games

MDPI

Article

The Impact of Discretionary Measurement Criteria on Investors' Judgement and Decisions

Tiago Cruz Gonçalves

Advance/CSG, ISEG—Lisbon School of Economics & Management, Universidade de Lisboa, 1200-781 Lisboa, Portugal; tiago@iseg.ulisboa.pt; Tel.: +351-213-925-800

Abstract: This study investigates the effect on nonprofessional investors' judgements and decisions of discretionary measurement choices. Using a paper-and-pencil experience, we collect and analyze information regarding investment amounts as well as past and future financial performance judgements of firms' earnings by manipulating fair value (mark-to-market and mark-to-model) criteria and benchmarking it with historical cost-based financial statements. We proxy nonprofessional investors with graduate students from a business school. Our results show evidence that nonprofessional investors view fair value changes as permanent. We argue for a cashflow volatility factor. Contrary to previous research, we do not find evidence of any effect on investors' willingness to invest (average budget amounts invested) or performance judgments (past and future). We corroborate previous evidence that investors rank measurement concepts' relevance differently for different classes, although, on average, mark-to-market fair values and historical cost are rated more relevant and reliable than mark-to-model fair values.

Keywords: measurement theory; nonprofessional investors; judgement and decision; fair value; mark-to-market vs. mark-to-model

Citation: Gonçalves, T.C. The Impact of Discretionary Measurement Criteria on Investors' Judgement and Decisions. *Games* **2022**, *13*, 3. https://doi.org/10.3390/g13010003

Academic Editors: Anna Conte, Patrizia Sbriglia, Marco Stimolo and Ulrich Berger

Received: 31 October 2021
Accepted: 21 December 2021
Published: 23 December 2021

1. Introduction

When market prices reflect all value-relevant information, significant advantages of fair value accounting emerge as market prices (fair value) equal value in use, but only under perfect and complete markets assumption. Accordingly, if we recognize all assets and liabilities on the balance sheet and measure them at market price (fair value), the book value of net assets reports the market value of equity. So, under perfect and complete markets, investors do not need to estimate equity value because the balance sheet reports the equity value through fair value accounting. When the market is imperfect, on the other hand, it is necessary to determine a value with a method whose reliability must be proven.

In fact, the discussion about the measurement of fair value makes relevant the distinction between price and value. The question is whether accounts should reflect prices or values.

Recently, financial crises, especially the one initiated in 2008, led to a considerable debate on the pros and cons of using a full mark-to-market accounting system. Contemporaneously, the US Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) took steps in this direction in an attempt to globalize accounting standards and reduce management (strategic) discretionary. The accounting standards SFAS 157 and International Accounting Standard (IAS) 39 adapt the fair value approach and attempt to use only market prices where appropriate. For example, SFAS 157 distinguishes between different levels of input to the valuation process. Level 1 input are quoted prices in active markets for identical assets or liabilities. In cases where market prices are not appropriate, level 2 inputs should be used if possible. Examples include quoted prices for similar assets and interest rate and yield curves or other market corroborated inputs. Finally, if this kind of information is also unavailable, then level 3 inputs

can be used, consisting of unobservable prices that reflect firms' own assumptions and information about the asset (mark-to-model). IAS 39 has similar provisions.

Measurement theory is key for financial investing and reporting. In fact, identifying the measurement criteria most adequate is of interest to practitioners, standard setters and academics as well. Standard setters face the (mostly political) problem of identifying accounting measurement concepts that provide the needs of information for a group of heterogeneous users and settings [1].

This study investigates the effect on nonprofessional investors' judgements and decisions of different measurement discretionary choices. Using a paper-and-pencil experience, we collect and analyze information regarding investment amounts as well as past and future financial performance judgements of firms' earnings by manipulating the aforementioned fair value (multiple level) criteria and benchmarking it with historical cost-based financial statements. We proxy nonprofessional investors with graduate students from a top business school in Lisbon (ISEG). We refer to nonprofessional investors, according to New York Stock Exchange (NYSE), as any natural person who receives market data solely for his/her personal, non-business use and who is not a "Securities Professional," meaning that the person is not registered or qualified with the Securities and Exchange Commission, the Commodities Futures.

We find evidence that nonprofessional investors' investment decision is affected vis-à-vis a cashflow estimation factor but not in their willingness to invest. Investment amounts of the total budget remain statistically unchanged when we manipulate the measurement of assets by using multi-level fair values where discretion is allowed. Contrarily, participants view fair value changes as permanent. Consequently, we argue that fair value changes recognition will induce volatility on future cashflows forecasted to evaluate an investment's fundamental value.

We do not find evidence that past or future performance judgements are affected by our manipulation of measurement criteria. Contrary to prediction, nonprofessional investors remain unaffected by any fair value recognition when assessing past performance. They also do not feel less confident in predicting future earnings when presented with fair value-based financial statements. A potential explanation may be that, as argued in the extant literature, familiarity and expertise may mitigate the predicted effects.

Finally, we find that nonprofessional investors view measurement criteria differently in terms of reliability and relevance. We find that regardless of between-group differences, participants view historical cost as most relevant. Additionally, there seem to be different relevance assessments for different classes of assets, across the same measurement criteria. As Gassen and Schwedler [2] argue, the decision usefulness of a specific measurement criterium is also influenced by the class of assets to be measured. Regarding reliability, our results show no effect for different measurement concepts. Finally, we find that participants rate lower and similarly mark-to-model reliability and relevance, but they distinguish those two attributes when assessing historical cost and mark-to-market measurements.

