MDPI

Article

Søren Johansen and Katarina Juselius: A Bibliometric Analysis of Citations through Multivariate Bass Models

Fragiskos Archontakis [1,*] **and Rocco Mosconi** [2]

[1] School of Economics, Business Administration & Legal Studies, International Hellenic University, 57001 Thermi, Greece
[2] Dipartimento di Ingegneria Gestionale Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milano, Italy; rocco.mosconi@polimi.it
[*] Correspondence: f.archontakis@ihu.edu.gr

Abstract: We showcase the impact of Katarina Juselius and Søren Johansen's contribution to econometrics using bibliometric data on citations from 1989 to 2017, extracted from the Web of Science (WoS) database. Our purpose is to analyze the impact of KJ and SJ's ideas on applied and methodological research in econometrics. To this aim, starting from WoS data, we derived two composite indices whose purpose is to disentangle the authors' impact on applied research from their impact on methodological research. As of 2017, the number of applied citing papers per quarter had not yet reached the peak; conversely, the peak in the methodological literature seem to have been reached around 2000, although the shape of the trajectory is very flat after the peak. We analyzed the data using a multivariate dynamic version of the well known Bass model. Our estimates suggest that the methodological literature is mainly driven by "innovators", whereas "imitators" are relatively more important in the applied literature: this might explain the different location of the peaks. We also find that, in the literature referring to KJ and SJ, the "cross-fertilization" between methodological and applied research is statistically significant and bi-directional.

Keywords: bass diffusion model; bibliometrics; cointegration

1. Introduction

Using bibliometric methods in order to value the quantity and quality of knowledge produced by researchers is increasingly the standard practice in most disciplines (Garfield et al. 1978; Redner 1998). In the field of economics, Kalaitzidakis et al. (1999) provided a ranking of European departments based on ten top journals, which was later updated and expanded to include, amongst others, also a ranking of academic journals in economics (Kalaitzidakis et al. 2003). At the same time, Coupé (2003) published a paper including rankings for researchers based on publications and citations; there, he explicitly mentions the highly-cited work by Søren Johansen and Katarina Juselius on cointegration, stating that "first in the citation ranking is Søren Johansen. Thanks to his top cited papers on cointegration written at the beginning of the 1990s, he is first on the three different citation rankings"(Coupé 2003, p. 1336).

The aim of this paper is to showcase, through a bibliometric analysis, the impact of Katarina Juselius (KJ) and Søren Johansen's (SJ) contribution to the field of econometrics. An important distinctive trait of their scientific production is to combine methodological and applied research, placing their work in the so-called "Pasteur's Quadrant" (Stokes 1997), characterized by use-inspired basic research, where applied objectives are chased in parallel with fundamental scientific creativity. This motivates our main research question: what is the influence of KJ and SJ's work on applied and methodological research in econometrics? We believe that, from this analysis, we can learn something about the mechanisms of scientific discovery in general. Although the methodology used in this paper is different, our analysis has some resemblance to the work by Stigler (1994), who analyzed citation data

in the journals of statistics and probability, investigating the mechanisms of knowledge diffusion within and across fields. Among other findings, he observed that "there is a tendency for influence to flow from theory to applications to a much greater extent than in the reverse direction" (see Stigler 1994, p. 94). As we will show in this paper, this is, to some extent, confirmed also in the abundant literature inspired by KJ and SJ, although the flow running from applied econometrics toward econometric methodology is also clearly visible in this case. We think that this depends on the peculiar approach to empirical research inspired by KJ and SJ work, "in which data would be allowed to speak freely without being silenced by prior restriction and in which basic hypotheses could be adequately tested and empirically relevant structures estimated"—the quote is taken from Juselius (2021, p. 6), in the same Special Issue of *Econometrics* hosting this paper. This approach requires a continuous dialogue with methodologists, posing to them challenging requests for appropriate statistical models and suitable probability results allowing for correct inference within such models.

Our empirical investigation is based on citation data collected through the Web of Science (WoS) database, based on which we derived two new composite indices whose purpose is to disentangle the citations originated in the applied econometric research from those coming from the methodological research. Our analysis reveals that the majority of citations (about 85%) arise from applied research. Of course, to put this figure into perspective, one should compare it with the share of methodological research in econometrics in general: unfortunately, we do not have this information (our impression is that the share is somewhat lower than 15%). Interestingly, the dynamic pattern of the two indices is quite different: the citation peak in the applied literature does not seem to be reached yet, whereas the peak in the methodological literature seems to have occurred around the turn of the century. To analyze these bibliometric data, we resorted to a multivariate dynamic version of the well known Bass (1969) model, proposed by Boswijk et al. (2009) building on Franses (2003), Boswijk and Franses (2005) and Fok and Franses (2007). Bibliometric evidence suggests that Bass-type models provide a useful way to fit most Nobel in Economics prize winner citation trajectories; see Bjork et al. (2014). This fact might indicate that, up to a point, economic knowledge could follow the well-known product life cycle, which is usually characterized by the following phases: introduction, growth, maturity (including peak) and decline, within the context of a scholar's professional lifetime and beyond. An interesting aspect of Bass models is that they describe the diffusion pattern as dependent on two key parameters, p and q, measuring the relevance of innovation and imitation, respectively: these two parameters are shown in Min et al. (2018) to have an important role in the growth and decay of citation counts in several scientific disciplines. In this paper, we will show that, according to our estimates, the relative importance of imitative and innovative mechanisms is quite different for methodological and applied econometric research: this seems to be responsible for the different trajectories of the two research strands.

The paper is organized as follows: Section 2 describes the data collection and management process to support the analysis. Section 3 presents the univariate and multivariate Bass model, and Section 4 illustrates our empirical findings. Finally, Section 5 concludes and provides directions for further research.

A word on notation used in the paper. The backshift operator L is defined as $LX_t = X_{t-1}$, where X_t is a time series; the difference operator is defined as $\Delta = 1 - L$, so that $\Delta X_t = X_t - X_{t-1}$. I_n is the $n \times n$ identity matrix, $u_{n,i}$ is the i-th column of I_n, $\mathbf{1}_n = \sum_{i=1}^{n} u_{n,i}$, $\mathbf{0}_{m,n}$ is an $m \times n$ matrix of zeros, $diag\{a_i\}$ is the block diagonal matrix whose generic diagonal block is the matrix a_i (of course any of the a_i's could also be a scalar).

2. The Data

This section describes the line of thought and the data collection process, including the source and sample size, while providing some preliminary analysis through stylized facts.

To the purpose of this study, we consider the scientific production on cointegration by KJ and SJ as an indissoluble whole, where economic questions motivate the development of econometric theory and the development of econometric theory sharpens the economic questions. Their papers on cointegration are, therefore, analyzed together, whether single authored or coauthored and whether the main focus is on methodology (with just an illustrative example) or application (with a pedagogical effort to illustrate how the methodology can be applied to a real problem).

On 9 April 2018, we collected from the Web of Science (WoS) the data about the citations received by KJ and SJ for papers between 1989:Q1 and 2017:Q3.[1] For practical reasons, we limit the analysis to the 10 most quoted papers, which are presented in Table 1, sorted by publication date.

Table 1. The 10 "Top Cited" papers by S. Johansen and K. Juselius, in chronological order (data collected on 9 April 2018).

Order (Time)	Paper	Citations (WoS)	New Citations (WoS)
1	Johansen (1988)	4008	4008
2	Johansen and Juselius (1990)	2567	1060
3	Johansen (1991)	2256	997
4	Johansen (1992c)	170	45
5	Johansen and Juselius (1992)	477	90
6	Johansen (1992b)	249	40
7	Johansen (1992a)	251	69
8	Johansen and Juselius (1994)	196	32
9	Johansen et al. (2000)	167	60
10	Hendry and Juselius (2001)	112	56
		10,453	6457

The total number of citations received by the top ten papers amounted at that time to 10,453, whereas the number of citing papers was 6457,[2] so that every citing paper cites, on average, 1.62 papers, with a maximum of 7 observed four times.[3] In terms of the number of citations, well ahead of the rest of the publications, are the papers by Johansen (1988) in the *Journal of Economic Dynamics & Control* with 4008 citations, the joint paper by Johansen and Juselius (1990) in the *Oxford Bulletin of Economics & Statistics* with 2567 citations, and the paper by Johansen (1991) in *Econometrica* with 2256 citations. For each paper, the last column in Table 1, "new citations", indicates the number of citing papers referring to that paper and to none of the earlier ones: for example, paper number 7 is cited by 251 papers, but only 69 of them cite paper number 7 and none of the earlier. A high "new citations"/citations ratio suggests that the paper has broken new ground in the field: for example, the paper by Hendry and Juselius (2001) treats data from the field of energy, and as a result, energy-related papers often cite Hendry and Juselius (2001), rather than the earlier papers. Notice that the first three papers account for 84.5% of the citations and 93.9% of the citing papers.

To avoid double counting, we focus on the number of citing papers rather than on the number of citations,[4] and we define by c_t the number of citing papers published in quarter t (t ranges from 1989:Q1 to 2017:Q3, i.e., 115 quarters). Based on the WoS data, we split c_t into two composite indices aimed at measuring the impact of KJ and SJ ideas on applied and methodological econometric research, respectively.[5] To this aim, we have analyzed each of the 6457 citing papers, classifying them according to their methodological or applied nature. The classification is essentially based on the title of the citing paper.[6] We adopted the following classification:

- Purely applied (PA) papers: the title refers to an application, with no reference to an econometric method, technique or issue. We have found $n_{PA} = 4198$ such papers.

- Mainly applied (MA) papers: the title refers both to an econometric method, technique or issue and an application, and the focus seems to be on the latter (e.g., "Does exchange-rate volatility affect import flows in G-7 countries? Evidence from cointegration models"). We have found $n_{MA} = 1451$ such papers.
- Purely methodological (PM) papers: the title refers to an econometric method, technique or issue, with no reference at all to an application. We have found $n_{PM} = 716$ such papers.
- Mainly methodological (MM) papers: the title refers both to an econometric method, technique or issue and an application, and the focus seems on the first (e.g., "Robust cointegration testing in the presence of weak trends, with an application to the human origin of global warming"). We have found $n_{MM} = 92$ such papers.

We have, therefore, derived four quarterly time series, labeled $c_{PA,t}$, $c_{MA,t}$, $c_{PM,t}$ and $c_{MM,t}$, counting the citing papers of each group in each quarter; of course, $c_t = c_{PA,t} + c_{MA,t} + c_{PM,t} + c_{MM,t}$. The four time series are reported in Figure 1, where one can observe that the behavior of $c_{PA,t}$ and $c_{MA,t}$ is quite similar, steadily increasing over time, with some low frequency fluctuations, which seem to be shared by both series (the correlation is 69.5%). Conversely, $c_{PM,t}$ has a peak around the year 2000 with about 10 papers per quarter, and then it declines until 2005, seeming to stabilize at around 5 papers per quarter. The series $c_{MM,t}$ is irregular, due to the small number of MM papers, but resembles $c_{PM,t}$ to some extent, as it shows a higher frequency around the year 2000; then, the frequency seems to slightly decline.
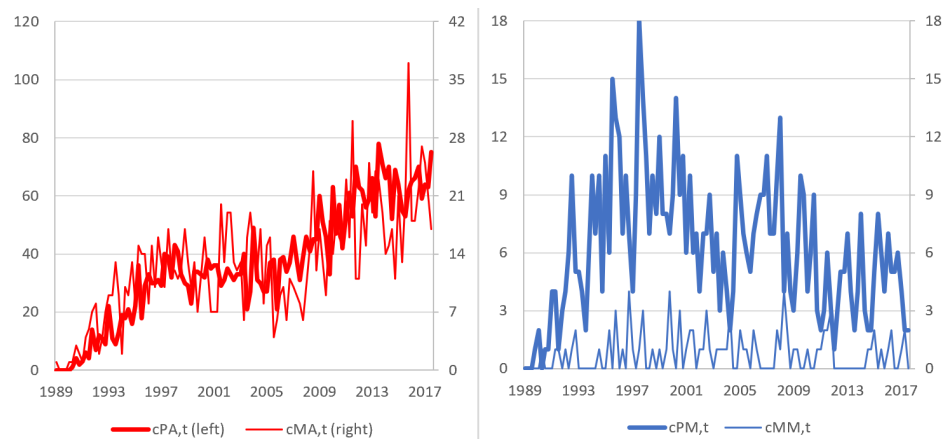

Figure 1. Time series plot of $c_{PA,t}$, $c_{MA,t}$, $c_{PM,t}$ and $c_{MM,t}$, quarterly data from 1989:1 to 2017:3.

Finally, by combining the four series with suitable weights, we obtained two composite indicators, whose purpose is to measure the impact of KJ and SJ ideas on applied ($c_{1,t}$) and methodological ($c_{2,t}$) research:

$$c_{1,t} = c_{PA,t} + \omega c_{MA,t} + (1 - \omega) c_{MM,t} \tag{1}$$

$$c_{2,t} = c_{PM,t} + \omega c_{MM,t} + (1 - \omega) c_{MA,t} \tag{2}$$

Of course $c_{1,t} + c_{2,t} = c_t$ for any ω by construction. Composite indicators have several pros and cons, as illustrated for example in Nardo et al. (2008) and Kuc-Czarnecka et al. (2020): they allow to summarize complex, multi-dimensional realities, reducing the dimensionality. On the other hand, they might simplify too much, and, even more importantly, the selection of indicators and weights could be the subject of dispute. It is, therefore, important to motivate clearly one's weighting choice and to provide an extensive sensitivity analysis. We provide a thorough discussion of both aspects in Appendix B. In short, the baseline results presented in this paper are based on $\omega = 0.85$. This choice is motivated by two main reasons: (i) ω should be in the range from 0.5 to 1, extremes excluded, since the papers classified as MA (or MM) should contribute mainly ($\omega > 1/2$)

to the applied (or methodological) index but also, to a lesser extent $(1 - \omega > 0)$, to the methodological (or applied) index; (ii) $\omega = 0.85$ would be approximately equal to $n_{PA}/(n_{PA} + n_{PM}) = 0.8545$—in practice, this corresponds to the assumption that the share of "applied research" of an MA paper is similar, on average, to the share of applied research in the econometric literature referring to KJ and SJ papers in general. Notice, however, that, as illustrated in Appendix B, the main results of our econometric analysis are robust to the choice of ω in the range from 0.5 to 1.

