

# Endogenous selection into single and coauthorships by surnames in economics and management

PRELIMINARY!!! PLEASE DO NOT DISTRIBUTE.

December 16, 2014

BY DAVID ONG<sup>1</sup>, HO FAI CHAN<sup>2</sup>, BENNO TORGLER<sup>3</sup> AND YU (ALAN) YANG<sup>4</sup>

*Many prior studies suggest that alphabetic ordering confers professional advantages to authors with earlier surname initials. However, these assume that authors are select into coauthorships, without regard to the incentives identified. We consider the alternative and develop a model of endogenous selection into single and coauthorships for economics and tested it using management as a benchmark. We predict that lower citation ranked (“ability”) authors with earlier surnames would be less desirable as coauthors, while higher ability authors with later surnames would have a lower desire to coauthor. Both are therefore more likely to single author. Furthermore, higher ability authors with earlier surnames should have more and better coauthoring options, all the more so, for authors of non-alphabetically ordered papers. Consistent with our predictions, we found citation ranks were increasing on surnames for single authored works and decreasing for coauthored in economics both absolutely and compared to management, and that the effect was even more severe for non-alphabetically ordered papers. We show that economists take into account both the likely contribution of their coauthors and their own share of the credit, when deciding to coauthor, and thus, offer an alternative explanation for prior findings. We also show that the alphabetical convention could have important consequences for research productivity.*

*JEL codes:* J01, J1, J15

*Keywords:* Alphabetic order effect, endogenous teams, contests

\*

<sup>1</sup> Email corresponding author: [dvdong@gmail.com](mailto:dvdong@gmail.com), Tel: +86-755-2603-2655, Fax: +86-755-2603-5344, Address: 708, Peking University HSBC Business School, University Town, Shenzhen, China, 518055

<sup>2</sup> Queensland University of Technology, Brisbane, Australia

<sup>3</sup> Queensland University of Technology, Brisbane, Australia

<sup>4</sup> Peking University HSBC Business School, Shenzhen, China

## Introduction

Coauthoring is increasingly prevalent in many of the sciences including economics (Hudson, 1996), marketing (Brown, Chan, & Lai, 2006), and finance (Brown, Chan, & Chen, 2011). This trend could be due to a number of reasons. Authors could be exploiting the gains from specialization in increasingly specialized fields, or hedging against the risks of rejection, or delays in review, or changing the tradeoff between quantity and quality (Hudson, 1996) etc. See Bruno (2014) for a recent review of the theories of coauthorship.

There are two main citation conventions for authorship order: alphabetical by the initial of surnames (from this point forward referred to merely as “alphabetical” order or “by surname”) and by relative contributions. Economics is among the fields that uses alphabetical ordering. 86 percent of coauthored papers in economics journals for five major journals for the last two decades (Engers, Gans, Grant, & King, 1999) and 92 percent of the top three finance journals (Brown et al., 2011) listed authors alphabetically. In contrast, only 30 percent of papers published in the major biological journals use alphabetical listings. See Waltman (2012) for a ranking by alphabetization of 25 subject categories of the mathematical, the social, and the hard sciences, as well as, the humanities.

This difference in citation convention has important consequences for researchers and research. Citation indices have generally only included the names of the first author. Second authors onwards may often be listed as “et al.” within articles (Van Praag & Van Praag, 2008). Not surprisingly, prior work finds that being first-author increases salience and attributions of credit (Nudelman & Landers, 1972). Efthyvoulou (2008) show that authors in economics with surnames with the initial “A” are significantly more likely to have abstract views and downloads than authors with surnames with the initial “Z”. Huang (2014) finds that papers of first-authors with earlier surnames get more citations. There

seems little dispute that more citations lead to higher pay (Hamermesh, Johnson, & Weisbrod, 1982; Hilmer, Hilmer, & Ransom, 2012; Moore, Newman, & Turnbull, 2001; Sauer, 1988).

More importantly, and perhaps as a consequence of these advantages, earlier surnames are promoted more quickly. In economics, faster promotions and greater likelihood of tenure (but curiously only in top 10 departments), fellowships of the Econometric Society, the Clark Medal, and the Nobel Prize for economics accrue to economists with earlier surnames. Surnames have no effect on promotion in psychology, which uses relative contributions ordering (Einav & Yariv, 2006). Efthyvoulou (2008) confirmed the higher rate of promotion for a larger sample of highly ranked research departments in US and in UK. He also found that the rate of career advancement of 1,500 chemists at British universities was decreasing on the position of authors' surname in the alphabet.

There is some evidence that authors react to these incentives. Efthyvoulou (2008) demonstrate how authors manipulate their names in order to gain precedence, using prefixes like "De" and suppressing prefixes like "Van". Van Praag and Van Praag (2008) find that higher inequality of reputation among the authors increases the probability of non-alphabetical ordering, while higher reputation of the coauthors lowers that probability.

There has been some work on endogenous selection into coauthorships by surnames. Einav and Yariv (2006) find that though the relative frequency of authors' surnames in single-authored, two-authored and three-authored papers did not differ significantly, authors with later surnames are significantly less likely to participate in four- and five-author projects. However, Einav and Yariv (2006) focused on the rate of participation as a function of surnames rather than quality of participants as a function of surnames.

In summary, the prior empirical literature seems to have largely established that alphabetic ordering confers professional advantages to authors with earlier

surnames. This literature has assumed that authors do not endogenously select into coauthorships by the incentives identified. We address this omission by considering the possibility that researchers coauthor or single author based upon the expected contributions of potential coauthors (“ability”) and their own share of the total credit.

Our key assumption is that if earlier surnamed coauthors get more of the credit for the quality of a coauthored paper, they are more likely to make a larger contribution. This leads to the following observations.

*Observation 1:* Economists with later surnames have an ex-ante (i.e., before they meet a specific coauthor) lower surplus from coauthoring than those with earlier surnames, because they can expect less of the credit for the quality of any coauthored work.

*Observation 2:* The ex-ante deficit in credit for economists with later surnames is increasing on their ability, as measured by citation ranks.

These observations have the following implications for economics.

- a) Authors of lower ability and earlier surnames will be less preferred as coauthors.
- b) Authors of higher ability and later surnames would have a weaker incentive to coauthor.
- c) Authors of the highest ability and earlier surnames would face a thicker market for potential coauthors than authors of the highest ability and later surnames.

Based on a) and b), we predict

The quality (as measured by citation ranks) of authors in coauthored papers should be decreasing on the surnames of the authors. The quality of authors in single authored papers should be increasing on the initial of the surnames of the authors. Furthermore, due to authors abstaining from coauthorships, economics

should have a relatively lower frequency of coauthored works than comparable fields like management that do not have an alphabetical convention. As a consequence of c), the effect should be stronger with lower tier journals where the market for coauthors should be thinner. Authors with earlier surnames would have higher quality coauthored works because they have more options on higher quality coauthors, and advance more quickly. This effect would be exacerbated when for authors of articles who are not ordered alphabetically.

