Online publishing and citation success in the business and economic history of Spain, 1997-2011

By

Bernardo Báñez-Lazo*+
Rasol Eskandari**
John Goddard*

*Bangor Business School
** Salford Business School
+ corresponding author

May, 2013
Online publishing and citation success in the business and economic history of Spain, 1997-2011

Abstract

This paper examines the determinants of citation success among authors who have published on the economic and business history of Spain. It departs from the dominant cross section approach to the quantitative assessment of citation success by enabling a 15-year time series analysis of peer-reviewed Spanish and Latin American outlets. Moreover, it considers working papers published online and assesses the role of Spanish as a medium to communicate with an international audience. Our results suggest a high concentration of publications and citations in a small number of authors (including non-residents), the number of years since publication and the importance of international outlets in citation success. Dissemination of online publications was not statistically significant in this sample.

Keywords: knowledge diffusion, electronic publishing, citation indexes, bibliometrics (publication scores), impact, Spain

JEL codes: A11, N0, N8, M4, O31
Online publishing and citation success in the business and economic history of Spain, 1997-2011

1 Introduction

The number of citations of an author’s work by other academic actors has been widely adopted as a measure of academic impact. This notion has led to a notable increase in the use of quantitative methods such as journal rankings and impact factors to ascertain not only the quality of an individual’s contribution but also the quality of academic publications, to the extent that in some circles quantitative measures to assess research quality now determine job promotion, university reputation, and even research project funding. Although measuring citation success is relatively easy, its links with quality of research are doubtful as a higher success rate could be determined by factors other than its added value (Di Vaio et al., 2012). For instance, empirical studies tend to attract more citation than theoretical contributions (Johnston et al., 2012). There is also evidence to question the motivation of those making citations to the extent that, if true, ‘the phenomenon of citation would lose its role as a reliable measure of impact’ (Bormann and Daniel, 2008) and could even be considered a futile exercise (Chang and McAlee, 2012; Crespo et al., 2011; Hoepner and Unerman, 2012; Hussain, 2012; Johnston et al., 2012; Lozano, 2012; Vanclay, 2012). Links between citation success and journal ranking have also been questioned on the grounds that quantification strangles specialists fields and inhibits innovation (Editorial, 2009; Wilson, 2012), as high impact contributions are not the exclusive remit of high ranking, well established outlets.

Arguments about citation success can be particularly poignant to knowledge areas such as History where the diversity of topics and emphasises on documentary evidence (rather than agenda setting) result in most outlets having low citation impact scores and more so, for academic production outside of Anglo-Saxon countries. Yet some of these arguments are largely based on anecdotal evidence. Little is known of trends and directions of citations in the broad fields of economic and business history and, as noted by Baten and Muschalli (2012), about the scholars who are representing it.

Alongside international studies, there has been an interest in bibliometric research exploring Spanish scientific production since at least 1992.¹ These studies include the pioneering contribution by Lafuente Félez and Oro (1992), which was quickly followed by others such as those by Oriol Amat (Amat et al., 1998; Amat and Oliveras Sobrevias, 1999, 2001; Amat et al., 2001), FUNCAS (1999), Boyns and Carmona (2002), Tedde de Lorca (2004), Tirado Fabregat and Pons Novella (2006), and more recently Buela-Casal et al. (2011); and Crespo et al. (2011); Gutierrez-Hidalgo and Baños-Sánchez Matamoros (2010). In this body of work there is a clear interest in economic, business and accounting research but seldom has any attention been given to business history while accounting and economic history are dealt in isolation and appear as distinct subject areas.

¹ Sardá-Dexeus, J., 1947. Spanish prices in the nineteenth century. Quarterly Journal of Economics 62, 143-159, is widely considered as the first Spanish contribution published in a top international journal and more so with an article in economic history. However and according to Bagues, M., 2012. Publicaciones en Economía: ¿quién, dónde, cómo y para qué?. Nada es gratis. Un blog de economía casi siempre en español., Salvador Barberá was the first Spanish resident author to publish in a top five ranked journal by ISI Web of Science in 1977. Regardless of this debate, the 25 years since 1987 saw an increase to an accumulated 180 contributions in Web of Science.
While there has been no comprehensive study of the broad but related areas of accounting, business and economic history of Spain, there has been a significant increase in the number of working paper, pre-publication and actual publication items that appear online. Anderson et al. (2001) already compared the performance of printed and online articles from the early days of the Internet (i.e. 1997 to 1999). Presumably citation and impact factor patterns have changed since then. Evans (2008) argues that as more journal articles were digitalised and became available online, references tended to be more recent, and more of those citations were made to fewer journals and articles. However, evidence in Lozano (2012) claims that the best (i.e., most cited) work now comes from increasingly diverse sources, irrespective of the journals’ impact factor. Hence there is room to further explore trends and directions of the business and economic history of Spain but also the role of online publications in citation patterns.