Our paper extends previous investigations. Similar to Warne [3] we analyze the impact of fair value measurement on nonprofessional investors' judgement and decisions. Extant research [3], using an experiment from which we adapt part of our research instrument, shows that fair value recognition of non-current assets has an impact on those investors. The author finds evidence that investors are less willing to invest and are less confident about their performance judgement of the firms that report fair value recognitions of non-current assets, as opposed to historical cost disclosures. We do not find evidence of any impact on willingness to invest. Furthermore, our paper extends this research twofold: we research that impact under a completely different financial reporting environment. International Financial Reporting Standards (IFRS) extensively allow for fair value recognition not only on the same non-current assets but other assets and liabilities. In Warne [3], Generally Agreed Accounting Principals (GAAP) didn't allow fair value (directly) for those items. This can have competing effects of (lack of) familiarity on participants' answers, that we are able to rule out by using graduate students already familiar with IFRS. Additionally,

we are able to conduct our experiment on a setting where multiple levels of Fair values (mark-to market or model) are already in place and can be fine-tuned to test differences for those levels on nonprofessional investors' perceptions and decisions. In addition, a paper by Gassen and Schwedler [2] surveys professional investors to identify the decision usefulness of different accounting measurement concepts. They find that respondents distinguish between mark-to-market and mark-to-model fair values. Furthermore, professional investors rank mark-to-market fair values as most decision-useful. They also show evidence that respondents rank as least decision-useful mark-to-model fair values. Results differed across asset classes. We make some bridging to their conclusions by assessing the relevance and reliability of different classes of assets and by surveying familiarity with mark-to-market and mark-to-model fair values. We extend their analysis by researching the effect of measurement criteria on relatively less sophisticated investors, which have been declared by supervisors and regulators as the main concern when looking for improvement on regulation. We also differ from the aforementioned study by employing an experiment methodology as opposed to their survey.

Results are important for a broad group of individuals. Financial statements' preparers (and users) learn that several competing consequences underly their measurement concepts choices and that those discretionary choices bear additional unattended (and probably unwanted) results on valuation volatility and investors' confidence. Standard setters and regulators may find that our results present effects on judgement and decisions of nonprofessional investors that are statistically and economically relevant and, thus, should be balanced in their work. Finally, academics face additional layers of research that deem the debate about fair value measurement advantages yet not fully explored.

The paper is organized as follows. Section 2 reviews extant literature and presents the research hypothesis; Section 3 concerns research design and data; Section 4 describes the main results; and Section 5 concludes the paper.

2. Literature Review and Research Hypothesis

The latest financial crises had prevailing effects on global economies. In the Financial Reporting realm, this led to a considerable debate on the pros and cons of using a full mark-to-market accounting system for banks and insurance companies. Contemporaneously, the US FASB and the IASB took steps in this direction in an attempt to globalize accounting standards. The accounting standards SFAS 157 and IAS 39 adapt the fair value approach and attempt to use only market prices where appropriate. For example, SFAS 157 distinguishes between different levels of input to the valuation process. Level 1 input are quoted prices in active markets for identical assets or liabilities that the reporting entity has the ability to access at the measurement date. In this context, an active market is one with sufficient frequency and volume to provide pricing information on an ongoing basis. In cases where market prices are not appropriate, level 2 inputs should be used if possible. Examples include quoted prices for similar assets and interest rate and yield curves or other market corroborated inputs. Finally, if this kind of information is also unavailable, then level 3 inputs can be used, consisting of unobservable prices that reflect the firm's own assumptions and information about the asset. IFRS have similar provisions.

Thus, fair value accounting is one that updates the measurement of balance sheet items to the most recent data, as opposed to the historical cost measurement criterium. Fair value measurement can be applied at three different levels—the full mark-to-market model where assets are valued at liquid market prices; the mixed model where market prices are used to assess fair values of items not traded on liquid markets; and, finally, level 3 mark-to-model criterium where companies use their best estimates to update item's value. We will refer to fair value accounting bearing in mind that those three levels are ranked and companies are only allowed to relax pure market prices where there is no liquid markets' information available but lead to different perceptions (in terms of judgements and confidence) when information users are faced with financial reporting data.

Measurement theory is key for financial reporting research and standards-setting [4,5]. In fact, most theoretical foundations used by accounting researchers come from the fields of economics; psychology and other social sciences. One exception is the former theory. Previous literature and regulators associate historical cost measurement with greater reliability and fair value-based measurement with increasing relevance. Relevance requires that the financial accounting information should be such that the users need it and it is expected to affect their decisions. Reliability requires that the information should be accurate and true and fair. None-the-less, some call for external verifiability of measurement models (i.e., [6]). Adding to that, different settings (in terms of both the users and the context) deem measurement concepts ranked differently. Gassen and Schwedler [2] provide evidence of this phenomenon by surveying professional investors and their advisors, about their opinions on the decision usefulness of different accounting measurement concepts. They find that respondents clearly distinguish between mark-to-market and mark-to-model fair values. While they consistently rank mark-to-market fair values as most decision-useful, they generally rank mark-to-model fair values as least decision-useful. In addition, the ranking differs across asset classes.

Measurement theory is also key for financial reporting. In fact, identifying the measurement criteria most adequate is of interest to practitioners, standard setters and academics as well. Standard setters face the (mostly political) problem of identifying accounting measurement concepts that provide the needs of information for an ex-ante unknown group of heterogeneous users and settings [1,4,5].

The debate on fair value accounting raises issues that tend to improperly consider the role of accounting lumped together with the suitability of fair value and its measurement. In fact, the discussion about the measurement of fair value makes relevant the distinction between price and value. The question is whether accounts should reflect prices or values.

Keynes [7] defines the value of an asset as resulting from the discounting of its cash flows generated by ownership (intrinsic value). Price, contrarily, is the outcome of the law of supply and demand. Although based on the value of the company, the price also reflects subjective elements such as the negotiating power of buyers and sellers, their relationship, the desire to complete the transaction, and their interest in doing so, among other elements.

In light of this, only when markets are perfect and complete, market value is fair value. Barth and Landsman [8] argue that under that condition, the balance sheet includes all the information useful to a valuation of a company. In this scenario, management and market are capable of ascertaining the necessary elements of assets to come up with a fair value, and thus making the observation of an income is not necessary to the valuation of a company. When the market is imperfect, on the other hand, it is necessary to determine a value with a method whose reliability must be proven.