In order to fix ideas, we provide a short example based on the first few citations in the WoS data. For the year 1989, we have to split the only two existing citations by Baillie and Bollerslev (classified as MA) and Gilbert (classified as PM).[7] Thus, for this example we have obtained the series illustrated in Table 2.

Table 2. Illustration of the classification scheme: citations in 1989.

Author(s)	Quarter	c_t	$c_{PA,t}$	$c_{MA,t}$	$c_{PM,t}$	$c_{MM,t}$	$c_{1,t}$	$c_{2,t}$
Baillie-Bollerslev	1989:Q1	1	0	1	0	0	0.85	0.15
	1989:Q2	0	0	0	0	0	0	0
	1989:Q3	0	0	0	0	0	0	0
Gilbert	1989:Q4	1	0	0	1	0	0	1

The cumulative applied index at time $T = 115$, i.e., 2017:Q3, is equal to the following:

$$
\begin{aligned}
C_{1,T} &= \sum_{t=1}^{T}(c_{PA,t} + \omega c_{MA,t} + (1-\omega)c_{MM,t}) \\
&= 4198 + 0.85 \times 1451 + 0.15 \times 92 = 5445.2,
\end{aligned}
$$

while the cumulative methodological index will be equal to the following:

$$
\begin{aligned}
C_{2,T} &= \sum_{t=1}^{T}(c_{PM,t} + \omega c_{MM,t} + (1-\omega)c_{MA,t}) \\
&= 716 + 0.85 \times 92 + 0.15 \times 1451 = 1011.8.
\end{aligned}
$$

This shows that the majority of citations originates from applied research: defining $C_t = C_{1,t} + C_{2,t}$, the ratio $\frac{C_{1,T}}{C_T}$ is 84.4%, whereas $\frac{C_{2,T}}{C_T}$ is 15.6%. To check the appropriateness of our classification scheme, we analyzed how these ratios vary by publishing journal. Tracking down the 6457 citing papers, we obtained from the WoS database that they appeared in 696 distinct journals. Table 3 provides the ranked list of the top 20 journals by the number of citing papers: these journals hosted 2676 citing papers, i.e., 41.4%.

The evidence in Table 3 seems to confirm the validity of our classification: the average $c_{1,t}$ for the papers that appeared in mainly applied journals (for example, *Energy Policy*, *Journal of International Money and Finance*, *Journal of Policy Modelling*) is above 90%. To the other extreme, the average $c_{2,t}$ is above 90% for the *Journal of Econometrics* and for *Econometric Theory* (but also for *Econometrica*, which hosted 24 citing papers). Other journals, such as *Oxford Bulletin of Economics and Statistics*, *Journal of Applied Econometrics*, *Journal of Forecasting* are more balanced, with an average $c_{2,t}$ around 50%. We believe that the evidence in Table 3 supports the idea that classifying based on the title and the abstract is more accurate than classifying based on the publishing journal.

The time series $c_{1,t}$ and $c_{2,t}$ are illustrated in Figure 2. The plot shows some evidence of a "second wind" especially in the applied index $c_{1,t}$ but to some extent also in the methodological index $c_{2,t}$: both series seem to have a peak around 1998, after which they start decreasing very slowly, but around 2004 the citations start increasing again, especially for the applied research index, whereas the references found in methodological papers remain rather steady. A possibility/conjecture is that the second wind was triggered by the 2003 Nobel Prize in Economics, which popularized the concept of cointegration in a wider

variety of scientific disciplines. The trajectory of $c_{1,t}$ resembles the cases of Friedrich Hayek, referred to in Bjork et al. (2014) as "bi-modal", whereas the trajectory of $c_{2,t}$ resembles more closely the cases of Kenneth Arrow and Milton Friedman, called "staying power" in Bjork et al. (2014). Boswijk et al. (2010) also claim the same with different wording: they report evidence of a "second life" for the famous Engle and Granger (1987) paper in *Econometrica* after the authors were awarded the Nobel prize in 2003, an event which is likely to have revamped the interest in the work of KJ and SJ as well.

Table 3. The top 20 journals supplying citations to S. Johansen and K. Juselius' works.

Rank	Journal	C_T	$\frac{C_{1,T}}{C_T}$	$\frac{C_{2,T}}{C_T}$
1	APPLIED ECONOMICS	539	92.0%	8.0%
2	APPLIED ECONOMICS LETTERS	254	89.2%	10.8%
3	ENERGY ECONOMICS	208	94.9%	5.1%
4	ECONOMIC MODELLING	204	92.7%	7.3%
5	ENERGY POLICY	172	95.1%	4.9%
6	JOURNAL OF ECONOMETRICS	154	2.9%	97.1%
7	J. OF INTERNATIONAL MONEY & FINANCE	135	95.7%	4.3%
8	JOURNAL OF POLICY MODELING	121	96.4%	3.7%
9	ECONOMICS LETTERS	104	62.0%	38.0%
10	JOURNAL OF MACROECONOMICS	96	92.0%	8.0%
11	OXFORD BULLETIN OF ECON. & STAT.	92	46.5%	53.5%
12	ECONOMETRIC THEORY	86	0.2%	99.8%
13	EMPIRICAL ECONOMICS	81	88.6%	11.4%
14	JOURNAL OF FUTURES MARKETS	70	96.1%	3.9%
15	JOURNAL OF APPLIED ECONOMETRICS	68	54.7%	45.3%
16	MANCHESTER SCHOOL	66	94.6%	5.4%
17	ENERGY	61	92.5%	7.5%
18	JOURNAL OF BANKING & FINANCE	58	94.8%	5.2%
19	JOURNAL OF BUSINESS & ECON. STAT.	55	38.4%	61.6%
20	JOURNAL OF FORECASTING	52	51.4%	48.6%
	All Journals	6457	84.4%	15.6%
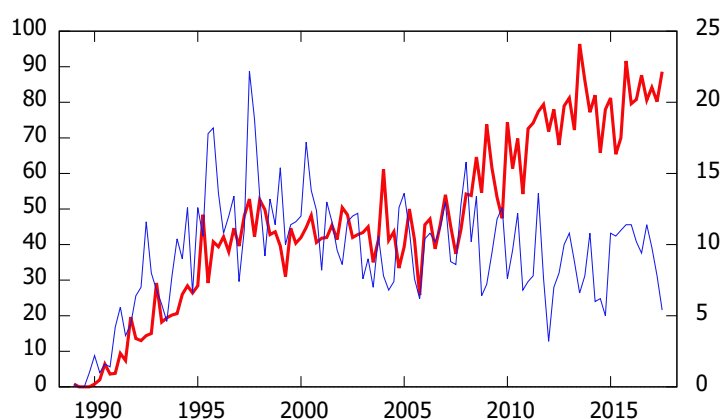
Figure 2. The composite citation indices $c_{1,t}$ (thick red, left scale) and $c_{2,t}$ (thin blue, right scale), quarterly data from 1989:1 to 2017:3.

3. The Bass Diffusion Model

The Bass diffusion model (Bass 1969) is widely used in many fields. Originally developed for marketing applications, the model has since been adopted also in other fields, such as the analysis of the diffusion of technological innovation (see Guseo and Guidolin 2008), bibliometric analysis (see Bjork et al. 2014) and epidemiology (see Eryarsoy et al. 2021).

The continuous time Bass model assumes a population of m potential adopters. Let us define by $t > 0$ the time of adoption of a randomly picked potential adopter: t is therefore a random variable. Define by $f(t)$ the corresponding density, and by $F(t) = \int_0^t f(u)du$ the cumulative density function, i.e., the probability that adoption occurs before t. Notice that the expected number of adopters at time t is given by the following:

$$\bar{C}(t) = mF(t) \tag{3}$$

and the corresponding "adoption intensity" is given by the following:

$$\bar{c}(t) = mf(t). \tag{4}$$

Bass assumes that the hazard rate $\frac{f(t)}{1-F(t)}$ is a linear function of the expected number of previous adopters:

$$\frac{f(t)}{1-F(t)} = p + qF(t), \tag{5}$$

where q is defined as the "imitation parameter" (or internal influence, or word-of-mouth effect) since it represents the idea that some potential adopters (imitators) tend not to adopt initially, but are more likely to adopt when the innovation is widespread. Conversely, p is defined as the "innovation parameter" (or external influence or advertising effect) since it represents the idea that some potential adopters (innovators) decide to adopt the innovation regardless of the level of diffusion. It is interesting to observe that when $q = 0$, Equation (5) implies a constant hazard, and therefore the Bass model collapses into the exponential distribution. In other words, in the absence of imitators, the adoption peak, as in the exponential distribution, would occur at the beginning of the process.[8]

Using (3) and (4), the differential Equation (5) can be rewritten as follows:

$$\bar{c}(t) = mp + (q - p)\bar{C}(t) - \frac{q}{m}\bar{C}(t)^2. \tag{6}$$

The solution to (6) with $\bar{C}(0) = 0$ is the following:

$$\bar{C}(t) = m\frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}, \tag{7}$$

so that

$$\bar{c}(t) = \frac{\partial \bar{C}(t)}{\partial t} = m\frac{p(p+q)^2 e^{-(p+q)t}}{\left(p + qe^{-(p+q)t}\right)^2}. \tag{8}$$

Starting from the latter equation, one can easily find the timing of the adoption peak t^P (i.e., the inflection point of the diffusion curve), the corresponding peak $\bar{c}^P$, and the level of adoption at the peak $\bar{C}^P$:

$$t^P = \frac{1}{p+q}\ln\left(\frac{q}{p}\right), \tag{9}$$

$$\bar{c}^P = \bar{c}\left(t^P\right) = m\frac{(p+q)^2}{4q}, \tag{10}$$

$$\bar{C}^P = \bar{C}\left(t^P\right) = m\frac{q-p}{2q}. \tag{11}$$

Formula (9) shows that the location of the peak depends on the innovation parameter p and the imitation parameter q through the sum $(p+q)$ and the ratio $\frac{q}{p}$: as clear in Formula (5), when either innovators or imitators or both are very active so that the sum

$(p + q)$ is large, then the hazard is large, which leads to a rapid exhaustion of the population at risk and therefore to an early peak.

3.1. Bass Discrete Time Model

A number of estimation procedures have been proposed to estimate the parameters m, p and q (see for example Satoh 2001). Bass (1969) suggested a simple estimation strategy based on Ordinary Least Squares (OLS) applied to a discretized version of (6) where essentially the expected adoption stock $\bar{C}(t)$ and the expected adoption flow $\bar{c}(t)$ are replaced by the observed counterpart C_t and $c_t = C_t - C_{t-1}$, and an error term is added. This leads to the following:

$$c_t = mp + (q - p)C_{t-1} - \frac{q}{m}C_{t-1}^2 + u_t. \tag{12}$$

In the standard discrete time Bass model, u_t is assumed to be $iid N(0, \sigma^2)$, so that OLS is the natural candidate for estimation. To apply OLS, (12) is then reparameterized as follows:

$$c_t = \beta_0 + \beta_1 C_{t-1} + \beta_2 C_{t-1}^2 + u_t. \tag{13}$$

The "reduced form" parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$ are related to the "structural form" parameters $\boldsymbol{\theta} = (m, p, q)'$ by the following:

$$
\begin{aligned}
\beta_0 &= mp, \\
\beta_1 &= q - p, \\
\beta_2 &= -\frac{q}{m}.
\end{aligned} \tag{14}
$$

and these relations can be inverted:[9]

$$
\begin{aligned}
m &= \frac{-\beta_1 - \sqrt{\beta_1^2 - 4\beta_0\beta_2}}{2\beta_2}, \\
p &= \frac{\beta_0}{m} = \frac{-\beta_1 + \sqrt{\beta_1^2 - 4\beta_0\beta_2}}{2}, \\
q &= -m\beta_2 = \frac{\beta_1 + \sqrt{\beta_1^2 - 4\beta_0\beta_2}}{2}.
\end{aligned} \tag{15}
$$

Assuming that u_t is uncorrelated, homoskedastic and normal, ML estimates of the parameters vector, say $\hat{\boldsymbol{\beta}}$, can be obtained by OLS, and the corresponding variance–covariance matrix $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}$ can be obtained as usual.[10] Replacing $\hat{\boldsymbol{\beta}}$ in (15) instead of $\boldsymbol{\beta}$ gives $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}(\hat{\boldsymbol{\beta}})$. Defining by

$$J_{\theta.\beta} = \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\beta}'}$$

and using the delta method, the variance–covariance matrix associated to $\hat{\boldsymbol{\theta}}$ is given by the following:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\theta}}} = \hat{J}_{\theta.\beta}\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}}\hat{J}'_{\theta.\beta}, \tag{16}$$

where $\hat{J}_{\theta.\beta}$ is the estimated counterpart of $J_{\theta.\beta}$. Tedious computation shows the following:

$$J_{\theta.\beta} = \frac{\partial \theta}{\partial \beta'} = \frac{1}{\delta_\beta} \begin{bmatrix} 1 & -\frac{\beta_1 + \delta_\beta}{2\beta_2} & \frac{\beta_1(\beta_1 + \delta_\beta) - 2\beta_0\beta_2}{2\beta_2^2} \\ -\beta_2 & \frac{\beta_1 - \delta_\beta}{2} & -\beta_0 \\ -\beta_2 & \frac{\beta_1 + \delta_\beta}{2} & -\beta_0 \end{bmatrix}, \tag{17}$$

where $\delta_\beta = \sqrt{\beta_1^2 - 4\beta_0\beta_2}$. $\hat{J}_{\theta.\beta}$ is therefore obtained by replacing $\hat{\beta}$ in (17).