What follows sheds a first light regarding the importance of citation success in Spanish business and economic history circles and whether there is enough evidence to suggest that online publication results in higher citation success (see further Bátiz-Lazo and Eskandari, 2013). The remainder of the article maps as follows the next two sections describe the dataset and regression analysis. Here it is noted that the unit of analysis is Spain rather than the production by Spanish scholars. The fourth section offers comments on inferential statistics. The fifth and final section concludes.

2 Dataset

An initial hurdle to explore whether research published online had an impact on citation success was establishing the boundaries of contributions to the economic, business and accounting history of Spain. All potential objective measures (such as JEL codes) suffered from shortcomings as they were asked to combine overlapping knowledge areas with a specific geography. Instead the research adopted an inductive approach to data collection. It started by identifying relevant items distributed by New Economic Papers in Business, Economic and Financial History (nep-his, http://nep.repec.org) between 1997 and 2011. Nep-his is a weekly report that was launched in May 1997 as part of a digital library called Research Papers in Economics (RePEc, http://repec.org) (Bátiz-Lazo and Krichel, 2012). Data collection started in October 2011. By then nep-his had grown to over 6,500 subscribers and was effectively positioned as the premier outlet to distribute, search and recall recent additions to the broad literatures of accounting, business and economic history. This weekly report had the advantage of having being edited by the same person throughout the analysis period. A search of Spain or Spanish2 in the title, abstract or keyword was the main criteria to determine the boundaries of accounting, business and economic history of Spain. The selection criteria also considered studies dealing with events prior to the 1700s and formation of Spain as a nation state as well as items dealing with the colonial period. To be included part of the sample the latter were required to have a clear reference to the study of some form of interaction between the colonies and activities within the metropolis. Similar search criteria in the title, abstract or keyword were then applied to identify items in a second digital library, namely the Social Science Research Network (SSRN.com). Together these sources resulted in 301 items online.

2 Searches terms also included España, español, española as well as specific regions (such as Castilla, Castile, Cataluña, Catalunya and Catalonia). But the latter didn’t result in any significant increase in the number of items.
The next step was to identify relevant peer-reviewed outlets. This set built upon the 14 outlets on economic history used by Di Vaio et al. (2012)\(^3\), of which nine were found to carry items that met our selection criteria. The list of 14 outlets was used as a prompt in a survey of chief editors of accounting and business history journals. They then identified 38 potential outlets of which 27 carried items that met our selection criteria. In tandem and with the aim of eliciting suggestions of other potential relevant outlets, the list of 14 outlets was posted to the main mailing lists of economic and business history in Spain\(^4\) and Latin America\(^5\) as well as personal communications (email) to economic history societies in Latin America (namely Argentina, Brazil, Colombia, Mexico, and Uruguay) and the Academy of Accounting Historians. These mailings resulted in seven innovations, which included two other digital libraries\(^6\) and from which 18 peer-reviewed outlets were found to have published relevant research. As a result, 864 items were identified within 53 peer-reviewed outlets, of which only 56 items overlapped with online working papers (49 sourced from nep-his and 7 from SSRN). In total the dataset encompassed 1,109 items: 245 online working papers without a journal paper and 864 journal papers of which 56 had a matching online working paper.

This dataset compares handsomely with other systematic studies of citation success. For instance, the study of economic history by Di Vaio et al.(2012) encompassed 657 citations from 217 research articles published in 2007 within 14 international peer-reviewed outlets with general-interest in economic history. They collected information for 450 authors and sourced their citation data from the survey of major economics journals by Kalaitzidakis et al. (2010).

### 3 Empirical Estimates

An ‘ad hoc’ model of citation success assumed it would be a function of three elements, namely author characteristics, outlet characteristics and time lag or the number of years between publication and data collection. The dependent variable in all estimations of a tobit regression is the natural logarithm of the number of citations, where observations with zero citations were treated as left-censored. See Table 1, where t-statistics are reported in parentheses and numbers 1 to 9 denote alternative estimates of the model.

In ascertaining the values in Table 1, Harzing’s ‘Publish or Perish’ software provided a measure for the number of citations (Harzing, 2010). Independent variables were extracted from individual records and included the number of years since initial publication (age), the number of authors (authors), percentage of female authors (female), percentage of authors resident outside of Spain (foreign).