In conclusion, the reliability and relevance of the attribute measured are key points of measuring assets especially during increased uncertainty environments such as the recent financial and subsequent economic crises.

Nobes [9] provides the first major analysis of fair value accounting. Plantin et al. [10] and Penman [6] argue the pros and cons of contemporary fair value accounting more deeply. Conceptually, fair value accounting should provide information with a higher degree of decision usefulness and relevance of accounting data that would mitigate information asymmetries that investors face in the market. Additionally, fair value also decreases incentives to increase gains on trading and assets securitization, providing more credibility to financial reporting. Conversely, if fair value cannot be determined unambiguously it loses objectivity. As Ryan [11] argues, when active markets are missing, fair value can only be measured according to subjective assumptions and thus become a black box tool for discretionary earning management.

Existing literature focus attention on the role of fair value disclosure on market wide consequences (i.e., [12]) but lacks analysis in terms of individual investors' decision usefulness. An additional stream of literature studies the impact of unrealized gains and losses on judgements and decisions. Both Hirst and Hopkins [13] and Maines and McDaniel [14]

find evidence that unrealized gains and losses included in a statement of comprehensive income affect judgment about firms' performance by analysts and non-professionals, respectively. Bloomfield et al. [15] concluded that unrealized gains and losses increase price (and returns) volatility when the correlation between those unrealized gains and losses and firms' performance is high.

The aforementioned studies focus on investors' reactions to unrecognized gains and losses regarding changes in the value of financial assets and liabilities for which liquid markets already provide mark-to-market fair values and under a financial reporting that requires mandatory recognition of those changes in its values. We extend those studies by providing evidence for additional items where fair value changes are optional and under a financial reporting environment where firms are able to use level 3 (mark-to-model) fair values.

Our study also relates to another recent stream of literature that evaluates investors' judgement and perceptions regarding financial reporting disclosed by firms of pro-forma earnings. Elliott [16] shows that nonprofessional investors are influenced by the emphasis placed on pro-forma profit relative to GAAP loss. The presentation of a reconciliation between those two figures doesn't seem to mitigate that evidence unless a side-by-side format is adopted. Contrarily, professional investors' judgements are not influenced by the pro forma disclosure unless there is such a side-by-side reconciliation of both numbers. Frederickson et al. [17] find similar results.

According to Maines and McDaniel [14], nonprofessional investors engage in sequential information search strategies while using financial reporting data. Additionally, this group looks for cues from management to determine the relative importance of information. Research on judgment and decision making has also shown that the mere order of information, regardless of its relevance to the current task, may have effects on information processing. Tversky and Kahneman [18] show evidence that individuals correlate importance with serial position. When uncertain about the estimate they want to report, the first piece of evidence serves as an anchor for the judgement task. Finally, research that analyzed both professional and nonprofessional investors consistently finds that nonprofessional investors are more susceptible to irrelevant information and, thus, engage in non-normative judgement and decisions [16]. Consequently, the effects of different measurement choices should likely be more pronounced in nonprofessional investors.

Extant literature also shows that accounting choices can have an impact on stock prices volatility. Previous research on unrealized gains and losses (UGL) shows evidence that nonprofessional investors are affected by UGL, especially when the latter are correlated with prior returns [15]. Barth et al. [19] also argue that recognizing assets at fair values increases more volatility than historical cost-based measurement. Finally, as this volatility is a key component of non-systematic risk, the discretionary adoption of fair value-based measurement should affect the equity risk [20]. Gonçalves et al. [4] further argue that if investors are not awarded additional risk premium for increased volatility then investment in more volatile equity will decrease.

As part of this decision investment, investors are faced with the task of predicting future cash flows that will justify fair value for the stock traded. To do that, they need to evaluate accounting income. At our setting, bottom line income will include two components: transitory and persistent income. According to Ou and Penman [21], transitory income will have no predictive value of future income. Only persistent income can be predictive of future wealth growth unless a liquidation view is adopted.

None-the-less, psychological theory of causal stability (e.g., [22]) predicts that individuals, when faced with changes, will look for and evaluate the sources of those changes to determine its recurrence into the future. Consequently, if nonprofessional investors depart from rational economic analysis, they will allow spillover effects of transitory income when predicting future earnings.

Similarly, we hypothesize the following:

Hypothesis 1 (H1): *Nonprofessional investors analyzing financial statements produced under fair value-based measurement criteria will be willing to invest less than investors receiving historical cost-based financial statements.*

Hypothesis 2 (H2): *Nonprofessional investors analyzing financial statements produced under fair value-based measurement criteria will view changes from fair values as persistent/ non-transitory income.*

Contemporaneous to this process of decision, we can identify moderating factors that will affect nonprofessional investors' decision to invest in a stock. In fact, investors' decisions will be based not only on predicted future financial performance per si but also on judgements including confidence both on their past performance assessments and on their future performance predictions. Given that, as mentioned before, extant research shows that nonprofessional investors use simple models when making decisions, they will likely see increases to income due to fair values as positive (and conversely, decreases as negative). On top of that, due to "spillover effects" documented by psychology research, their assessment of future performance will probably suffer influence from past fair value adjustments.

Research in accounting shows also that confidence decreases when the complexity of a judgment increases [23]. By that token, judging the future performance of a firm would probably be a more complex task for a nonprofessional investor. Adding to that, additional volatility introduced by fair value changes will likely cause additional complexity if individuals see those changes as non-transitory.

Hypothesis 3 (H3): *Nonprofessional investors analyzing financial statements produced under fair value-based measurement criteria will judge less (more) favorable past performance, if fair value changes decrease (increase) income, than investors receiving historical cost-based financial statements.*

Hypothesis 4 (H4): *Nonprofessional investors analyzing financial statements produced under fair value-based measurement criteria will be less (more) confident about future performance, if fair value changes decrease (increase) income, than investors receiving historical cost-based financial statements.*

In this paper, we decided to test also nonprofessional investors' perception of different measurement concepts in terms of reliability and relevance. We survey participants about the reliability and relevance of different classes of assets, on which measurement choices are manipulated. We similarly extend previous literature [2] by explicitly introducing the distinction between mark-to-market and mark-to-model fair values.

