0.36)	0.02	90
Forecast Horizon	**Ethanol (2005:06–2020:12)**				**Coal (2009:01–2020:12)**				**Natural Gas (1990:05–2020:12)**			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	−1.17 (0.94)	0.20 (0.98)	0.000	62	−7.47 (0.71)	3.59 (0.63)	0.01	48	6.34 (0.65)	−2.57 (0.68)	0.001	122
2	−2.25 (0.84)	0.01 (0.99)	0.000	61	−8.79 (0.64)	3.94 (0.58)	0.01	47	4.53 (0.72)	−1.80 (0.75)	0.001	121
3	−1.75 (0.85)	−0.51 (0.92)	0.0003	60	−12.05 (0.48)	4.97 (0.44)	0.03	46	3.33 (0.77)	−1.46 (0.78)	0.001	120
4	−4.24 (0.58)	0.12 (0.98)	0.000	59	−10.28 (0.54)	4.16 (0.52)	0.03	45	1.17 (0.91)	−0.28 (0.95)	0.000	119
Panel B												
Forecast Horizon	**Oil (1986:01–2003:12)**				**Silver (1986:01–2003:12)**				**Gold (1986:01–2003:12)**			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	13.76 (0.20)	−6.08 (0.22)	0.01	71	−7.65 (0.11)	**6.70** ** **(0.04)**	0.05	71	−5.00 (0.20)	**4.81** ** **(0.02)**	0.04	70
2	10.50 (0.28)	−3.63 (0.41)	0.01	71	−9.78 ** (0.03)	**9.03** *** **(0.01)**	0.15	71	−5.45 * (0.08)	**5.21** *** **(0.002)**	0.13	71
3	8.50 (0.33)	−2.22 (0.58)	0.01	71	−9.99 *** (0.01)	**8.81** *** **(0.001)**	0.23	71	−5.83 ** (0.03)	**5.09** *** **(0.0001)**	0.21	71
4	6.63 (0.41)	−0.73 (0.86)	0.001	71	−9.64 *** (0.01)	**8.55** *** **(0.000)**	0.29	71	−5.68 ** (0.02)	**4.87** *** **(0.00)**	0.23	71
Forecast Horizon	**Platinum (1986:04–2003:12)**				**Palladium (1986:04–2003:12)**				**Zinc (1997:08–2003:12)**			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	−0.44 (0.95)	3.41 (0.37)	0.01	71	14.97 (0.25)	−8.85 (0.27)	0.02	71	−15.13 (0.20)	7.27 (0.16)	0.08	25
2	−2.22 (0.74)	4.96 (0.14)	0.06	71	10.69 (0.40)	−4.78 (0.54)	0.01	71	−14.62 (0.14)	**8.84** * **(0.06)**	0.19	25
3	−3.44 (0.57)	**5.33** * **(0.07)**	0.10	70	7.10 (0.56)	−2.52 (0.72)	0.005	71	−14.04 (0.12)	**8.85** ** **(0.02)**	0.26	25
4	−3.99 (0.46)	**5.77** ** **(0.02)**	0.14	71	4.21 (0.72)	−0.32 (0.96)	0.000	71	−13.54 * (0.09)	**8.84** *** **(0.004)**	0.34	25

Table 10. *Cont.*

Panel B

Forecast Horizon	Natural Gas (1990:05–2003:12)			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	13.81 (0.55)	−2.31 (0.84)	0.001	54
2	8.26 (0.69)	0.71 (0.94)	0.0001	54
3	4.46 (0.81)	3.12 (0.69)	0.005	54
4	2.21 (0.89)	5.08 (0.43)	0.02	54

Panel C

Forecast Horizon	Oil (2004:01–2020:12)				Silver (2004:01–2020:12)				Gold (2004:01–2020:12)			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	4.49 (0.79)	−1.07 (0.87)	0.000	68	15.00 (0.17)	−3.28 (0.59)	0.004	68	15.76 *** (0.01)	−3.55 (0.25)	0.02	68
2	2.74 (0.85)	−0.59 (0.92)	0.0001	67	17.44 * (0.07)	−5.04 (0.36)	0.02	67	16.72 *** (0.004)	−3.98 (0.18)	0.05	67
3	3.61 (0.78)	−1.27 (0.82)	0.001	66	16.49 * (0.06)	−4.63 (0.37)	0.03	66	16.89 *** (0.002)	−3.98 (0.15)	0.07	66
4	1.10 (0.93)	−0.82 (0.88)	0.001	65	14.29 * (0.07)	−3.83 (0.42)	0.03	65	16.95 *** (0.00)	−4.05 (0.11)	0.10	65
Forecast Horizon	**Platinum (2004:01–2020:12)**				**Palladium (2004:01–2020:12)**				**Zinc (2004:01–2020:12)**			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	7.76 (0.35)	−3.24 (0.37)	0.01	68	21.83 * (0.05)	−3.74 (0.54)	0.004	68	13.28 (0.52)	−3.69 (0.64)	0.01	68
2	9.88 (0.19)	−4.73 (0.19)	0.02	67	25.47 *** (0.01)	−6.15 (0.28)	0.02	67	9.62 (0.63)	−2.25 (0.76)	0.003	67
3	10.02 (0.16)	−4.84 (0.16)	0.04	66	25.20 *** (0.01)	−5.94 (0.27)	0.03	66	6.23 (0.75)	−0.61 (0.93)	0.0003	66
4	8.70 (0.17)	−4.35 (0.14)	0.05	65	23.84 *** (0.01)	−5.08 (0.29)	0.03	65	3.33 (0.86)	0.67 (0.92)	0.001	65

Forecast Horizon	Natural Gas (2004:01–2020:12)			
h	**C**	**(Y_{30}-Y_2)**	**R^2**	**N**
1	−3.61 (0.84)	−0.66 (0.92)	0.000	68
2	−1.65 (0.92)	−1.69 (0.78)	0.001	67
3	−0.53 (0.97)	−2.79 (0.62)	0.005	66
4	−3.15 (0.79)	−1.85 (0.71)	0.003	65

Notes: The forecast horizon (h) is in quarters. Y_{30}-Y_2 denotes the yield spread calculated as the difference between the yield rates on 10-year and 2-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

6.2. Commodity Prices following Periods of Non-Positive Yield Spreads

What happens to the price of commodities in the months following downward-sloped or flat yield curves? To answer this question, we track the evolution in the price of the sampled commodities throughout the sample period. The period between 1986 and December 2020 witnessed eight periods during which the yield spread (10-year minus 1-year yields) was negative or equal to zero. The left-hand side of Table 11 lists the dates when the yield spread became non-positive. For each commodity, we calculate the returns accumulated 1, 2, 3, 4, 5 and 6 quarters following these non-positive yield spread periods.

Table 11. Commodity prices in the periods following non-positive yield spread. Panel A: oil. Panel B: silver. Panel C: gold. Panel D: platinum. Panel E: palladium. Panel F: zinc. Panel G: ethanol. Panel H: coal. Panel I: natural gas.

Panel A							
Periods Associated with Non-Positive Yield Spreads		**Return 1Q Later**	**Return 2 Qs. Later**	**Return 3Qs. Later**	**Return 4Qs. Later**	**Return 5Qs. Later**	**Return 6Qs. Later**
Start	**End**						
25/01/1989	30/06/1989	−0.69%	7.65%	0.05%	−15.79%	82.98%	40.31%
04/08/1989	13/10/1989	10.72%	−14.89%	−12.11%	90%	47.34%	2.82%
17/03/2000	28/04/2000	9.48%	27.2%	12.9%	9.83%	4.97%	−13.95%
27/12/2005	29/03/2006	10.64%	−5.33%	−8.13%	−0.63%	6.37%	22.89%
05/06/2006	05/06/2007	15.42%	33.35%	59.3%	94.77%	61.91%	−37.8%
20/07/2007	08/08/2007	32.31%	27.19%	71.43%	59.67%	−15.4%	−45.17%
23/05/2019	03/06/2019	1.3%	5.35%	−11.4%	−30.87%	−22.31%	−14.29%
05/08/2019	08/10/2019	13.26%	−52.33%	−22.29%	−21.74%		
	Average	11.56%	3.52%	11.22%	23.15%	23.69%	−6.46%

Panel B							
Periods Associated with Non-Positive Yield Spreads		**Return 1 Q Later**	**Return 2 Qs. Later**	**Return 3Qs. Later**	**Return 4Qs. Later**	**Return 5Qs. Later**	**Return 6Qs. Later**
Start	**End**						
25/01/1989	30/06/1989	3.01%	2.29%	−3.39%	−5.2%	−8.63%	−17.61%
04/08/1989	13/10/1989	4.82%	−0.62%	−4.51%	−15.27%	−17.18%	−22.41%
17/03/2000	28/04/2000	0.81%	−3.63%	−2.76%	−12.16%	−15.28%	−14.46%
27/12/2005	29/03/2006	−6.68%	4.22%	16.82%	20.47%	11.56%	25.71%
05/06/2006	05/06/2007	−11.74%	3.32%	49.76%	24.31%	−11.35%	−31.94%
20/07/2007	08/08/2007	17.81%	29.92%	27.52%	16.4%	−24.35%	−2.58%
23/05/2019	03/06/2019	29.44%	16.18%	16.21%	23.88%	81.38%	63.41%
05/08/2019	08/10/2019	2.64%	−14.1%	7.8%	34.89%		
	Average	5.01%	4.7%	13.43%	10.92%	2.31%	0.02%

Panel C							
Periods Associated with Non-Positive Yield Spreads		**Return 1 Q Later**	**Return 2 Qs. Later**	**Return 3Qs. Later**	**Return 4Qs. Later**	**Return 5Qs. Later**	**Return 6Qs. Later**
Start	**End**						
25/01/1989	30/06/1989	−2.68%	7.39%	−1.88%	−4.61%	2.46%	5.01%
04/08/1989	13/10/1989	14.15%	3.47%	0.66%	6.66%	10.54%	−0.55%
17/03/2000	28/04/2000	1.09%	−3.75%	−4.33%	−3.79%	−2.66%	1.6%
27/12/2005	29/03/2006	2.72%	4.41%	11.29%	15.38%	13.54%	29.57%
05/06/2006	05/06/2007	2.2%	19.1%	47.63%	30.2%	19.3%	12.08%
20/07/2007	08/08/2007	24.17%	36.16%	30.78%	27.18%	8.85%	32.31%
23/05/2019	03/06/2019	17.15%	11.76%	24.32%	30.43%	45.93%	38.87%
05/08/2019	08/10/2019	4.21%	11.23%	21.6%	26.14%		
	Average	7.88%	11.22%	16.26%	15.95%	13.99%	16.98%