We remark that when one considers n "seemingly unrelated" equations such as (13), i.e.,

$$c_{i,t} = \beta_{0,i} + \beta_{1,i} C_{i,t-1} + \beta_{2,i} C_{i,t-1}^2 + u_{i,t}, \qquad i = 1, ..., n, \tag{18}$$

and the variance–covariance matrix of $u_t = (u_{1,t}, ..., u_{n,t})'$, say Ω_u, is not diagonal, then equation by equation OLS is no longer equivalent to ML. In this case, the likelihood can be maximized by iterated Seemingly Unrelated Regression Equations (SURE), obtaining $\hat{\beta}_i = (\hat{\beta}_{0,i}, \hat{\beta}_{1,i}, \hat{\beta}_{2,i})', i = 1, ..., n, \hat{\Omega}_u$, and the variance–covariance matrix of $\hat{\beta} = \left(\hat{\beta}_1', ..., \hat{\beta}_n'\right)'$, i.e.,

$$\hat{\Sigma}_{\hat{\beta}} = \begin{bmatrix} \hat{\Sigma}_{\hat{\beta}_1} & \cdots & \hat{\Sigma}'_{\hat{\beta}_n, \hat{\beta}_1} \\ \vdots & \ddots & \vdots \\ \hat{\Sigma}_{\hat{\beta}_n, \hat{\beta}_1} & \cdots & \hat{\Sigma}_{\hat{\beta}_n} \end{bmatrix}.$$

Then, applying (15) and (16) to each pair $(\hat{\beta}_i, \hat{\Sigma}_{\hat{\beta}_i})$ it is easy to obtain the ML estimates of the structural parameters $\theta_i = (m_i, p_i, q_i)'$ and the associated variance–covariance matrices.

3.2. Boswijk and Franses Model

Boswijk and Franses (2005), henceforth BF, emphasize two major problems in the model (13):

- The assumption that u_t is uncorrelated is at odds with the empirical evidence that deviations of the observed adoption path with respect to the ideal equilibrium path are persistent.
- The assumption that u_t is homoskedastic is disputable since, at the beginning and at the end of the diffusion process, when c_t is expected to be close to zero, the variance of c_t likely to be much smaller than around the peak; related to this, simulating (13) with an homoskedastic and Gaussian error is likely to produce negative values of c_t in the initial and final phases of the diffusion.

To deal with the first problem, they propose the following alternative model:[11]

$$\Delta c_t = \alpha\left(c_{t-1} - mp - (q - p)C_{t-1} + \frac{q}{m}C_{t-1}^2\right) + u_t. \tag{19}$$

To understand the relationship between (12) and (19) it is interesting to observe that, adding and subtracting $\alpha\left(-(q - p)C_{t-2} + \frac{q}{m}C_{t-2}^2\right)$ to the right hand side of (19), and rearranging, one obtains the following:

$$\Delta c_t = \alpha\left(c_{t-1} - mp - (q - p)C_{t-2} + \frac{q}{m}C_{t-2}^2\right) - \alpha\left((q - p)\Delta C_{t-1} - \frac{q}{m}\Delta C_{t-1}^2\right) + u_t. \tag{20}$$

To interpret (20), define the following:

$$c_t^* = mp + (q - p)C_{t-1} - \frac{q}{m}C_{t-1}^2,$$

and notice that c_t^* is the expected value or c_t according to the Bass discrete time model (12). Using the notation c_t^*, (20) can be rewritten as the following:

$$\Delta c_t = \alpha\left(c_{t-1} - c_{t-1}^*\right) - \alpha\Delta c_t^* + u_t. \tag{21}$$

The parameter α is expected to be negative. The first term in (21), i.e., $\alpha\left(c_{t-1} - c_{t-1}^*\right)$, can be thought of as an Error-Correction Mechanism: if $c_{t-1} = c_{t-1}^*$, the ECM is ineffective; if instead $c_{t-1} > c_{t-1}^*$, then the ECM term partly corrects the disequilibrium by reducing c_t with respect to c_{t-1}; conversely, if $c_{t-1} < c_{t-1}^*$, then, through the negative α, c_t will increase with respect to c_{t-1}. The second term in (21), i.e., $-\alpha\Delta c_t^*$, can be thought of as a "Target Seeking" Mechanism, which induces dynamics in c_t, even if $c_{t-1} = c_{t-1}^*$ and $u_t = 0$: in fact Δc_t^* will be zero when ΔC_{t-1} (and therefore ΔC_{t-1}^2) is zero, which happens when the target level m is reached and therefore $C_{t-1} = C_{t-2} = m$. Another viewpoint on the BF model, seen as an AR(2) model for C_t with state dependent parameters is given in Appendix A.

It is important to remark that the standard Bass model (12) is a special case of (19) with $\alpha = -1$, so that one can set up a test $H_0 : \alpha = -1$ to decide which model is preferable. The interpretation of the parameters m, p and q is exactly the same in both models since (19) is a generalized version of the original Bass model, where the "adjustment intensity", instead of being fixed at -1, is represented by the unrestricted parameter α. For example, when $\alpha = -0.5$, only half of the disequilibrium observed at the end of a time unit is adjusted within the subsequent time unit: this gives rise to some persistence in the disequilibrium.

To deal with the second problem (heteroskedasticity), BF propose to model u_t as the following:

$$u_t = c_{t-1}^\phi \varepsilon_t, \tag{22}$$

where ε_t is assumed to be uncorrelated and homoskedastic with variance σ^2, so that the variance of u_t is assumed to be proportional to $c_{t-1}^{2\phi}$; the authors do not consider ϕ as a parameter to be estimated, but they rather fix it heuristically to either $1/2$ or 1, finding that $1/2$ is preferable in their application. In the application, we will use the residuals of the homoskedastic model to test for homoskedasticity vs. heteroskedasticity of the proposed type.

Model (19) can be reparameterized in different ways:

$$\Delta c_t = \alpha\left(c_{t-1} - \beta_0 - \beta_1 C_{t-1} - \beta_2 C_{t-1}^2\right) + u_t, \tag{23}$$

$$= \alpha c_{t-1} + \gamma_0 + \gamma_1 C_{t-1} + \gamma_2 C_{t-1}^2 + u_t. \tag{24}$$

The parametrization (24) is suited for estimation, either with OLS when u_t is assumed to be uncorrelated and homoskedastic, or by WLS (dividing left and right by c_{t-1}^ϕ), if u_t is assumed to follow (22). Conversely, the parametrization (23) is useful because the parameters in $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$ are related to the parameters $\boldsymbol{\theta} = (m, p, q)'$ as in (15): therefore if we obtain estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}}$, we can map them into estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}}$ using (15) and (16) directly.[12]

Let us define $\boldsymbol{\gamma}' = (\gamma_0, \gamma_1, \gamma_2)$. Assuming that $u_t \sim iidN(0, \sigma^2)$, ML estimates of $\boldsymbol{\pi}' = (\alpha, \boldsymbol{\gamma}')$ can be obtained by OLS in (24), obtaining $\hat{\alpha}$, $\hat{\boldsymbol{\gamma}}$ and the corresponding variance–covariance matrix:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\pi}} = \begin{bmatrix} \hat{\sigma}_\alpha^2 & \hat{\boldsymbol{\Sigma}}_{\hat{\gamma}.\hat{\alpha}}' \\ \hat{\boldsymbol{\Sigma}}_{\hat{\gamma}.\hat{\alpha}} & \hat{\boldsymbol{\Sigma}}_{\hat{\gamma}} \end{bmatrix}.$$

Notice that $\boldsymbol{\beta} = -\frac{1}{\alpha}\boldsymbol{\gamma}$; therefore, ML estimates of $\boldsymbol{\beta}$ are given by the following:

$$\hat{\boldsymbol{\beta}} = -\frac{1}{\hat{\alpha}}\hat{\boldsymbol{\gamma}}. \tag{25}$$

We then obtain the following:

$$J_{\beta.\pi} = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\pi}'} = \frac{1}{\alpha^2} \begin{bmatrix} \gamma_0 & -\alpha & 0 & 0 \\ \gamma_1 & 0 & -\alpha & 0 \\ \gamma_2 & 0 & 0 & -\alpha \end{bmatrix} = \frac{1}{\alpha^2}(\boldsymbol{\gamma}, -\alpha \boldsymbol{I}_3),$$

and, using the delta method, we have the following:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\beta}} = \hat{J}_{\beta.\pi} \hat{\boldsymbol{\Sigma}}_{\hat{\pi}} \hat{J}'_{\beta.\pi} = \frac{1}{\hat{\alpha}^2} \left(\frac{\hat{\boldsymbol{\gamma}}\hat{\boldsymbol{\gamma}}'}{\hat{\alpha}^2} \hat{\sigma}_\alpha^2 - \frac{\hat{\boldsymbol{\Sigma}}_{\hat{\gamma}.\hat{\alpha}}\hat{\boldsymbol{\gamma}}' + \hat{\boldsymbol{\gamma}}\hat{\boldsymbol{\Sigma}}'_{\hat{\gamma}.\hat{\alpha}}}{\hat{\alpha}} + \hat{\boldsymbol{\Sigma}}_{\hat{\gamma}} \right), \tag{26}$$

where $\hat{J}_{\beta.\pi}$ is the estimated counterpart of $J_{\beta.\pi}$. Starting from (25) and (26) one can obtain $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}_{\hat{\theta}}$ using (15) and (16). In particular, replacing $\hat{\boldsymbol{\Sigma}}_{\hat{\beta}} = \hat{J}_{\beta.\pi} \hat{\boldsymbol{\Sigma}}_{\hat{\pi}} \hat{J}'_{\beta.\pi}$ in (16), one obtains the following:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\theta}} = \hat{J}_{\theta.\beta} \hat{\boldsymbol{\Sigma}}_{\hat{\beta}} \hat{J}_{\theta.\beta}' = \hat{J}_{\theta.\beta} \hat{J}_{\beta.\pi} \hat{\boldsymbol{\Sigma}}_{\hat{\pi}} \hat{J}_{\beta.\pi}' \hat{J}_{\theta.\beta}'.$$

Additionally, in this case, when one considers n "seemingly unrelated" equations such as (24), i.e.,

$$c_{i,t} = \alpha_i c_{i,t-1} + \gamma_{0,i} + \gamma_{1,i} C_{i,t-1} + \gamma_{2,i} C_{i,t-1}^2 + u_{i,t}, \qquad i = 1, ..., n, \tag{27}$$

and the variance–covariance matrix of $\boldsymbol{u}_t = (u_{1,t}, ..., u_{n,t})'$, for example $\boldsymbol{\Omega}_{\boldsymbol{u}}$, is not diagonal, then equation by equation OLS is no longer equivalent to ML. In this case, the likelihood can be maximized by iterated SURE, obtaining $\hat{\boldsymbol{\pi}}_i = (\alpha_i, \hat{\gamma}_{0,i}, \hat{\gamma}_{1,i}, \hat{\gamma}_{2,i})'$, $i = 1, ..., n$, $\hat{\boldsymbol{\Omega}}_{\boldsymbol{u}}$, and the variance-covariance matrix of $\hat{\boldsymbol{\pi}} = (\hat{\boldsymbol{\pi}}_1', ..., \hat{\boldsymbol{\pi}}_n')'$, i.e.,

$$\hat{\boldsymbol{\Sigma}}_{\hat{\pi}} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\hat{\pi}_1} & \cdots & \hat{\boldsymbol{\Sigma}}'_{\hat{\pi}_n, \hat{\pi}_1} \\ \vdots & \ddots & \vdots \\ \hat{\boldsymbol{\Sigma}}_{\hat{\pi}_n, \hat{\pi}_1} & \cdots & \hat{\boldsymbol{\Sigma}}_{\hat{\pi}_n} \end{bmatrix}.$$

Then, starting from each pair $(\hat{\boldsymbol{\pi}}_i, \hat{\boldsymbol{\Sigma}}_{\hat{\pi}_i})$ one can obtain the ML estimates of the structural parameters $\boldsymbol{\theta}_i = (m_i, p_i, q_i)'$ and the associated variance–covariance matrices as illustrated above.

As for the asymptotic properties of ML estimates of the structural parameters, Boswijk and Franses (2005) prove that $\hat{m}$ is consistent in T (as the time span increases C_T ideally coincides with m), whereas $\hat{p}$ and $\hat{q}$ are not; moreover, they show that the asymptotic distribution cannot be proved to be normal. However, they demonstrate with an extensive simulation that when the frequency is allowed to go to infinity along with the time span, then $\hat{m}$, $\hat{p}$ and $\hat{q}$ are essentially unbiased and asymptotically normal; they also show that this is approximately valid, even with a fixed time span, at least if it includes the inflection point $\frac{\ln q - \ln p}{p+q}$. In other words, if the observed time span includes the inflection point and the sampling frequency is reasonably high, their results suggest that using the standard normal and the χ^2 for making inference on the parameters is a reasonable approximation.

3.3. Boswijk et al. Multivariate Model

Boswijk et al. (2009), henceforth BFF, propose a multivariate generalization of (19). The BFF model is made up of n equations, and can be written as follows:

$$\Delta c_{i,t} = \sum_{j=1}^{n} \alpha_{ij} \left(c_{j,t-1} - (p_j + q_j) C_{j,t-1} + \frac{q_j}{m_j} C_{j,t-1}^2 - p_j m_j \right) + u_{i,t}, \qquad i = 1, ..., n, \tag{28}$$

where, in a simplified homoskedastic version of the model, we might assume that $\boldsymbol{u}_t = [u_{1,t}, ..., u_{n,t}] \sim iidN_n(\boldsymbol{0}, \boldsymbol{\Omega})$.[13] Along the lines of the BF model, (28) may be reparametrized as follows:

$$\Delta c_{i,t} = \sum_{j=1}^{n} \alpha_{ij} \left(\boldsymbol{\beta}_j' X_{j,t-1} - \beta_{0j} \right) + u_{it}, \qquad i = 1, \dots, n, \tag{29}$$

with

$$X_{jt}' = \left[c_{j,t}, C_{j,t}, C_{j,t}^2 \right], \quad \boldsymbol{\beta}_j' = \left[1, -\beta_{1j}, -\beta_{2j} \right]$$

or, more compactly,

$$Y_t = \boldsymbol{\alpha}\boldsymbol{\beta}' X_{t-1} + \boldsymbol{u}_t, \tag{30}$$

where

$$
\underset{n\times 1}{Y_t} = \begin{bmatrix} \Delta c_{1,t} \\ \vdots \\ \Delta c_{n,t} \end{bmatrix}, \quad
\underset{(3n+1)\times 1}{X_t} = \begin{bmatrix} X_{1,t} \\ \vdots \\ X_{n,t} \\ 1 \end{bmatrix}, \quad
\underset{n\times 1}{\boldsymbol{u}_t} = \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{n,t} \end{bmatrix},
$$

$$
\underset{n\times n}{\boldsymbol{\alpha}} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}, \quad
\underset{(3n+1)\times n}{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{\beta}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{\beta}_n \\ \beta_{01} & \cdots & \beta_{0n} \end{bmatrix}.
$$

Since this paper's main goal is to celebrate Søren Johansen and Katarina Juselius, it is nice to remark that, apart from the exclusion restrictions in $\boldsymbol{\beta}$, and the fact that the rank of $\boldsymbol{\alpha}\boldsymbol{\beta}'$ is actually full, (30) has the mathematical form of the "reduced rank regression" popularized by Søren and Katarina; therefore, in estimating and interpreting the model, we can benefit directly from the results inspired by their work, in particular Hansen (2003). Notice that the (exclusion) restrictions on the matrix $\boldsymbol{\beta}$ can be written as the following:

$$\mathrm{vec}(\boldsymbol{\beta}) = H_\beta \boldsymbol{\varphi}_\beta + h_\beta \tag{31}$$

for suitable restriction matrices H_β and h_β.[14] It might be also interesting to consider restrictions on $\boldsymbol{\alpha}$ of the following type:

$$\mathrm{vec}(\boldsymbol{\alpha}) = H_\alpha \boldsymbol{\varphi}_\alpha, \tag{32}$$

for example to test the hypothesis that the matrix $\boldsymbol{\alpha}$ is diagonal, under which (28) would collapse into n "seemingly unrelated" BF equations such as (19).[15] Of course, when $\boldsymbol{\alpha}$ is unrestricted, we have that $H_\alpha = I_{n^2}$ and $\boldsymbol{\varphi}_\alpha = \mathrm{vec}(\boldsymbol{\alpha})$.

