What's in a Surname? The Effects of Surname Initials on Academic Success*

Liran Einav† Stanford University

Leeat Yariv‡ UCLA and Caltech

March 10, 2005

Abstract

We present evidence that a variety of proxies for success in the U.S. economics labor market (tenure at highly ranked schools, fellowship in the Econometric Society, and to a lesser extent, Nobel Prize and Clark Medal winnings) are correlated with surname initials, favoring economists with surname initials earlier in the alphabet. These patterns persist even when controlling for country of origin, ethnicity, and religion. We suspect that these effects are related to the existing norm in economics prescribing alphabetical ordering of authors' credits. Indeed, there is no significant correlation between surname initials and tenure at departments of psychology, where authors are credited roughly according to their intellectual contribution. The economics market participants seem to react to this phenomenon. Analyzing publications in the top economics journals since 1980, we note two consistent patterns: authors participating in projects with more than three authors have significantly earlier surname initials, and authors writing papers in which the order of credits is non-alphabetical have significantly higher surname initials.

Journal of Economic Literature classification numbers: A11, A13, J23, J70, Z13.
Keywords: Norms, Economics Job Market, Alphabetical Discrimination.

*We thank David Laibson, David Levine, Enrico Moretti, Muriel Niederle, and Nicola Persico for many helpful comments and suggestions. Ted Chang, Shipra Kaul, Shuhei Kurizaki, and Sujey Subramanian provided outstanding research assistance.

†Department of Economics, Stanford University, Stanford, CA 94305-6072, leinav@stanford.edu

‡Department of Economics, UCLA, Los Angeles, CA 90095-1477 and Division of Humanities and Social Sciences, Caltech, Pasadena, CA 91125, lyariv@econ.ucla.edu

"When I was growing up, I got used to being near the end of the alphabet. While I didn't have it as bad as the people with surnames ending in Y or Z, I was still envious of the Allens and Browns. I spent my days in public schools sitting with the same people, always near the back of the classroom. The Sullivans, the Smiths, a Stroup and even a Strupeck."

- Dave Stroup, "The Stroup Effect," *Georgetown Voice*, 4.15.2004.

1 Introduction

There is abundant research identifying external characteristics (race, gender, adolescent height) that affect labor market outcomes. In this paper, we focus on the effects of surname initials on professional outcomes in the academic labor market for economists.

We analyze profile data of faculty in all top 35 U.S. economics departments and detailed characteristics of publications in five of the top economics journals in the years 1980-2002. The contribution of the paper is three-fold. First, we identify significant and robust consequences of faculty surname initials on proxies of professional success. Faculty with lower (earlier) surname initials are significantly more likely to receive tenure at higher ranked departments, are significantly more likely to become fellows of the Econometric Society, and, to a lesser extent, are more likely to win the Clark Medal and the Nobel Prize. Second, we provide multi-layered evidence suggesting that this weak form of discrimination is driven by an "innocent" and widely accepted professional norm. Namely, the norm prescribing the alphabetical ordering of collaborators on written publications. Third, we demonstrate the reactions of the economics labor market to these effects by ways of faculty publication patterns.

In more detail, the first part of our investigation considers the long-run effects of alphabetical placement using the comparison of name distributions of senior and junior economics faculty. We find a significant effect of alphabetical ordering on tenure in departments ranked as top 5 and top 10 departments. This effect fades when considering our entire set of top 35 departments. As an illustration, when looking at averages, we find that the gap between the average seniors' surname initials and the average juniors' surname initials increases with the ranking of schools. For example,

while the aggregate average corresponding to senior faculty in top 35 departments is not significantly different from the aggregate average corresponding to junior faculty in those same departments, this difference is significant when restricting attention to the top 5 departments, and equals 2.3. That is, juniors' last names start with letters that are, on average, 2.3 letters later in the alphabet than seniors' in the top five departments. These significant differences are virtually the same even after we control for country of origin, ethnicity, religion, or departmental fixed effects. We find similar patterns for other proxies for academic accomplishment, such as fellowship in the Econometric Society, receipt of the Nobel Prize, and winning of the Clark Medal.

We suspect the "alphabetical discrimination" reported in this paper is linked to the norm in the economics profession prescribing alphabetical ordering of credits on publications. Indeed, coauthored papers are very common in the economics profession. In the years 1980-2002, five of the most prominent economics journals published 50% multi-authored papers. In 88% of these articles, the authors were listed alphabetically. In contrast, in many of the widely read journals of neighboring disciplines[1] the rate of coauthorship stands similar to economics, but 40-50% of the corresponding coauthors are listed alphabetically (see tables 1 and 2 in Engers et al. (1999)). Thus, the institutional structure of promotions in the economics labor market treats participants with different surname initials in an inherently asymmetric way. As a test, we replicate our analysis for faculty in the top 35 U.S. psychology departments, for which coauthorships are not normatively ordered alphabetically.[2] We find no significant effects of alphabetical placement on tenure status.