We tested these predictions for single and two coauthored papers in the top 23 economics journals from 1900-2000, using the top 30 management journals as a benchmark. (See Table 6 in the Appendix for the full list.) Our empirical results are consistent with our predictions. We replicate the stylized facts of the prior literature of disproportionate credit to earlier surname authors discussed above, but now explain them as the possible consequences of endogenous selection into single and coauthored work.

Other than motivating the revaluation of prior results showing disproportionate advantage to authors with earlier surnames, our findings could have implications for endogenous teams with asymmetric surplus (Wuchty, Jones, & Uzzi, 2007). Our findings suggest that among the many possible incentives that people have for scholarship, e.g., intrinsic motivations, are contest incentives. See Dechenaux, Kovenock, and Sheremeta (2012) for a recent survey.

### *Theories of coauthorship conventions*

Laband and Tollison (2000) suggest that alphabetization is a form of pay compression which encourages collaboration of the form seen in industrial settings (Lazear & Oyer, 2012). Brown et al. (2011) explain the correlation between higher quality and alphabetization as being due to the greater difficulty of determining contribution due to the higher degree of effort required, or because

authors are less worried about getting credit proportional to their effort for higher quality papers, or because the authors are prominent.

To our knowledge, there is only one formal theoretical paper about surname order. Engers et al. (1999) model of effort in coauthorships proves that it is never an equilibrium for authors always to be listed in the order of relative contribution. This is a consequence of the fact of the market drawing stronger inferences about relative contributions of authors to the paper when the authors are not in alphabetical order. This asymmetry entails the second author losing more credit than the first-author gains when they appear in non-alphabetical order. They demonstrate that alphabetical ordering is inefficient and that higher effort will be elicited from authors if the relative contribution convention were adopted. Hence, this theory predicts that quality of authors should be *decreasing* on alphabetization.

However, Brown, Chan and Lai (2006) find that quality, as measured by citations in 19 leading marketing journals is positively correlated with the quality of article. Joseph, Laband and Patil (2005) illustrate through simulations of authors with stochastic quality realizations that the rate of alphabetization increases the publication hurdle, and conditional on clearing the hurdle, quality is increasing on alphabetization. They argue that inter- and intra-disciplinary differences in conventions can be explained by publication hurdles. This is due to the fact that both authors must be good to get into top journals, and that one was of a significantly lower quality than the other if the surname order was not alphabetical. They propose that the higher effort that is required by the higher hurdle lowers the likelihood of a large discrepancy at the right tail of distribution.

To our knowledge, no paper addresses selection into single and coauthored papers by surname and ability.

### Theory: the coauthor game

Our analysis is based upon the already stated assumption that economists would accept earlier surnamed coauthors getting more of the credit for the quality of the paper if they are likely to make a larger contribution. From this, Observations 1 and 2 in the introduction follow immediately. However, we can illustrate them, for the interested reader, with a simple example of how potential matches might be formed. This example can easily be generalized.

Suppose Nature draws two authors from a set of three with surnames initials  $\{A, B, C\}$  of three types of qualities  $\{H, M, L\}$ , which we think of as being based on past citations and fixed at the moment of choice. Potential matches require both authors to say  $\mathcal{Y}$ . Our assumption implies that later surnames would have more potential to match with higher types of earlier surnames, e.g.,  $B_M$  would say  $\mathcal{Y}$ =yes to match with  $A_H$ , but  $\mathcal{N}$ =no with  $A_L$ . See Table 1<sup>5</sup>.

TABLE 1: POTENTIAL MATCHES OF EACH TYPE OF AUTHORS

<b>AB</b>	<i>H</i>	<i>M</i>	<i>L</i>	<b>AC</b>	<i>H</i>	<i>M</i>	<i>L</i>	<b>BC</b>	<i>H</i>	<i>M</i>	<i>L</i>
<i>H</i>		$\mathcal{Y}\mathcal{Y}$	$\mathcal{Y}\mathcal{Y}$	<i>H</i>		$\mathcal{Y}\mathcal{Y}$	$\mathcal{Y}\mathcal{Y}$	<i>H</i>		$\mathcal{Y}\mathcal{Y}$	$\mathcal{Y}\mathcal{Y}$
<i>M</i>	$\mathcal{Y}\mathcal{N}$		$\mathcal{Y}\mathcal{Y}$	<i>M</i>	$\mathcal{Y}\mathcal{N}$		$\mathcal{Y}\mathcal{Y}$	<i>M</i>	$\mathcal{Y}\mathcal{N}$		$\mathcal{Y}\mathcal{Y}$
<i>L</i>	$\mathcal{Y}\mathcal{N}$	$\mathcal{Y}\mathcal{N}$		<i>L</i>	$\mathcal{Y}\mathcal{N}$	$\mathcal{Y}\mathcal{N}$		<i>L</i>	$\mathcal{Y}\mathcal{N}$	$\mathcal{Y}\mathcal{N}$	

Notes: AB, AC, BC are potential matches for two authors of surnames A, B or C. *H, M, L* are quality types.  $\mathcal{Y}$ =yes to a match.  $\mathcal{N}$ =no. Realized matches require  $\mathcal{Y}\mathcal{Y}$ .

Table 2 summarizes the incentive compatible potential matches after low ability As ( $A_L$ ) and high ability Cs ( $C_H$ ) have selected to single authorships. It is, thus,

<sup>5</sup> Note that these are potential matches only. A model of realized matches would require knowledge of the distribution of names and of capacities of authors. We exclude symmetric matches since they would not change the average quality of single and coauthors.

immediate that the average quality of single authors is increasing on the surname initial. Because of this, the average quality of the authors who could coauthor will be decreasing on their surname initials. The average quality of As who could coauthor is  $H$ . The average quality of the Bs who could coauthor is  $M$ . The average quality of the Cs who could coauthor is  $L$ .

TABLE 2: AVERAGE QUALITY OF INCENTIVE COMPATIBLE COAUTHORS

Average quality by surname	Incentive compatible coauthors			
$A_H$	$(A_H, B_M)$	$(A_H, B_L)$	$(A_H, C_M)$	$(A_H, C_L)$
	$(A_M, B_L)$	$(A_M, C_L)$	$A_L$ single authors	
$B_M$	$(B_H, C_M)$	$(B_H, C_L)$		
	$(B_M, C_L)$	$(B_M, A_H)$		
	$(B_L, A_H)$	$(B_L, A_M)$	$C_H$ single authors	
$C_L$	$(C_M, A_H)$	$(C_M, B_H)$		
	$(C_L, A_H)$	$(C_L, A_M)$	$(C_L, B_M)$	$(C_L, B_H)$

*Notes:* The average quality of incentive compatible authors is decreasing on the coauthors' surnames because authors with later surnames and higher ability will not want to coauthor with  $A_L$ , while  $C_H$  will not want to coauthors with others with earlier surnames and lower ability.