Assuming that $\boldsymbol{u}_t \sim iid\mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$, the log-likelihood function is given by $\ell(\boldsymbol{\varphi}, \boldsymbol{\Omega}) = -\frac{T}{2}\left[n\ln(2\pi) + \ln|\boldsymbol{\Omega}| + \mathrm{tr}(\boldsymbol{\Omega}^{-1} M_{uu}) \right]$, where $M_{uu} = T^{-1}\sum_{t=1}^{T} \boldsymbol{u}_t \boldsymbol{u}_t'$. Since the log-likelihood score is bi-linear in the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, one can employ the generalized reduced rank regression algorithm proposed by Hansen (2003) for likelihood maximization of I(1) VAR models under linear restrictions. This provides maximum likelihood estimates of the parameters $\boldsymbol{\varphi} = \left[\boldsymbol{\varphi}_\alpha', \boldsymbol{\varphi}_\beta' \right]'$ and $\boldsymbol{\Omega}$, for example, $\hat{\boldsymbol{\varphi}}$ and $\hat{\boldsymbol{\Omega}}$.[16]

To work out the variance–covariance matrix associated to $\hat{\boldsymbol{\varphi}}$, notice that the model (30) under the restriction (31) and (32) is a sub-model of the following regression model:

$$Y_t = \boldsymbol{\Pi} X_{t-1} + \boldsymbol{u}_t,$$

where $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}' = \boldsymbol{\Pi}(\boldsymbol{\varphi})$ is a smooth function of the vector of the parameters in $\boldsymbol{\varphi}$. The second derivatives of the log-likelihood with respect to $\mathrm{vec}(\boldsymbol{\Pi})$ are given by $-T\, M_{XX} \otimes \boldsymbol{\Omega}^{-1}$, see e.g., Johansen (2006, Equation (13)), where $M_{XX} = T^{-1}\sum_{t=1}^{T} X_{t-1} X_{t-1}'$. Because the

parameters in $\boldsymbol{\varphi}$ and in $\boldsymbol{\Omega}$ are asymptotically independent, one finds that the Hessian with respect to $\boldsymbol{\varphi}$ equals the following:

$$\mathcal{H}_{\boldsymbol{\varphi}} = \frac{\partial^2 \ell(\boldsymbol{\varphi}, \boldsymbol{\Omega})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}'} = -T \, J'_{\Pi.\boldsymbol{\varphi}} \left(M_{XX} \otimes \boldsymbol{\Omega}^{-1} \right) J_{\Pi.\boldsymbol{\varphi}'} \tag{33}$$

where

$$J_{\Pi.\boldsymbol{\varphi}} = \frac{\partial \, \mathrm{vec} \, \boldsymbol{\Pi}(\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}'}.$$

In order to describe $J_{\Pi.\boldsymbol{\varphi}}$ in more detail observe that, in the present case, one has $\mathrm{vec} \, \boldsymbol{\Pi}(\boldsymbol{\varphi}) = \mathrm{vec} \, \boldsymbol{\alpha}\boldsymbol{\beta}' = (\boldsymbol{\beta} \otimes I_n)\mathrm{vec} \, \boldsymbol{\alpha} = \left(I_{(3n+1)} \otimes \boldsymbol{\alpha} \right)\mathrm{vec} \, \boldsymbol{\beta}'$. Therefore, using (31) and (32), one finds the following:

$$\partial \mathrm{vec}\left(\boldsymbol{\alpha}\boldsymbol{\beta}' \right) / \partial \boldsymbol{\varphi}'_{\boldsymbol{\alpha}} = (\boldsymbol{\beta} \otimes I_n) H_{\boldsymbol{\alpha}},$$
$$\partial \mathrm{vec}\left(\boldsymbol{\alpha}\boldsymbol{\beta}' \right) / \partial \boldsymbol{\varphi}'_{\boldsymbol{\beta}} = \left(I_{(3n+1)} \otimes \boldsymbol{\alpha} \right) \mathcal{K}_{(3n+1),n} H_{\boldsymbol{\beta}}.$$

where $\mathcal{K}_{mn}$ is a commutation matrix, which satisfies $\mathcal{K}_{mn} \mathrm{vec}(M) = \mathrm{vec}(M')$ when M is $m \times n$. Therefore

$$J_{\Pi.\boldsymbol{\varphi}} = \mathrm{blkdiag}\left((\boldsymbol{\beta} \otimes I_n) H_{\boldsymbol{\alpha}}, \left(I_{(3n+1)} \otimes \boldsymbol{\alpha} \right) \mathcal{K}_{(3n+1),n} H_{\boldsymbol{\beta}} \right).$$

The variance–covariance matrix of $\hat{\boldsymbol{\varphi}}$ can be then estimated by the following:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\varphi}}} = -\hat{\mathcal{H}}_{\boldsymbol{\varphi}}^{-1}$$

where $\hat{\mathcal{H}}_{\boldsymbol{\varphi}}$ is obtained by plugging the ML estimates $\hat{\boldsymbol{\varphi}}$ and $\hat{\boldsymbol{\Omega}}$ instead of $\boldsymbol{\varphi}$ and in $\boldsymbol{\Omega}$ in (33).

4. Results

Our statistical analysis is based on two equations, headed to $c_{1,t}$ and $c_{2,t}$, respectively (see Section 2 for a definition of the indices). In this section, we will first discuss the estimates of the reduced form models and then the corresponding estimates of the structural form models.

4.1. Analysis of the Reduced Form—Comparing Bass, BF, BFF

As illustrated in the previous section, the two univariate Bass Equation (18) may be seen as a restricted version of the bivariate BFF model (29), with four restrictions: $\alpha_{11} = \alpha_{22} = -1$ and $\alpha_{12} = \alpha_{21} = 0$. Similarly, the univariate dynamic BF model (27) may also be seen as a restricted version of the bivariate BFF model (29), with only two restrictions: $\alpha_{12} = \alpha_{21} = 0$. In all cases, assuming that the errors in the two equations are simultaneously correlated, i.e.,

$$\boldsymbol{\Omega}_u = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

efficient estimates of all models may be obtained by maximum likelihood, using the Hansen (2003) algorithm as illustrated in Section 3.3.[17] The results are shown in Table 4.

Table 4. ML estimates of the reduced form parameters.

		Model (18)		Model (27)		Model (29)	
		Estimate	t-Ratio	Estimate	t-Ratio	Estimate	t-Ratio
APP	$\hat{\alpha}_{1,1}$	-1		-0.435	-5.74	-0.508	-6.26
	$\hat{\alpha}_{1,2}$	0		0		0.715	2.81
	$\hat{\beta}_{0,1}$	17.55	9.00	19.40	5.28	19.81	4.78
	$\hat{\beta}_{1,1}$	0.0196	9.62	0.0183	4.76	0.0190	4.61
	$\hat{\beta}_{2,1}$	-1.33×10^{-6}	-3.22	-1.054×10^{-6}	-1.35	-1.32×10^{-6}	-1.60
	$\hat{\sigma}_1$	10.40		8.42		8.15	
MET	$\hat{\alpha}_{2,1}$	0		0		0.0635	2.34
	$\hat{\alpha}_{2,2}$	-1		-0.492	-6.28	-0.547	-6.44
	$\hat{\beta}_{0,2}$	5.90	8.28	6.27	5.05	6.74	4.65
	$\hat{\beta}_{1,2}$	0.0205	5.77	0.0199	3.21	0.0187	2.73
	$\hat{\beta}_{2,2}$	-2.06×10^{-5}	-5.71	-2.07×10^{-5}	-3.29	-1.97×10^{-5}	-2.86
	$\hat{\sigma}_2$	3.30		2.79		2.72	
	$\hat{\rho}$	0.416		0.140		0.161	
	$\log L$	-715.81		-682.06		-675.26	

Table 5 reports some misspecification tests based on the residuals illustrated in Figure 3.

The two rows, headed AC, in the table report the results of tests for auto correlation of the residuals of the applied and methodological equation, respectively. Specifically, we tested for serial correlation up to $k = 20$ lags using the Ljung-Box Q-statistic, whose null hypothesis is that the errors are uncorrelated.[18] The p-value is zero for the standard Bass model (18): therefore, residuals serial correlation is a major problem for that model. Conversely, in models (27) and (29), the white noise assumption is not rejected for the methodological equation, while for the applied equation, there is a clear improvement over model (18), but some autocorrelation seems to remain for both models, which suggests to invest more on the dynamic specification, which is left for further research.

The four rows headed HSK in the table report the results of two different types of Breusch–Pagan tests for heteroskedasticity for the applied and methodological equation, respectively. In all cases, the null hypothesis is that the errors are homoskedastic, but we introduced two different alternatives. In fact, as seen in Equation (22), Boswijk and Franses (2005) suggest that the standard deviation should be proportional to $c_{i,t-1}^{\phi}$ (with $\phi = 1/2$ or $\phi = 1$); therefore, we introduced the constant and $c_{i,t-1}^{2\phi}$ in the auxiliary regression, with two alternative values for ϕ. For model (18), the null is rejected in most cases.[19] Conversely, in spite of the very convincing argument supporting heteroskedasticity made by the cited authors, we did not find statistically significant evidence in this sense for this data set in (27) and (29); therefore, for the analysis in this paper, we did not consider the heteroskedastic versions of BF and BFF models.

The log-likelihood increases by 33.75 from model (18) to model (27): the LR test is therefore $\chi_2^2 = 67.50$, and the p-value is essentially zero. According to this result, the standard Bass model seems unable to capture the persistent swings clearly visible in Figure 2 and in the first plot of Figure 3: notice in fact that both parameters α_{11} and α_{22} estimated in (27) are approximately -0.5 and statistically different from -1, which implies that only half of the distance from the ideal Bass path is corrected within one quarter, giving rise to persistent disequilibria. However, even model (27) is not satisfactory: in fact, the log-likelihood of model (29) is significantly higher (the LR test is $\chi_2^2 = 14.6$, p-value 0.00111). This result is interesting since it suggests the existence of Granger causality running from the methodological research to applied research and/or vice-versa.

Table 5. Residual based tests for autocorrelation and heteroskedasticity. AC: Ljung-Box Q-statistics up to $k = 20$ lags (see footnote 25 for p). HSK: Breusch–Pagan test, including the constant and $c_{1,t-1}^{2\phi}$ (or $c_{2,t-1}^{2\phi}$) in the regression where the dependent variable is $\frac{\hat{u}_{1,t}^2}{\hat{\sigma}_1^2}$ (or $\frac{\hat{u}_{2,t}^2}{\hat{\sigma}_2^2}$).

		Model (18)		Model (27)		Model (29)	
		Test	***p*-Value**	**Test**	***p*-Value**	**Test**	***p*-Value**
	$AC\ \chi_{k-p}^2$	247.3	0.000	32.64	0.013	31.74	0.016
APP	$HSK\ (\phi = 1/2)\ \chi_1^2$	3.96	0.047	0.88	0.348	2.79	0.095
	$HSK\ (\phi = 1)\ \chi_1^2$	3.86	0.047	0.33	0.566	1.27	0.260
	$AC\ \chi_{k-p}^2$	101.1	0.000	27.48	0.051	26.31	0.069
MET	$HSK\ (\phi = 1/2)\ \chi_1^2$	1.41	0.235	1.77	0.184	2.96	0.085
	$HSK\ (\phi = 1)\ \chi_1^2$	6.30	0.012	0.91	0.340	1.46	0.226
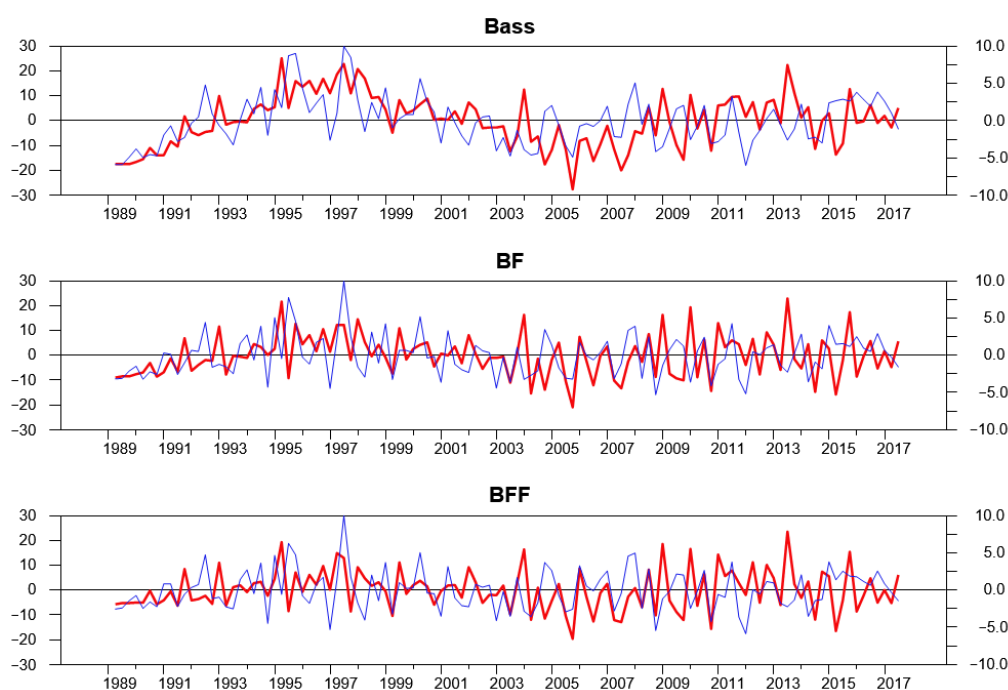
Figure 3. Residuals of different models: Bass = model (18), BF = model (27), BFF = model (29). Thick red line = "Applied" (left scale). Thin blue line = "Methodological" (right scale).