Hypothesis 5 (H5): *Nonprofessional investors will judge mark-to-market fair value-based measurement as most relevant and mark-to-model fair values as least relevant.*

Hypothesis 6 (H6): *Nonprofessional investors will judge historical cost-based measurement more reliable than fair value measurements.*

3. Research Design and Data

Experimental methods are particularly interesting to manipulate variables of interest, while controlling for other irrelevant variables. We choose to manipulate measurement criteria used to report different classes of assets and liabilities in order to evaluate the effects of these choices on nonprofessional investors' judgements and decisions. This research design allows us to rule out alternative explanations and overcome previous research caveats about the effects of fair value reporting on financial decisions. In fact, the extant literature on this topic presents conflicting or inconclusive results due to low power or measurement error [12]. Thomas [24] also notes that results reflect, at best, influential equity

investors. As so, it is of interest to focus our attention on nonprofessional investors, since they remain largely unknown to academics, standard setters and business community.

Previous research uses MBAs as surrogates for nonprofessional investors. Examples include [12,25,26]. Additionally, Elliott et al. [27] provide evidence that graduate students are a reasonable proxy for nonprofessional investors.

One hundred and fifteen graduate students from a Master of Science in finance and accounting and executive education on accounting and finance from a top business school in Lisbon (ISEG) participated in this experiment as proxies for nonprofessional investors. All participants already successfully completed one or more intermediate and advanced courses in financial accounting. In fact, when asked to classify their knowledge of financial reporting and accounting on a 5 point scale with endpoints labeled 1—"unfamiliar" to 5—"very familiar", the average grading was 3.08. As control questions, they were asked to rate their familiarity with several measurement criteria (including historical cost and mark-to-market and mark-to-model fair values) on the same 5 point scale. Average answers ranged from 3.93 for the historical cost to 2.54 for mark-to-model fair values. Two-thirds of the participants were women. The average age was 24.9 years old, which might account for the fact that 12% declared that already invested in equity instruments. Additionally, the average work experience is 2.2 years and almost 30% of the participants stated that they have used financial statements in the context of job tasks.

Participants were randomly assigned to one of the groups designed to test the use of financial statements prepared with or without fair value measurement criteria where discretion is allowed, (some classes of financial assets and liabilities are mandatorily measured at fair value). Results show that groups do not present any statistically significant differences in what concerns any of the demographics collected.

All participants viewed a set of financial statements (balance sheet and income statement along with additional notes regarding assets and liabilities measurement to avoid unintended demand effects). The financial statements were preceded by an introduction where participants were informed that they were about to analyze data from a hypothetical firm modeled after the data found for those financial statements on all non-financial firms traded on Lisbon Euronext Stock Exchange Index—PSI.

We use a two-group between-subject design. The manipulated variable for each group is the measurement criteria used to evaluate some assets and liabilities for which IFRS allows a choice to use fair value with the corresponding effects on comprehensive income. The historical cost (HC) condition serves as a benchmark to examine the effect of fair value multiple levels criterium disclosure on investors' decision to invest and performance and confidence judgment when analyzing financial statements.

We collected data on two earnings performance-dependent variables—current earnings performance and future earning potential judgements. Additionally, we collected information regarding investment decisions as a dependent variable—the investment amount each participant would place on the hypothetical firm. Regarding past performance judgement, participants were asked to rank their opinion on an 11-point scale with endpoints labeled 0 (very weak) to 100 (very strong). For their assessment of future performance potential, an estimated amount of operating income and another for comprehensive income were requested. Additionally, they were asked to rate their confidence in producing those estimates on the same 11-point scale. Participants made an investment decision based on the following instructions: "Assume you have 10,000 € to invest in this stock. Assume also that each stock is currently traded at 2 € per share, immediately after the disclosure of the attached financial statements. How much of your initial budget would you invest on the stocks of this firm?"

To make our conclusions comparable to previous literature (e.g., [2]), we also asked participants to evaluate both the reliability and relevance of the different measurement criteria manipulated in this experiment. To prevent drawing attention to the distinction between different criteria, participants were provided with the definition of relevance and reliability as defined on Portuguese accounting standards (based on IFRS) and then asked

to rank those attributes for a class of assets produced according to different measurement criteria. In an 11-point scale for which endpoints are labeled 0 (not at all) to 100 (very), participants ranked both reliability and relevance for Cash and Marketable Securities and Investment Buildings (Fair value—mark-to-market); for Production Equipment (Fair value—mark-to-model) and Accounts Receivables (Historical Cost).

4. Empirical Results

Descriptive statistics and test results for the dependent variables defined in the several research hypotheses (H) are presented in a set of tables presented below.

H1 predicts that when participants are presented with financial statements prepared with fair value-based measurements, investment amounts they are willing to invest will be lower than when given historical cost-based financial statements. Table 1 corroborates this assertion. Panel A shows that the average amount invested by a participant in the fair value (FV) based statements group is 4108.33 € of the initial 10,000 € budget, whereas for historical cost (HC) based financial statements group participants' average amount is 4527.27 €. In panel B of Table 1, we show the results of a planned comparisons test according to H1 (HC > FV). Test statistic t = 1.036 is not significant at a 5% level (p-value = 0.152 (1T)). Contrary to previous research, we do not find a statistically significant difference between the budget spend by nonprofessional investors in firms with fair values reported versus firms with historical cost-based financial reports. Our results show that the FV group shows a decrease in average amount but an increase in the standard deviation of those amounts when compared with the HC group (2297.92 vs. 2014.80, respectively). These results might decrease the power of our test. An alternative explanation is that familiarity and expertise reduce the use of irrelevant information. Smith and Kida [28] find less evidence of anchoring as familiarity and expertise increase. As already mentioned, our participants are graduate students that already completed successfully at least one intermediate or advanced course in financial accounting and financial reporting, where they were exposed to measurement theory.