While Table 2 exhibits the average quality of potential coauthors with a given surname as a function of surname initials, Table 3 exhibits the pairs in which a certain surname is likely to get more credit for a coauthored work. The pattern of earlier surnames getting more credit is similar to what would be predicted by other theories about coauthors discussed in the introduction. However, we illustrate here that the first author in fact deserves the extra credit because these coauthorships were formed in the anticipation of this unequal distribution of credit.

Our key result is authors with earlier surnames have more options of higher quality. The  $A_H$ s have four options, twice as many as the  $B_H$ s both in terms of numbers and quality. The  $A_M$ s have twice the number of options as  $B_M$ . As mentioned, the  $C_H$ s have abandoned coauthoring. The  $C_M$ s could coauthor, but



there are no more letters which would want to give  $C_M$  more credit because  $C_M$  contributes more. Similarly with  $C_L$  and  $B_L$ .

TABLE 3: COAUTHOR OPTIONS BY SURNAME AND ABILITY

Surname	Preferred by				# of Coauthor options
A	$(A_H, B_M)$	$(A_H, B_L)$	$(A_H, C_M)$	$(A_H, C_L)$	6
	$(A_M, B_L)$	$(A_M, C_L)$			
B	$(B_H, C_M)$	$(B_H, C_L)$			3
	$(B_M, C_L)$				
C					0

*Notes:* Authors with later surnames have fewer options in terms of coauthors.

Thus, if coauthorships are not fully discounted or if coauthored works exploit synergies, authors with surnames of A would enjoy a disproportionate advantage against Bs, who would in turn enjoy an advantage against Cs in the market for reputation.

From this coauthor game, we make the following predictions for our empirical analysis with respect to the economics, which has an alphabetical convention, and management literature, which does not.

*Predictions:*

- P1. Citations of authors of coauthored papers will be decreasing on surnames in economics relative to management, and possibly absolutely as well.
- P2. Citations of authors of single authored papers will be increasing on surnames in economics relative to management.
- P3. The effect found in P1 and P2 will be stronger for lower quality journals.
- P4. Economics will have a higher probability of single authorships than management.

- P5. Authors with earlier surnames will have more and higher quality coauthored works and will be promoted more quickly. (This has already been found in the literature. We discuss it after our empirical results.)
- P6. First authors of non-alphabetical ordered articles will have a lower citation rank, but a similar trend due to the fact that their authors must be of lower quality in order to tolerate losing more credit when the first author is non-alphabetically ordered (Engers et al., 1999).

### **Data results**

We now test our theoretical predictions P1 - P4 on citations of papers from 23 Economics journals and 30 Management journals obtained from Thomson Reuters Web of Science<sup>6</sup> on 7<sup>th</sup> November, 2012 (for top 23 Economics journals) and 7<sup>th</sup> March, 2013 (for top 30 Management journals). These publication and citation records are from 1900-2000 and include 43,013 economics and 52,765 management publications. 35 percent (or 15,110) of economics and 43 percent (or 22,871) of management publications are co-authored. 83.41 percent of economics and 47.82 percent management co-authored articles are alphabetically ordered. We exclude any post-publication activities (e.g. reply, corrections etc.), conference paper and book reviews. Information on articles from economics, finance and management journals are recorded (if available) up until 2011 December and 2012 December, respectively (due to date of data collection). Table 4 gives a breakdown by the number of authors for each field.

<sup>6</sup> See <http://wokinfo.com/>.

TABLE 4: DESCRIPTIVE STATISTICS

# of Auth ors	Variables	Management					Economics				
		N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
1	Average Citations	29892	0.71	2.27	0.00	79.50	27903	1.12	2.67	0.00	71.10
	Alphabetical Ordering	29892	1.00	0.00	1.00	1.00	27903	1.00	0.00	1.00	1.00
2	Average Citations	15829	1.72	3.50	0.00	119.50	12048	2.34	4.42	0.00	175.10
	Alphabetical Ordering	15829	0.58	0.49	0.00	1.00	12048	0.86	0.35	0.00	1.00
3	Average Citations	5336	2.21	4.11	0.00	143.90	2653	2.84	4.66	0.00	65.10
	Alphabetical Ordering	5336	0.28	0.45	0.00	1.00	2653	0.77	0.42	0.00	1.00
4	Average Citations	1262	2.38	3.85	0.00	51.60	338	3.90	8.99	0.00	118.60
	Alphabetical Ordering	1262	0.14	0.34	0.00	1.00	338	0.64	0.48	0.00	1.00
≥5	Average Citations	444	2.29	3.31	0.00	27.10	71	1.80	3.15	0.00	20.70
	Alphabetical Ordering	444	0.10	0.30	0.00	1.00	71	0.37	0.49	0.00	1.00

We follow Huang (2014) in using citation rank instead of citation. As he points out, raw citations follows a power law distribution (Gupta, Campanha, & Pesce, 2005; Redner, 2005), which could create econometric problems using OLS. Frequently cited papers may drive estimates, even if they are few. For each publication year, define the citation rank of a paper with  $c$  times of 10 years citations as

$$Citation\ rank = \frac{N_{citation < c} + 1}{N_{total}}$$

$N_{citation < c}$  is the number of papers with citations less than  $c$ .  $N_{total}$  is the total number of papers (both economics and management) published in the same year. This measure ranges from nearly zero to one hundred. For a particular paper, it can be understood as the proportion of articles published in the same year that have fewer citations, in percentage terms.

$$\begin{aligned}
& \textit{Citation rank}_{aij} \\
& = \beta_0 + \beta_1 \textit{Econ} + \beta_2 \textit{Initial}_a + \beta_3 \textit{Initial}_a * \textit{Econ} + \theta X_{aij} + \gamma Z_{jt} + \varepsilon
\end{aligned}$$

- $\textit{Citation}_{aij}$  = citations for author with surname  $a$  in article  $i$  published in year  $j$ .
- $\textit{Econ}$  = economics
- $\textit{initial}_a$  = a number from 1 to 26 representing the initial of the surname  $a$ .
- $X_{aij}$  = control variables for author with surname  $a$  in article  $i$  published in year  $j$ .
- $Z_{jt}$  = interaction between year  $j$  dummies and  $\textit{Econ}$  dummy.

It is important to note that though citations are both on the left hand side to be explained and implicitly on the right hand side as the basis of the choice of researchers in choosing coauthors, we assume that researchers take theirs and their potential coauthors past citations (and estimated abilities) as fixed when they choose coauthors. This avoids simultaneity and endogeneity issues.

Figure 1 illustrates the average citations for authors of single and two authored works for both management and economics. There is a slight decreasing trend for coauthored management papers, which is the benchmark. Citations of management single authored papers are nearly parallel with management coauthored. The citations for authors of coauthored economics papers are decreasing both absolutely and with respect to that trend. Citations for single authored economics papers are increasing both absolutely and with respect to authors of management coauthored papers. Table 5- Table 7 indicate that these trends for economics are in fact significant.