To shed some light on this, we observe that the estimates of α_{12} and α_{21} in (29) are both positive and statistically significant, suggesting that an increase in the methodological research leads to expect more applications in the future, and that an increase in applications stimulates further methodological research, with a continuous dialogue between the economic problems and econometric methods, which is exactly in the spirit of KJ and SJ's main message to the profession.

To provide a visual illustration of the relevance of the dynamic interaction between methodological and applied research in this field, we carry on a simulation exercise, similar in spirit to impulse response analysis. Impulse response functions, being the reactions of the variables to shocks entering the system, are useful for studying the interactions between variables in a vector autoregressive model (Lütkepohl 2016). In a more general non-linear setting, Potter (2000) and Koop et al. (1996) remark that nonlinear models produce impulse responses that are history- and shock-dependent; to overcome this problem, they introduce the notion of "generalized impulse response functions", based on a stochastic simulation, which can be applied in both the linear and non-linear case. We considered this tool, but since the non-linearity is relatively mild in our case, we opted for a tailored solution that is closer to the traditional deterministic impulse response analysis.

We initialize $C_{i,0} = c_{i,0} = 0$,[20] and then we compute two alternative trajectories for $c_{i,t}$ based on the estimated counterpart of (29). In the first dynamic simulation, $u_{i,t}$ is set to zero for all i and t: this leads to the "unshocked" paths $c_{i,t}^U$, corresponding to the deterministic trajectory that would take place in the absence on any innovation, starting from the assumed initial conditions. In the second dynamic simulation, we set $u_{j,1} = \hat{\sigma}_j$ in the j-th equation, whereas all other innovations (different equations and/or different times) are set to zero so that the impulse corresponds to one standard deviation in just one of the equations;[21] this exercise leads to the shocked paths $c_{j,i,t}^S$, where the first subscript, j, indicates which equation has been shocked. The standardized response of the i-th equation to an impulse on the j-th equation are then given by the difference of the two trajectories, standardized by the standard deviation of the output variable as follows:

$$IR_{j,i,t} = \frac{\left(c_{j,i,t}^S - c_{i,t}^U\right)}{\hat{\sigma}_i} \qquad j = 1, ..., n; \quad i = 1, ..., n; \quad t = 1, ...$$

The IRs therefore isolate that part of the trajectory $c_{j,i,t}^S$, which can be attributed to the shock. The first 20 IRs are illustrated in Figure 4. Notice that, by construction, $IR_{j,i,1}$ is equal to 1 for $j = i$, 0 otherwise. According to Figure 4, in the short run, the (standardized) response of the methodological literature to a (standardized) impulse in the applied literature appears qualitatively very similar to the (standardized) response of the applied literature to a (standardized) impulse in the methodological literature.
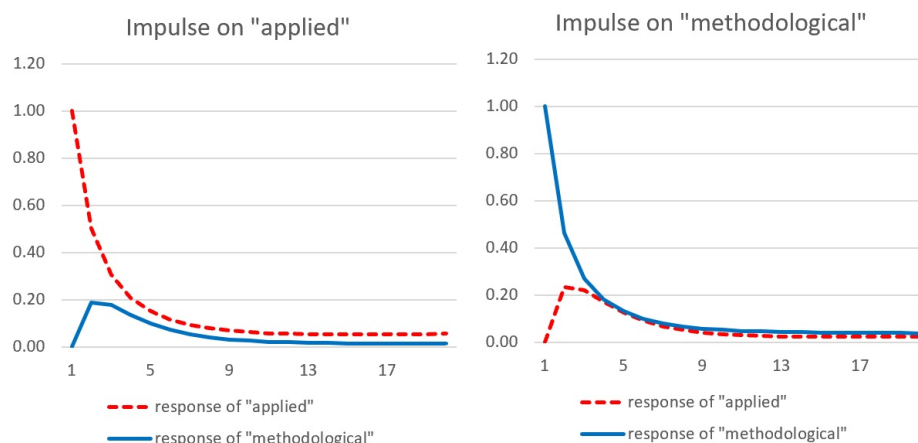

Figure 4. Standardized impulse responses based on model (29). Initialization: $C_{i,0} = c_{i,0} = 0$, $i = 1, 2$.

Some more insight on the relationship between methodological and applied research can be obtained by analyzing the cumulative IRs. It is important to remark that, given the mathematical nature of the model, the shocks do not have permanent effects. In fact, as t goes to infinity, the cumulative citations $C_{i,t}$ will eventually reach the saturation point m_i irrespective of the initial conditions and/or the shocks they undergo: this implies that $\sum_{t=1}^{\infty} c_{i,t}^U = \sum_{t=1}^{\infty} c_{j,i,t}^S = m_i - C_{i,0}$ for any i, j and $C_{i,0}$, and therefore $\sum_{t=1}^{\infty} IR_{j,i,t} = 0$. As a consequence, although the first IRs illustrated in Figure 4 are positive, at some point they turn negative (although with a very small magnitude) so that, in the limit, the cumulative sum is zero. This behavior is better illustrated through the cumulative IRs, illustrated in Figure 5, for a much longer period (500 quarters).
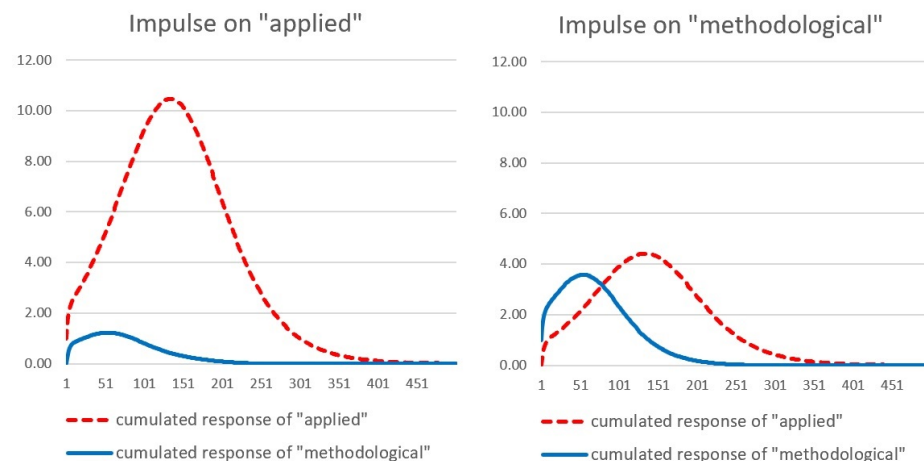
Figure 5. Cumulative standardized impulse responses based on model (29). Initialization: $C_{i,0} = c_{i,0} = 0$, $i = 1, 2$.

Figure 5 shows that an impulse equal to $\hat{\sigma}_1$ (i.e., 8.15 papers) in the applied literature is strongly "self exciting", giving rise to a very long sequence of positive IRs in the applied literature itself, adding up to $10.5\hat{\sigma}_1$ (about 85 papers) in the subsequent 150 quarters (almost 40 years), before it starts fading away. Conversely, the cumulative impact on the methodological literature of the same impulse is shorter living, and way less relevant ($1.2\hat{\sigma}_2$, i.e., about 3 papers). On the other hand, an impulse equal to $\hat{\sigma}_2$ (i.e., 2.72 papers) in the methodological literature is not so "self-exciting" (the peak of the cumulative IRs is only $3.6\hat{\sigma}_2$—10 papers—about 50 quarters after the impulse), whereas the cumulative impact on the applied literature seems very important (the peak is equal to $4.4\hat{\sigma}_1$—35 papers—about 150 quarters after the impulse). This evidence seems to suggest that, although in the short run, the cross fertilization is rather balanced, in the long run, the methodological literature triggers the applications more than the other way around.[22]

Actually, the extremely long sequence of positive IR's, well beyond the observed period of 115 quarters, casts some doubt on the validity of the implicit assumption that the impulses do not have a permanent effect. We think that a hint for future research arising from the current study is to develop an alternative model where the saturation point m_i is not already set at the beginning of the process, but it is to some extent "path dependent". In fact, if an idea appears more successful than what was initially assumed (i.e., we observe some unexpected citations), we should reconsider the expected total number of citations in the long run, leading to an upward revision. Conversely, when an idea is suddenly abandoned, possibly in favor of an alternative paradigm (i.e., we observe an unexpected reduction in the number of citations), we should reasonably revise downwards the expected total number of citations in the long run.

4.2. Analysis of the Structural Form

Table 6 reports the "structural" parameters m, p and q in the models (12), (19) and (28), which are based on the ML estimates of (18), (27) and (29), respectively. The associated standard errors are computed using the delta method, as illustrated in Sections 3.1–3.3. It is important to remark that the standard errors reported for models (12) are not reliable: they appear to be much lower than in the other two models, but the assumptions for applying ML—in particular, the absence of serial correlation—are clearly invalid for that model as illustrated in Table 5.

Table 6. Estimates of the structural parameters of the Bass, BF and BFF models. The standard errors of the $\hat{t}^P$'s measured in quarters.

	Coefficient	Model (12) Estimate	Std.err.	Model (19) Estimate	Std.err.	Model (28) Estimate	Std.err.
APP	$\hat{m}_1$	15,619.7	3289.6	18,332.5	9687.4	15,350.6	6396.2
	$\hat{p}_1$	0.00112	1.98×10^{-4}	0.00106	4.79×10^{-4}	0.00129	4.65×10^{-4}
	$\hat{q}_1$	0.0207	0.00216	0.0193	0.00421	0.0203	0.00442
	$\hat{t}_1^P$	2022:2	16.2	2024:3	42.1	2020:4	32.4
	$\widehat{\tilde{c}}_1^P$	89	9.9	99	30.0	88	18.4
	$\widehat{\tilde{C}}_1^P$	7386	1578.9	8664	4688.9	7188	3050.6
MET	$\hat{m}_2$	1229.3	61.3	1212.2	100.7	1227.7	119.7
	$\hat{p}_2$	0.00480	5.38×10^{-4}	0.00517	9.50×10^{-4}	0.00549	4.91×10^{-4}
	$\hat{q}_2$	0.0253	0.00333	0.0251	0.00580	0.0242	0.00643
	$\hat{t}_2^P$	2002:4	4.4	2002:1	7.2	2001:2	8.4
	$\widehat{\tilde{c}}_2^P$	11	0.5	11	0.8	11	1.0
	$\widehat{\tilde{C}}_2^P$	498	23.7	481	40.2	475	49.9

The timing of the citations peaks $\hat{t}_1^P$ and $\hat{t}_2^P$, the associated peaks $\widehat{\tilde{c}}_1^P$ and $\widehat{\tilde{c}}_2^P$, and the corresponding cumulative number of citations at the peak $\widehat{\tilde{C}}_1^P$ and $\widehat{\tilde{C}}_2^P$, are obtained by plugging the estimated structural parameters in (9)–(11), and the associated standard errors are computed using the delta method.[23]

According to the evidence provided in Table 6, the estimates of the structural parameters are rather robust to the model used. Our comments are focused on the results based on model (28), which is statistically preferable.

It is interesting, and not surprising, that the "innovation parameter" p is much higher for the methodological literature, whereas the "imitation parameter" is quite similar in the two strands of the literature: this makes imitation relatively more important than innovation in the applied literature. As for the citation peaks, it seems that the peak in the methodological literature (11 papers per quarter) was reached in 2001, whereas the peak in the applied literature (88 papers per quarter) is expected in 2020, 12 quarters after the end of the estimation sample (although the associated standard error is extremely large—32 quarters). Based on the discussion of the properties of the estimates provided in Boswijk and Franses (2005), the estimates of the methodological equation should, therefore, be regarded as more reliable since the inflection point of the diffusion curve appears to be within the sample; this is less so for the estimates of the applied equation, where the estimated inflection point is outside the sample (of course we do not know the "true" inflection point). Not surprisingly, the standard error of $\hat{m}_1$ is quite large (the coefficient of variation $\hat{\sigma}_{\hat{m}_1}/\hat{m}_1$ is about 42%), while the standard error of $\hat{m}_2$ is much smaller (the coefficient of variation $\hat{\sigma}_{\hat{m}_2}/\hat{m}_2$ is less than 10%).