Table 11. *Cont.*

Panel D

Periods Associated with Non-Positive Yield Spreads Start	End	Return 1 Q Later	Return 2 Qs. Later	Return 3Qs. Later	Return 4Qs. Later	Return 5Qs. Later	Return 6Qs. Later
25/01/1989	30/06/1989	−0.42%	−1.73%	−4.32%	−2.18%	−13.93%	−17.58%
04/08/1989	13/10/1989	3.49%	−0.83%	−0.99%	−13.79%	−13.63%	−17.16%
17/03/2000	28/04/2000	13.43%	13.6%	19.99%	18.93%	1%	−15.14%
27/12/2005	29/03/2006	11.96%	5.97%	5.79%	15.39%	18.77%	29.56%
05/06/2006	05/06/2007	−1.99%	13.05%	75.25%	54.95%	5.31%	−39.39%
20/07/2007	08/08/2007	14.14%	45.92%	58.18%	20.8%	−34.01%	−22.86%
23/05/2019	03/06/2019	16.41%	11.05%	5.9%	5.82%	8.37%	26.52%
05/08/2019	08/10/2019	8.32%	−17.59%	−1.28%	−3.06%		
	Average	8.17%	8.68%	19.82%	12.11%	−4.02%	−8.01%

Panel E

Periods Associated with Non-Positive Yield Spreads Start	End	Return 1 Q Later	Return 2 Qs. Later	Return 3Qs. Later	Return 4Qs. Later	Return 5Qs. Later	Return 6Qs. Later
25/01/1989	30/06/1989	−8.94%	−13.75%	−17.41%	−25.11%	−38.73%	−48.28%
04/08/1989	13/10/1989	−0.93%	−6.91%	−16.05%	−32.23%	−35.82%	−29.44%
17/03/2000	28/04/2000	24.56%	21.82%	69.84%	10.91%	−26.72%	−44.09%
27/12/2005	29/03/2006	−6.35%	−5.45%	1.15%	6.38%	10.12%	5.17%
05/06/2006	05/06/2007	−9.21%	−4.44%	52.74%	15.44%	−26.12%	−55.75%
20/07/2007	08/08/2007	3.72%	21.13%	19.77%	−9.12%	−38.45%	−43.44%
23/05/2019	03/06/2019	16.71%	39.16%	83.09%	48.74%	74.85%	75.24%
05/08/2019	08/10/2019	24.88%	26.93%	17.42%	45.57%		
	Average	5.55%	9.81%	26.32%	7.57%	−11.55%	−20.08%

Panel F

Periods Associated with Non-Positive Yield Spreads Start	End	Return 1 Q Later	Return 2 Qs. Later	Return 3Qs. Later	Return 4Qs. Later	Return 5Qs. Later	Return 6Qs. Later
17/03/2000	28/04/2000	0.02%	−9.09%	−9.56%	−17.29%	−27%	−34.66%
27/12/2005	29/03/2006	19.53%	27.26%	64.03%	24.16%	28%	17.24%
05/06/2006	05/06/2007	−25.82%	−36.91%	−26.28%	−49.13%	−54.42%	−72.24%
20/07/2007	08/08/2007	−20.16%	−30.17%	−37.26%	−52.14%	−69.52%	−66.48%
23/05/2019	03/06/2019	−14.05%	−14.01%	−23.99%	−21.75%	−3.41%	6.25%
05/08/2019	08/10/2019	4.25%	−17.65%	−8.67%	1.14%		
	Average	−6.04%	−13.43%	−6.96%	−19.17%	−25.27%	−29.98%

Panel G

Periods Associated with Non-Positive Yield Spreads Start	End	Return 1 Q Later	Return 2 Qs. Later	Return 3Qs. Later	Return 4Qs. Later	Return 5Qs. Later	Return 6Qs. Later
27/12/2005	29/03/2006	35.07%	−29.26%	−0.08%	−9.02%	−21.84%	−37.88%
05/06/2006	05/06/2007	−22.17%	−7.31%	11.04%	10.47%	2.36%	−34.06%
20/07/2007	08/08/2007	0.54%	14.36%	40.65%	10.57%	−6.67%	−12.47%
23/05/2019	03/06/2019	−12.49%	−2.91%	−16.66%	−23.33%	−9.78%	−7.47%
05/08/2019	08/10/2019	−7.42%	−38.07%	−5.69%	−3.26%		
	Average	−1.29%	−12.64%	5.85%	−2.91%	−8.98%	−22.97%

Table 11. *Cont.*

Periods Associated with Non-Positive Yield Spreads		Return 1 Q Later	Return 2 Qs. Later	Return 3Qs. Later	Return 4Qs. Later	Return 5Qs. Later	Return 6Qs. Later
Panel H							
Start	**End**						
23/05/2019	03/06/2019	−9.22%	−7.99%	−9.62%	−24.16%	−32.15%	0.41%
05/08/2019	08/10/2019	0.3%	−6.35%	−21.45%	−13.3%		
	Average	−4.46%	−7.17%	−15.54%	−18.73%	−32.15%	0.41%
Panel I							
Start	**End**						
17/03/2000	28/04/2000	22.41%	44.57%	100.35%	54.95%	0.83%	1.94%
27/12/2005	29/03/2006	−15.18%	−22.3%	−12.91%	5.2%	−6.36%	−5.02%
05/06/2006	05/06/2007	−28.01%	−10.9%	20.8%	55.25%	−7.63%	−28.79%
20/07/2007	08/08/2007	24%	33.46%	81.08%	32.6%	8.63%	−22.72%
23/05/2019	03/06/2019	−1.87%	1.58%	−25.09%	−26.05%	3.5%	4.33%
05/08/2019	08/10/2019	−6.42%	−22.07%	−20.28%	14.82%		
	Average	−0.85%	4.06%	23.99%	22.79%	−0.21%	−10.05%

Notes: the tables report the commodity returns accumulated after 1, 2, 3, 4, 5 and 6 quarters following downward-sloped or flat yield curves.

Oil, silver, gold, platinum, palladium, and natural gas prices surged strongly in the quarters following the periods associated with equality in long and short-term Treasury yields. For example, tracking the prices of these commodities three quarters after an end in the zero slope in the bond term structure reveals significant positive returns on average (oil 11.22%; silver 13.43%; gold 16.26%; platinum 19.82%; palladium 26.32%; ethanol 5.85%; natural gas 24%). In contrast, coal and zinc prices present a mixed and inconclusive picture with a tendency to negative returns. This finding emphasizes that investors should note that flat or downward-sloped yield curves seem to be reasonable points at which to take long positions in several commodities that they plan to hold for a relatively long period of time.

Our findings are even more pronounced if we consider the recent relatively flat yield curve observed during the last week in February 2020 due to the outbreak of the coronavirus. However, we did not include the findings in the table because the difference between the 10-year and 1-year bond interest rates was 0.03% (0.0003). While quite small, it is not a non-positive yield. In addition, prices recovered sharply after two to four quarters. Nevertheless, the findings in Table A1 in the online Appendix A lend support to our conjecture.

7. Conclusions

We investigated an important, yet barely discussed, issue: Can yield spreads forecast future innovations in the commodity market? If so, is this long-term correlation stable over time? Despite the extensive research linking economic real activity to lagged yield spreads, the predictive ability of the yield curve has not been proven with regard to commodities often used in as raw materials.

Our findings can be summarized as follows. First, the prediction ability of the yield curve is evident mainly in the period before the financialization of commodities era, but is absent between 2004 and 2020. Second, structural break tests confirm the changes in the correlation between the six yield spreads proposed and future commodity prices. Third, the findings of the dynamic conditional correlation confirm the time-varying nature of the yield spread in predicting the future evolution in commodity prices. One explanation might be the increased flow of money into the commodity market and the increased correlation between it and equity markets. These changes disconnected the prices of commodities from the economic cycle.

The structural breaks and the fading correlation between the variables of interest are critical for those involved in risk management and investment diversification. Furthermore, our results may be useful for policy makers who must make decisions about policies to target and control inflation. Future research can extend the standing literature by addressing the interplay between the shape of the term structure and future evolution of asset prices in the wake of pandemic outbreaks and the massive monetary intervention conducted by central banks under severe economic conditions.

Author Contributions: Conceptualization, Y.I.-B. and M.Q.; methodology, Y.I.-B. and M.Q.; software, Y.I.-B. and M.Q.; validation, Y.I.-B. and M.Q.; formal analysis, Y.I.-B. and M.Q.; investigation, Y.I.-B. and M.Q.; resources, M.Q.; data curation, Y.I.-B. and M.Q.; writing—original draft preparation, Y.I.-B. and M.Q.; writing—review and editing, Y.I.-B. and M.Q.; visualization, Y.I.-B. and M.Q.; supervision, M.Q.; project administration, M.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Evolution in the commodity prices following the relatively flat curve witnessed in February 2020.

	Accumulated Returns after:									
	1 Month	2 Months	3 Months	4 Months	5 Months	6 Months	7 Months	8 Months	9 Months	10 Months
Oil	−50.9%	−66.1%	−33.4%	−22.4%	−17.3%	−13.1%	−19.3%	−22.7%	−8.4%	−1.6%
Silver	−18.4%	−16.1%	−2.9%	−1.6%	25.4%	44.4%	26.5%	34.2%	28.4%	42.6%
Gold	−0.8%	4.7%	5.4%	7.0%	15.2%	16.1%	12.8%	15.5%	9.6%	14.5%
PLTNM	−20.0%	−17.0%	−4.9%	−13.9%	1.7%	0.2%	−9.7%	−6.4%	4.0%	11.9%
PLDM	−14.0%	−25.0%	−25.3%	−31.1%	−13.4%	−18.0%	−16.6%	−10.5%	−11.5%	−10.8%
Zinc	−9.4%	−7.1%	−1.6%	1.3%	9.9%	22.2%	16.9%	25.4%	35.9%	41.7%
Ethanol	−29.6%	−27.5%	−14.3%	−12.0%	−10.9%	−0.4%	1.6%	21.5%	6.2%	2.3%
Coal	−2.3%	−10.0%	−23.5%	−21.6%	−23.4%	−26.7%	−22.2%	−14.6%	−6.1%	17.8%
Nat.Gas	−10.2%	−5.5%	−6.3%	−19.8%	−2.1%	34.8%	15.8%	63.7%	56.8%	46.2%

Notes: With the outbreak of the COVID−19 pandemic in February 2020, the U.S. 1-year yield was 1.43% while the 10-year was 1.46%. That is, the yield spread was 0.03%. The table reports the evolution in the prices of commodities in the few months following this relatively flat curve. Though we discuss one case, the overall picture is clear and shows that the current yield spread is a relatively good predictor of the future evolution in commodity prices.