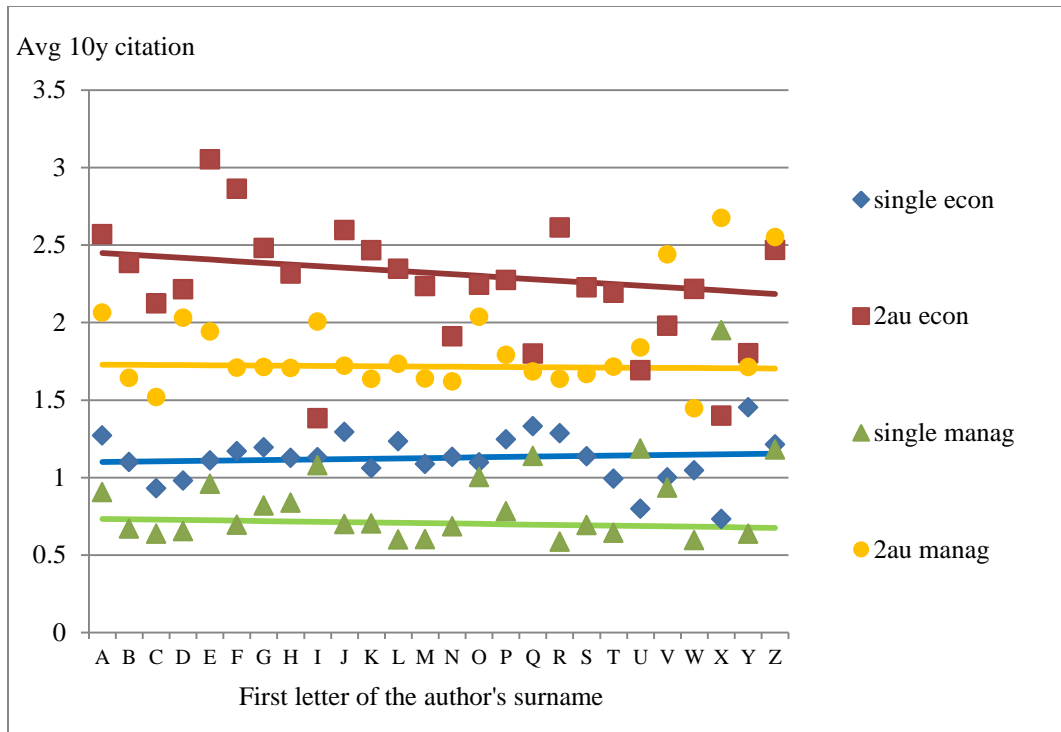


FIGURE 1: AVERAGE FOR 10 YEARS CITATIONS OF AUTHORS OF SINGLE AND TWO-AUTHORED WORKS. 113,549 AUTHOR-ARTICLE OBSERVATIONS FROM 1900 TO 2000.

*Notes:*

1. single econ=average citations for single authored economics papers.
2. single management=average citations for single authored economics papers.
3. 2au econ =average citations for two coauthored economics papers.
4. 2au manag =average citations for two coauthored management papers.

Figure 3 and Figure 4 in the Appendix also show that the rate of single authorships decreased dramatically after 1960. We, therefore, perform robustness checks for the regressions by restricting the samples from 1960 onwards for Table 5-Table 7

Table 5 presents the regression of the average rank of single authored papers with authors grouped by the initial of surnames for the top 23 economics with the top 30 management journals a benchmark. The average rank of papers can then be understood as the average rank of each letter group of authors. Authors of single

authored management papers with a surname of A have an average rank of 30.184 percentile among all management and economics articles. Authors of single authored economics papers with a surname of A are ranked 10.148 percentage points higher. The insignificant coefficient of -0.043 suggests that management authors' ranks are not affected by their surnames. Economists ranks increases by 0.125 percent for every increase in the initial of their surnames. Thus, for every 1,000 economists of single authored papers, one economist is “mis-ranked” with respect to economists with an immediately earlier surname. However, this grows to  $0.125 \times 25 = 31.25$  for those with surnames starting with Z with respect to those with surnames starting with A.

For every year that either an economist or a management academic stays in the field, 0.353 percentage points is added to their rank. For every citation of an author over their lifetime, their rank also goes up by 0.019 percentage points. Every page of the paper adds 2.384 percentage points to the rank the authors' rank. The significance of these findings does not change if we look at the sample after 1960, when coauthorships were much more likely.

TABLE 5: REGRESSION OF CITATION RANK FOR ECONOMICS AND MANAGEMENT SINGLE AUTHORED PAPERS

Dependent variable	Citation rank (0 – 100)					
	1900-2000			1960-2000		
	(1)	(2)	(3)	(4)	(5)	(6)
Initials	-0.043 (0.028)	-0.038 (0.025)	-0.041* (0.024)	-0.052* (0.031)	-0.047* (0.026)	-0.054** (0.025)
<b>Initials*econ</b>	<b>0.125***</b> (0.041)	<b>0.067*</b> (0.035)	<b>0.066*</b> (0.034)	<b>0.070</b> (0.044)	<b>0.101***</b> (0.039)	<b>0.103***</b> (0.038)
Econ	10.148*** (0.547)	2.714* (1.503)	1.642 (1.433)	17.679*** (0.595)	-0.037 (1.512)	-0.586 (1.444)
Academic age		0.353*** (0.034)	0.055 (0.034)		0.474*** (0.038)	0.125*** (0.036)
Academic age <sup>2</sup>		-0.010*** (0.001)	-0.003*** (0.001)		-0.012*** (0.001)	-0.005*** (0.001)
Pages		2.384*** (0.179)	2.254*** (0.181)		2.582*** (0.194)	2.422*** (0.195)

Pages <sup>2</sup>		-0.026***	-0.025***		-0.027***	-0.026***
		(0.005)	(0.005)		(0.005)	(0.005)
Lifetime citation			0.019***			0.017***
			(0.001)			(0.000)
Constant	30.184***	6.827***	7.100***	32.092***	4.540***	5.505***
	(0.378)	(1.345)	(1.323)	(0.419)	(1.394)	(1.374)
Year dummy	N	Y	Y	N	Y	Y
Econ*Year dummy	N	Y	Y	N	Y	Y
Observations	57,795	57,795	57,795	41,630	41,630	41,630
R-squared	0.029	0.270	0.307	0.081	0.285	0.332

*Notes:* Management single-authored paper is the benchmark. Econ=economics. Initial=order of initial of surname. Academic age= average scientific age of coauthors. Number of pages= number of pages of the paper. Life time citation = total 10 years citations of all papers of the author. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 6 displays the comparable regression for coauthored papers. Authors of two author coauthored management papers with a surname of A have an average rank of 50.002 percentile among all management and economics paper citations. Authors of two authored economics papers with a surname of A are ranked 7.186 percentage points higher. Although we see a significant coefficient of 0.064 for the effect of management authors' surnames on their ranks, it becomes insignificant when we add more controls. The initial of the surname has the opposite effect for coauthored works in economics. Economists ranks now decrease by a significant -0.134 percentage points for every increase in the initial of their surnames. Thus, for every 1,000 economists of coauthored papers, one economist is mis-ranked with respect to economists with an immediately earlier surname. However, this grows to  $-0.134 \times 25 = -33.35$  for those with surnames starting with Z with respect to those with surnames starting with A.