Figure 6 illustrates the observed time series along with the estimated unconditional expectation obtained by plugging the estimated structural parameters in Equation (8). Strictly speaking, since at the end of the sample (2017:Q3) we observe $C_{1,T} = 5445.2$ and $C_{2,T} = 1011.8$, our point estimates would ideally imply that we should expect $\hat{m}_1 - C_{1,T} = 9905$ applied WoS papers and $\hat{m}_2 - C_{2,T} = 216$ methodological WoS papers citing KJ and SJ in the future. We think that this interpretation is hazardous, to say the least. It is worth observing that the estimates of the structural parameters, especially m_1 and m_2, are very unstable as observed among others in Chandrasekaran and Tellis (2018), and they mainly seem to represent the history of the process in a descriptive sense rather than being a reliable forecasting tool in an inferential sense. For example, if we re-estimate the parameters based on the sub sample 1989:Q1-2005:Q4, so that the end of the sample occurs right before the "second wind" clearly visible in the plot, we would obtain $\hat{m}_1 = 2710.6$

$(\hat{\sigma}_{\hat{m}_1} = 142.5)$, $\hat{m}_2 = 684.3$ $(\hat{\sigma}_{\hat{m}_2} = 34.7)$, $\hat{t}_1^P$ =1999:Q1 and $\hat{t}_2^P$ =1997:Q4.[24] Therefore, estimating the same model 15 years ago, one would be convinced of the following: (i) that the citations peak was already reached several years before (and then the estimates would be regarded as reliable); (ii) that the potential for this literature was about one fifth of what it appears now; and (iii) that by 2020, the interest in KJ and SJ work will have disappeared (see Figure 7). However, as this Special Issue confirms, no prediction could have proved more wrong!
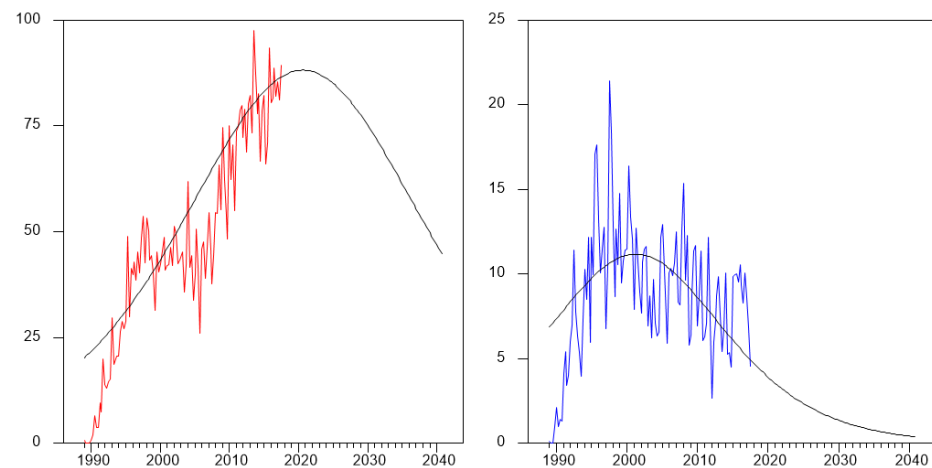
Figure 6. Observed time series along with the estimated unconditional expectation based on Equation (8). **Left**—applied index; **right**—methodological index.
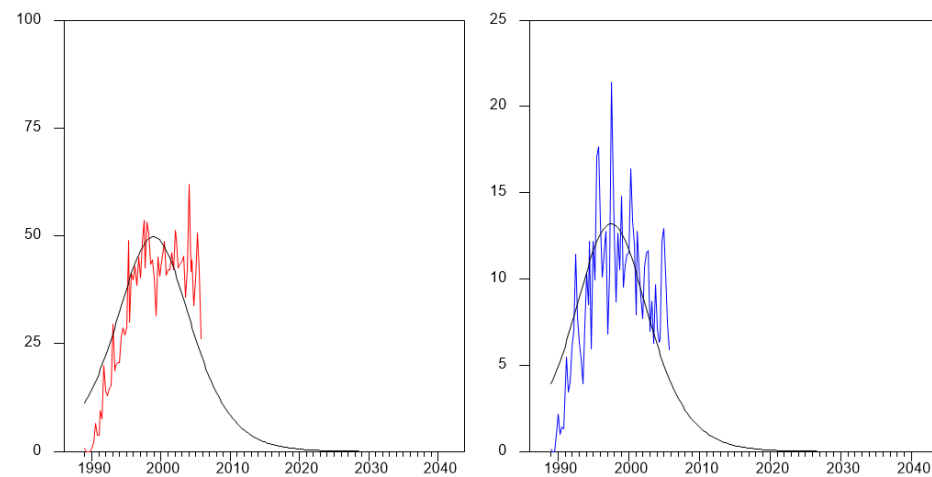

Figure 7. Forecasting fallacy: observed time series along with the estimated unconditional expectation based on (8), parameters re-estimated based on the trimmed sample 1898:Q1-2005:Q4. **Left**—applied index; **right**—methodological index.

5. Conclusions and Suggestions for Further Research

Our main purpose in writing this paper was to contribute to the Festschrift in honor of Katarina Juselius and Søren Johansen as a sign of gratitude for their being for us a constant source of inspiration. We tried to find a way to show how profoundly they contributed to the development of economic ideas, emphasizing one key aspect of their approach, namely, the dialogue between empirical economics and econometric methodology. To this aim, we have proposed an operational way to disentangle, as much as possible, their contribution to applied and methodological econometric research, through the development of two indices based on the Web of Science database. We hope that this can also be a contribution

to bibliometric studies since a similar approach to assess in a quantitative way the impact of new ideas on methodological and applied research, and on the interaction between them, can be used for other areas. Ideally, similar analyses might be employed to investigate even more general epistemological issues, such as the relationship between theoretical and empirical research.

We think that the data we describe in Section 2 are very interesting per se. They show that KJ and SJ's influence on the literature is extremely important: their top 10 papers sum up to about 10,500 WoS citations (more than 50,000 in GS) from about 6500 citing papers, an average of more than 200 papers per year. Based on our indicators, 85% of the citing papers are essentially applied, whereas 15% are methodological: we do not have a benchmark for comparison, but we have the impression that the share of methodology is somewhat larger than in the econometric literature in general. As of 2017, the number of applied citing papers per quarter had not yet reached the peak (although a "false peak" seems to have occurred around 2000); conversely, the peak in the methodological literature seems to have been reached around 2001, although the shape of the trajectory is very flat after the peak, similar to what Bjork et al. (2014) has identified in a minority of Nobel prize winners and defined as "staying power".

To model the data, we resorted to an innovative dynamic multivariate version—proposed in Boswijk et al. (2009)—of the well-known Bass (1969) model. It was a pleasure for us to observe and emphasize that this model resembles so closely the Vector ECM model popularized by KJ and SJ; in particular, the bilinear nature of the model allows to use the Hansen (2003) algorithm to maximize the likelihood, which generalizes Johansen's ML algorithm, adapting it to a rather general class of restrictions, which includes our case.

The estimated model conveys very interesting information. As seen in Formula (9), the location of the citations peaks depends on the relative importance of the "innovation parameter" p and the "imitation parameter" q. Our estimates suggest that the different location of the peaks might be explained by the higher value of the parameter q_2 with respect to q_1, whereas p_1 and p_2 are quite similar: using the standard terminology in the Bass model literature, the difference in the parameters suggests that the methodological literature is mainly driven by "innovators", whereas "imitators" are relatively more important in the applied literature.

Another interesting finding is that, in the literature referring to KJ and SJ, the "cross-fertilization" between methodological and applied research is statistically significant and bi-directional (although possibly more effective from methodology to applications than the other way round). According to our impulse response analysis, rounding our figures, 8 unexpected applied papers in one quarter lead to predict that 3 methodological papers will follow, whereas 3 unexpected methodological papers lead to predict that 40 applied papers will follow (this is not so unbalanced as it seems at first sight since the scale of the two strands of literature is different). These results testify that one of the most important messages that Katarina Juselius and Søren Johansen have emphasized in their writings—i.e., that the applications should pose challenging problems to the methodology and that the methodology should sharpen the ability of applied researchers to ask meaningful questions to the data—has become a common heritage in this literature.

As for the estimated dimension of KJ and SJ influence, as measured by the parameters m_1 and m_2 (often called "saturation point" or "ceiling"), a word of caution is in order. Our estimates, $\hat{m}_1 = 15{,}351$ and $\hat{m}_2 = 1228$, imply that we should expect about 10,000 applied WoS papers and 200 methodological WoS papers citing KJ and SJ in the future. We do not consider these figures very reliable. Indeed, early in the literature, it was pointed out by Heeler and Hustad (1980) and others (e.g., Hyman 1988) that the predicting ability of the Bass model depends on the generation of accurate estimates of m. Srinivasan and Mason (1986) report problems with convergence when the data set does not contain the peak time period (i.e., the inflection point of the curve). The parameter m is, again, under attack in Van den Bulte and Lilien (1997): there is evidence of downward bias in the estimation of the saturation point. Finally, in their review article Chandrasekaran

and Tellis (2018), point out the overall poor forecasting ability, the unstable parameter estimates and the difficulty to define a clear stopping rule for the time window regarding data collection for the Bass model (since the data should, in theory, end when the entire market has adopted). We add one more critique to the list: the parameter m, in the logic of the Bass model, appears to be in the DNA of the process since the onset and to be immutable over time. In all versions of the model that we have considered, the "shocks" (i.e., the unexpected citing papers) have no permanent effect in the sense that they determine, at most, a persistent (but transitory) departure from a path, which eventually leads to m. In the spirit of the unit roots literature, so much inspired by the contribution of Katarina Juselius and Søren Johansen, our suggestion for further methodological research is to try and conceive a new model, where the shocks are allowed to have a permanent effect on the "ceiling". We think that this is absolutely needed in the applications, such as the bibliometric ones, where the notion of "population at risk" or "potential" is not obvious. However, also in marketing, or epidemiology, or in the analysis of technological innovation, the final diffusion is likely to be influenced in a crucial way by events that are largely unpredictable; therefore, pretending that the same differential equation—where m is fixed since $t = 0$—drives the dynamics of the process along its entire history might not be a realistic representation of the observed phenomena.

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

Data Availability Statement: The data are available from the authors upon request.

Appendix A. Bass vs. Autoregressive Models

It is interesting to observe that (12) can be regarded as a univariate AR(1) model for C_t with state dependent parameters. In fact, the model can be rewritten as the following:

$$\Delta C_t = \mu + \pi_t C_{t-1} + u_t \tag{A1}$$

with

$$
\begin{aligned}
\mu &= pm \\
\pi_t &= -p + q\left(1 - \frac{C_{t-1}}{m}\right)
\end{aligned}
$$

m can be seen as a steady state for C_t: in fact, if $C_{t-1} = m$, then $\pi_t = -p$, so that in the absence of shocks (i.e., $u_t = 0$), we have that $\Delta C_t = pm - pm = 0$. The state dependent parameter π_t controls the strength of the adjustment to the steady state.

- If $q = 0$, $p > 0$, and $m > 0$, then $\pi_t = -p$ and the model collapses into a standard stationary AR(1) with unconditional expectation m.
- If $q > p > 0$, and $m > 0$, then $\pi(0) > 0$ so that initially the system behaves like an explosive AR(1) with a positive drift mp. When $C_{t-1} = m\left(1 - \frac{p}{q}\right)$, then $\pi_t = 0$ so that the system locally behaves like a random walk with drift. When $C_{t-1} > m\left(1 - \frac{p}{q}\right)$, then $\pi_t < 0$ so that the system starts adjusting. An illustrative example, based on the estimated parameters for the methodological index, is given in Figure A1.
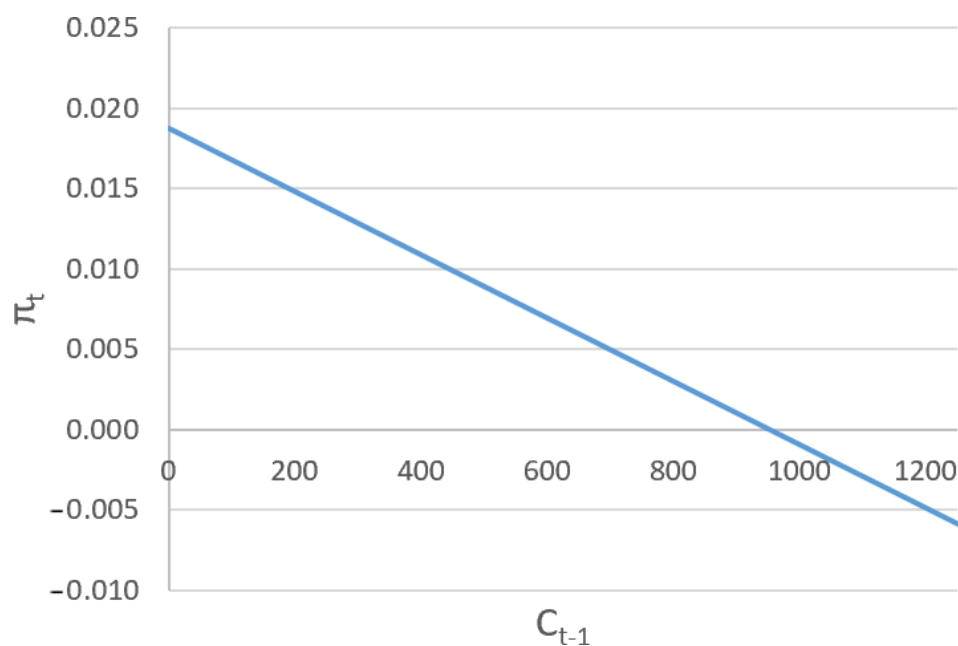
Figure A1. Illustration of π_t as a function of C_{t-1}, with $p = 0.00549$, $q = 0.0242$, $m = 1227.7$ as in the estimated equation for $C_{M,t}$.

Similarly, the model (19) can be seen as a univariate AR(2) model for C_t with state dependent parameters:

$$\Delta C_t = \mu + \pi_t C_{t-1} + \gamma \Delta C_{t-1} + u_t \tag{A2}$$

with:

$$
\begin{aligned}
\mu &= -\alpha m p, \\
\pi_t &= \alpha\left(p - q\left(1 - \frac{C_{t-1}}{m}\right)\right), \\
\gamma &= 1 + \alpha,
\end{aligned}
$$

so that when $\alpha = -1$ (A2) collapses into (A1). We remark that as far as α is negative the sign of π_t is the same in both models, and depends only on the sign of $C_{t-1} - m\left(1 - \frac{p}{q}\right)$. The magnitude of π_t instead is affected by α: everything else being fixed, when $-1 < \alpha < 0$, the process is less explosive at the beginning, and the strength of adjustment is weaker in the end, as compared to the case $\alpha = -1$.

Finally, the model (28) can be seen as a VAR(2) model for $C_t = [C_{1,t}, ..., C_{n,t}]'$ with state-dependent parameters:

$$\Delta C_t = \mu + \alpha \beta_t' C_{t-1} + \Gamma \Delta C_{t-1} + u_t \tag{A3}$$

where

$$
\underset{n\times 1}{C_t} = \begin{bmatrix} C_{1,t} \\ \vdots \\ C_{n,t} \end{bmatrix}, \quad
\underset{n\times 1}{u_t} = \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{n,t} \end{bmatrix}, \quad
\underset{n\times 1}{\mu} = -\begin{bmatrix} \sum_{j=1}^n \alpha_{1j} p_j m_j \\ \vdots \\ \sum_{j=1}^n \alpha_{nj} p_j m_j \end{bmatrix}
$$

$$
\underset{n\times n}{\alpha} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}, \quad
\underset{n\times n}{\beta_t} = diag\left\{p_i - q_i\left(1 - \frac{C_{i,t-1}}{m_i}\right)\right\}, \quad \Gamma = I_n + \alpha.
$$

It is easily seen that, when $\boldsymbol{\alpha}$ is diagonal, (A3) collapses into n seemingly unrelated equations such as (A2), while when $\boldsymbol{\alpha} = -\boldsymbol{I_n}$, (A3) collapses into n seemingly unrelated equations such as (A1).