The second part of our investigation aims at analyzing the extent to which the effects of alphabetical placement are internalized by potential authors in their choices of the number of coauthors as well as in their willingness to follow the alphabetical ordering norm. We find that the distribution of authors' surnames in single-authored, double-authored, and triple-authored papers is, in fact, not significantly different from one another. Nonetheless, the distribution of authors' surnames in four- and five-author papers differs significantly and corresponds to authors that are, on average, lower in the alphabet. Furthermore, reversal of authors' names is correlated with the placement of the

[1] *American Journal of Sociology*, *American Psychologist*, *Angewandte chemie*, and *New England Journal of Medicine*.

[2] Generally, the order of authors in psychology journals is determined according to intellectual contribution. An exception to this rule is the head of the lab, who sometimes appears last.

authors' surnames in the alphabet. That is, surnames of authors writing papers in which credits do not follow the lexicographic ordering start with letters that are statistically higher than those of surnames belonging to authors of papers in which credits are specified lexicographically.

In Section 3 we provide a benchmark model that captures the stylized facts observed in the data. We consider authors who choose between writing alone or joining a pool of coauthors. The value of authoring alone is independent of the author surname's initial, while the value of entering the pool of coauthors depends on the author's surname initial and the distribution of surnames amongst the potential coauthors. Such a model produces an equilibrium surname distribution of coauthors and a threshold surname initial, below which authors would choose to coauthor, and above which they will choose to author alone. This qualitative pattern is consistent with our empirical observation. Furthermore, under mild conditions, this equilibrium is unique.

Recently, some attention has been given to the study of the publication process in the economics profession. Engers et al. (1999) analyze a theoretical model in which authors participate in a market. They show that when two authors bargain over their placement in a paper's credits and receive a final payoff that depends on the market's perception of each author's contribution, alphabetical ordering of names arises as an equilibrium. Ellison (2002) provides a review of the trends in publication in top economic journals and illustrates how that process has changed (namely slowed down) over the past three decades. A general overview of recent trends in the economics profession appears in Gans (2001).

In the hard sciences, Shevlin and Mark (1997) found a correlation between citation rates (in the *1994 Science Citation Index*) and authors' alphabetical placement. Significantly more papers written by authors with earlier initials in the alphabet were cited. However, this correlation disappears when controlling for the base rate distribution of names using the London phone book. Over and Smallman (1970) looked at *The Journal of Physiology*, in which alphabetical ordering was mandatory. They found less collaborative publication by scientists with surnames starting with letters later in the alphabet (P-Z) than in other journals in the field. Zuckerman (1968) conducted interviews with Nobel laureates in the hard sciences. Zuckerman notes that laureates often exercise their *noblesse oblige* by giving credit to less eminent coauthors increasingly as their own eminence grows, particularly

after winning the prize. This *noblesse oblige* has its limits; laureates' contributions to prize-winning research are more visible than contributions to their other research.[3]

In the broader scheme of things, the current paper falls within the rubric of work trying to detect channels by which labor market outcomes differ according to participants' external characteristics (see, e.g., Bertrand and Mullainathan (2004) and Persico et al. (2004)). In that respect, we identify one such institutional channel pertaining to last names within the economics profession.

The paper is organized as follows. Section 2 describes the data used for the analysis and the results. Section 3 provides a simple model that is consistent with the stylized facts we find. Section 4 concludes by suggesting possible mechanisms through which "alphabetical discrimination" operates, and potential policy implications.

2 Data and Results

In this section we trace the following expositional route. We start by describing the data sets that have been constructed for this study. We then identify consequences of alphabetical ordering and their robustness, as well as explore several potential channels by which these effects were generated. We conclude by demonstrating the market responses to these identified effects.

2.1 Faculty Data: Is alphabetical ordering consequential?

We collected demographic data regarding faculty at the top 35 Economics and Psychology Departments in the U.S. Table A1 provides the full lists and their sources. The vast majority of the faculty data was collected from departmental web sites and faculty home pages. For all faculty, we recorded their names, tenure status (untenured, tenured, and emeritus), nationalities, whether they are fellows of the Econometric Society (from the society's web page, as of January 2004), and the year they obtained their Ph.D. when available.[4]

[3] It is worth noting that economics and the hard sciences differ in the dimensions in which intellectual collaboration takes place. Indeed, Laband and Tollison (2000) perform a comparative study of intellectual collaboration in economics and biology. They find that while the incidence and extent of formal intellectual collaboration through coauthorships are greater in biology than in economics, the incidence and extent of informal intellectual collaboration (through, e.g., discussions at conferences) are greater in economics than in biology.