We predict in P1-P2 that single and coauthored papers in economics have opposite trends on surname initials with respect to the corresponding management benchmarks. The regressions in Table 5 and Table 6 indicate both trends are significantly different from zero and in the correct directions. Table 7 provides a direct comparison between these two trends and shows they are significantly

different. The magnitude of coefficients provides consistent evidence about the opposite direction.

TABLE 6: REGRESSION OF CITATION RANK FOR ECONOMICS AND MANAGEMENT TWO AUTHORS COAUTHORED PAPERS

Dependent variable	Citation rank (0 – 100)					
	1900-2000			1960-2000		
	(1)	(2)	(3)	(4)	(5)	(6)
Initials	0.064** (0.025)	0.040* (0.024)	0.036 (0.023)	0.046* (0.025)	0.025 (0.024)	0.021 (0.023)
<b>Initials*econ</b>	<b>-0.134***</b> (0.037)	<b>-0.122***</b> (0.034)	<b>-0.139***</b> (0.033)	<b>-0.121***</b> (0.037)	<b>-0.100***</b> (0.034)	<b>-0.116***</b> (0.033)
Econ	7.186*** (0.495)	-5.753*** (1.152)	-4.811*** (1.097)	7.595*** (0.496)	-6.388*** (1.152)	-5.396*** (1.097)
Academic age		0.486*** (0.036)	-0.003 (0.034)		0.514*** (0.036)	0.016 (0.034)
Academic age <sup>2</sup>		-0.013*** (0.001)	-0.004*** (0.001)		-0.014*** (0.001)	-0.004*** (0.001)
Pages		1.885*** (0.047)	1.732*** (0.044)		1.995*** (0.052)	1.833*** (0.049)
Pages <sup>2</sup>		-0.019*** (0.001)	-0.019*** (0.001)		-0.021*** (0.001)	-0.020*** (0.001)
Lifetime citation			0.016*** (0.000)			0.016*** (0.000)
Constant	50.002*** (0.339)	28.201*** (0.937)	29.347*** (0.895)	50.276*** (0.343)	26.929*** (0.954)	28.321*** (0.910)
Year dummy	N	Y	Y	N	Y	Y
Econ*Year dummy	N	Y	Y	N	Y	Y
Observations	55,754	55,754	55,754	51,864	51,864	51,864
R-squared	0.009	0.133	0.188	0.011	0.134	0.194

Notes: Management coauthored papers is the benchmark. Econ=economics. Initial=order of initial of surname. Academic age= average scientific age of coauthors. Number of pages= number of pages of the paper. Life time citation = total 10 years citations of all papers of the author. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



TABLE 7: CITATION RANK FOR ECONOMICS AND MANAGEMENT SINGLE AND TWO AUTHORS COAUTHORED PAPERS

Dependent variable Time period	Citation rank (0 – 100)					
	1900-2000			1960-2000		
	(1)	(2)	(3)	(4)	(5)	(6)
Initials	0.064** (0.025)	0.048** (0.024)	0.045* (0.023)	0.046* (0.025)	0.023 (0.024)	0.020 (0.023)
Initials*single	-0.107*** (0.038)	-0.085** (0.035)	-0.086** (0.034)	-0.098** (0.040)	-0.072** (0.036)	-0.075** (0.035)
Initials*econ	-0.134*** (0.037)	-0.131*** (0.034)	-0.151*** (0.033)	-0.121*** (0.037)	-0.097*** (0.035)	-0.115*** (0.033)
<b>Initials*econ*single</b>	<b>0.259***</b> (0.055)	<b>0.198***</b> (0.049)	<b>0.218***</b> (0.048)	<b>0.192***</b> (0.057)	<b>0.196***</b> (0.052)	<b>0.215***</b> (0.051)
Single	-19.818*** (0.508)	-14.040*** (0.487)	-13.446*** (0.474)	-18.184*** (0.541)	-12.398*** (0.503)	-11.799*** (0.486)
Econ	7.186*** (0.495)	-5.185*** (0.966)	-4.739*** (0.923)	7.595*** (0.496)	-6.077*** (0.966)	-5.380*** (0.922)
Single*econ	2.962*** (0.738)	7.105*** (0.686)	6.436*** (0.667)	10.083*** (0.775)	5.564*** (0.718)	4.865*** (0.694)
Academic age		0.401*** (0.025)	0.009 (0.024)		0.481*** (0.026)	0.051** (0.025)
Academic age <sup>2</sup>		-0.011*** (0.001)	-0.003*** (0.001)		-0.012*** (0.001)	-0.004*** (0.001)
Pages		2.169*** (0.089)	2.034*** (0.091)		2.308*** (0.101)	2.151*** (0.103)
Pages <sup>2</sup>		-0.023*** (0.002)	-0.023*** (0.002)		-0.024*** (0.002)	-0.024*** (0.002)
Lifetime citation			0.017*** (0.000)			0.016*** (0.000)
Constant	50.002*** (0.339)	24.097*** (0.954)	24.339*** (0.945)	50.276*** (0.343)	21.701*** (1.027)	22.550*** (1.018)
Year dummy	N	Y	Y	N	Y	Y
Econ*Year dummy	N	Y	Y	N	Y	Y
Observations	113,549	113,549	113,549	93,494	93,494	93,494
R-squared	0.090	0.259	0.300	0.089	0.240	0.291

Notes: Management coauthored papers is the benchmark. Econ=economics. Initial=order of initial of surname. Academic age= average scientific age of coauthors. Number of pages= number of pages of the paper. Life time citation = total 10 years citations of all papers of the author. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

By testing for the difference in citations between authors of single and coauthored papers in economics using citations of single and coauthored papers in

management as a benchmark, we diminished the importance of many of the econometric issues, such as self-citations, unknown and/or non-stationary distribution of names after entry of many non-Westerners into academia, which can make identification so difficult in this literature. Presumably, these confounds are similar among single and coauthored works. In this respect, we also make a methodological contribution.

Table 8 shows the effect of surnames on ranking for single authored papers is driven by the bottom five journals for economics. The effect of surnames on ranking for two-authored economics papers is also larger in the bottom five journals. This is consistent with our prediction P3 that the effect of endogenous selection should be stronger with lower tier journals, where the market for coauthors should be thinner.