Appendix B. Sensitivity to ω

As discussed in Section 2, the parameter ω in (1) and (2) controls for the weight of the "Mainly Applied" (MA) and "Mainly Methodological" (MM) papers on the aggregate indices $c_{1,t}$ (applied) and $c_{2,t}$ (methodological). Meaningful values of ω are in the range $(0.5; 1)$: with $\omega = 0.5$ the papers classified as MA and MM are essentially pooled together, and allowed to contribute evenly to both indices. In the opposite polar case, $\omega = 1$, MA (or MM) is considered equivalent to PA (or PM). We observe that, in principle, instead of a single weight ω, it would be possible to consider two different weights for MA and MM papers, for example, ω_A and ω_M, defining the following:

$$
\begin{aligned}
c_{1,t} &= c_{PA,t} + \omega_A c_{MA,t} + (1 - \omega_M)c_{MM,t} \\
c_{2,t} &= c_{PM,t} + \omega_M c_{MM,t} + (1 - \omega_A)c_{MA,t}
\end{aligned}
$$

We remark, however, that in our dataset, any value $0.5 < \omega_M < 1$ would leave the two indexes essentially unchanged since there are only 92 papers classified as MM in front of 716 classified as PM and 4198 classified as PA; therefore, our choice to set $\omega_M = \omega_A$ is a minor problem. Conversely, in our dataset, the critical issue is ω_A, mainly because of its impact on $c_{2,t}$: in fact, there are 1451 papers classified as MA and 716 classified as PM, so that setting $\omega_A = 0.5$, the MA papers would be as influential as the PM papers in the index $c_{2,t}$. This argument induced us to set $\omega_A = \omega_M = 0.85$. With this choice, $(1 - \omega_A)$ is relatively close to 0 so that $c_{2,t}$ reflects mainly the 716 PM papers and, therefore, is a more reliable measure of the methodological research. Notice that this choice has a minor impact on the reliability of the applied index $c_{1,t}$ for two reasons: (i) the 4198 papers classified as PA outnumber the 1451 MA papers, and (ii) the correlation between $c_{PA,t}$ and $c_{MA,t}$ it quite high, 69.5% (see Figure 1) (conversely the correlation between $c_{PM,t}$ and $c_{MA,t}$ is only 1.7%). Finally, notice that setting $\omega_A = 0.85$ would make it approximately equal to $n_{PA}/(n_{PA} + n_{PM}) = 0.8545$: in practice, this corresponds to the assumption that the share of "applied research" of an MA paper is similar, on average, to the share of applied research in the econometric literature referring to KJ and SJ papers in general. Let us now discuss how a different choice of ω would affect our results.

As illustrated in Table A1, changing ω affects quite relevantly the magnitude of the indices (especially $c_{2,t}$), as well as the correlation among them. To explain the impact on the magnitude, remember that, when $\omega = 1$, the 1451 MA papers are treated *de facto* as the "Purely Applied"(PA) ones, whereas when $\omega = 0.5$, only half of them (725.5) is treated as applied, while the other half is treated as methodological, and therefore, contribute also to the methodological index $c_{2,t}$.[25] The impact on the correlation is instead explained by the fact that the correlation between $c_{PA,t}$ and $c_{MA,t}$ it quite high (69.5%), whereas the correlation between $c_{PM,t}$ and $c_{MA,t}$ is negligible (1.7%); see Figure 1.

Table A1. Sensitivity to ω: impact of ω on some characteristics of the composite citation indices.

	$\omega = 0.5$	$\omega = 0.85$	$\omega = 1$
$C_{1,T}$	4969.5	5445.2	5649
$C_{2,T}$	1487.5	1011.8	808
$corr(c_{1,t}, c_{2,t})$	0.543	0.253	0.048

Given this impact of ω on the composite indices, it is interesting to analyze to which extent the results of the econometric model depend on it. The analysis is limited to the general model (29) since our analysis shows that it is preferable with respect to the restricted

counterparts (27) and (18). In this appendix, we show that our results are essentially robust to changes in ω.

Table A2 shows how the estimates of the reduced for changes when ω is changed.

Table A2. Sensitivity to ω: ML estimates of the reduced form parameters based on model (29).

| | | $\omega = 0.5$ | | $\omega = 0.85$ | | $\omega = 1$ | |
		Estimate	**t-Ratio**	**Estimate**	**t-Ratio**	**Estimate**	**t-Ratio**
APP	$\hat{\alpha}_{1,1}$	-0.620	-7.17	-0.508	-6.26	-0.459	-5.92
	$\hat{\alpha}_{1,2}$	0.782	4.00	0.715	2.81	0.624	2.33
	$\hat{\beta}_{0,1}$	16.95	4.58	19.81	4.78	21.11	4.88
	$\hat{\beta}_{1,1}$	0.0209	5.13	0.0190	4.61	0.0180	4.33
	$\hat{\beta}_{2,1}$	-1.86×10^{-6}	-2.06	-1.32×10^{-6}	-1.60	-1.06×10^{-6}	-1.33
	$\hat{\sigma}_1$	7.46		8.15		8.62	
MET	$\hat{\alpha}_{2,1}$	0.0902	2.25	0.0635	2.34	0.0590	2.38
	$\hat{\alpha}_{2,2}$	-0.545	-6.12	-0.547	-6.44	-0.596	-7.04
	$\hat{\beta}_{0,2}$	8.87	4.81	6.74	4.65	5.64	4.42
	$\hat{\beta}_{1,2}$	0.0161	2.66	0.0187	2.73	0.0223	2.96
	$\hat{\beta}_{2,2}$	-9.54×10^{-6}	-2.23	-1.97×10^{-5}	-2.86	-3.15×10^{-5}	-3.39
	$\hat{\sigma}_2$	3.46		2.72		2.71	
$\hat{\rho}$		0.253		0.161		0.034	

The sign and significance of the parameters are essentially the same irrespective of ω. Interestingly, the difference in the correlation between the two indices induced by ω, illustrated in Table A1, are reflected in different estimates of ρ, whereas the estimates of $\alpha_{1,2}$ and $\alpha_{2,1}$ (i.e., the parameters controlling for the dynamic interaction among the two processes) remain quite stable. Due to this, the (unreported) pattern of the standardized IRs and cumulative IRs computed with $\omega = 0.5$ and $\omega = 1$ are very similar to those illustrated in Figures 4 and 5 for $\omega = 0.85$. Unreported results show that also the misspecification tests are qualitatively unchanged for all ωs with respect to those reported in Table 5 for model (29): homoskedasticity and uncorrelatedness appear acceptable for any value of ω.

Table A3 illustrate how the structural parameters change when ω is changed. The influence on the m has an obvious interpretation: as ω increases, a larger share of the MA papers is removed from the methodological index (so that m_2 declines) and added to the applied index (so that m_1 increases). As for the ps and the qs, we observe that as ω increases, p_1 and q_1 decrease, whereas p_2 and q_2 increase. As a consequence of these changes, the timing of the peaks, obtained by formula (9), change: specifically, as ω increases, the peak in the applied literature moves to the right, whereas the peak in the methodological peak moves to the left. The distance between the peaks is 13 years with $\omega = 0.5$, and about 24 years when $\omega = 1$. This is not surprising: as illustrated in Figure 1, the dynamic behavior of the MA paper resembles closely the PA papers, and therefore, when 50% of them are considered methodological, $c_{1,t}$ and $c_{2,t}$ become more similar, and the two peaks become closer (although they still remain quite far away from each other).

An interesting consequence of the fact that, increasing ω, the peak of the applied literature moves ahead is that the quality of the structural parameters for the applied curve (already quite poor with $\omega = 0.5$) decreases considerably: when $\omega = 1$, the standard error associated to $\hat{m}_1$ is as large as 9516, and the standard error associated to the estimated timing of the peak $\hat{t}_1^P$ turns out to be 42 quarters, more than 10 years. It is a well-known fact in the literature that the estimates of the Bass model are quite poor if the sample period does not include the inflection point, which is quite likely the case for the applied literature if we trust the point estimates, and even more so when $\omega = 1$.

Table A3. Sensitivity to ω: ML estimates of the structural form parameters based on model (29). The standard errors of the $\hat{t}^P$ measured in quarters.

	Coefficient	$\omega = 0.5$ Estimate	Std.err.	$\omega = 0.85$ Estimate	Std.err.	$\omega = 1$ Estimate	Std.err.
APP	$\hat{m}_1$	11,994.6	3644.9	15,350.6	6396.2	18,058.0	9516.0
	$\hat{p}_1$	0.00141	3.94×10^{-4}	0.00129	4.65×10^{-4}	0.00117	5.31×10^{-4}
	$\hat{q}_1$	0.0224	0.00425	0.0203	0.00442	0.0192	0.00456
	t_1^P	2018:1	22.6	2020:4	32.4	2023:2	42.4
	c_1^P	76	10.3	88	18.4	98	28.0
	C_1^P	5618	1719.8	7188	3050.6	8479	4578.2
MET	m_2	2129.1	351.5	1227.7	119.7	907.1	62.3
	p_2	0.00416	8.77×10^{-4}	0.00549	4.91×10^{-4}	0.00622	0.00133
	q_2	0.0203	0.00604	0.0242	0.00643	0.0285	0.00682
	t_2^P	2005:1	11.7	2001:2	8.4	1999:4	7.2
	c_2^P	16	1.2	11	1.0	10	0.8
	C_2^P	846	122.6	475	49.9	355	34.2

Notes

[1]   Many thanks for the provision of the initial Web of Science data to Evi Sachini, Antonis Kardasis and Penny Nikolaidou of the National Documentation Centre/N.H.R.F. based in Athens, Greece.

[2]   Around the same time Google Scholar (GS) reported more than 50,000 citations for the same 10 papers. We opted for WoS instead of GS because, to avoid double counting, the analysis carried on in this paper is based on the citing papers instead of the citations, and working out the citing papers from GS is not easy. Admittedly, one drawback with using WoS instead of GS is that books cannot be considered; we think however that this would not substantially change the picture. In fact, according to GS, the book Johansen (1995) would rank 4th in terms of citations, the book Juselius (2006) would rank 6th, and adding both books the total citations count would be about 15% higher; however, since many papers citing one of the books will cite also some of the older papers, the impact of the books on the citing papers is likely to be way less than 10%.

[3]   Among the "super-citing" papers, there are also 14 papers with six citations and 53 papers with five citations. We remark that 40 of the 6457 papers (i.e., 0.62%) are authored or coauthored by SJ and/or KJ: given the small share we did not correct for self-citations.

[4]   An alternative way of measuring the influence of a paper could be based on counting the authors instead of the papers. We could then consider the number of authors citing KJ or SJ in each quarter, or preferably the number of "new authors", i.e., the number of authors citing KJ or SJ for the first time in each quarter, who never cited them before (this would avoid double counting, and would be a more precise measure of "contagion"). We do not explore this alternative in the present paper, leaving it for future research.

[5]   Classifying an econometric paper as "methodological" or "applied" is clearly arbitrary to some extent. A general discussion, although related to the 'delineation of scientific areas' may be found in Zitt (2006); he states that fields may be defined at various levels (e.g., institutional setting of academic actors; shared topics and possibly shared journals; shared terminology; close connections of collaboration or citation, etc.) and concludes that "... natural borders, generally speaking, are an illusion" (Zitt 2006, p. 6). In fact, a more scientific-bibliometric related methodological approach could be the analysis based on networks, as for instance in Vieira and Teixeira (2010), although this is outside the scope of the present paper.

[6]   When unsure regarding the screening, we proceeded following Katsaliaki and Mustafee (2011, p. 1434): "The two authors independently and critically reviewed all the abstracts of the (...) papers and read the full text when necessary." Notice that an alternative classification scheme could be based on the publishing journal since some journals are more oriented toward applications, while others are more methodological. As discussed below, we believe that our approach provides a more accurate measure.

[7]   The title of the MA paper by Baillie and Bollerslev is "Common stochastic trends in a system of exchange rates", while the title of the PM paper by Gilbert is "Economic theory and econometric models"

[8]   Actually, at the individual level, the term "innovator" associated to a constant hazard is somewhat misleading, and not exactly a synonym of "early adopter". In fact, an individual with constant hazard rate might well be a laggard, especially if his/her individual hazard rate is low. The parameter p is hardly interpretable in epidemiology, where the notion of "innovator" is essentially limited to the "patient zero".

[9]   We remark that that the solution is not unique. The formulae in (15) are the ones giving positive values of m, p and q with our estimated β's.

[10] Maintaining the assumption that u_t is i.i.d. normal, an alternative estimation strategy could be based on Non Linear Least Squares (NLLS). Estimates of m, p and q would be based on the following:

$$\min_{m,p,q} \sum_{t=2}^{T} \left(c_t - mp - (q-p)C_{t-1} + \frac{q}{m}C_{t-1}^2 \right)^2.$$

 The advantage of NLLS is that it provides directly the estimates of the parameters of interest (m, p and q) and the corresponding standard error, without having to resort to the delta method. The disadvantage is that convergence of the numerical optimization routines is sometimes not easy: this is partly due to the strong collinearity, and partly to the fact that the optimization problem has two solutions. In the following, we opt for OLS and the delta method.

[11] We decided to adopt slightly different symbols with respect to BF. In particular our α has opposite sign with respect to theirs.

[12] The parametrization (23) is also better suited than (24) for multivariate generalizations, as illustrated in the next subsection.

[13] Actually, Boswijk et al. (2009) propose an heteroskedastic version of the model, where $u_t = \text{diag}\left\{c_{i,t-1}^{\phi}\right\}\varepsilon_t$, with $\varepsilon_t = [\varepsilon_{1,t}, ..., \varepsilon_{n,t}] \sim iidN_n(\mathbf{0}, \mathbf{\Omega}_\varepsilon)$ and ϕ fixed to either 1/2 or 1. In this paper we only briefly discuss the heteroskedastic BFF model, since in our application suitable heteroskedasticity tests seem to accept the hypothesis of homoskedasticity.