Table A2. Estimation results of model 5 with the Y10-Y3M indicator. Panel A: entire sample. Panel B: sample period 2004–2020.

Panel A																
Forecast Horizon	Oil (1986:01–2020:12)								Silver (1986:01–2020:12)							
h	C	$(Y_{10}\text{-}Y_{3M})$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{10}\text{-}Y_{3M})$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	5.92	−0.79	−0.55	9.32	**−14.07b**	0.11	0.03	139	1.67	1.06	−0.85	−10.29	−4.90	0.06	−0.01	139
	(0.45)	(−0.12)	(−0.08)	(0.56)	(−2.43)	(1.08)			(0.21)	(0.28)	(−0.21)	(−1.04)	(−1.44)	(0.99)		
2	7.33	−1.74	−5.23	2.47	−3.76	0.04	−0.01	138	3.12	1.40	−3.50	−6.35	**−4.05c**	−0.05	0.02	138
	(0.84)	(−0.41)	(−1.16)	(0.22)	(−0.99)	(0.66)			(0.58)	(0.54)	(−1.26)	(−0.93)	(−1.73)	(−1.10)		
3	4.84	−0.54	−5.93	−9.86	−5.02	−0.02	0.04	137	−0.37	2.76	−2.12	−8.83	**−5.06a**	−0.03	0.05	137
	(0.75)	(−0.17)	(−1.81)	(−1.23)	(−1.75)	(−0.35)			(−0.09)	(1.32)	(−0.96)	(−1.64)	(−2.62)	(−0.77)		
4	4.49	−1.48	−2.40	−5.33	6.23	−0.01	−0.01	136	0.69	2.56	−3.33c	−0.55	−0.82	−0.04	0.02	136
	(0.84)	(−0.56)	(−0.85)	(−0.80)	(1.17)	(−0.25)			(0.19)	(1.42)	(−1.71)	(−0.12)	(−0.23)	(−1.38)		

Table A2. *Cont.*

Forecast Horizon																
Panel A																
Forecast Horizon	**Gold (1986:01–2020:12)**								**Platinum (1986:04–2020:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	3.73	0.54	1.61	−1.05	−2.52	0.01	−0.02	138	3.32	−0.56	−3.60	−8.25	−1.92	0.02	−0.02	139
	(0.86)	(0.26)	(0.73)	(−0.20)	(−1.37)	(0.45)			(0.46)	(−0.16)	(−0.96)	(−0.91)	(−0.61)	(0.4)		
2	5.19c	0.43	−2.09	2.21	−1.25	−0.03	0.01	137	3.83	−0.31	**−6.05b**	−5.14	0.50	−0.03	0.02	138
	(1.76)	(0.30)	(−1.37)	(0.60)	(−0.97)	(−1.14)			(0.79)	(−0.13)	(−2.43)	(−0.84)	(0.24)	(−0.87)		
3	4.68c	0.33	−0.91	1.45	−1.05	−0.02	−0.01	136	0.63	1.01	−2.32	**−7.80c**	−1.71	−0.04	0.02	137
	(1.84)	(0.27)	(−0.71)	(0.46)	(−0.93)	(−0.85)			(0.17)	(0.55)	(−1.21)	(−1.66)	(−1.02)	(−1.44)		
4	5.25b	0.38	−1.91	3.31	−2.83	−0.02	0.01	135	1.19	0.56	−1.71	−0.17	4.16	**−0.04c**	0.02	136
	(2.34)	(0.34)	(−1.61)	(1.19)	(−1.27)	(−1.22)			(0.38)	(0.36)	(−1.03)	(−0.04)	(1.34)	(−1.8)		
Forecast Horizon	**Palladium (1986:04–2020:12)**								**Zinc (1997:08–2020:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	**19.89c**	−6.40	**−12.07b**	−7.36	**8.63c**	0.03	0.03	139	−1.27	0.56	−6.60	−15.05	0.07	**0.17b**	0.01	93
	(1.94)	(−1.29)	(−2.27)	(−0.57)	(1.93)	(0.37)			(−0.13)	(0.11)	(−1.3)	(−1.07)	(0.02)	(2.03)		
2	**16.52b**	−4.10	**−7.36c**	−10.51	4.02	−0.05	0.02	138	−2.35	2.75	−2.46	−11.65	−0.71	0.03	−0.03	92
	(2.16)	(−1.11)	(−1.87)	(−1.1)	(1.21)	(−0.92)			(−0.3)	(0.7)	(−0.61)	(−1.05)	(−0.23)	(0.45)		
3	**11.65c**	−1.77	−0.97	−5.21	2.92	−0.03	−0.02	137	−6.15	5.27	−2.60	−5.86	−2.95	0.02	−0.01	91
	(1.78)	(−0.56)	(−0.29)	(−0.64)	(1)	(−0.69)			(−0.91)	(1.59)	(−0.77)	(−0.63)	(−1.07)	(0.43)		
4	**9.71c**	−1.60	0.49	−1.63	**13.19b**	−0.03	0.01	136	−6.83	**6.04b**	−1.16	7.09	3.45	0.02	0.01	90
	(1.72)	(−0.57)	(0.16)	(−0.23)	(2.35)	(−0.72)			(−1.14)	(2.05)	(−0.37)	(0.86)	(0.57)	(0.38)		
Forecast Horizon	**Ethanol (2005:06–2020:12)**								**Coal (2009:01–2020:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	−1.39	−0.65	3.71	**47.44c**	−7.37	0.20	0.07	62	4.07	−0.18	−3.93	−14.93	**9.82b**	−0.12	0.10	48
	(−0.07)	(−0.07)	(0.35)	(1.95)	(−1.12)	(1.35)			(0.26)	(−0.03)	(−0.54)	(−0.84)	(2.45)	(−1.11)		
2	−5.25	−0.63	1.61	−2.20	**−7.50c**	0.13	0.02	61	−1.02	2.98	−5.38	−16.73	**5.05c**	**−0.20b**	0.15	47
	(−0.43)	(−0.1)	(0.25)	(−0.15)	(−1.85)	(1.47)			(−0.09)	(0.55)	(−1.02)	(−1.29)	(1.72)	(−2.42)		
3	−5.53	0.68	−0.35	−8.83	**−5.12c**	0.01	−0.03	60	−13.96	7.81	−1.61	−12.06	−1.60	**−0.13c**	0.07	46
	(−0.63)	(0.16)	(−0.08)	(−0.83)	(−1.7)	(0.12)			(−1.34)	(1.66)	(−0.36)	(−1.1)	(−0.61)	(−1.84)		
4	−3.89	0.66	−3.32	2.19	0.06	−0.02	−0.07	59	−12.55	6.48	−1.08	−7.51	1.56	−0.06	−0.01	45
	(−0.58)	(0.2)	(−0.87)	(0.27)	(0.01)	(−0.33)			(−1.3)	(1.48)	(−0.24)	(−0.74)	(0.23)	(−0.93)		
Forecast Horizon	**Natural gas (1990:05–2020:12)**															
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**								
1	5.82	−3.09	15.11c	15.35	1.91	0.00	0.00	122								
	(0.36)	(−0.4)	(1.87)	(0.74)	(0.29)	(−0.01)										
2	−0.62	0.29	2.48	−14.15	−3.37	0.02	−0.03	121								
	(−0.06)	(0.05)	(0.45)	(−1)	(−0.73)	(0.22)										
3	−6.55	2.69	−0.25	**−26.80b**	−5.08	0.06	0.02	120								
	(−0.75)	(0.65)	(−0.06)	(−2.46)	(−1.38)	(0.85)										
4	−5.93	1.55	1.67	**−16.60c**	**12.51c**	0.04	0.03	119								
	(−0.86)	(0.47)	(0.47)	(−1.93)	(1.85)	(0.7)										
Panel B																
Forecast Horizon	**Oil (2004:01–2020:12)**								**Silver (2004:01–2020:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	−0.10	−0.53	0.86	16.77	**−15.95b**	0.17	0.04	68	11.98	−4.06	−2.96	−13.37	−4.65	0.11	−0.03	68
	(0.00)	(−0.05)	(0.07)	(0.61)	(−2.11)	(1.01)			(0.86)	(−0.6)	(−0.4)	(−0.77)	(−0.98)	(1.09)		
2	3.69	−2.15	0.72	9.54	−5.95	0.08	−0.03	67	**16.38c**	−4.37	−7.96	−11.18	−3.03	−0.06	0.06	67
	(0.25)	(−0.3)	(0.09)	(0.52)	(−1.2)	(0.73)			(1.77)	(−0.98)	(−1.63)	(−0.98)	(−0.97)	(−0.9)		
3	−0.32	0.45	−6.34	−16.19	**−6.44c**	−0.02	0.03	66	9.86	−1.47	−4.98	−11.62	−4.22	−0.03	0.05	66
	(−0.03)	(0.08)	(−1.1)	(−1.21)	(−1.68)	(−0.27)			(1.29)	(−0.4)	(−1.26)	(−1.27)	(−1.6)	(−0.54)		
4	−0.25	−0.17	−1.22	−1.81	8.51	−0.03	−0.05	65	**10.70c**	−0.62	**−6.94c**	3.41	−0.20	−0.06	0.03	65
	(−0.03)	(−0.04)	(−0.24)	(−0.17)	(1.04)	(−0.44)			(1.7)	(−0.21)	(−1.98)	(0.45)	(−0.03)	(−1.23)		
Forecast Horizon	**Gold (2004:01–2020:12)**								**Platinum (2004:01–2020:12)**							
h	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**	**C**	**(Y_{10}-Y_{3M})**	**ΔSP**	**ΔEX**	**ΔIP**	**ΔEPU**	**R^2**	**N**
1	**13.40b**	−2.83	−0.20	3.97	−1.68	0.02	−0.03	68	7.99	−5.32	−5.15	0.76	−1.60	0.13	−0.02	68
	(2.15)	(−0.94)	(−0.06)	(0.51)	(−0.79)	(0.45)			(0.66)	(−0.9)	(−0.79)	(0.05)	(−0.39)	(1.41)		
2	**15.01a**	−2.91	−3.22	0.05	−0.50	−0.03	0.02	67	11.86	−5.76	**−10.43b**	−1.19	1.81	−0.01	0.05	67
	(3.24)	(−1.31)	(−1.31)	(0.01)	(−0.32)	(−0.96)			(1.44)	(−1.45)	(−2.4)	(−0.12)	(0.65)	(−0.17)		
3	**14.24a**	−2.59	−1.62	1.16	−0.16	−0.02	−0.02	66	7.39	−3.57	−4.49	−9.43	−1.74	−0.05	0.05	66
	(3.44)	(−1.31)	(−0.76)	(0.23)	(−0.11)	(−0.67)			(1.16)	(−1.17)	(−1.36)	(−1.23)	(−0.79)	(−1.04)		
4	**15.03a**	−2.38	−2.94	5.51	−1.90	−0.03	0.06	65	7.35	−3.13	−3.70	2.13	2.23	−0.05	0.02	65
	(4.32)	(−1.42)	(−1.52)	(1.31)	(−0.6)	(−1.29)			(1.43)	(−1.26)	(−1.29)	(0.34)	(0.47)	(−1.31)		