Table 1. Analysis of Investment Decisions.

Panel A—Descriptive Statistics			
Group	**N° Obsv.**	**Mean Investment Amount**	**Std Deviation**
HC	55	4527.27	2014.80
FV	60	4108.33	2297.92
Total	115	4308.70	2167.93
Panel B—Planned Comparisons Test Results			
Test	**df**	**t-statistics**	**Sig.**
HC > FV	113	1.036	0.152 (1T)

Participants made an investment decision based on the following instructions: "Assume you have 10,000 € to invest in this stock. Assume also that each stock is currently traded at 2 € per share, immediately after the disclosure of the attached financial statements. How much of your initial budget would you invest in the stocks of this firm?" Panel A presents Investment amounts. Panel B presents results of a planned comparisons test as defined by research hypothesis 1. Dependent variable is mean investment amount by group. HC = participants presented with financial statements produced under historical cost-based measurement. FV = participants presented with financial statements produced under fair value-based measurement.

H2 required participants to present earnings forecasts. In this research hypothesis, we predict that nonprofessional investors will assess fair value earnings to be permanent, contrary to economic theory. To test H2, participants predicted next year's operating earnings and comprehensive income. Similar to Warne [3], to avoid demand effects and conceal the objectives of this study, we did not ask directly for an estimation of next year's fair value changes. Instead, we infer that from decomposing comprehensive income into three components: operating earnings, non-operating earnings and fair value changes. Given that we only manipulate fair value changes, we can infer a forecast for those fair value changes by holding fix the other components.

Table 2 presents the descriptive statistics and test results for H2. Panel A presents average forecasted operating earnings (OP) for both groups and for the total sample. FV group forecasted a smaller amount of OP compared to the HC group (54,109.39 versus 59,140.56, respectively). Panel C shows planned comparisons test results for that forecast. A t-stat of 1.285 with a *p*-value of 0.101 (1T) does not allow us to infer (at a 5% level) that those two amounts are significantly different.

Table 2. Analysis of Earnings Forecast.

Panel A—Descriptive Statistics (Operating Earnings)			
Group	**N° Obsv.**	**Mean Earnings Forecast**	**Std Deviation**
HC	54	59,140.56	19,142.417
FV	59	54,109.39	22,175.885
Total	113	56,513.66	20,843.215
Panel B—Descriptive Statistics (Comprehensive Income)			
Group	**N° Obsv.**	**Mean Earnings Forecast**	**Std Deviation**
HC	53	40,092.00	11,334.393
FV	59	17,978.31	17,638.856
Total	112	28,442.82	18,594.534
Panel C—Planned Comparisons Test Results (Operating Earnings)			
Test	**df**	**t-statistics**	**Sig.**
HC > FV	113	1.285	0.101 (1T)
Panel D—Planned Comparisons Test Results (Comprehensive Income)			
Test	**df**	**t-statistics**	**Sig.**
HC > FV	113	7.971 (*)	0.000 [a] (1T)

Participants were asked to make two earnings forecast for the following year: operating earnings and bottom line comprehensive income. Forecasted Fair value changes is implicitly estimated by differences between the two groups' comprehensive income forecast, to avoid drawing attention to research topic. (Assuming benchmark group HC average estimated non-operating earnings of (19,048.56), then average fair value changes predicted by FV group are (17,082.52); assuming an average fixed percentage (32.21%) of operating earnings is predicted to be non-operating earnings by benchmark group HC, then average fair value changes predicted by FV group are (18,702.45)). Panels A and B present Operating Earnings and Comprehensive Income forecasted amounts, respectively. Panels C and D present results of a planned comparisons tests as defined by research hypothesis 2. Dependent variables are mean forecasted amounts by group. HC = participants presented with financial statements produced under historical cost-based measurement. FV = participants presented with financial statements produced under fair value-based measurement. (*) Variances unequal; [a] Statistically significant at 5% level.

Panel B reports descriptive statistics for the second forecasted earnings number—comprehensive income. FV group clearly forecasts a lower comprehensive income (17,978.31) than the HC group (40,092.00). Panel D shows that this difference is statistically significant. Planned comparison t-test presented a stat. of 7.794 with a significance of 0.000. Since we manipulated only fair value changes between groups, we find significant evidence that confirms H2. Nonprofessional investors assess fair value changes as permanent.

We compute fair value changes in two ways. First, we hold fix an amount of non-operating earnings, and since we required forecasts for OP and comprehensive income, we can infer predicted fair value changes. Using implicit forecasted non-operating earnings (−19,048.56) from the HC group, we obtain fair value changes estimated at −17,082.52. Alternatively, we hold fix a percentage of OP (67.29%) that nets out into comprehensive income in the HC group forecast. We estimate a forecasted fair value change of −18,702.45. Both values are statistically significant at a t-test with a zero hypothesized value. We can conclude that, as predicted by H2, nonprofessional investors view fair value changes as permanent and predictive of future cash flows.

We test, as a robustness check, alternative strategies to estimate forecasted fair value changes. We computed changes holding back a fixed amount and a percentage as non-operating earnings based on provided financial statements (which only differs between groups in fair value changes recognized). Results remained significant.

Taking together H1 and H2, we can conclude that the adoption of fair value measurement concepts affects nonprofessional investors' decisions in what concerns investment. Although we did not confirm a statistically significant difference in investment amounts (H1), we did find a forecasted cash flow (and estimated fundamental equity value) factor (H2). We cannot infer that nonprofessional investors are less willing to invest when firms report under fair value measurement concepts. However, we find a statistically significant belief from nonprofessional investors that fair value changes are permanent and can predict a firm's future income, thus affecting investments' cash flows. In conclusion, measurement concepts choice can induce increased volatility vis-à-vis forecasted cash flows.