[4] The year of the completion of Ph.D. is only available for about 80% of Economics faculty and 50% of Psychology faculty.

As already mentioned, our goal is to assess whether faculty's last names have any noticeable effect on their professional success. We concentrate on two such measures: whether the faculty are tenured, and whether they are fellows of the Econometric Society.[5]

Tables 1-4 report our main findings. We coded surname initials into numbers between 1 and 26 lexicographically (A corresponding to 1, B to 2, etc.). Overall, we find some suggestive evidence that faculty with lower initials are more likely to succeed in the profession. In particular, we find that tenured faculty at the top 5 Economics departments have significantly lower last names than junior faculty at the same departments. This negative relationship remains significant for the top 10 Economics Departments, but gradually disappears as we look at the set of top 20 and top 35 departments (Table 1). At top 5 and top 10 departments, the estimated magnitude of the effect is quite big: one lower letter in the alphabet increases one's tenure probability by almost one percent.

Table 1: Probit Regressions: Dependent Variable - 1 if Tenured (Economics)

Sample	Top 5 Econ	Top 10 Econ	Top 20 Econ	Top 35 Econ
Number of Obs.	208	405	799	1,233
Number of Tenured (%)	147 (70.7%)	293 (72.3%)	585 (73.2%)	911 (73.9%)
Last Name Initial[a]	-0.0295**	-0.0208**	-0.0078	-0.0047
	(-2.17)	(-2.08)	(-1.12)	(-0.84)
Pseudo R^2	0.0189	0.0092	0.0014	0.0005
Predicted Effect (A-Z)[b]	0.807-0.552	0.793-0.618	0.758-0.693	0.754-0.716

**, * Statistically significant at the 5% and 10% confidence level, respectively.

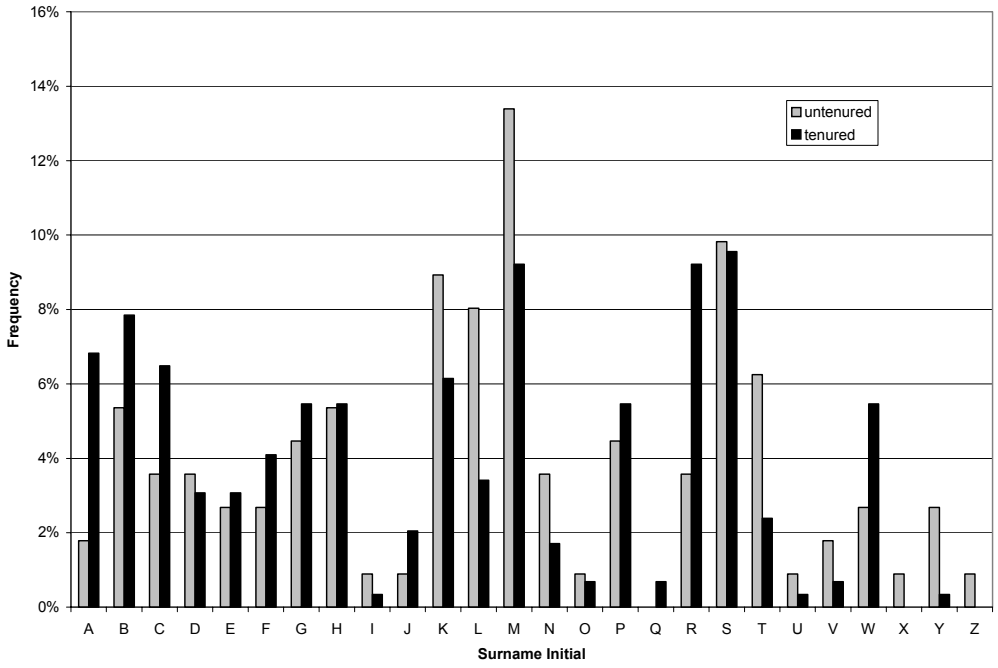
[a] Probit coefficients. Z-Stats below coefficients.

[b] The magnitude of the effect is reported as the predicted Pr(Tenure|'A') and Pr(Tenure|'Z').

In order to help the reader visualize the differences in distributions of initials among junior and senior faculty producing the observations reported in Table 1, Figure 1 below portrays the distributions of initials corresponding to faculty at the top 10 Economics departments. The distributions presented in Figure 1 are based on 293 tenured and 112 untenured faculty (as in the second column of Table 1). The means and standard deviations of these distributions are 11.13 (6.86) and 12.69 (6.44), respectively.