Table A2. *Cont.*

Panel B

Forecast Horizon	Palladium (2004:01–2020:12)								Zinc (2004:01–2020:12)							
h	C	(Y_{10}-Y_{3M})	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_{3M})	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	25.75 (1.66)	−8.22 (−1.09)	−8.27 (−0.99)	−5.39 (−0.28)	6.03 (1.14)	0.15 (1.26)	−0.01	68	8.68 (0.66)	−4.06 (−0.63)	**−15.79b** (−2.23)	−15.38 (−0.93)	2.54 (0.56)	**0.22b** (2.19)	0.06	68
2	**29.00a** (2.64)	−8.33 (−1.58)	−8.79 (−1.51)	−15.25 (−1.13)	2.57 (0.69)	−0.06 (−0.73)	0.03	67	5.83 (0.53)	−0.67 (−0.13)	−6.85 (−1.19)	−9.64 (−0.72)	0.78 (0.21)	0.03 (0.37)	−0.05	67
3	**23.67b** (2.54)	−5.24 (−1.18)	−2.35 (−0.49)	−12.46 (−1.11)	0.92 (0.29)	−0.07 (−1.06)	−0.01	66	−0.61 (−0.06)	3.31 (0.73)	−5.75 (−1.17)	−3.17 (−0.28)	−1.63 (−0.5)	0.01 (0.13)	−0.04	66
4	**21.04a** (2.65)	−3.07 (−0.8)	−3.58 (−0.81)	−4.10 (−0.43)	1.08 (0.15)	−0.07 (−1.26)	−0.02	65	−2.05 (−0.24)	4.53 (1.12)	−2.82 (−0.6)	10.35 (1.02)	3.08 (0.4)	0.00 (0)	−0.03	65

Forecast Horizon	Natural gas (2004:01–2020:12)							
h	C	(Y_{10}-Y_{3M})	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	4.42 (0.23)	−5.06 (−0.54)	**17.53c** (1.68)	29.15 (1.2)	0.44 (0.07)	−0.03 (−0.2)	−0.01	68
2	1.56 (0.11)	−5.31 (−0.77)	2.90 (0.38)	−6.76 (−0.38)	−4.44 (−0.92)	0.06 (0.58)	−0.05	67
3	−1.68 (−0.15)	−4.13 (−0.75)	−1.47 (−0.25)	−18.22 (−1.32)	−5.65 (−1.43)	0.04 (0.51)	−0.01	66
4	−0.06 (−0.01)	−5.29 (−1.23)	−0.68 (−0.14)	−10.39 (−0.96)	11.01 (1.35)	0.03 (0.42)	−0.01	65

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_{3M} denotes the yield spread calculated as the difference between the yield rates on 10-year and 3-month government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. a, b and c denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A3. Estimation results of Equation (1) with the Y_{10}-Y_1 indicator. Panel A: sample period 1986–2003. Panel B: sample period 2004–2020.

Panel A

Forecast Horizon	Oil (1986:01–2003:12)								Silver (1986:01–2003:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	10.68	−5.04	0.38	−2.13	10.18	0.04	−0.06	72	−12.36	**10.33 **** **(0.03)**	0.09	−11.41	−3.41	0.06	0.02	72
2	9.29	−2.61	**−9.56 *** **(0.07)**	−5.23	9.22	0.00	0.00	72	**−10.04 *** **(0.05)**	**8.40 ***** **(0.01)**	0.54	−6.30	−1.82	−0.01	0.05	72
3	7.75	−0.90	−5.37	−4.32	1.16	0.03	−0.03	72	**−10.75 ***** **(0.01)**	**8.55 ***** **(0.00)**	0.40	**−9.64 *** **(0.09)**	−2.06	0.00	0.13	72
4	6.47	−0.83	−3.00	−7.62	1.38	0.04	−0.03	72	**−10.23 ***** **(0.00)**	**8.30 ***** **(0.00)**	−0.13	**−7.98 *** **(0.09)**	−1.94	0.00	0.18	72

Forecast Horizon	Gold (1986:01–2003:12)								Platinum (1986:04–2003:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−8.47	5.17	4.41	−10.64	3.55	0.01	0.03	70	−3.87	4.24	−1.97	−22.85	7.19	−0.07	0.05	72
2	−4.65	**5.43 ***** **(0.01)**	−1.55	2.89	−4.56	−0.01	0.07	70	−4.31	**4.94 *** **(0.08)**	−1.66	**−14.42 **** **(0.03)**	3.87	−0.03	0.07	72
3	**−4.99 *** **(0.05)**	**4.71 ***** **(0.00)**	−0.59	0.26	−2.62	−0.01	0.09	70	−5.63	**5.03 **** **(0.02)**	0.32	**−10.42 *** **(0.05)**	6.87	−0.01	0.11	72
4	**−5.03 **** **(0.03)**	**4.71 ***** **(0.00)**	−0.98	−0.86	−3.38	−0.01	0.12	70	−5.16	**5.25 ***** **(0.01)**	−0.22	−5.71	5.58	−0.02	0.11	72

Forecast Horizon	Palladium (1986:04–2003:12)								Zinc (1997:08–2003:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	11.59	−7.16	**−14.61 *** **(0.05)**	−12.90	**33.00 **** **(0.04)**	−0.09	0.08	72	**−21.33 *** **(0.06)**	9.16	8.23	−25.35	11.31	0.03	0.10	26
2	7.05	−5.81	−3.70	−11.63	**31.04 ***** **(0.01)**	−0.07	0.06	72	**−16.97 **** **(0.04)**	7.48	6.01	−27.80	6.11	0.01	0.16	26
3	3.32	−4.93	3.38	−4.13	**36.35 ***** **(0.00)**	−0.02	0.14	72	**−15.35 **** **(0.02)**	**7.55 **** **(0.04)**	3.84	**−25.00 *** **(0.09)**	−0.50	0.08	0.23	26
4	1.79	−3.31	3.58	−4.30	**30.21 ***** **(0.00)**	−0.02	0.11	72	**−15.03 ***** **(0.01)**	**8.66 ***** **(0.01)**	2.34	−12.80	3.54	0.09	0.28	26

Forecast Horizon	Natural gas (1990:05–2003:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	4.59	−4.38	18.18	−12.43	36.19	0.01	−0.04	55
2	−3.53	3.48	5.06	−33.81	22.93	−0.03	−0.03	55
3	−12.73	7.90	5.30	**−49.78 ***** **(0.00)**	**23.53 *** **(0.09)**	0.11	0.15	55
4	−9.98	8.53	4.29	**−31.91 **** **(0.02)**	14.60	0.09	0.10	55

Table A3. *Cont.*

Panel B																
Forecast Horizon	Oil (2004:01–2020:12)								Silver (2004:01–2020:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	0.01	−0.71	0.87	16.72	−15.96 ** (0.04)	0.17	0.04	67	13.07	−6.08	−2.52	−14.07	−4.75	0.08	−0.03	67
2	3.06	−2.22	0.81	9.33	−6.02 * (0.09)	0.07	−0.03	66	15.64 * (0.09)	−4.56	−7.99	−11.44	−3.10	−0.06	0.06	66
3	2.19	−1.56	−6.05	−16.46	−6.38	−0.04	0.04	65	10.99	−2.58	−4.89	−11.82	−4.20	−0.04	0.06	65
4	1.64	−1.67	−1.05	−1.99	8.65	−0.04	−0.05	64	11.94 * (0.05)	−1.64	−6.86 * (0.06)	3.29	−0.10	−0.06	0.03	64
Forecast Horizon	Gold (2004:01–2020:12)								Platinum (2004:01–2020:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	12.91 ** (0.03)	−2.92	−0.23	3.82	−1.73	0.02	−0.03	67	6.12	−5.08	−5.07	0.34	−1.76	0.12	−0.03	67
2	14.64 *** (0.00)	−3.10	−3.25	−0.10	−0.54	−0.03	0.02	66	9.76	−5.21	−10.51 ** (0.02)	−1.47	1.68	−0.01	0.04	66
3	14.14 *** (0.00)	−2.91	−1.65	1.01	−0.19	−0.02	−0.02	65	7.63	−4.39	−4.44	−9.72	−1.78	−0.05	0.06	65
4	15.12 *** (0.00)	−2.78	−2.98	5.36	−2.01	−0.03	0.07	64	7.83	−4.00	−3.69	1.85	2.06	−0.05	0.03	64
Forecast Horizon	Palladium (2004:01–2020:12)								Zinc (2004:01–2020:12)							
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	24.57 * (0.09)	−9.53	−7.76	−6.41	5.78	0.10	−0.02	67	9.45	−5.33	−15.72 ** (0.03)	−15.73	2.51	0.21 ** (0.04)	0.06	67
2	26.18 ** (0.02)	−7.80	−8.83	−15.73	2.36	−0.06	0.02	66	6.95	−1.46	−6.89	−9.66	0.83	0.03	−0.05	66
3	23.33 *** (0.01)	−5.92	−2.31	−12.83	0.82	−0.07	−0.01	65	2.55	1.68	−5.68	−3.06	−1.41	0.01	−0.05	65
4	21.35 *** (0.00)	−3.68	−3.65	−4.26	1.01	−0.07	−0.02	64	1.63	2.63	−2.57	10.56	3.89	−0.01	−0.05	64
Forecast Horizon	Natural Gas (2004:01–2020:12)															
h	C	(Y_{10}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N								
1	1.86	−3.96	17.37	28.99	0.32	−0.02	−0.01	67								
2	−0.63	−4.66	2.84	−7.03	−4.58	0.06	−0.05	66								
3	−2.01	−4.66	−1.41	−18.54	−5.74	0.04	−0.01	65								
4	−1.42	−5.19	−0.79	−10.76	10.45	0.03	−0.01	64								

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_1 denotes the yield spread calculated as the difference between the yield rates on 10-year and 1-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A4. Estimation results of Equation (1) with the Y_{10}-Y_2 indicator. Panel A: sample period 1986–2003. Panel B: sample period 2004–2020.