H3 predicts that fair value changes will influence past performance judgements in the direction of its sign. Participants were asked to rate past performance based on the financial statements handed to them on a scale of 0 (very weak) to 100 (very strong). Table 3 presents the descriptive statistics and test results for H3. Panel B shows that planned comparisons test stat t of −0.566 is not significant at 5% (p-value = 0.287(1T)). We cannot conclude that fair value changes recognition affects nonprofessional investors' judgment of a firm's performance.

Table 3. Analysis of Past Performance.

Panel A—Descriptive Statistics			
Group	**N° Obsv.**	**Past Financial Performance**	**Std Deviation**
HC	54	51.76	18.69
FV	60	53.75	18.81
Total	114	52.81	18.70
Panel B—Planned Comparisons Test Results			
Test	**df**	**t-statistics**	**Sig.**
HC > FV	112	−0.566	0.287 (1T)

Participants were asked to rank past financial performance judgement on an 11-point scale with endpoints labeled 0 (very weak)—100 (very strong). Panel A presents average responses (scale 0–100). Panel B presents results of a planned comparisons test as defined by research hypothesis 3. HC = participants presented with financial statements produced under historical cost-based measurement. FV = participants presented with financial statements produced under fair value-based measurement.

In Table 4, we take further performance judgement analysis by testing H4. We predict that participants in group FV will be less confident about future performance predictions than the ones from the HC group. We asked participants to rate the confidence they have on their earnings forecast on an 11-point scale with endpoints labeled 0—(not at all confident) to 100—(very confident). Panel A shows very similar average confidence ratings for both groups with FV slightly below HC (39.83 vs. 42.59, respectively). Panel B presents the results of a planned comparisons test. The average rating does not significantly change regarding nonprofessional investors' perception of firm's performance for different measurement concepts (t-stat = 0.786 with a sig. of 0.22). We cannot confirm H4.

Taking together H3 and H4, we did not find evidence of any effect on investors' judgement of past and future performance derived from fair value changes recognition.

Finally, we asked participants to identify the reliability and relevance of values of four different classes of assets—Cash and Marketable Securities; Buildings; Operating Equipment; and Accounts Receivables. We manipulate measurement policies for two of them (Buildings—mark-to-market fair value vs. historical cost-, and Equipment—mark-to-model vs. historical cost) between groups. Hold the other two equal for both groups (Cash and Marketable Securities—mandatory mark-to-market fair values—, and Accounts Receivables—historical cost) in order to make the experience closer to real data and avoid demand effects. We provide participants with Portuguese standards (IFRS-based) official definition of relevance and reliability. We then request them to rate each class of assets on an 11-point scale from 0—(not at all) to 100—(very) both for reliability and relevance.

Table 4. Analysis of Confidence about Future Performance.

Panel A—Descriptive Statistics			
Group	**N° Obsv.**	**Mean Confidence Rate**	**Std Deviation**
HC	54	42.59	19.32
FV	59	39.83	18.05
Total	113	41.15	18.64
Panel B—Planned Comparisons Test Results			
Test	**df**	**t-statistics**	**Sig.**
HC > FV	111	0.786	0.220 (1T)

Participants were asked to rate their confidence in future financial performance judgements on an 11-point scale with endpoints labeled 0 (not confident)—100 (very confident). Panel A presents average responses (scale 0–100). Panel B presents results of a planned comparisons test as defined by research hypothesis 4. HC = participants presented with financial statements produced under historical cost-based measurement. FV = participants presented with financial statements produced under fair value-based measurement.

Table 5 presents average ratings and test statistics for relevance assessments of the different classes of assets. Panel A shows that both groups rate Accounts Receivable (Historical cost) as the most relevant whereas the other classes are ranked least relevant. We conducted a repeated-measures analysis of variance with group interactions. Panel B presents test results. We find that there is a significant difference between the relevance assessments of Receivables' historical cost versus the remaining classes of assets (Z-stat = 34.805 with sig. = 0.000). Puzzlingly, we do not find any group interaction (RELEV*GROUP Z-stat = 0.012 with sig. = 0.455 (1T)).

Table 5. Analysis of Relevance of Measurement Concepts.

Panel A—Descriptive Statistics (Variable = Average Relevance Rating (Std Dev))				
Group	**Cash**	**Buildings ***	**Equipment ***	**Accts. Receiv.**
HC	61.09 (22.02)	63.09 (20.15)	60.36 (20.07)	74.09 (18.81)
FV	60.17 (22.61)	64.50 (18.29)	61.67 (18.54)	73.67 (17.97)
Total	60.61 (22.23)	63.83 (19.13)	61.04 (19.21)	73.87 (18.29)
Panel B—Repeated Measures ANOVA Test Results (with group interactions)				
Variable	**df**	**Z-statistics**	**Sig.**	
RELVC	1	34.805	0.000 [a] (1T)	
RELVC*GROUP	1	0.012	0.455 (1T)	

We provide participants with Portuguese standards' official definition of Relevance (based on IFRS) and asked them to rate relevance of four different classes of assets: cash and marketable securities (mandatory fair value—mark-to-market); Buildings (manipulated between groups—historical cost vs. mark-to-market fair values); Production Equipment (manipulated between groups—historical cost vs. mark-to-model fair values); and Accounts Receivables (historical cost). Answers are presented in an 11-point scale with endpoints labeled 0 (not at all)—100 (very). Panel A presents average responses (scale 0–100). Panel B presents results of a repeated-measures ANOVA test with group interactions as defined by research hypothesis 5. HC = participants presented with financial statements produced under historical cost-based measurement. FV = participants presented with financial statements produced under fair value-based measurement. RELVC = average relevance rating. GROUP = dummy variable that takes value 1 if participant is from group FV. [(*)] manipulated in the experiment; [(a)] Statistically significant at 5% level.

Since we manipulate two classes of assets between groups, we find it intriguing that no effect is detected between groups for those classes of assets. One potential explanation links our results with those of Gassen and Schwedler [2]. They provide evidence of professional investors ranking measurement criteria differently across classes. They show that despite a general classification of fair value mark-to-market as most decision-useful and mark-to-model fair values as least useful, investors perceive assets measurement desirable attributes different across classes.