[5] Since not all departments are consistent in posting their emeriti faculty on their web pages, all the reported results are based on a sample that does not include emeriti faculty. Adding the available emeriti faculty, however, does not change any of the results.

Figure 1: Distribution of surname initials in Top 10 Economics Departments



In addition, the pattern, magnitude, and statistical significance of these effects do not change if we control for departmental fixed effects. Controlling for the number of publications slightly reduces the reported relationship.[6] This should not be surprising, as publications are endogenous, to the extent that they are likely to be correlated with professional success, even if the latter is driven solely by a lower surname initial (for an analysis of the publication choice made by authors, see Section 2.2).

Are the results driven by nationality, race, or religion? Our analysis is cross-sectional. Thus, one may be concerned that our findings are an artifact of the age gap between the average tenured faculty and the average junior faculty. Indeed, the observed relationship could conceivably be the consequence of a trend in the profession. For example, if (i) the fraction of non-Americans

[6] We use publication counts at five top economics journals between 1980 and 2002 (see Section 2.2 below). Out of three publication counts we tried – simple count, count of papers in which the author is the first author, and count normalized by the number of co-authors – only the latter had a meaningful effect.

at the junior faculty rank is higher at higher ranked universities; and (ii) foreign names have on average higher initials, then our interpretation of Table 1 would be misleading. In order to control for such potential trends, we control for American nationality, as well as for the origin of the name. For the latter, we used the help of two undergraduate research assistants, who independently and subjectively classified last names as Jewish, Indian, and other Asian. As the overlap between the two classifications produced by the research assistants was not perfect, we separately added each of them as a control, resulting in six different dummy variables (referred to as *origin controls* in all the tables that follow). Table 2 illustrates the persistence of the effects of alphabetical ordering even after controlling for the origin of the name or for the nationality of the individual. The results are very similar to those presented in Table 1.

Table 2: Probit Regressions: Dependent Variable - 1 if Tenured (Economics)

Sample	Top 5 Econ	Top 10 Econ	Top 20 Econ	Top 35 Econ
Number of Obs.	208	405	799	1,233
Number of Tenured (%)	147 (70.7%)	293 (72.3%)	585 (73.2%)	911 (73.9%)
Last Name Initial[a]	-0.0279*	-0.0205*	-0.0053	-0.0033
	(-1.88)	(-1.94)	(-0.73)	(-0.58)
American Nationality[a]	0.7339**	0.7044**	0.6701**	0.4988**
	(3.59)	(4.59)	(5.82)	(5.60)
Six origin controls	yes	yes	yes	yes
Pseudo R^2	0.1061	0.0998	0.0832	0.0604
Predicted Effect (A-Z)[b]	0.819-0.584	0.811-0.644	0.767-0.725	0.761-0.735

**, * Statistically significant at the 5% and 10% confidence level, respectively.

[a] Probit coefficients. Z-Stats below coefficients.

[b] The magnitude of the effect is reported as the predicted Pr(Tenure|'A') and Pr(Tenure|'Z'), both evaluated at the mean value of the other regressors.

Table 2 also illustrates the significant relationship between American nationality and a tenured position. This is probably driven by two factors. First, there is a general trend towards higher fraction of foreign economists who study and obtain academic positions in the U.S. Second, there is a higher propensity of non-Americans to move back to their countries of origin later on in their careers.

As a further control for any trends in name distributions, we restricted the sample to include only economists who obtained their Ph.D. between 1991 and 2000, thereby reducing the age gap

between a representative tenured faculty and a representative junior faculty and restricting attention to economists who are "just after" tenure and those who are "just before." The pattern of the effects reported in Tables 1 and 2 does not change. In fact, the magnitude of the effects at top 5 and top 10 departments increases by 60-70%. The statistical significance of these effects is, of course, lower, as sample sizes are about four times smaller; z-statistics for top 5 and top 10 departments are about 1.5 (*p-value* of 0.12 and 0.15, respectively).[7]

Other proxies for professional success Table 3 provides another layer of suggestive evidence, by reporting a similar analysis for fellowship in the Econometric Society. Given that out of the 252 Econometric Society fellows in our sample only two are non-tenured, this set of results is almost orthogonal to the results provided in tables 1 and 2.