[14] Precisely,

$$\underset{(3n^2+n)\times 3n}{\mathbf{H}_{\boldsymbol{\beta}}} = \text{diag}\{\mathbf{H}_i\} \quad , \qquad \underset{(3n^2+n)\times 1}{\mathbf{h}_{\boldsymbol{\beta}}} = \text{diag}\{h_i\}\mathbf{1}_n,$$

 with

$$\underset{(3n+1)\times 3}{\mathbf{H}_i} = \begin{bmatrix} \mathbf{u}_{n,i} \otimes [\mathbf{u}_{3,2}, \mathbf{u}_{3,3}] & \mathbf{0}_{3n,1} \\ \mathbf{0}_{1,2} & 1 \end{bmatrix} \quad , \qquad \underset{(3n+1)\times 1}{\mathbf{h}_i} = \begin{bmatrix} \mathbf{u}_{n,i} \otimes \mathbf{u}_{3,1} \\ 0 \end{bmatrix} \qquad i = 1, ..., n.$$

[15] A diagonal $\boldsymbol{\alpha}$ corresponds to $\mathbf{H}_{\boldsymbol{\alpha}} = diag\{\mathbf{u}_{n,i}\}$. As already observed, in this case ML would not correspond to equation by equation OLS, due to the correlation of the error terms. One might maximize the likelihood either by iterated SUR as illustrated in Section 3.2, or equivalently using the algorithm illustrated here.

[16] Minor modifications are needed if instead we assume heteroskedasticity of the type postulated in Boswijk et al. (2009), where $u_t = W_t\varepsilon_t$, with $W_t = \text{diag}\left\{c_{i,t-1}^{\phi}\right\}$ (ϕ is assumed to be known) and $\varepsilon_t \sim iidN(\mathbf{0}, \mathbf{\Omega}_\varepsilon)$. Notice that, premultiplying (30), left and right, by W_t^{-1}, using the properties of the vec operator, one obtains either the following:

$$W_t^{-1}Y_t = \left[\left(X_{t-1}' \otimes W_t^{-1} \right)(\boldsymbol{\beta} \otimes I_n) \right]\text{vec}(\boldsymbol{\alpha}) + u_t$$

 or the following:

$$W_t^{-1}Y_t = \left[\left(X_{t-1}' \otimes W_t^{-1} \right)(I_{3n+1} \otimes \boldsymbol{\alpha}) \right]\text{vec}(\boldsymbol{\beta}') + u_t$$

 The first equation allows to estimate $\boldsymbol{\alpha}$ by GLS when $\boldsymbol{\beta}$ and $\mathbf{\Omega}$ are known, while the second allows to estimate $\boldsymbol{\beta}$ by GLS when $\boldsymbol{\alpha}$ and $\mathbf{\Omega}$ are known. A "switching" iterative algorithm similar to Hansen (2003) is therefore possible also in this case. Of course, linear restrictions on $\text{vec}(\boldsymbol{\alpha})$ or $\text{vec}(\boldsymbol{\beta})$ are easily dealt with also in this case.

[17] Models (18) and (27) are also estimated using iterated SURE, obtaining exactly the same results.

[18] We also considered different values of k, from 4 to 20, and the results remain essentially unchanged. Regarding the number of degrees of freedom, as illustrated in Appendix A, the standard Bass model can be seen as an AR(1) with state dependent parameters, while the BF and BFF models can be seen as AR(2): therefore we considered heuristically $k - p$ degrees of freedom in the Q test, with $p = 1$ for the standard Bass model and $p = 2$ for BF and BFF models.

[19] Actually, the slope in the auxiliary regression is negative in some cases, which is exactly the opposite of BF intuition. We think that the result might reflect the neglected autocorrelation rather than heteroskedasticity: the ample swings in the residuals clearly visible in Figure 3 are misinterpreted by the test as heteroskedasticity.

[20] Alternative initializations are possible: this point is further discussed in footnote 25.

[21] For simplicity, we do not "orthogonalize" the shocks by assuming some direction for the simultaneous relationship: we believe that this is justified in this case, given the modest correlation between the residuals (16.1%). As a robustness check we also tried to orthogonalize in either direction, and to apply the "ordering invariant" method proposed in Pesaran and Shin (1998) but, as expected given the low correlation, the results are essentially unchanged. For a discussion of the simultaneous correlation, see also Appendix B.

[22] It is important to remark that, given the nonlinear dynamics implied by (29), the impulse responses will change according to the initial conditions. We also considered alternative initializations, starting in different points of the diffusion path: we observed that when the impulse is given further ahead along the diffusion path, the shape of the responses changes in a rather

intuitive way: the peak of the cumulative IRs occurs earlier, and the intensity becomes weaker. This can be explained in the light of the discussion presented in Appendix A: in the initial stages of the process, when both $C_{1,t}$ and $C_{2,t}$ are close to zero and much lower than m_1 and m_2, respectively, the processes behave as explosive AR(2), and therefore, the shocks are initially amplified; however, as $C_{1,t}$ and $C_{2,t}$ grow, the processes become less and less explosive, until eventually they start adjusting and the cumulative impact of the shock is driven down to zero. However, some characteristics of the cumulative IRs do not change, even when the initial conditions are modified: the cumulative cross impact seems to be relatively stronger from the methodological to the applied literature than vice versa.

[23]   Defining $\psi = [t^P, \bar{c}^P, \bar{C}^P]'$, starting from (9)–(11), we have the following:

$$J_{\psi.\theta} = \frac{\partial \psi}{\partial \theta'} = \begin{bmatrix} (p+q)^2 & 0 & 0 \\ 0 & 4q & 0 \\ 0 & 0 & 2q \end{bmatrix}^{-1} \begin{bmatrix} 0 & \ln p - \ln q - 1 - \frac{q}{p} & \ln p - \ln q + 1 + \frac{p}{q} \\ (p+q)^2 & 2m(p+q) & \frac{m}{q}(p+q)(q-p) \\ q-p & -m & m\frac{p}{q} \end{bmatrix}$$

The variance-covariance matrix for $\hat{\psi}$ is then obtained as the following:

$$\hat{\Sigma}_{\hat{\psi}} = \hat{J}_{\psi.\theta} \hat{\Sigma}_{\hat{\theta}} \hat{J}'_{\psi.\theta}.$$

[24]   These estimates are based on model (27) since we could not achieve convergence in model (29). The estimates based on model (18) and the same sample are almost identical.

[25]   Similarly, when $\omega = 1$, the 92 MM papers are entirely treated as methodological, whereas, when $\omega = 0.5$, only half of them (46) are treated as methodological, while the other half is treated as applied. Given the small number of MM papers, their influence on the indices is negligible, and that is why in our discussion we emphasize the role of the MA papers.

References

Bass, Frank M. 1969. A new product growth for model consumer durables. *Management Science* 15: 215–27. [CrossRef]

Bjork, Samuel, Avner Offer, and Gabriel Söderberg. 2014. Time series citation data: The nobel prize in economics. *Scientometrics* 98: 185–96. [CrossRef]

Boswijk, H. Peter, Dennis Fok, and Philip Hans Franses. 2009. *A New Multivariate Product Growth Model*. Technical Report, Econometrics Discussion Paper 2009/07. Amsterdam: University of Amsterdam.

Boswijk, H. Peter, and Philip Hans Franses. 2005. On the econometrics of the bass diffusion model. *Journal of Business & Economic Statistics* 23: 255–68. [CrossRef]

Boswijk, H. Peter, Philip Hans Franses, and Dick van Dijk. 2010. Cointegration in a historical perspective. *Journal of Econometrics* 158: 156–59. [CrossRef]

Chandrasekaran, Deepa, and Gerard J. Tellis. 2018. *A Summary and Review of New Product Diffusion Models and Key Findings*. Cheltenham: Edward Elgar Publishing, chapter 14, pp. 291–312. [CrossRef]

Coupé, Tom. 2003. Revealed performances: Worldwide rankings of economists and economics departments, 1990–2000. *Journal of the European Economic Association* 1: 1309–45. [CrossRef]

Engle, Robert F., and Clive W. J. Granger. 1987. Co-integration and error correction: Representation, estimation, and testing. *Econometrica* 55: 251–71. [CrossRef]

Eryarsoy, Enes, Dursun Delen, Behrooz Davazdahemami, and Kazim Topuz. 2021. A novel diffusion-based model for estimating cases, and fatalities in epidemics: The case of covid-19. *Journal of Business Research* 124: 163–87. [CrossRef]

Fok, Dennis, and Philip Hans Franses. 2007. Modeling the diffusion of scientific publications. *Journal of Econometrics* 139: 376–90. [CrossRef]

Franses, Philip Hans. 2003. The diffusion of scientific publications: The case of econometrica, 1987. *Scientometrics* 56: 29–42. [CrossRef]

Garfield, Eugene, Morton V. Malin, and Henry Small. 1978. Citation data as science indicators. In *Toward a Metric of Science: The Advent of Science Indicators*. Edited by Yehuda Elkana, Joshua Lederberg, Robert K. Merton, Arnold Thackray and Harriet Zuckerman. New York: John Wiley & Sons, pp. 179–207.

Guseo, Renato, and Mariangela Guidolin. 2008. Cellular automata and riccati equation models for diffusion of innovations. *Statistical Methods and Applications* 17: 291–308. [CrossRef]

Hansen, Peter Reinhard. 2003. Structural changes in the cointegrated vector autoregressive model. *Journal of Econometrics* 114: 261–95. [CrossRef]

Heeler, Roger M., and Thomas P. Hustad. 1980. Problems in predicting new product growth for consumer durables. *Management Science* 26: 1007–20. [CrossRef]

Hendry, David, and Katarina Juselius. 2001. Explaining cointegration analysis. part ii. *The Energy Journal* 22: 75–120. [CrossRef]

Hyman, Michael R. 1988. The timeliness problem in the application of bass-type new product-growth models to durable sales forecasting. *Journal of Business Research* 16: 31–47. [CrossRef]

Johansen, Søren. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12: 231–54. [CrossRef]

Johansen, Søren. 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica* 59: 1551–80. [CrossRef]

Johansen, Søren. 1992a. Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics* 52: 389–402. [CrossRef]

Johansen, Søren. 1992b. Determination of cointegration rank in the presence of a linear trend. *Oxford Bulletin of Economics and Statistics* 54: 383–97. [CrossRef]

Johansen, Søren. 1992c. Testing weak exogeneity and the order of cointegration in uk money demand data. *Journal of Policy Modeling* 14: 313–34. [CrossRef]

Johansen, Søren. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press. [CrossRef]

Johansen, Søren. 2006. Statistical analysis of hypotheses on the cointegrating relations in the i(2) model. *Journal of Econometrics* 132: 81–115. [CrossRef]

Johansen, Søren, and Katarina Juselius. 1990. Maximum likelihood estimation and inference on cointegration—With applications to the demand for money. *Oxford Bulletin of Economics and Statistics* 52: 169–210. [CrossRef]

Johansen, Søren, and Katarina Juselius. 1992. Testing structural hypotheses in a multivariate cointegration analysis of the ppp and the uip for uk. *Journal of Econometrics* 53: 211–44. [CrossRef]

Johansen, Søren, and Katarina Juselius. 1994. Identification of the long-run and the short-run structure an application to the islm model. *Journal of Econometrics* 63: 7–36. [CrossRef]

Johansen, Søren, Rocco Mosconi, and Bent Nielsen. 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. *The Econometrics Journal* 3: 216–49. [CrossRef]

Juselius, Katarina. 2006. *The Cointegrated VAR Model: Methodology and Applications*. Oxford: Oxford University Press.

Juselius, Katarina. 2021. Searching for a theory that fits the data: A personal research odyssey. *Econometrics* 9: 5. [CrossRef]

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas, and Thanasis Stengos. 1999. European economics: An analysis based on publications in the core journals. *European Economic Review* 43: 1150–68. [CrossRef]

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas, and Thanasis Stengos. 2003. Rankings of academic journals and institutions in economics. *Journal of the European Economic Association* 1: 1346–66. [CrossRef]

Katsaliaki, Korina, and Navonil Mustafee. 2011. Applications of simulation within the healthcare context. *Journal of the Operational Research Society* 62: 1431–51. [CrossRef]

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter. 1996. Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74: 119–47. [CrossRef]

Kuc-Czarnecka, Marta, Samuele Lo Piano, and Saltelli Andrea. 2020. Quantitative storytelling in the making of a composite indicator. *Social Indicators Research* 1: 775–802. [CrossRef]

Lütkepohl, Helmut. 2016. *Impulse Response Function*. London: Palgrave Macmillan UK, pp. 1–5. [CrossRef]

Min, Chao, Ying Ding, Jiang Li, Yi Bu, Lei Pei, and Jianjun Sun. 2018. Innovation or imitation: The diffusion of citations. *Journal of the Association for Information Science and Technology* 69: 1271–82. [CrossRef]

Nardo, Michela, Michaela Saisana, Andrea Saltelli, Stefano Tarantola, Anders Hoffmann, and Enrico Giovannini. 2008. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. Paris: OECD Publishing.

Pesaran, H. Hashem, and Yongcheol Shin. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* 58: 17–29. [CrossRef]

Potter, Simon M. 2000. Nonlinear impulse response functions. *Journal of Economic Dynamics and Control* 24: 1425–46. [CrossRef]

Redner, Sidney. 1998. How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems* 4: 131–34. [CrossRef]

Satoh, Daisuke. 2001. A discrete bass model and its parameter estimation. *Journal of the Operations Research Society of Japan* 44: 1–18. [CrossRef]

Srinivasan, V., and Charlotte H. Mason. 1986. Nonlinear least squares estimation of new product diffusion models. *Marketing Science* 5: 169–78. [CrossRef]

Stigler, Stephen M. 1994. Citation patterns in the journals of statistics and probability. *Statistical Science* 9: 94–108. [CrossRef]

Stokes, Donald E. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington: Brookings Institution Press.

Van den Bulte, Christophe, and Gary L. Lilien. 1997. Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Science* 16: 338–53. [CrossRef]

Vieira, Pedro, and Aurora Teixeira. 2010. Are finance, management, and marketing autonomous fields of scientific research? An analysis based on journal citations. *Scientometrics* 85: 627–46. [CrossRef]

Zitt, Michel. 2006. Scientometric indicators: A few challenges. data mine-clearing; knowledge flows measurements; diversity issues. Invited plenary talk. Paper presented at the International Workshop on Webometrics, Informetrics and Scientometrics & Seventh COLLNET Meeting, Nancy, France, December 12–15.