These were followed by six measures of academic impact, namely a dichotomous variable identified whether the item was published in an outlet considered by Thomson’s Social Science Citation Index (SSCI?) and if so, the index’s value (SSCI-rank). A dichotomous variable identified whether the item was published in an outlet considered by Elsevier’s Scopus database (Scopus?) and if so, the index’s value (Scopus-rank). Two

---

\(^3\) The journals were: Annales: Histoire, Sciences Sociales; Australian Economic History Review; Cliometrica: Journal of Historical Economics and Econometric History; Economic History Review; European Review of Economic History; Explorations in Economic History; Indian Economic and Social History Review; Irish Economic and Social History; JahrbuchfürWirtschaftsgeschichte; Journal of Economic History; Revista de Historia Económica /Journal of Iberian and Latin American Economic History; Rivista di Storia Economica; Scandinavian Economic History Review.

\(^4\) aeh@listserv.rediris.es; hist_empresa@listserv.rediris.es

\(^5\) h-mexico@servidor.unam.mx; and http://www.h-net.org/~latam/

alternative and ‘ad hoc’ measures of output quality were manufactured. One labelled ‘Anglo Saxon’ averaged the ranking for each of the 53 outlets given by the journal ranking of the Association of Business Schools in the UK\(^7\) together with the ranking of the Australian Research Council.\(^8\) The second construct was labelled ‘Spanish’ and it resulted from averaging the rankings by IN-RECS at Granada University\(^9\) and Carhus Plus by the Generalitat de Catalunya.\(^10\) The chief reason to combine these indexes was that both Scopus and the Social Science Citation Index exclude outlets we have described as ‘domestic’ (most of which published in and where manual inspection suggested a large number of items had been published. ‘Domestic’ outlets mostly published in Spanish and thus, enabled an opportunity to assess the potential of impact of non-English communications in citation success. Admittedly, however, the assessment of quality is a categorical variable and performing an arithmetic operation is not defined. This operation also reduced degrees of freedom of inferential statistics. But manual inspection suggested there was little agreement between the four rankings, thus potentially increasing variance, while, at the same time, there was a degree of harmony in the ranking of outlets when grouped through our innovations.

Independent variables also identified whether at least one of the authors was registered in RePEc digital library (REPEC-reg?) and also whether it was within the top five per cent of citations in that database (top 5%). Dichotomous variables identified whether there was an online working paper in RePEc (REPECwp), SSRN (SSRNwp) or both (wp?) as well as whether the paper was written in Spanish (language). Estimates also included the difference in years between publication in a peer reviewed outlet and the online working paper (wp-gap).

In Table 1 equation (1) is the preferred model, with all coefficients significant at the 0.1 level or lower. Equation (2) added the proportion on foreign authors (insignificant). Equations (3) and (4) substituted ‘Anglo ranking’ and ‘Spanish ranking’ respectively for the Thomson’s SSCI ranking (both were not significant). Equation (5) substituted the Scopus ranking for the Thomson’s SSCI ranking, both were significant, but Thomson’s SSCI ranking was more significant. Equation (6) substituted the REPEC-registered dichotomous variable for the top 5% dummy, both of which were significant but the ‘top 5%’ was more significant. Equation (7) includes the dichotomous variables for RePEc and SSRN working papers, both were not significant. Equation (8) included the dichotomous variable for previous publication as a working paper, and the gap between working paper and article publication but neither was significant. Equation (9) included the language dummy which was also not significant.

There were two likely sources of bias in the regression estimates: self-citation and the age of the paper. With regards to the first, Harzing’s ‘Publish or Perish’ only offers a raw estimate of citations because she considers that excluding them is ‘normally not worthwhile’ (Harzing, 2010). Estimates by Aksnes (2006) suggest a strong positive correlation between the number of authors and self-citations but that the percentage of self-citation reduces considerably and monotonically for assessment periods of three years or longer. These estimates also observe strong variations of self-citation across disciplines. However, they record the greatest number of self-citation amongst least cited papers from others whereas in a follow up study, Fowler and Aksnes (2007) find strong evidence that self-citation can increase citation from others. Hence the causality of self-citation in citation success is complex and available evidence inconclusive.