Table 3: Probit Regressions: Dependent Variable - 1 if fellow of the Econometric Society

Sample	Tenured faculty at top 5 Econ	Tenured faculty at top 10 Econ	Tenured faculty at top 35 Econ	All faculty at top 35 Econ
Number of Obs.	147	293	911	1,233
Number of ES Fellows (%)	89 (60.5%)	153 (52.2%)	250 (27.4%)	252 (20.4%)
Last Name Initial[a]	-0.0229	-0.0200*	-0.0037	-0.0058
	(-1.49)	(-1.82)	(-0.58)	(-0.99)
Pseudo R^2	0.0114	0.0082	0.0003	0.0008
Predicted Effect (A-Z)[b]	0.691-0.470	0.600-0.407	0.287-0.256	0.221-0.180
Last Name Initial[a]	-0.0262	-0.0184*	-0.0047	-0.0063
	(-1.56)	(-1.64)	(-0.70)	(-1.04)
American Nationality[a]	-0.0591	-0.0140	0.1720*	0.2829**
	(-0.27)	(0.09)	(1.86)	(3.29)
Six origin controls	yes	yes	yes	yes
Pseudo R^2	0.0340	0.0162	0.0117	0.0246
Predicted Effect (A-Z)[b]	0.705-0.457	0.596-0.414	0.288-0.250	0.215-0.173

**, * Statistically significant at the 5% and 10% confidence level, respectively.

[a] Probit coefficients. Z-Stats below coefficients.

[b] The magnitude of the effect is reported as the predicted Pr(Fellow|'A') and Pr(Fellow|'Z'), both evaluated at the mean value of the other regressors.

Table 3 reveals a strikingly similar pattern to that reported in Table 1. Lower initial tenured faculty at top 5 and top 10 Economics Departments are significantly more likely to be fellows of

[7] One should note that we cannot control for age. First, age information is not available for many faculy members. Second, any proxy for age (such as the year of the Ph.D., which we have) will mechanically explain a large portion of the variation in tenure status, leaving only little variation to be explained by other variables.

the Econometric Society. The magnitude of this effect is, again, almost one percent per letter. As before, the effect gradually vanishes as we expand the set of faculty to include faculty in top 20, top 35, and junior faculty.

Table 4 presents similar results for the Nobel Prize and the Clark medal, as additional proxies for professional success. As one can see, while we obtain, again, a negative relationship between surname initials and the likelihood of winning these honors, these results are not statistically significant. This may be due to the small number of observations: there are 13 (7) Nobel Laureates and 14 (13) Clark Medal winners in our sample of top 35 (10) departments.

Table 4: Probit Regressions: Dependent Variable - 1 if won the Nobel Prize or Clark Medal

Sample	Tenured faculty at top 10 Econ		Tenured faculty at top 35 Econ	
Measure	Nobel Prize	Clark Medal	Nobel Prize	Clark Medal
Number of Obs.	293	293	911	911
Number of winners (%)	7 (2.4%)	13 (4.4%)	13 (1.4%)	14 (1.5%)
Last Name Initial[a]	-0.0370	-0.0079	-0.0144	-0.0011
	(-1.39)	(-0.41)	(-0.86)	(-0.07)
Pseudo R^2	0.0323	0.0016	0.0055	0.0000
Predicted Effect (A-Z)[b]	0.048-0.005	0.052-0.034	0.020-0.008	0.016-0.015

[a] Probit coefficients. Z-Stats below coefficients.

[b] The magnitude of the effect is reported as the predicted Pr(Prize|'A') and Pr(Prize|'Z')

Beyond the direct evidence provided by these results, they also help in boosting our confidence that the tenure results were not generated by recent time trends in the profession. While, on average, there is probably an age gap between Econometric Society fellows and other tenured faculty who are not fellows, this gap is different from the gap between tenured and junior faculty. Therefore, if there are some name trends in the profession that generate our results, such trends should have spanned multiple age gaps, which is even less likely. The fact that alphabetical placement affects this alternative array of success proxies, pertaining to economists at different stages in their careers, indicates that the reported alphabetical discrimination is not a mere artifact of a trend in the name distribution of more able economists.

Is it sorting? The preceding results may raise the suspicion that the reported findings are a consequence of a simple sorting mechanism: out of a representative pool of economists who enter the

labor market, faculty with lower surname initials are more likely to be promoted. Faculty who are not promoted are getting tenured in lower ranked schools, thereby creating the biggest tenured-junior difference at the top departments.

Unfortunately, this simple sorting story is not entirely consistent with the data. The distribution of surname initials of tenured faculty is quite similar among top 10 and lower ranked economics departments. The effects described in Table 1 are almost entirely driven by differences at the junior level. Junior faculty at top 10 departments have significantly higher surnames than junior faculty at other departments.

We should note that the results we find regarding Econometric Society fellows (Table 3), for which the sorting hypothesis is less relevant, are driven by both populations. About half of the effect is driven by fellows of the Econometric Society at top 10 departments having lower surname initials than those at lower ranked departments. The other half of the effect arises from top 10 faculty who are not fellows of the Econometric Society having higher surname initials than the corresponding faculty at lower ranked departments.