Panel A																
Forecast Horizon	Oil (1986:01–2003:12)								Silver (1986:01–2003:12)							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	10.51	−6.94	0.47	−2.11	9.52	0.04	−0.06	72	−11.02	13.14 ** (0.02)	−0.01	−11.22	−1.76	0.06	0.03	72
2	8.54	−2.87	−9.57 * (0.07)	−5.37	8.69	0.00	0.00	72	−9.15 * (0.05)	10.90 *** (0.00)	0.44	−6.19	−0.53	−0.01	0.07	72
3	6.92	−0.37	−5.42	−4.51	0.81	0.03	−0.03	72	−10.21 *** (0.01)	11.50 *** (0.00)	0.26	−9.61 * (0.08)	−0.86	0.00	0.17	72
4	5.07	0.34	−3.11	−7.94	0.88	0.04	−0.03	72	−9.70*** (0.00)	11.15 *** (0.00)	−0.26	−7.95 * (0.08)	−0.77	0.00	0.23	72
Forecast Horizon	Gold (1986:01–2003:12)								Platinum (1986:04–2003:12)							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−8.04	6.88 * (0.08)	4.32	−10.59	4.28	0.01	0.04	70	−3.70	5.81	−2.05	−22.86 ** (0.03)	7.76	−0.07	0.05	72
2	−4.12	7.13 *** (0.00)	−1.63	2.96	−3.76	−0.01	0.09	70	−3.63	6.24 * (0.07)	−1.71	−14.32 ** (0.03)	4.67	−0.03	0.07	72
3	−4.85 ** (0.03)	6.54 *** (0.00)	−0.69	0.25	−2.02	−0.01	0.13	70	−5.00	6.41 ** (0.02)	0.27	−10.32 * (0.05)	7.67 * (0.07)	−0.01	0.12	72
4	−4.94 *** (0.01)	6.59 *** (0.00)	−1.08	−0.89	−2.80	−0.01	0.18	70	−4.48	6.69 *** (0.01)	−0.27	−5.61	6.42 * (0.09)	−0.02	0.12	72
Forecast Horizon	Palladium (1986:04–2003:12)								Zinc (1997:08–2003:12)							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	9.00	−7.31	−14.69	−13.43	31.39	−0.09	0.07	72	−19.80 * (0.06)	10.45	8.30	−24.92	12.87	0.03	0.10	26
2	6.24	−7.35	−3.65	−11.75	30.10	−0.07	0.06	72	−16.40 ** (0.04)	9.25 * (0.08)	5.98	−27.08	7.58	0.00	0.17	26
3	2.72	−6.32	3.43	−4.21	35.57	−0.02	0.14	72	−14.85 ** (0.02)	9.42 ** (0.03)	3.79	−24.23 * (0.09)	1.01	0.08	0.26	26
4	1.38	−4.23	3.61	−4.36	29.68	−0.02	0.11	72	−14.53 *** (0.01)	10.91 *** (0.00)	2.28	−11.87	5.29	0.08	0.33	26

Table A4. *Cont.*

Panel A

Forecast Horizon	Natural Gas (1990:05–2003:12)							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	4.28	−5.76	18.25	−12.59	35.60	0.01	−0.04	55
2	−0.05	1.51	5.22	−32.94	23.99	−0.04	−0.03	55
3	−8.32	6.74	5.43	−48.61 *** (0.01)	25.29 * (0.07)	0.10	0.13	55
4	−6.13	8.14	4.38	−30.86 ** (0.03)	16.34	0.08	0.08	55

Panel B

Forecast Horizon	Oil (2004:01–2020:12)								Silver (2004:01–2020:12)							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	0.64	−1.32	0.90	16.67	−15.94 ** (0.04)	0.17	0.04	67	12.44	−6.54	−2.47	−14.23	−4.82	0.08	−0.03	67
2	2.39	−2.04	0.81	9.29	−6.07	0.07	−0.03	66	15.43 * (0.09)	−5.12	−7.94	−11.58	−3.15	−0.06	0.05	66
3	2.44	−2.01	−6.02	−16.53	−6.38 * (0.09)	−0.04	0.04	65	12.12	−3.87	−4.83	−11.96	−4.17	−0.04	0.06	65
4	2.98	−2.99	−0.97	−2.09	8.80	−0.04	−0.05	64	13.44 ** (0.03)	−3.08	−6.78 * (0.06)	3.18	0.08	−0.06	0.04	64
Forecast Horizon	**Gold (2004:01–2020:12)**								**Platinum (2004:01–2020:12)**							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	12.82 ** (0.04)	−3.31	−0.20	3.73	−1.76	0.02	−0.03	67	4.78	−4.82	−5.04	0.25	−1.84	0.12	−0.03	67
2	14.76 *** (0.00)	−3.68	−3.21	−0.21	−0.56	−0.03	0.02	66	8.53	−5.06	−10.48 ** (0.02)	−1.57	1.59	−0.01	0.04	66
3	14.61 *** (0.00)	−3.74	−1.60	0.89	−0.20	−0.02	−0.01	65	7.96	−5.35	−4.37	−9.88	−1.82	−0.05	0.06	65
4	15.67 *** (0.00)	−3.66 * (0.09)	−2.93	5.24	−2.01	−0.03	0.08	64	8.58 * (0.09)	−5.23	−3.62	1.69	2.05	−0.05	0.04	64
Forecast Horizon	**Palladium (2004:01–2020:12)**								**Zinc (2004:01–2020:12)**							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	23.20	−9.94	−7.69	−6.65	5.65	0.10	−0.02	67	10.90	−7.32	−15.61 ** (0.03)	−15.98	2.51	0.21 ** (0.04)	0.06	67
2	25.65 ** (0.02)	−8.61	−8.75	−15.96	2.27	−0.06	0.02	66	9.24	−3.49	−6.80	−9.83	0.89	0.03	−0.05	66
3	24.84 *** (0.01)	−8.04	−2.20	−13.10	0.83	−0.08	0.00	65	5.34	−0.22	−5.64	−3.14	−1.25	0.01	−0.05	65
4	23.42 *** (0.00)	−5.90	−3.52	−4.47	1.21	−0.07	−0.01	64	4.68	0.63	−2.45	10.55	4.42	−0.01	−0.06	64

Forecast Horizon	Natural Gas (2004:01–2020:12)							
h	C	(Y_{10}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	2.17	−4.83	17.42	28.85	0.30	−0.02	−0.01	67
2	−0.89	−5.19	2.89	−7.17	−4.63	0.06	−0.05	66
3	−1.55	−5.77	−1.34	−18.72	−5.77	0.04	−0.01	65
4	−1.25	−6.14	−0.74	−10.95	10.32	0.03	−0.01	64

Notes: The forecast horizon (h) is in quarters. Y_{10}-Y_2 denotes the yield spread calculated as the difference between the yield rates on 10-year and 2-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A5. Estimation results of Equation (1) with the Y_{30}-Y_{3M} indicator. Panel A: sample period 1986–2003. Panel B: sample period 2004–2020.

Panel A

Forecast Horizon	Oil (1986:01–2003:12)								Silver (1986:01–2003:12)							
h	C	(Y_{30}-M_3)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{30}-M_3)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	12.77	−4.45	0.56	−1.91	11.52	0.04	−0.06	72	−13.32	7.41 ** (0.03)	−0.04	−11.21	−4.69	0.05	0.01	72
2	9.63	−1.93	−9.52 * (0.07)	−5.25	9.60	0.00	0.00	72	−11.32 ** (0.04)	6.28 *** (0.01)	0.39	−6.24	−3.08	−0.02	0.05	72
3	8.10	−0.78	−5.34	−4.29	1.39	0.03	−0.03	72	−12.25 *** (0.01)	6.50 *** (0.00)	0.24	−9.61 * (0.09)	−3.43	−0.01	0.14	72
4	6.17	−0.41	−3.02	−7.71	1.32	0.04	−0.03	72	−11.34 *** (0.00)	6.13 *** (0.00)	−0.26	−7.89 * (0.09)	−3.12	−0.01	0.17	72
Forecast Horizon	**Gold (1986:01–2003:12)**								**Platinum (1986:04–2003:12)**							
h	C	(Y_{30}-M_3)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{30}-M_3)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−7.98	3.22	4.39	−10.34	3.31	0.01	0.02	70	−5.10	3.47	−2.09	−22.94	6.29	−0.07	0.05	72
2	−4.92	3.77 *** (0.01)	−1.60	3.03	−5.13	−0.02	0.06	70	−5.62	3.98 * (0.06)	−1.79	−14.49 ** (0.03)	2.88	−0.04	0.08	72
3	−5.12 * (0.06)	3.22 *** (0.01)	−0.63	0.41	−3.06	−0.02	0.07	70	−6.89 * (0.08)	4.02 *** (0.01)	0.20	−10.48 ** (0.04)	5.88	−0.01	0.12	72
4	−4.96 ** (0.04)	3.11 *** (0.00)	−1.00	−0.68	−3.73	−0.01	0.09	70	−6.42 * (0.07)	4.17 *** (0.01)	−0.34	−5.76	4.59	−0.02	0.12	72