| Table 1: Tobit estimates of citation success, 1997-2011 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | (1)             | (2)             | (3)             | (4)             | (5)             | (6)             | (7)             | (8)             | (9)             |
| age             | 0.770           | 0.771           | 0.695           | 0.691           | 0.691           | 0.747           | 0.783           | 0.772           | 0.770           |
| age-square      | (9.42)*         | (9.43)*         | (8.52)*         | (8.38)*         | (8.56)*         | (9.17)*         | (9.51)*         | (9.44)*         | (9.40)*         |
| authors         | 0.442           | 0.442           | 0.467           | 0.469           | 0.464           | 0.426           | 0.414           | 0.446           | 0.442           |
| female          | (-0.390)        | (-0.407)        | (-0.349)        | (-0.341)        | (-0.339)        | (-0.388)        | (-0.365)        | (-0.389)        | (-0.390)        |
| foreign         | -0.249          | -0.249          | (-0.94)         | (-0.94)         | (-0.94)         | (-0.94)         | (-0.94)         | (-0.94)         | (-0.94)         |
| SSCI?           | -0.491          | -0.489          | -0.469          | -0.436          | -0.471          | -0.491          | -0.491          | -0.491          | -0.491          |
| SSCI -rank      | 2.192           | 2.211           | 2.190           | 2.233           | 2.132           | 2.189           | 2.189           | 2.189           | 2.189           |
| Anglo-rank      | (5.68)*         | (5.72)*         | (5.62)*         | (5.76)*         | (5.14)*         | (5.53)*         | 0.038           | 0.038           | (0.44)          |
| Spanish-rank    | 0.032           | 0.032           | 0.032           | 0.032           | 0.032           | 0.032           | 0.032           | 0.032           | 0.032           |
| Scopus?         | 0.222           | 0.222           | 0.222           | 0.222           | 0.222           | 0.222           | 0.222           | 0.222           | 0.222           |
| top 5%          | 1.698           | 1.745           | 2.070           | 2.079           | 1.990           | 1.647           | 1.651           | 1.696           | 1.696           |
| REPEC-reg?      | 0.913           | 0.913           | 0.913           | 0.913           | 0.913           | 0.913           | 0.913           | 0.913           | 0.913           |
| REPECwp         | 0.232           | 0.232           | 0.232           | 0.232           | 0.232           | 0.232           | 0.232           | 0.232           | 0.232           |
| SSRNwp          | 0.472           | 0.472           | 0.472           | 0.472           | 0.472           | 0.472           | 0.472           | 0.472           | 0.472           |
| wp?             | -0.414          | -0.414          | -0.414          | -0.414          | -0.414          | -0.414          | -0.414          | -0.414          | -0.414          |
| wp-gap          | 0.380           | 0.380           | 0.380           | 0.380           | 0.380           | 0.380           | 0.380           | 0.380           | 0.380           |
| language        | 0.005           | 0.005           | 0.005           | 0.005           | 0.005           | 0.005           | 0.005           | 0.005           | 0.005           |

* = t-student significant at the 1% level or lower.

4 Discussion

There were two likely sources of bias in the regression estimates: self-citation and the age of the paper. With regards to the first, Harzing’s ‘Publish or Perish’ only offers a raw estimate of citations because she considers that excluding them is ‘normally not worthwhile’ (Harzing, 2010). Estimates by Aksnes (2006) suggest a strong positive correlation between the number of authors and self-citations but that the percentage of self-citation reduces considerably and monotonically for assessment periods of three years or longer. These estimates also observe strong variations of self-citation across disciplines. However, they record the greatest number of self-citation amongst least cited papers from others whereas in a follow up study, Fowler and Aksnes (2007) find strong evidence that self-citation can increase citation from others. Hence the causality of self-citation in citation success is complex and available evidence inconclusive.
On the other hand, close examination of the long-term behaviour for aggregate citation success, peer-reviewed, and online papers suggested that citation success reached its highest point seven years after publication (see further Bátiz-Lazo and Eskandari, 2013). In this regard it is important to note that 19 of the top 25 cited items were published before 2005, and only six out of the 25 top cited items were published after 2005. This suggest that the longer a paper has been around, the greater the possibility of having an impact and hence the significance of this variable in the regression. Indeed, the average age in the sample was 7 years old (st. dev. of 4.1, mode and median equal to 1). There was an attempt to control this effect by introducing a proxy to the age of the researcher but the information was not readily available.

Manual examination also suggested there was an increase in the number of online working papers since 2006. But there was only a 5% chance to find an online working paper for the average paper in the sample. The recent rise in online items helps to explain the poor overlap between online and peer-reviewed observed before (i.e. 56 out of the 1,109 items) as well as the poor performance of these variables in equations 7 and 8.

Of the 1,109 items, 442 (40%) were written in other medium (mainly English). In other words, most papers in the sample were written in Spanish (667 items or 60%) by resident authors. Yet these were mainly made out of papers with zero citations. Indeed, that was the case for 579 papers (52% of all papers in the sample). Here thus lies the explanation for the lack of statistical significance of the language variable in equation 9.