Is this a new phenomenon? In order to check whether the reported effects are new, we collected data on faculty at the top 5 departments from the past, for academic years 1979-1980 and 1989-1990. Repeating the same exercise for those groups, we find no significant relationship between last names and seniority. Note, however, that coauthorships have flourished in the past two decades or so. Indeed, as Hudson (1996) notes, in the period 1950-1965, the highest proportion of multi-authored papers for any of the journals discussed in this paper was 15.6 percent (in *Econometrica*), while in the years 1966-1970, the average proportion of multi-authored papers stood around 23 percent, which monotonically increased to the 50 percent level observed in the period this paper studies. In fact, Rosenblat and Mobius (2004) document a steep rise in coauthorships in the period right after the invention of the Internet in 1991. Consequently, if the ordering of authors' names is the channel by which alphabetical discrimination operates, one would expect the effects to be much weaker for past periods, in which many of the senior faculty had created a career based on predominantly single-authored papers.

Does alphabetical discrimination happen without alphabetical author ordering? The corresponding analysis for faculty at Psychology Departments results in smaller, insignificant, and often reversed relationship between last names and seniority status (Table 5). As the convention in Psychology is not to order coauthors alphabetically, but otherwise may be one of the closest disciplines to Economics, we interpret this finding as suggestive that the mechanism through which the last names may operate is, indeed, the alphabetical ordering of coauthors in Economics.

Table 5: Probit Regressions: Dependent Variable - 1 if Tenured (Psychology)

Sample	Top 5 Psych	Top 10 Psych	Top 20 Psych	Top 35 Psych
Number of Obs.	392	556	904	1,466
Number of Tenured (%)	320 (81.6%)	446 (80.2%)	733 (81.1%)	1,200 (81.9%)
Last Name Initial[a]	0.0101	0.0097	-0.0024	-0.0020
	(0.94)	(1.07)	(-0.34)	(-0.36)
Pseudo R^2	0.0024	0.0021	0.0001	0.0001
Predicted Effect (A-Z)[b]	0.787-0.853	0.774-0.840	0.817-0.801	0.824-0.811
Last Name Initial[a]	0.0114	0.0108	-0.0008	-0.0001
	(1.03)	(1.16)	(-0.10)	(-0.01)
Six origin controls	yes	yes	yes	yes
Pseudo R^2	0.0366	0.0459	0.0275	0.0262
Predicted Effect (A-Z)[b]	0.791-0.861	0.782-0.853	0.818-0.813	0.825-0.825

**, * Statistically significant at the 5% and 10% confidence level, respectively.

[a] Probit coefficients. Z-Stats below coefficients.

[b] The magnitude of the effect is reported as the predicted Pr(Tenure|'A') and Pr(Tenure|'Z'), both evaluated at the mean value of the other regressors.

2.2 Publications Data: Do authors respond?

The second part of our data set is comprised of dates, authors, and paper length for all publications at *The American Economic Review (AER), Econometrica*, *The Journal of Political Economy (JPE)*, *The Quarterly Journal of Economics (QJE)*, and *The Review of Economic Studies (REStud)*, from 1980 until 2002. For the purposes of the paper, we excluded notes and comments, as well as unrefereed publications.[8] Table 6 contains a summary of this part of the data. As can be seen from Table 6, about half of the papers are multi-authored, and in the vast majority of them (88 percent) authors are ordered alphabetically.

[8] in particular, publications in the *AER* May issue were not included

Table 6: Publications - Descriptive Statistics

Number of Authors[b]	Obs.	% of Total	% Alphabetically[a]	Mean Initial	Std. Dev.
1	3,378	49.8%	-	11.38	6.95
2	2,691	39.6%	91.4%	11.43	4.97
3	628	9.3%	80.7%	11.60	4.08
4	84	1.2%	31.0%	10.55	3.38
5	8	0.1%	0%	7.17	3.15
Total	6,789	100%	87.7%	11.40	5.96

[a] Alphabetical order refers to alphabetical ordering of all authors.
[b] No paper in the data set has more than five authors.

Table 7 presents the main findings from these data. The two regressions we report use the average initial of all authors as the dependent variable. As explanatory variables, Regression 1 uses the number of authors and a dummy variable which equals to 1 when the order of authors is non-alphabetical. Regression 2 allows the non-alphabetical ordering to have separate effect for each number of coauthors by including the interaction terms between the number of authors and the non-alphabetical ordering dummy. Coauthors with higher surname initials are, of course, more likely to be listed last in the credits list. We use the average surname initial for *all* coauthors in order to eliminate such effects. Thus, by basing the analysis on the average initial, we do not have to control for the relative position of each coauthor within each particular paper.