TABLE 8: REGRESSION OF TOP FIVE AND BOTTOM FIVE JOURNALS

	Single-authored		Two-authored	
	1900-2000	1960-2000	1900-2000	1960-2000
<i>Top 5 econ journals</i>				
<b>Initial*econ</b>	<b>0.045</b>	<b>0.088</b>	<b>-0.149***</b>	<b>-0.106**</b>
	(0.044)	(0.054)	(0.047)	(0.047)
<i>Bottom 5 econ journals</i>				
<b>Initial*econ</b>	<b>0.132**</b>	<b>0.150**</b>	<b>-0.150**</b>	<b>-0.132**</b>
	(0.064)	(0.065)	(0.062)	(0.062)
<u>Covariates controlled for:</u>				
Author-level	Y	Y	Y	Y
Journal-level	Y	Y	Y	Y
Year, discipline fixed effect and their interactions	Y	Y	Y	Y

*Notes:* Management coauthored papers are the benchmark. Econ=economics. Initial=order of initial of surname. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Top five econ journals: American Economic Review, Quarterly Journal of Economics, Journal of Political Economy, Econometrica, Review of Economic Studies. Bottom five econ journals (in our sample of 23 econ journals, according to 5 years impact factor in 2012: International Economic Review, Journal of Law and Economics, Journal of Economic Theory, Economic Theory, Games and Economic Behaviors. All 30 management journals is benchmark.

P6 is born out in Table 9, which shows that non-alphabetically listed two authored works have a significantly lower citation rank (-7.071 percentage points for model (1)), but an insignificantly lower slope (-0.061).

TABLE 9: TWO AUTHOR COAUTHORS OF NON-ALPHABETICAL VS ALPHABETICAL ECONOMICS PAPERS

Dependent variable	Citation rank (0 – 100)					
	1900-2000			1960-2000		
Time period	(1)	(2)	(3)	(4)	(5)	(6)
Letter	-0.059** (0.028)	-0.065** (0.026)	-0.080*** (0.025)	-0.066** (0.028)	-0.062** (0.026)	-0.077*** (0.025)
<b>Letter*non-alpha</b>	<b>-0.061</b> (0.081)	<b>-0.100</b> (0.073)	<b>-0.094</b> (0.072)	<b>-0.043</b> (0.080)	<b>-0.068</b> (0.075)	<b>-0.059</b> (0.073)
Non-alpha	<b>-7.071***</b> (1.104)	<b>-4.179***</b> (1.008)	<b>-3.170***</b> (0.985)	<b>-5.927***</b> (1.102)	<b>-4.443***</b> (1.030)	<b>-3.436***</b> (1.005)
Academic age		0.408*** (0.054)	-0.015 (0.054)		0.412*** (0.054)	-0.020 (0.054)
Academic age <sup>2</sup>		-0.012*** (0.002)	-0.005** (0.002)		-0.012*** (0.002)	-0.004** (0.002)
Pages		1.902*** (0.067)	1.743*** (0.062)		1.930*** (0.071)	1.764*** (0.065)
Pages <sup>2</sup>		-0.019*** (0.001)	-0.018*** (0.001)		-0.020*** (0.001)	-0.019*** (0.001)
Lifetime citation			0.012*** (0.000)			0.012*** (0.000)
Constant	58.179*** (0.383)	22.719*** (1.154)	25.468*** (1.108)	58.634*** (0.381)	22.247*** (1.173)	25.174*** (1.125)
Year dummy	N	Y	Y	N	Y	Y
Econ*Year dummy	N	Y	Y	N	Y	Y
Observations	24,096	24,096	24,096	23,244	23,244	23,244
R-squared	0.009	0.167	0.217	0.007	0.148	0.200

Notes: Only two-authored econ papers included. Non-alpha=the ordering of authors is non-alphabetical. Initial=order of initial of surname. Econ=economics. Academic age= average scientific age of coauthors. Number of pages= number of pages of the paper. Life time citation = total 10 years citations of all papers of the author. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Thus, people with later surname initials should still suffer from having fewer collaborators as predicted by Table 3. They suffer even more when their collaborators must in equilibrium be of lower quality to tolerate losing more credit to a first author with a later surname (Engers et al., 1999). This equilibrium decrease in quality is the extra “welfare loss” to both authors, but we show that it could be entirely due to endogenous selection.

### Robustness checks

Table 10 shows that the baseline results of economics single and two-authored papers hold even when we do not use management journals as a benchmark. For management itself, we found a decreasing trend (-0.043) with initials for single-authored papers here, and it’s marginally significant in Table 5, which is consistent the natural tendency of authors with earlier surnames to get more citations due to the fact that reference lists are usually alphabetical (Huang, 2014). The increasing trend for two-authored management papers in Table 10 shows that the baseline results of economics single and two-authored papers, which also appear in Column (1) and (4) in Table 6, seems to contradict the natural tendency mentioned above (Huang, 2014). But it becomes insignificant with more controls, as can be seen in other columns of Table 6.

TABLE 10: REGRESSION OF CITATION RANK OF SINGLE AND COAUTHORED ECONOMICS AND MANAGEMENT PAPERS

	Citation rank (0 – 100)			
	Econ single	Econ 2-author	Management single	Management 2-author
Initial	0.082*** (0.029)	-0.070*** (0.027)	-0.043 (0.028)	0.064** (0.025)
Constant	40.332*** (0.395)	57.188*** (0.361)	30.184*** (0.378)	50.002*** (0.339)
Observations	27,903	24,096	29,892	31,658
R-squared	0.000	0.000	0.000	0.000

*Note:* Regressions without benchmark. Initial=order of initial of surname. Robust standard errors in parentheses: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Significance for Initial alone is lost if we restrict the sample to after 1960, when coauthored works overtakes single authored works. We also lose significance if we include pages, academic age and lifetime citation controls for the whole sample period.

However, we do lose significance for economics single authored papers in Table 10 if we restrict the sample to after 1960 or include more controls. Again, this is most likely due to the natural tendency of authors with earlier surnames to get more citations.

We found no significant result for 3-authored and 4-authored econ papers for either economics or management. Due to the increasing marginal cost of coordinating more authors and what we would expect are the smaller marginal gains for each extra person, we think it likely that a third or fourth author was invited to join when the paper with two or three authors, respectively, ran into problems. In that case, the credit by surname order may be less important than getting the best author.

Lastly, Figure 2 exhibits a possible consequence of the diminished incentive to coauthor in economics: a 10 percent higher probability of single authorship than management across all surnames. This gap is merely suggestive of the welfare loss however, because there could be other differences in the incentives to coauthor in management.