Finally, in Table 6, we show the descriptive analysis and test results for H6 regarding the reliability of different classes of assets. In Panel A, we can identify that participants, on average, rank as more reliable historical cost measured assets and fair value mark-to-market. They also rank, on average, as least reliable, mark-to-model fair value measured

assets. Panel B presents the repeated measures analysis of variance test results. We do not find significant effects both at the assets classes' level and on the interaction with the groups. A Z-stat of 0.270 with a sig. of 0.303 does not allow us to corroborate predicted results. Nonprofessional investors, proxied by graduate students in this experience, rate measurement concepts similar reliability. Furthermore, participants allocated to different groups rate similarly reliable those classes of assets that were manipulated to test the effect of fair value (multi-level) measurement concepts (Z-stat = 0.887 with sig. = 0.174 (1T)).

Table 6. Analysis of Reliability of Measurement Concepts.

Panel A—Descriptive Statistics (Variable = Average Reliability Rating (Std Dev))				
Group	**Cash**	**Buildings ***	**Equipment ***	**Accts. Receiv.**
HC	67.71 (22.24)	64.45 (20.25)	64.64 (20.16)	68.64 (19.30)
FV	64.00 (23.95)	61.58 (19.86)	61.33 (18.22)	60.67 (17.59)
Total	65.77 (23.12)	62.96 (20.01)	62.91 (19.16)	64.48 (18.78)
Panel B—Repeated Measures ANOVA Test Results (with group interactions)				
Variable	**df**	**Z-statistics**	**Sig.**	
RELIAB	1	0.270	0.303 (1T)	
RELIAB*GROUP	1	0.887	0.174 (1T)	

We provide participants with Portuguese standards' official definition of Reliability (based on IFRS) and asked them to rate reliability of four different classes of assets: cash and marketable securities (mandatory fair value—mark-to-market); Buildings (manipulated between groups- historical cost vs. mark-to-market fair values); Production Equipment (manipulated between groups- historical cost vs. mark-to-model fair values); and Accounts Receivables (historical cost). Answers are presented in an 11-point scale with endpoints labeled 0 (not at all)—100 (very). Panel A presents average responses (scale 0–100). Panel B presents results of a repeated-measures ANOVA test with group interactions as defined by research hypothesis 6. HC = participants presented with financial statements produced under historical cost-based measurement. FV = participants presented with financial statements produced under fair value-based measurement. RELIAB = average reliability rating. GROUP = dummy variable that takes value 1 if participant is from group FV. (*) manipulated in the experiment.

We conducted additional robustness checks of H5 and H6 together. In non-tabulated results, we find evidence that participants distinguish between reliability and relevance of pure mark-to-market fair value and historical cost from those of mark-to-model fair value. Consistent with previous research [2], participants rate, on average, similarly lower reliability and relevance to mark-to-model, while they clearly distinguish reliability and relevance of historical cost and mark-to-market fair values.

We also test for the robustness of our main results across different statistical tests. In non-tabulated results, we test a multinomial logit model prediction of willingness to invest across different measurements (FV vs. HC) and individuals' characteristics (gender, age and familiarity with measurement theory). Results remain similar.

5. Conclusions

We use a paper-and-pencil experience to analyze the effects of different measurement concepts on nonprofessional investors' investment decisions and judgements. In this experience, we proxy nonprofessional investors by graduate students from a Master of Science in accounting and finance and executive education program students of the same area who were already exposed to intermediate and advanced courses in financial reporting. Previous research argues that the latter group is a valid surrogate for the former [27].

We find evidence of interesting effects: firstly, nonprofessional investors' investment decision is affected vis-à-vis a cashflow estimation factor but not in their willingness to invest. In fact, investments' amounts of the total budget remain statistically unchanged when we manipulate the measurement of assets by using multi-level fair values where discretion is allowed. Contrarily, participants view fair value changes as permanent. Consequently, we argue that consistent with previous research on unrealized gains and losses (e.g., [15]), fair value changes recognition will induce volatility on future cashflows forecasted to evaluate the investment's fundamental value.

Secondly, we do not find evidence that past or future performance judgements are affected by our manipulation of measurement criteria. Contrary to prediction, nonprofessional investors remain unaffected by any fair value recognition when assessing past performance. They also do not feel less confident in predicting future earnings when presented with fair value-based financial statements. A potential explanation may be that, as argued in the extant literature, familiarity and expertise may mitigate the predicted effects [28].

Finally, consistent with previous research [2], we find that nonprofessional investors view measurement criteria differently in terms of reliability and relevance. We find that regardless of between-group differences, participants view historical cost as most relevant. Additionally, there seem to be different relevance assessments for different classes of assets, across the same measurement criteria. As previously argued [2], decision-usefulness is also influenced by the class of assets to be measured. Regarding reliability, our results do not show evidence that investors view reliability differences across measurement concepts. Nor are found effects in group assignment. But we do find evidence that participants distinguish reliability and relevance of mark-to-market and historical cost measurement, but not for mark-to-model measurement, which they rate consistently lower for both attributes.

Results are important for a broad group of individuals. Financial statements' preparers (and users) learn that several competing consequences underly their measurement concepts choices and that those discretionary choices bear additional unattended (and probably unwanted) results on valuation volatility and investors' judgements. Standard setters and regulators may find that our results present effects on judgement and decisions of nonprofessional investors that are statistically and economically relevant and, thus, should be balanced in their work. Finally, academics face additional layers of research that deem the debate about fair value measurement advantages yet not fully explored.

This research presents several limitations. First, we limited the amount of information participants received to a set of financial statements reporting last year's performance so that they could complete the experience in a reasonable amount of time. Regular activities demand investors evaluate the financial performance of a firm based on a more complex set of information. Nevertheless, as pointed out in Elliott [16], reducing the complexity of the information environment allows for stronger inferences about the factors that influence nonprofessional investors' judgement and decisions.