Table 7: OLS Regressions: Dependent Variable - Average initial of author/s last name/s

	Regression 1[a]		Regression 2[a]	
	Coefficient	t-stat	Coefficient	t-stat
Two-Author	-0.0152	-0.10	0.0130	0.08
Three-Author	0.0631	0.31	0.0656	0.30
Four-Author	-1.4100**	-3.24	-2.7801**	-4.00
Five-Author	-5.0476**	-4.43	-4.2013**	-3.78
Non-alphabetical	0.8463**	3.40	-	
Two-Author × Non-alphabetical	-		0.5177	1.52
Three-Author × Non-alphabetical	-		0.8337**	2.04
Four-Author × Non-alphabetical	-		2.8306**	3.44
Constant	11.37**	94.69	11.38**	94.72
R^2	0.0048		0.0058	
Number of Obs.	6,789		6,789	

**, * Statistically significant at the 5% and 10% confidence level, respectively.
[a] We correct for heteroskedasticity: the dependent variable is the average initial, so its variance is inversely related to the number of authors.

The results indicate that there is no significant effect on coauthorship patterns among single-authored, double-authored, or triple-authored papers. Moreover, one can test whether the initials of authors of joint papers are independent draws from the distribution of initials of authors of single-authored papers. For two-author and three-author papers we cannot reject the null that the surname initial of authors participating in such papers are independent draws from the distribution of surname initials of single-authored papers. This observation, however, ceases to be true for the four-author and five-author papers in our sample, in which authors with significantly lower initials are more likely to participate. This is shown by the negative and significant effect of the four-author and five-author dummy variables in both regressions of Table 7. It implies that, on average, authors with lower initials are more likely to select themselves into four- and five-author projects. The effect is quite big: the average initial of four- and five-author papers is about half a standard deviation lower than that of other papers. Presumably, this is because authors with higher initials will find themselves consistently as the fourth or fifth listed author, and will not get as much credit for their work. This idea is formalized in the next section.

Finally, we find significant evidence that coauthors with later surname initials are more likely to reverse the order in which coauthors are listed. On average, non-alphabetical ordering is more prevalent in papers authored by economists with higher than average initials. The results of Regression 2 in Table 7 show that while this effect is most significant for four-author papers, it is also present in two- and three-author papers. This relationship cannot be explained by simple ordering of coauthors when they are not equal contributors, unless higher initial coauthors are more likely to be greater contributors, which seems unlikely. In our view, this effect can only be driven by the perceptions of authors that the order of authors is consequential. Thus, this finding by itself suggests that such authors perceive alphabetical discrimination to exist. In light of our previous findings it seems that such perception may indeed have some ground.[9]

It should be noted that while the reported results pool all five journals, the results are qualitatively similar for the *JPE*, *QJE*, and *REStud* when the same regressions are estimated separately for each

[9] To the extent that coauthorships allow an author to write more papers, this effect alone may make the resumes of higher-initial authors shorter. If coauthorships are not sufficiently discounted when making tenure decisions, this may be a mechanism which leads to alphabetical discrimination.

journal. The results for the *AER* are weaker, while *Econometrica* publications reveal no interesting pattern in the dimensions we analyze. All the reported results are fairly robust to the inclusion of a time trend.

3 A Simple Model

In this section we provide a simple benchmark model in an attempt to illustrate more formally the stylized facts suggested by our empirical results. Denote an initial by $x \in [0,1]$, and let $F(x)$ be the cumulative distribution function of available authors, which is assumed to be strictly increasing. Our key assumption is that authors do not decide about coauthorship on a project-by-project basis, depending on the potential coauthor, but make a longer term decision of how much they will invest in searching for joint projects. To simplify, we assume that each author, with an initial x, decides whether to make herself available for the pool of potential coauthors. It is straightforward to extend the model to allow for idiosyncratic shocks to this decision, or to allow for a more continuous strategy space.[10] The basic intuition, however, is given by the simplest model, which we present below.