Table A5. *Cont.*

Panel A

Forecast Horizon	Palladium (1986:04–2003:12)								Zinc (1997:08–2003:12)							
h	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	10.62	−4.30	−14.65 * (0.05)	−13.36	33.17 ** (0.04)	−0.08	0.07	72	−21.63 * (0.07)	6.28	7.90	−25.51	9.97	0.03	0.09	26
2	7.04	−3.89	−3.67	−11.85	31.52 *** (0.01)	−0.07	0.06	72	−18.24 ** (0.04)	5.68 * (0.09)	5.59	−27.52	5.02	0.01	0.17	26
3	3.00	−3.14	3.38	−4.38	36.62 *** (0.00)	−0.01	0.13	72	−17.30 *** (0.01)	6.09 ** (0.02)	3.31	−24.44 * (0.09)	−1.61	0.08	0.27	26
4	0.66	−1.64	3.51	−4.65	29.98 *** (0.00)	−0.02	0.11	72	−17.16 *** (0.01)	6.94 *** (0.00)	1.75	−12.20	2.27	0.09	0.34	26

Forecast Horizon	Natural Gas (1990:05–2003:12)							
h	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	7.71	−4.40	18.39	−11.95	37.52	0.01	−0.04	55
2	−3.67	2.41	5.02	−33.69	22.62	−0.03	−0.03	55
3	−14.50	6.16	5.13	−49.82 *** (0.00)	22.36	0.10	0.15	55
4	−13.07	7.20 * (0.07)	4.04	−32.21 ** (0.02)	12.97	0.09	0.11	55

Panel B

Forecast Horizon	Oil (2004:01–2020:12)								Silver (2004:01–2020:12)							
h	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−3.50	1.08	0.76	16.84	−16.06 ** (0.04)	0.17	0.04	67	15.85	−5.05	−2.31	−14.20	−4.77	0.08	−0.03	67
2	2.08	−0.98	0.81	9.35	−6.09	0.07	−0.03	66	17.66 * (0.08)	−3.76	−7.83	−11.53	−3.12	−0.06	0.06	66
3	0.66	−0.34	−6.05	−16.43	−6.48 * (0.09)	−0.04	0.04	65	13.20	−2.58	−4.77	−11.91	−4.17	−0.04	0.06	65
4	0.65	−0.63	−1.06	−1.98	8.48	−0.04	−0.05	64	13.85 * (0.05)	−1.87	−6.70 * (0.06)	3.26	0.23	−0.06	0.04	64

Forecast Horizon	Gold (2004:01–2020:12)								Platinum (2004:01–2020:12)							
h	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	15.19 ** (0.03)	−2.83	−0.11	3.73	−1.71	0.02	−0.03	67	9.44	−4.65	−4.86	0.20	−1.75	0.12	−0.02	67
2	16.84 *** (0.00)	−2.91	−3.12	−0.19	−0.52	−0.03	0.04	66	12.79	−4.61	−10.31 ** (0.02)	−1.60	1.68	−0.01	0.05	66
3	16.44 *** (0.00)	−2.83 * (0.08)	−1.52	0.92	−0.16	−0.02	0.01	65	9.74	−3.69	−4.28	−9.82	−1.80	−0.05	0.07	65
4	17.09 *** (0.00)	−2.63 * (0.05)	−2.79	5.32	−1.67	−0.03	0.09	64	9.50 * (0.09)	−3.28	−3.48	1.81	2.35	−0.05	0.04	64

Forecast Horizon	Palladium (2004:01–2020:12)								Zinc (2004:01–2020:12)							
h	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	29.84 * (0.07)	−8.31	−7.40	−6.65	5.77	0.10	−0.01	67	8.58	−2.99	−15.64 ** (0.03)	−15.73	2.41	0.21	0.05	67
2	30.47 *** (0.01)	−6.79	−8.53	−15.93	2.35	−0.06	0.03	66	5.39	−0.25	−6.91	−9.61	0.76	0.03	−0.05	66
3	26.31 *** (0.01)	−5.03	−2.10	−12.97	0.81	−0.07	0.00	65	−0.13	2.21	−5.79	−2.97	−1.49	0.01	−0.04	65
4	22.79 *** (0.01)	−2.97	−3.46	−4.30	1.26	−0.07	−0.02	64	−1.41	3.00	−2.82	10.61	3.38	−0.01	−0.04	64

Forecast Horizon	Natural gas (2004:01–2020:12))							
h	C	$(Y_{30}-M_3)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	0.34	−1.84	17.39	29.02	0.22	−0.02	−0.01	67
2	−1.16	−2.72	2.91	−7.04	−4.67	0.06	−0.06	66
3	−2.99	−2.55	−1.33	−18.55	−5.89	0.04	−0.01	65
4	−1.73	−3.17	−0.66	−10.78	10.40	0.03	−0.02	64

Notes: The forecast horizon (h) is in quarters. Y_{30}-Y_{3M} denotes the yield spread calculated as the difference between the yield rates on 10-year and 3-month government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A6. Estimation results of Equation (1) with the Y_{30}-Y_1 indicator. Panel A: sample period 1986–2003. Panel B: sample period 2004–2020.

Panel A																
Forecast Horizon	Oil (1986:01–2003:12)								Silver (1986:01–2003:12)							
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	11.73	−4.60	0.50	−1.93	10.11	0.04	−0.06	72	**−13.51 * (0.08)**	**8.82 *** (0.01)**	−0.08	−11.60	−2.95	0.06	0.03	72
2	9.08	−1.93	**−9.55 * (0.07)**	−5.29	8.95	0.00	0.00	72	**−11.34 ** (0.03)**	**7.39 *** (0.00)**	0.37	−6.54	−1.56	−0.01	0.08	72
3	7.17	−0.36	−5.41	−4.46	0.90	0.03	−0.03	72	**−12.32 *** (0.00)**	**7.68 *** (0.00)**	0.21	**−9.93 * (0.07)**	−1.88	0.00	0.19	72
4	5.40	−0.01	−3.08	−7.86	0.97	0.04	−0.04	72	**−11.52 *** (0.00)**	**7.31 *** (0.00)**	−0.30	**−8.22 * (0.07)**	−1.69	0.00	0.24	72
Forecast Horizon	Gold (1986:01–2003:12)								Platinum (1986:04–2003:12)							
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−8.48	4.09	4.35	−10.60	3.93	0.01	0.03	70	−4.59	3.77	−2.06	**−22.99 ** (0.03)**	7.30	−0.07	0.06	72
2	−4.99	**4.49 *** (0.00)**	−1.62	2.85	−4.25	−0.01	0.09	70	−4.77	**4.17 * (0.06)**	−1.74	**−14.49 ** (0.03)**	4.12	−0.03	0.08	72
3	**−5.71 ** (0.02)**	**4.16 *** (0.00)**	−0.69	0.14	−2.48	−0.01	0.13	70	**−6.35 * (0.08)**	**4.39 *** (0.01)**	0.22	**−10.54 ** (0.04)**	**7.05 * (0.09)**	−0.01	0.13	72
4	**−5.69 *** (0.01)**	**4.11 *** (0.00)**	−1.07	−0.97	−3.23	−0.01	0.16	70	**−5.93 * (0.07)**	**4.60 *** (0.00)**	−0.32	−5.85	5.76	−0.02	0.13	72
Forecast Horizon	Palladium (1986:04–2003:12)								Zinc (1997:08–2003:12)							
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	10.57	−5.02	**−14.64 * (0.05)**	−13.17	**32.11 ** (0.04)**	−0.09	0.08	72	**−21.84 * (0.06)**	6.90	8.09	−25.65	11.71	0.03	0.10	26
2	7.38	−4.78	−3.63	−11.59	**30.69 *** (0.01)**	−0.07	0.07	72	**−18.30 ** (0.03)**	**6.16 * (0.08)**	5.78	−27.69	6.57	0.00	0.18	26
3	3.17	−3.79	3.40	−4.19	**35.91 *** (0.00)**	−0.02	0.14	72	**−17.09 *** (0.01)**	**6.44 ** (0.02)**	3.55	**−24.73 * (0.08)**	0.02	0.08	0.28	26
4	1.10	−2.19	3.55	−4.48	**29.72 *** (0.00)**	−0.02	0.11	72	**−17.03 *** (0.00)**	**7.41 *** (0.00)**	2.01	−12.49	4.14	0.08	0.35	26
Forecast Horizon	Natural gas (1990:05–2003:12)															
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N								
1	6.68	−4.48	18.32	−12.18	36.38	0.01	−0.04	55								
2	−1.62	1.67	5.15	−33.25	23.58	−0.04	−0.03	55								
3	−11.31	5.34	5.33	**−49.12 *** (0.01)**	**24.34 * (0.08)**	0.10	0.14	55								
4	−8.91	6.01	4.30	**−31.30 ** (0.03)**	15.38	0.09	0.09	55								
Panel B																
Forecast Horizon	Oil (2004:01–2020:12)								Silver (2004:01–2020:12)							
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−3.42	1.15	0.77	16.91	**−16.04 ** (0.04)**	0.17	0.04	67	14.08	−4.73	−2.40	−14.42	−4.86	0.08	−0.03	67
2	0.85	−0.50	0.77	9.36	−6.13	0.07	−0.03	66	**16.73 * (0.09)**	−3.71	−7.89	−11.72	−3.19	−0.06	0.06	66
3	1.44	−0.75	−6.04	−16.50	**−6.46 * (0.09)**	−0.04	0.04	65	**13.60 * (0.09)**	−3.04	−4.77	−12.10	−4.19	−0.03	0.06	65
4	1.53	−1.12	−1.02	−2.06	8.57	−0.04	−0.05	64	**14.52 ** (0.03)**	−2.39	**−6.70 * (0.06)**	3.10	0.22	−0.06	0.04	64
Forecast Horizon	Gold (2004:01–2020:12)								Platinum (2004:01–2020:12)							
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	**14.67 ** (0.03)**	−2.87	−0.14	3.57	−1.76	0.02	−0.03	67	6.63	−3.78	−4.97	0.07	−1.86	0.12	−0.03	67
2	**16.43 *** (0.00)**	−3.02	−3.15	−0.36	−0.57	−0.03	0.04	66	10.72	−4.10	**−10.40 ** (0.02)**	−1.77	1.58	−0.01	0.04	66
3	**16.26 *** (0.00)**	**−3.04 * (0.07)**	−1.54	0.74	−0.21	−0.02	0.01	65	9.37	−3.91	−4.31	−10.04	−1.87	−0.05	0.07	65
4	**17.07 *** (0.00)**	**−2.89 ** (0.04)**	−2.83	5.14	−1.82	−0.03	0.10	64	**9.44 * (0.08)**	**−3.58 * (0.09)**	−3.53	1.58	2.15	−0.05	0.05	64
Forecast Horizon	Palladium (2004:01–2020:12)								Zinc (2004:01–2020:12)							
h	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_1)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	25.07	−6.88	−7.61	−6.89	5.57	0.11	−0.02	67	8.66	−3.34	**−15.66 ** (0.03)**	−15.93	2.38	**0.21 ** (0.04)**	0.05	67
2	**27.23 ** (0.02)**	−5.95	−8.68	−16.16	2.20	−0.06	0.02	66	6.70	−0.90	−6.87	−9.71	0.79	0.03	−0.05	66
3	**25.42 *** (0.01)**	−5.14	−2.15	−13.24	0.70	−0.07	0.00	65	2.66	1.13	−5.71	−2.99	−1.35	0.01	−0.05	65
4	**23.11 *** (0.00)**	−3.43	−3.49	−4.53	1.13	−0.07	−0.02	64	1.69	1.82	−2.61	10.69	4.00	−0.01	−0.05	64