Secondly, although previous literature shows evidence of graduate students as good surrogates for nonprofessional investors in terms of performance on these types of tasks [27], it is likely that the demographics of the participants do not fully reflect those of nonprofessional investors. Most likely, investment experience differs and that might affect the accuracy in assessing the opinions and decisions of nonprofessional investors.

To conclude, the aforementioned limitations can present directions for future research avenues, since the debate about measurement criteria seems to be far from fully explored.

Funding: This research was funded by FCT—Fundação para a Ciência e Tecnologia (Portugal), grant number UID/04521/2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Interested readers may contact the corresponding author for acquiring relevant datasets generated in this research study, under reasonable request.

Acknowledgments: I am grateful to José Almeida for his assistantship.

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

References

1. Watts, R.L. Corporate financial statements, a product of the market and political processes. *Aust. J. Manag.* **1977**, *2*, 53–75. [CrossRef]
2. Gassen, J.; Schwedler, K. The decision usefulness of financial accounting measurement concepts: Evidence from an online survey of professional investors and their advisors. *Eur. Account. Rev.* **2010**, *19*, 495–509. [CrossRef]
3. Warne, R.C. The Effects of Non-Financial Asset Fair-Value Recognition on Nonprofessional Investors' Judgments. *SSRN Electron. J.* **2008**. [CrossRef]
4. Gonçalves, T.; Gaio, C.; Lélis, C. Accrual mispricing: Evidence from European sovereign debt crisis. *Res. Int. Bus. Financ.* **2020**, *52*, 101111. [CrossRef]
5. Gaio, C.; Gonçalves, T.; Pereira, A. Financial crisis and impairment recognition in non-financial assets. *Rev. Bus. Manag.* **2021**, *23*, 370–387. [CrossRef]
6. Penman, S.H. Financial reporting quality: Is fair value a plus or a minus? *Account. Bus. Res.* **2007**, *37*, 33–44. [CrossRef]
7. Keynes, J.M. *General Theory of Employment, Interest, and Money*; Palgrave Macmillan: London, UK, 1936.
8. Barth, M.E.; Landsman, W.R. Fundamental issues related to using fair value accounting for financial reporting. *Account. Horiz.* **1995**, *9*, 97.
9. Nobes, C. *Asset Measurement Bases in UK and IASC Standards*; Certified Accountants Educational Trust: London, UK, 2001.
10. Plantin, G.; Sapra, H.; Shin, H.S. Marking-to-market: Panacea or Pandora's box? *J. Account. Res.* **2008**, *46*, 435–460. [CrossRef]
11. Ryan, S.G. Accounting in and for the subprime crisis. *Account. Rev.* **2008**, *83*, 1605–1638. [CrossRef]
12. Barth, M.E. Fair value accounting: Evidence from investment securities and the market valuation of banks. *Account. Rev.* **1994**, *69*, 1–25.
13. Hirst, D.E.; Hopkins, P.E. Comprehensive income reporting and analysts' valuation judgments. *J. Account. Res.* **1998**, *36*, 47–75. [CrossRef]
14. Maines, L.A.; McDaniel, L.S. Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. *Account. Rev.* **2000**, *75*, 179–207. [CrossRef]
15. Bloomfield, R.J.; Nelson, M.W.; Smith, S.D. Feedback loops, fair value accounting and correlated investments. *Rev. Account. Stud.* **2006**, *11*, 337–416. [CrossRef]
16. Elliott, W.B. Are investors influenced by pro forma emphasis and reconciliations in earnings announcements? *Account. Rev.* **2006**, *81*, 113–133. [CrossRef]
17. Frederickson, J.R.; Miller, J.S. The effects of pro forma earnings disclosures on analysts' and nonprofessional investors' equity valuation judgments. *Account. Rev.* **2004**, *79*, 667–686. [CrossRef]
18. Tversky, A.; Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science* **1974**, *185*, 1124–1131. [CrossRef]
19. Barth, M.E.; Landsman, W.R.; Wahlen, J.M. Fair value accounting: Effects on banks' earnings volatility, regulatory capital, and value of contractual cash flows. *J. Bank. Financ.* **1995**, *19*, 577–605. [CrossRef]
20. Hong, G.; Sarkar, S. Equity systematic risk (Beta) and its determinants. *Contemp. Account. Res.* **2007**, *24*, 423–466. [CrossRef]
21. Ou, J.A.; Penman, S.H. Accounting measurement, price-earnings ratio, and the information content of security prices. *J. Account. Res.* **1989**, *27*, 111–144. [CrossRef]
22. Weiner, B. Attributional thoughts about consumer behavior. *J. Consum. Res.* **2000**, *27*, 382–387. [CrossRef]
23. Chung, J.; Monroe, G.S. The effects of experience and task difficulty on accuracy and confidence assessments of auditors. *Account. Financ.* **2000**, *40*, 135–151. [CrossRef]
24. Thomas, J. Discussion: Financial Instrument Fair Values and Nonfinancial Firms. *J. Account. Audit. Financ.* **1999**, *14*, 275–278. [CrossRef]
25. Hodge, F.D. Hyperlinking unaudited information to audited financial statements: Effects on investor judgments. *Account. Rev.* **2001**, *76*, 675–691. [CrossRef]
26. Hirst, D.E.; Koonce, L.; Miller, J. The joint effect of management's prior forecast accuracy and the form of its financial forecasts on investor judgment. *J. Account. Res.* **1999**, *37*, 101–124. [CrossRef]
27. Elliott, W.B.; Hodge, F.D.; Kennedy, J.J.; Pronk, M. Are MBA students a good proxy for nonprofessional investors? *Account. Rev.* **2007**, *82*, 139–168. [CrossRef]
28. Smith, J.F.; Kida, T. Heuristics and biases: Expertise and task realism in auditing. *Psychol. Bull.* **1991**, *109*, 472–489. [CrossRef]