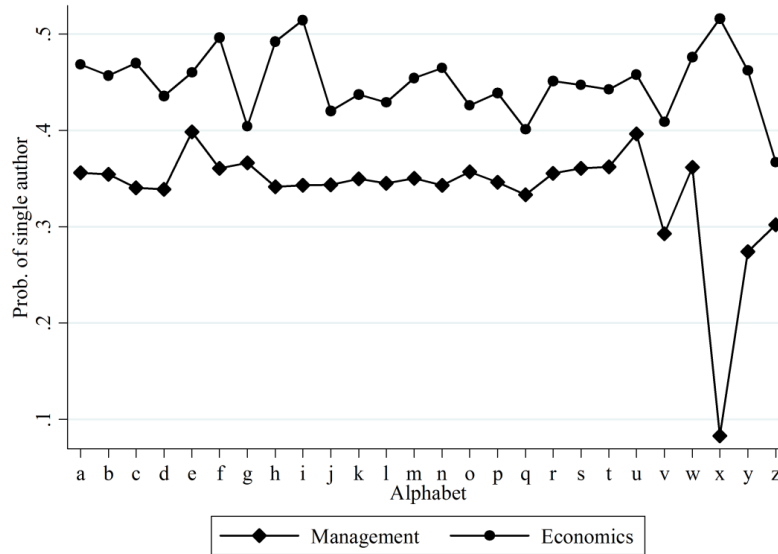


FIGURE 2: FREQUENCY OF SINGLE AUTHORSHIP IN MANAGEMENT AND ECONOMICS

*Notes:* This was calculated by dividing the number of single authored papers for each surname initial by the number of coauthored papers for the same initial.

## Discussion

We have shown evidence that economists take into account both the likely contribution of their coauthors and their own share of the credit, as a function of their respective surnames, when deciding to coauthor. Thus, we offer an alternative explanation to fact that has already been established by prior work, that first authors get a disproportionate share of the credit for the quality of papers. We show that this disproportionate share of the credit could be anticipated in the formation of the coauthorship, and thereby, not necessarily undeserved. Together with Ductor's (2013) finding that coauthored works engender more citations than single authored papers, our theoretical prediction P5 suggests an alternative explanation for the findings of Einav and Yariv (2006) that promotions, prizes, and higher pay are more likely to be accrued to people with earlier surnames in economics, or other fields in which alphabetical ordering is

used. P5 would also predict the results of Levitt and Thelwall's (2013) study of 20 social sciences subjects that there is a high correlation between alphabetical ordering of authors in a field and the proportion of first-authors near the beginning of the alphabet. Similarly, P5 predicts that two-author alphabetized articles would be more highly cited than non-alphabetized two-author articles in economics (Joseph et al., 2005) and agricultural economics (Laband & Tollison, 2006). P5 would also predict Brown et al.'s (2006) finding that the rate of coauthorships with alphabetical ordering was stronger for the top four journals than the rest of the 19 marketing journals mentioned.

Our findings may help in resolving the still controversial question of whether coauthoring is more productive of higher quality scholarship. Laband and Tollison (2000) find that coauthored papers are more likely to be accepted than single authored papers. Wuchty et al. (2007) and Chung, Cox, and Kim (2009) find that coauthored papers are more cited, when the citations are discounted by the number of coauthors. However, Medoff (2003) did not find increased citations, and Hollis (2001) even finds lower citations for coauthored papers. More recently, Ductor (2013) finds that coauthored papers are in fact more cited after controlling for one form of endogeneity in coauthor selection: common research interests. (See his paper for a review of the literature on the evidence for greater productivity of coauthorship.) These conflicting findings could be due to actual differences in the relative quality of coauthors as a function of endogenous selection into coauthorships due to whether fields use alphabetical or contribution ordering of coauthors as well as the number of coauthors.

Our findings motivate further research on whether some part of the aggregate citations of academic institutions, countries or ethnicities can be predicted by their draw of surnames. Freeman and Huang (2014) demonstrates that ethnicity can affect citations through the distribution of surnames. China's 1.4 billion people have only 4000 surnames. This is three percent of the 150,000 surnames of the

300 million people in the US<sup>7</sup>. Furthermore, one third of China's surnames are from the last four letters of the alphabet (Wang to Zou<sup>8</sup>). Universities in countries like China may foresee greater returns from fields like management rather than fields like economics, if the surnames of faculty tend to be from the end of the alphabet.

Our findings have been about coauthorships. However, they could also apply to selection into institutions with potential coauthorships if the potential for coauthorships is an important reason for joining an institution. For example, a junior candidate with a last name of Johnson, who has yet to prove his or her ability, might lean towards a department with established researchers with more surnames like Arrow or Becker rather than a department with more surnames like Wang and Zhang, due the complementarities that we have identified here. Departments may take the mirror incentives into account when considering candidates. Thus, Chinese universities could also have an incentive to focus on fields like economics, if for budget reasons, they tend to hire junior Chinese faculty who collaborate with senior non-Chinese faculty.

These incentives could operate even when people are selecting into fields, starting with the first paper, because the first paper could set a student onto a career. A Professor "Johnson" in economics or chemistry may be more disposed to mentoring a student with a surname of Zhang as a potential coauthor than a student with the surname of Clarke, whereas a Professor "Wang" in management may be more indifferent. This preference could be a contributing factor to the exceptional productivity of Chinese graduate students in chemistry, which also uses alphabetical ordering, documented in Gaule and Piacentini (2013).

<sup>7</sup> [http://en.wikipedia.org/wiki/List\\_of\\_common\\_Chinese\\_surnames](http://en.wikipedia.org/wiki/List_of_common_Chinese_surnames)

<sup>8</sup> <http://www.bloomberg.com/visual-data/best-and-worst/most-common-in-china-surnames>



Citation convention and initial distribution of surnames in departments within a certain local distribution of surnames could be important in predicting the level of talent which an academic field might draw. Though citations conventions are not policy variables in the usual sense, they do motivate the study of the incentive effects of prize sharing rules within and among contests.

### References

- Brown, C., Chan, K., & Chen, C. (2011). First-author conditions: evidence from finance journal coauthorship. *Applied Economics*, (937).
- Brown, C., Chan, K., & Lai, P. (2006). Marketing journal coauthorships: An empirical analysis of coauthor behavior. *Journal of Marketing Education*, 28(1), 17–25.
- Bruno, B. (2014). Economics of co-authorship. *Economic Analysis and Policy*.
- Chung, K. H., Cox, R. A. K., & Kim, K. A. (2009). On the relation between intellectual collaboration and intellectual output : Evidence from the finance academe. *The Quarterly Review of Economics and Finance*, 49, 893–916.
- Dechenaux, E., Kovenock, D., & Sheremeta, R. (2012). A Survey of Experimental Research on Contests , All-Pay Auctions and Tournaments.
- Ductor, L. (2013). Does Co-authorship Lead to Higher Academic Productivity?
- Efthyvoulou, G. (2008). Alphabet Economics: The link between names and reputation. *The Journal of Socio-Economics*, 37(3), 1266–1285.
- Einav, L., & Yariv, L. (2006). What’s in a surname? The effects of surname initials on academic success. *The Journal of Economic Perspectives*, 151(3712), 867–8.
- Engers, M., Gans, J., Grant, S., & King, S. (1999). First-author conditions. *Journal of Political Economy*, 107(4), 859–883.