The constant was negative yet strongly significant for all the nine formulations of the model. The reason for this lies behind the cumulative frequency for 9 or less citation from others was 91% of all papers in the sample. These values also quantify the expectation that the emphasis on documentary evidence and single author monographs would result in low citation values. However, less than 10% of the papers in the sample accrued 10 or more citations. In other words, there is evidence to suggest that a large portion of the production of systematic work around the accounting, business and economic history of Spain has no impact at all.

There were 870 individual authors in the sample of which 193 were females (22% of all authors) and 128 were non-residents (15%). Non-residents were mainly based in the USA (30 authors or 23% of non-residents) and the United Kingdom (30 authors or 23% of non-residents). But as a geographical block, the European Union housed 77 non-resident authors (60% of non-residents) whereas the Americas housed 48 (38%) and other locations 3 (2%). On average an individual contributed to 1.73 papers.11 There were 786 papers written by a single author and these represented 70% of the sample, while 323 papers were authored by two or more persons and represented only 30% of the sample. Yet in spite of this apparent lower proportion of multi-authored papers, the strong showing of the number of authors variable across all specifications suggest that collaboration can have a positive impact. As noted above, gender had no impact in any of the specifications.

Only a small number of authors were registered in RePEc as only 137 papers (12% of the total) had contributors who were amongst the 33,892 persons registered in RePEc. But consistently with the rankings of the latter, 52 papers (5%) resulted from contributions by RePEc’s top authors.12 Results in equations 5 and 6

---

11 Note that here there is no distinction between contributions to single and multiple authored items, all having the same weight.

12 See further http://ideas.repec.org/top/top.person.all.html.
showed statistical significance for, respectively, dichotomous variables relating to one of the authors being register in RePEc and for those within the top five percent of RePEc’s authors. The causality here is not clear because one would expect a highly cited author to be within the top of RePEc but at the same time, an online presence through RePEc can help showcase people’s work and thus contribute to citation success.

Only 19 of the 53 outlets in the sample were indexed by the Social Science Citation Index (averaging 1.02 impact factor) and 20 outlets were found in Scopus (averaging 0.22). The highest ranked outlet Social Science Citation Index in the sample was Economics & Human Biology (2.43) followed by Explorations in Economic History (1.22). Meanwhile in Scopus the highest ranked outlet was the Journal of Latin American Studies (0.49), followed by the Economic History Review (0.43). Both Scopus and the Social Science Citation Index exclude outlets we have described as ‘domestic’. Hence, two alternative and ‘ad hoc’ measures of output quality were manufactured. However and as shown in formulations 2 to 5 in Table 1, neither of these two measures were found to have statistical significance when ascertaining citation success. Thus suggesting the importance of publishing in international outlets to reach a wide audience.

5 Conclusions

Measuring the quality of research through citation success or impact factors are subject to several well-known weaknesses (Anderson et al., 2001; Vanclay, 2012). But in spite of limitations and trickery, this paper has ascertained the boundaries of research in business and economic history of Spain through citation success. There are clear publication strategies emerging from these results, such as the advantages of targeting international, multidisciplinary outlets through journal papers.

In spite of a large number of contributors, the production of knowledge in accounting, business and economic history of Spain is concentrated in a small number of individuals suggesting that less than a third of the sample is research active. This is an area strongly dominated by male authors and the evidence suggests that can have a significant positive impact. Clearly, more could be done to broaden international co-operation and gender diversity to explore the history of Spain.

Citation success in this discipline is not immediate as it can take up to seven years to reach its peak. And germane to this article is the role of online publications on citation success. However and rather disappointingly, online publishing seems to be a too recent a phenomenon to have had significant impact on citation success. This is not to negate its importance, rather that a different formulation is required to that adopted in this study. Manual inspection suggested that the production of both peer-reviewed and online items has increased since 2006. This might be associated with a reduction in the cost of publishing online. But its effect on citation success remains unsolved.

Finally, although statistical analysis gives greater weight to international outlets for citation success (which in turn seems to reinforce the perceived importance of a small number of outlets as being ‘world elite’),
a well-balanced research publications strategy should not disregard the importance of including ‘domestic’ journals. They might not be characterised by publishing leading, original research but they certainly have an audience. Moreover, in the case of the business and economic history of Spain, this audience spreads out through Latin America. In other words, for this area of knowledge, communication of research results in a medium other than English seems important for citation success and hints to the role of Spanish academic circles as intermediaries between Anglo-Saxon and Latin American research.

6 References