Table A6. *Cont.*

Panel B								
Forecast Horizon	Natural gas (2004:01–2020:12))							
h	C	(Y_{30}-Y_1)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−1.61	−1.10	17.31	29.03	0.16	−0.02	−0.01	67
2	−3.34	−1.97	2.82	−7.08	−4.75	0.06	−0.06	66
3	−3.84	−2.42	−1.37	−18.66	−5.96	0.04	−0.01	65
4	−3.23	−2.78	−0.79	−10.94	10.00	0.03	−0.02	64

Notes: The forecast horizon (h) is in quarters. Y_{30}-Y_1 denotes the yield spread calculated as the difference between the yield rates on 10-year and 1-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A7. Estimation results of Equation (1) with the Y_{30}-Y_2 indicator. Panel A: sample period 1986–2003. Panel B: sample period 2004–2020.

Panel A																
Forecast Horizon	Oil (1986:01–2003:12)								Silver (1986:01–2003:12)							
h	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	11.32	−5.65	0.56	−1.99	9.36	0.04	−0.06	72	**−11.97 * (0.09)**	**10.26 *** (0.01)**	−0.13	−11.33	−1.33	0.06	0.04	72
2	8.40	−1.97	**−9.57 * (0.07)**	−5.43	8.52	0.00	0.00	72	**−10.23 ** (0.03)**	**8.73 *** (0.00)**	0.31	−6.35	−0.24	−0.01	0.09	72
3	6.51	0.05	−5.46	−4.60	0.70	0.03	−0.03	72	**−11.45 *** (0.00)**	**9.29 *** (0.00)**	0.12	**−9.80 * (0.07)**	−0.58	0.00	0.22	72
4	4.35	0.81	−3.17	−8.10	0.73	0.04	−0.03	72	**−10.68 *** (0.00)**	**8.84 *** (0.00)**	−0.38	**−8.09 * (0.07)**	−0.44	0.00	0.28	72
Forecast Horizon	Gold (1986:01–2003:12)								Platinum (1986:04–2003:12)							
h	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−7.86	**4.87 * (0.09)**	4.29	−10.48	4.62	0.01	0.03	70	−4.23	4.62	−2.11	**−22.94 ** (0.03)**	7.92	−0.07	0.06	72
2	−4.30	**5.30 *** (0.00)**	−1.66	2.97	−3.45	−0.01	0.10	70	−4.00	**4.81 * (0.05)**	−1.76	**−14.35 ** (0.03)**	4.89	−0.04	0.08	72
3	**−5.34 ** (0.02)**	**5.13 *** (0.00)**	−0.75	0.19	−1.82	−0.01	0.16	70	**−5.61 * (0.09)**	**5.12 *** (0.01)**	0.19	**−10.41 ** (0.04)**	**7.85 * (0.06)**	−0.01	0.13	72
4	**−5.36 *** (0.00)**	**5.10 *** (0.00)**	−1.14	−0.92	−2.58	−0.01	0.21	70	**−5.15 * (0.09)**	**5.36 *** (0.00)**	−0.35	−5.71	**6.60 * (0.08)**	−0.02	0.14	72
Forecast Horizon	Palladium (1986:04–2003:12)								Zinc (1997:08–2003:12)							
h	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	8.32	−4.76	**−14.73 ** (0.04)**	−13.64	**30.87 ** (0.04)**	−0.08	0.07	72	**−20.44 * (0.07)**	7.38	8.17	−25.51	12.79	0.03	0.09	26
2	6.49	−5.51	−3.60	−11.75	**29.80 *** (0.01)**	−0.07	0.07	72	**−17.71 ** (0.03)**	7.05 * (0.07)	5.76	−27.27	7.72	0.00	0.19	26
3	2.48	−4.39	3.43	−4.32	**35.21 *** (0.00)**	−0.02	0.14	72	**−16.55 *** (0.01)**	**7.43 ** (0.02)**	3.52	**−24.26 * (0.09)**	1.24	0.08	0.30	26
4	0.64	−2.48	3.56	−4.56	**29.30 *** (0.00)**	−0.02	0.11	72	**−16.49 *** (0.00)**	**8.59 *** (0.00)**	1.96	−11.91	5.56	0.08	0.39	26
Forecast Horizon	Natural Gas (1990:05–2003:12)															
h	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N								
1	6.22	−5.39	18.38	−12.43	35.71	0.01	−0.04	55								
2	1.23	0.21	5.31	−32.64	24.24	−0.04	−0.03	55								
3	−7.52	4.24	5.48	**−48.20 *** (0.01)**	**25.63 * (0.07)**	0.10	0.13	55								
4	−5.53	5.36	4.41	**−30.44 ** (0.03)**	16.70	0.08	0.07	55								

Panel B																
Forecast Horizon	Oil (2004:01–2020:12)								Silver (2004:01–2020:12)							
h	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	(Y_{30}-Y_2)	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	−3.48	1.30	0.75	16.95	**−16.03 ** (0.04)**	0.17	0.04	67	13.47	−4.90	−2.36	−14.55	−4.93	0.08	−0.03	67
2	−0.15	−0.03	0.74	9.42	−6.15	0.07	−0.03	66	**16.55 * (0.09)**	−4.01	−7.84	−11.85	−3.23	−0.06	0.06	66
3	1.47	−0.84	−6.02	−16.53	**−6.47 * (0.09)**	−0.04	0.04	65	**14.86 * (0.07)**	−4.02	−4.68	−12.32	−4.20	−0.03	0.07	65
4	2.57	−1.79	−0.95	−2.18	8.67	−0.04	−0.05	64	**16.11 ** (0.02)**	−3.48	**−6.58 * (0.06)**	2.90	0.34	−0.06	0.05	64

Table A7. *Cont.*

Panel B																
Forecast Horizon	Gold (2004:01–2020:12)								Platinum (2004:01–2020:12)							
h	C	$(Y_{30}$-$Y_2)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_2)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	14.78 ** (0.03)	−3.23	−0.09	3.45	−1.79	0.02	−0.03	67	5.38	−3.52	−4.97	0.03	−1.93	0.12	−0.03	67
2	16.70 *** (0.00)	−3.48	−3.10	−0.50	−0.60	−0.03	0.04	66	9.64	−3.96	−10.39 ** (0.02)	−1.84	1.51	−0.01	0.04	66
3	16.88 *** (0.00)	−3.69 * (0.05)	−1.48	0.58	−0.24	−0.02	0.02	65	9.76	−4.53	−4.24	−10.22	−1.92	−0.05	0.07	65
4	17.75 *** (0.00)	−3.55 ** (0.03)	−2.76	4.97	−1.83	−0.03	0.11	64	10.18 * (0.06)	−4.35 * (0.07)	−3.45	1.39	2.12	−0.05	0.05	64
Forecast Horizon	Palladium (2004:01–2020:12)								Zinc (2004:01–2020:12)							
h	C	$(Y_{30}$-$Y_2)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N	C	$(Y_{30}$-$Y_2)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N
1	23.62	−6.84	−7.57	−7.03	5.47	0.11	−0.03	67	9.55	−4.16	−15.58 ** (0.03)	−16.12	2.35	0.21 ** (0.04)	0.05	67
2	26.66 ** (0.03)	−6.28	−8.62	−16.34	2.11	−0.06	0.02	66	8.53	−1.96	−6.79	−9.89	0.81	0.04	−0.05	66
3	26.79 *** (0.01)	−6.41	−2.03	−13.55	0.66	−0.07	0.01	65	5.05	0.00	−5.65	−3.13	−1.27	0.01	−0.05	65
4	24.98 *** (0.00)	−4.78	−3.35	−4.79	1.26	−0.07	−0.01	64	4.33	0.61	−2.48	10.59	4.39	−0.01	−0.06	64
Forecast Horizon	Natural Gas (2004:01–2020:12)															
h	C	$(Y_{30}$-$Y_2)$	ΔSP	ΔEX	ΔIP	ΔEPU	R^2	N								
1	−2.07	−0.97	17.31	29.02	0.14	−0.02	−0.01	67								
2	−4.18	−1.73	2.81	−7.08	−4.79	0.06	−0.06	66								
3	−3.91	−2.64	−1.33	−18.75	−6.00	0.04	−0.01	65								
4	−3.57	−2.90	−0.78	−11.03	9.87	0.03	−0.03	64								

Notes: The forecast horizon (h) is in quarters. Y_{30}-Y_2 denotes the yield spread calculated as the difference between the yield rates on 10-year and 2-year government bonds. The table reports the estimation results of Equation (1) with the Newey and West (1987) procedure. The sample period appears separately for each commodity. Figures in parentheses denote estimated standard errors. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Note

1 The conditional correlation between two random variables y1 and y2 is $\rho_{12,t} = E_{t-1}(y_{1t}y_{2t})/\sqrt{E_{t-1}(y_{1t}^2)E_{t-1}(y_{2t}^2)}$. It is acceptable to present returns as the conditional standard deviation times the standardized disturbance. $y_{it} = \sqrt{h_{it}}\varepsilon_{it}$. This is because $h_{it} = E_{t-1}(y_{it}^2)$. For each series i, ε_{it} is a standardized disturbance with a mean of zero and a variance of one. Accordingly, the conditional correlation can be presented as $\rho_{12,t} = E_{t-1}(\varepsilon_{1t}\,\varepsilon_{2t})/\sqrt{E_{t-1}(\varepsilon_{1t}^2)E_{t-1}(\varepsilon_{2t}^2)} = E_{t-1}(\varepsilon_{1t}\varepsilon_{2t})$. Hence, the conditional correlation is also the conditional covariance between the standardized disturbances. This is the spirit of the DCC method.