- Freeman, R., & Huang, W. (2014). Collaborating With People Like Me: Ethnic co-authorship within the US. *Journal of Labor Economics*.
- Gaule, P., & Piacentini, M. (2013). Chinese graduate students and US scientific productivity. *Review of Economics and Statistics*, (June 2009), 1–25.
- Gupta, H., Campanha, J., & Pesce, R. (2005). Power-law distributions for the citation index of scientific publications and scientists. *Brazilian Journal of Physics*.
- Hamermesh, D., Johnson, G., & Weisbrod, B. (1982). Citations and Salaries: Economic Scholarship , Rewards in Economics. *Southern Economic Journal*.
- Hilmer, C., Hilmer, M., & Ransom, M. (2012). Fame and the Fortune of Academic Economists: How the market rewards influential research in economics.
- Hollis, A. (2001). Co-authorship and the output of academic economists. *Labour Economics*, 8(4), 503–530.
- Huang, W. (2014). Do ABCs get mor citations than XYZs? *Economic Inquiry*, n/a–n/a.
- Hudson, J. (1996). Trends in multi-authored papers in economics. *The Journal of Economic Perspectives*, 10(3), 153–158.
- Joseph, K., Laband, D. N., Patil, V., Southern, S., Journal, E., & Jan, N. (2005). Author order and research quality. *Southern Economic Journal*, 71(3), 545–555.
- Laband, D. N., & Tollison, R. D. (2006). Alphabetized coauthorship. *Applied Economics*, 38(14), 1649–1653.
- Lazear, E., & Oyer, P. (2012). Personnel Economics. *The Handbook of Organizational Economics*.
- Levitt, J. M., & Thelwall, M. (2013). Alphabetization and the skewing of first authorship towards last names early in the alphabet. *Journal of Informetrics*, 7(3), 575–582.

- Medoff, M. H. (2003). Collaboration and the quality of economics research. *Labour Economics*, 10(5), 597–608.
- Moore, W., Newman, R., & Turnbull, G. (2001). Reputational capital and academic pay. *Economic Inquiry*, 39(4), 663–671.
- N. Laband, D., & D. Tollison, R. (2000). Intellectual Collaboration. *Journal of Political Economy*, 108(3), 632.
- Nudelman, A., & Landers, C. (1972). The failure of 100 divided by 3 to equal 33-1/3. *The American Sociologist*.
- Redner, S. (2005). Citation Statistics From More Than a Century of Physical Review. *Bulletin of the American Physical Society*.
- Sauer, R. D. (1988). Estimates of the Returns to Quality and Coauthorship in Economic Academia. *Journal of Political Economy*, 96(4), 855–66.
- Van Praag, C. M., & Van Praag, B. M. S. (2008). The Benefits of Being Economics Professor A (rather than Z). *Economica*, 75(300), 782–796.
- Waltman, L. (2012). An empirical analysis of the use of alphabetical authorship in scientific publishing. *Journal of Informetrics*, 6(4), 700–711.
- Wuchty, S., Jones, B., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036–1039.

## Appendix

TABLE 11: LIST OF JOURNAL AND THEIR ABBREVIATIONS

---

### Economics

American Economic Review  
 Brookings Papers On Economic Activity  
 Economic Journal  
 Econometrica  
 Econometric Theory  
 Games And Economic Behavior  
 International Economic Review  
 Journal of Business and Economic Statistics  
 Journal of Economic Literature  
 Journal of Economic Perspectives  
 Journal of Economic Theory  
 Journal of Financial Economics  
 Journal of Law and Economics  
 Journal of Monetary Economics  
 Journal of Econometrics  
 Journal of Finance  
 Journal of Political Economy  
 Journal of Public Economics  
 Quarterly Journal of Economics  
 Review of Economic Studies  
 Review of Economics And Statistics  
 Review of Financial Studies  
 Rand Journal of Economics

### Management

Academy of Management Journal  
 Academy of Management Review  
 Administrative Science Quarterly  
 California Management Review  
 Decision Sciences  
 Group and Organization Management  
 Harvard Business Review  
 Human Relations  
 Human Resource Management  
 Industrial and Labor Relations Review  
 Industrial Relations  
 Journal of Applied Psychology

Journal of Business Research  
 Journal of Business Venturing  
 Journal of Human Resources  
 Journal of International Business Studies  
 Journal of Management Studies  
 Journal of Organizational Behavior  
 Journal of Management  
 Journal of Occupational And Organizational Psychology  
 Journal of Vocational Behavior  
 Leadership Quarterly  
 Monthly Labor Review  
 Management Science  
 Organizational Behavior And Human Decision Processes  
 Organizational Research Methods  
 Organization Science  
 Personnel Psychology  
 Strategic Management Journal  
 Sloan Management Review

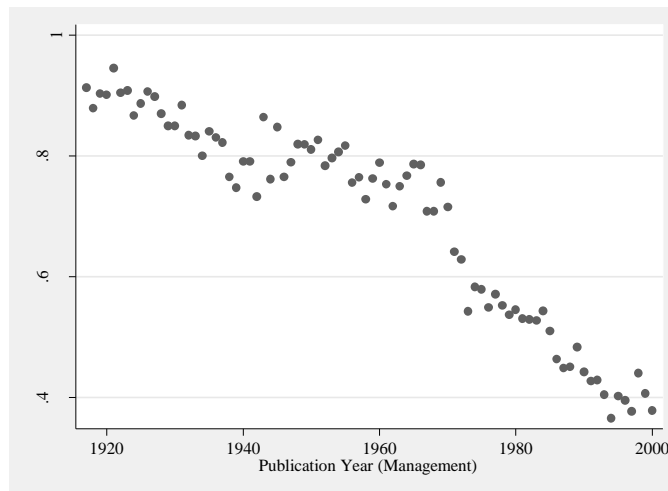


FIGURE 3: RATE OF SINGLE AUTHORSHIP IN MANAGEMENT

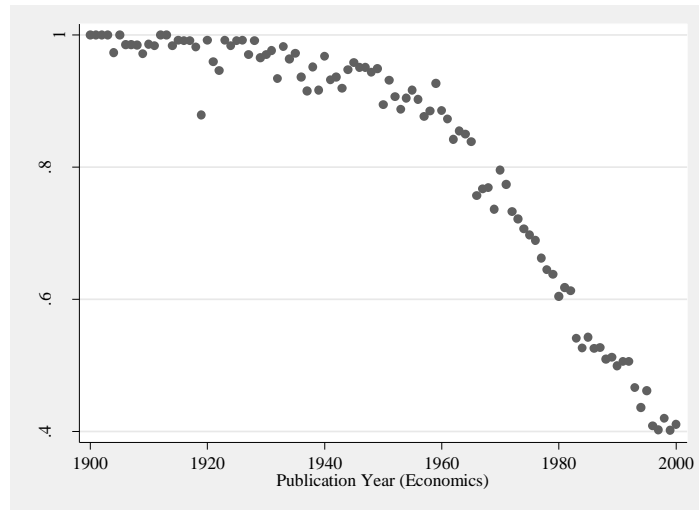


FIGURE 4: RATE OF SINGLE AUTHORSHIP IN ECONOMICS