References

Abdymomunov, Azamat. 2013. Predicting output using the entire yield curve. *Journal of Macroeconomics* 37: 333–44. [CrossRef]

Akram, Farooq. 2009. Commodity prices, interest rates and the dollar. *Energy Economics* 31: 838–51. [CrossRef]

Arango, Luis, Fernando Arias, and Adriana Flórez. 2011. Determinants of commodity prices. *Applied Economics* 44: 135–45. [CrossRef]

Bai, Jushan, and Pierre Perron. 2003. Critical values for multiple structural change tests. *Econometrics Journal* 6: 72–78. [CrossRef]

Bampinas, Georgios, and Theodore Panagiotidis. 2015. Are gold and silver a hedge against inflation? A two century perspective. *International Review of Financial Analysis* 41: 267–76. [CrossRef]

Batten, Jonathan, Cetin Ciner, and Brian Lucey. 2010. The macroeconomic determinants of volatility in precious metals markets. *Resources Policy* 35: 65–71. [CrossRef]

Beckmann, Joscha, and Robert Czudaj. 2013. Gold as an inflation hedge in a time-varying coefficient framework. *North American Journal of Economics and Finance* 24: 208–22. [CrossRef]

Bodie, Zvi, and Victor Rosansky. 1980. Risk and return in commodity futures. *Financial Analysts Journal* 36: 27–39. [CrossRef]

Bollerslev, Tim. 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics* 72: 498–505. [CrossRef]

Chevallier, Julien, Mathieu Gatumel, and Florian Ielpo. 2014. Commodity markets through the business cycle. *Quantitative Finance* 14: 1597–618. [CrossRef]

Chinn, Menzie, and Kavan Kucko. 2015. The predictive power of the yield curve across countries and time. *International Finance* 18: 129–56. [CrossRef]

Churchill, Sefa, John Inekwe, Kris Ivanovski, and Russell Smyth. 2019. Dynamics of oil price, precious metal prices and the exchange rate in the long-run. *Energy Economics* 84: 104508. [CrossRef]

Dai, Zhifeng, and Jie Kang. 2021. Bond yield and crude oil prices predictability. *Energy Economics* 97: 105205. [CrossRef]

Daskalaki, Charoula, George Skiadopoulos, and Nikolas Topaloglou. 2017. Diversification benefits of commodities: A stochastic dominance efficiency approach. *Journal of Empirical Finance* 44: 250–69. [CrossRef]

Demiralay, Sercan, Selcuk Bayraci, and Gaye Gencer. 2019. Time-varying diversification benefits of commodity futures. *Empirical Economics* 56: 1823–53. [CrossRef]
Duarte, Angelo, Wagner Gaglianone, Osmani de Carvalho Guillén, and João Issler. 2021. Commodity Prices and Global Economic Activity: A derived-demand approach. *Energy Economics* 96: 105120. [CrossRef]
Duffee, Gregory. 1998. The relation between treasury yields and corporate bond yield spreads. *Journal of Finance* 53: 2225–41. [CrossRef]
Engle, Robert. 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20: 339–50.
Estrella, Arturo. 2005. Why does the yield curve predict output and inflation? *Economic Journal* 115: 722–44. [CrossRef]
Estrella, Arturo, and Gikas Hardouvelis. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46: 555–76. [CrossRef]
Evgenidis, Anastasios, Stephanos Papadamou, and Costas Siriopoulos. 2020. The yield spread's ability to forecast economic activity: What have we learned after 30 years of studies? *Journal of Business Research* 106: 221–32. [CrossRef]
Fama, Eugene, and Kenneth French. 1988. Business cycles and the behavior of metals prices. *The Journal of Finance* 43: 1075–93. [CrossRef]
Fasanya, Ismail, and Crystal Awodimila. 2020. Are commodity prices good predictors of inflation? The African perspective. *Resources Policy* 69: 101802. [CrossRef]
Gagnon, Marie, Guillaume Manseau, and Gabriel Power. 2020. They're back! Post-financialization diversification benefits of commodities. *International Review of Financial Analysis* 71: 101515. [CrossRef]
Garner, Alan. 1989. Commodity prices: Policy target or information variable?: Note. *Journal of Money, Credit and Banking* 21: 508–14. [CrossRef]
Giacomini, Raffaella, and Barbara Rossi. 2006. How stable is the forecasting performance of the yield curve for output growth? *Oxford Bulletin of Economics and Statistics* 68: 783–95. [CrossRef]
Gogas, Periklis, Theophilos Papadimitriou, Maria Matthaiou, and Efthymia Chrysanthidou. 2015. Yield curve and recession forecasting in a machine learning framework. *Computational Economics* 45: 635–45.
Gorton, Gary, and Geert Rouwenhorst. 2006. Facts and fantasies about commodity futures. *Financial Analysts Journal* 62: 47–68. [CrossRef]
Granger, Clive. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society* 37: 424–38. [CrossRef]
Hamilton, James, and Dong Kim. 2002. A reexamination of the predictability of economic activity using the yield spread. *Journal of Money, Credit, and Banking* 34: 340–60. [CrossRef]
Hamilton, James, and Jing Wu. 2015. Effects of index-fund investing on commodity futures prices. *International Economic Review* 56: 187–205. [CrossRef]
Harvey, Campbell. 1989. Forecasts of economic growth from the bond and stock markets. *Financial Analysts Journal* 45: 38–45. [CrossRef]
Harvey, David, Neil Kellard, Jakob Madsen, and Mark Wohar. 2017. Long-run commodity prices, economic growth, and interest rates: 17th century to the present day. *World Development* 89: 57–70. [CrossRef]
Henderson, Brian, Neil Pearson, and Li Wang. 2015. New evidence on the financialization of commodity markets. *Review of Financial Studies* 28: 1285–311. [CrossRef]
Hu, Conghui, Zhibing Li, and Xiaoyu Liu. 2020. Liquidity shocks, commodity financialization, and market comovements. *Journal of Futures Markets* 40: 1315–36. [CrossRef]
Huang, Jianbai, Yingli Li, Hongwei Zhang, and Jinyu Chen. 2021. The effects of uncertainty measures on commodity prices from a time-varying perspective. *International Review of Economics & Finance* 71: 100–14.
Ioannidis, Christos, and Kook Ka. 2018. The impact of oil price shocks on the term structure of interest rates. *Energy Economics* 72: 601–20. [CrossRef]
Jahan, Sayeeda, and Apostolos Serletis. 2019. Business cycles and hydrocarbon gas liquids prices. *The Journal of Economic Asymmetries* 19: e00115. [CrossRef]
Kagraoka, Yusho. 2016. Common dynamic factors in driving commodity prices: Implications of a generalized dynamic factor model. *Economic Modelling* 52: 609–17. [CrossRef]
Kucher, Oleg, and Suzanne McCoskey. 2017. The long-run relationship between precious metal prices and the business cycle. *Quarterly Review of Economics and Finance* 65: 263–75. [CrossRef]
Kuosmanen, Petri, Jaana Rahko, and Juuso Vataja. 2019. Predictive ability of financial variables in changing economic circumstances. *North American Journal of Economics and Finance* 47: 37–47. [CrossRef]
Labys, Walter, Aymad Achouch, and Michael Terraza. 1999. Metal prices and the business cycle. *Resources Policy* 25: 229–38. [CrossRef]
Levine, Ari, Yao Ooi, Matthew Richardson, and Caroline Sasseville. 2018. Commodities for the long run. *Financial Analysts Journal* 74: 55–68. [CrossRef]
McMillan, David. 2021a. Predicting GDP growth with stock and bond markets: Do they contain different information? *International Journal of Finance and Economics* 26: 3651–75. [CrossRef]
McMillan, David. 2021b. When and Why Do Stock and Bond Markets Predict Economic Growth? *Quarterly Review of Economics and Finance* 80: 331–43. [CrossRef]

Newey, Whitney, and Kenneth West. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review* 28: 777–87. [CrossRef]

Orlowski, Lucjan. 2017. Volatility of commodity futures prices and market-implied inflation expectations. *Journal of International Financial Markets, Institutions and Money* 51: 133–41. [CrossRef]

Plosser, Charles, and Geert Rouwenhorst. 1994. International term structures and real economic growth. *Journal of Monetary Economics* 33: 133–55. [CrossRef]

Qadan, M., and H. Nama. 2018. Investor sentiment and the price of oil. *Energy Economics* 69: 42–58. [CrossRef]

Qadan, Mahmoud, David Aharon, and Ron Eichel. 2019. Seasonal patterns and calendar anomalies in the commodity market for natural resources. *Resources Policy* 63: 101435. [CrossRef]

Saar, Dan, and Yossi Yagil. 2015. Forecasting growth and stock performance using government and corporate yield curves: Evidence from the European and Asian markets. *Journal of International Financial Markets, Institutions and Money* 37: 27–41. [CrossRef]

Stock, James, and Mark Watson. 1989. New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual* 4: 351–94. [CrossRef]

Stock, James, and Mark Watson. 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 41: 788–829. [CrossRef]

Tang, Ke, and Wei Xiong. 2012. Index investment and the financialization of commodities. *Financial Analysts Journal* 68: 54–74. [CrossRef]

Umar, Zaghum, Syed Shahzad, and Dimitris Kenourgios. 2019. Hedging US metals & mining Industry's credit risk with industrial and precious metals. *Resources Policy* 63: 101472.

Wheelock, David, and Mark Wohar. 2009. Can the term spread predict output growth and recessions? A survey of the literature. *Federal Reserve Bank of St. Louis Review* 91: 419–40. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Journal of Risk and Financial Management Editorial Office
E-mail: jrfm@mdpi.com
www.mdpi.com/journal/jrfm

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34

www.mdpi.com

ISBN 978-3-0365-5846-2