Writing by oneself yields an expected net value of V, independent of the author's initial. Let $\mathcal{F}$ denote all cumulative distributions on $[0,1]$. Endow $\mathcal{F}$ with the weak* topology, and correspondingly $[0,1] \times \mathcal{F}$ with the product topology. If alphabetical ordering has an effect, entering the pool of coauthors depends on the author's relative alphabetical ranking within the pool of potential coauthors. Formally, let $G(\cdot)$ denote the cumulative distribution of initials amongst the pool of coauthors, which will be determined endogenously, and let $W(x, G(\cdot))$, $(x, G(\cdot)) \in [0,1] \times \mathcal{F}$ be the expected value of entering the pool of available coauthors for an author with an initial x. We assume that W is continuous in its arguments. Furthermore, W is decreasing in x due to the effects implied by alphabetical ordering of coauthors. More precisely, we assume that for any fixed distribution $G(\cdot)$, $x_1 < x_2$ implies that $W(x_1, G(\cdot)) \geqslant W(x_2, G(\cdot))$ and the latter inequality is strict whenever $G(x_2) > G(x_1)$. That is, the value of coauthoring is higher for authors with lower initials. Moreover, when comparing authors with different initials, the expected value from coauthoring is different for these authors if there is a positive probability of encountering a coauthor with an initial falling in between those of the two

[10] A more complete model could also take into account the possibility that effort choices by authors depend on their location in the list of credits by way of equating the corresponding value and cost of the joint project.

(thus, serving as a first author with x_2, and as second author with x_1).

Thus, for any fixed $G(\cdot)$, the author's optimal policy can be defined by a threshold $x(G) \in [0,1]$, so that if authors' initials fall below $x(G)$, they will enter the pool of coauthors, and if their initials fall above $x(G)$, they will author alone. An author of initial $x(G)$ exactly is indifferent between the two alternatives. In particular, for any behavior characterized by a threshold x, the distribution of initials within the pool of coauthors is given by:

$$G_x(y) = \begin{cases} \frac{F(y)}{F(x)} & y \leqslant x \\ 1 & y > x \end{cases} .$$

An *equilibrium* is thus a fixed point of the above distribution, namely a threshold x^* such that $x^* = x(G_{x^*})$.

Denote $G^0(y) \equiv 1$, so that $G^0(y)$ is the atomic distribution that puts all weight on the 0 initial. Note that if all authors choose to enter the pool of coauthors, then the payoff for each author of initial x is $W(x, F(\cdot)) \geqslant W(1, F(\cdot))$. This can be part of an equilibrium if and only if $W(1, F(\cdot)) \geqslant V$. Conversely, suppose that only authors of initial 0 enter the pool of coauthors. The payoff for an author of initial x contemplating becoming a coauthor is $W(x, G^0(\cdot)) \leqslant W(0, G^0(\cdot))$. From continuity, this can be part of an equilibrium if and only if $W(0, G^0(\cdot)) = V$.

Any intrinsic equilibrium, of the form $x^* \in (0,1)$, satisfies $W(x^*, G_{x^*}(\cdot)) = V$. A natural condition for equilibria to be intrinsic is then $W(1, F(\cdot)) < V < W(0, G^0(\cdot))$. Indeed, using the intermediate value theorem, the condition guarantees the existence of at least one such equilibrium. Furthermore, the condition assures the lack of any extreme equilibria of the form $x^* = 0$ or $x^* = 1$. To summarize,

Proposition 1 *Assume that* $W(0, F(y)) < V < W(1, G^0(y))$. *Then there exist equilibria of the authoring game and all are intrinsic.*

As it turns out, a natural condition for uniqueness can also be found. Indeed, suppose further that whenever $H_1, H_2 \in \mathcal{F}$ satisfy $H_1(y) \leqslant H_2(y)$ for all $y \in [0,1]$ then $W(x, H_1(\cdot)) \leqslant W(x, H_2(\cdot))$ for all x. Thus, W is monotonic in its second argument over the class of distributions that can be ordered through first order stochastic dominance. In words, this condition implies that joining the pool of coauthors is more valuable the more likely one is to meet an author of initial higher than hers

within the pool. Assume now that x_1^* and x_2^*, where $x_1^* < x_2^*$, correspond to two distinct equilibria. In particular,

$$W(x_1^*, G_{x_1^*}(\cdot)) = W(x_2^*, G_{x_2^*}(\cdot)) = V.$$

However, note that $G_{x_1^*}(y) \geqslant G_{x_2^*}(y)$ for all y, and $G_{x_2^*}(x_2^*) > G_{x_2^*}(x_1^*)$. From monotonicity,

$$W(x_1^*, G_{x_1^*}(\cdot)) \geqslant W(x_1^*, G_{x_2^*}(\cdot)) > W(x_2^*, G_{x_2^*}(\cdot))$$

in contradiction.

In summary, the simple model illustrates that authors with higher initials are less likely to enter coauthorships if two conditions apply. First, *ceteris paribus*, coauthors with lower initials benefit more from joint projects. Second, this effect is common knowledge among authors. Alternatively, authors with higher initials may enter coauthorships and, more often than others, reverse the order of listed authors. These are the two effects that are consistent with our empirical findings.

4 Conclusions

This paper presented several significant relationships between alphabetical placement and proxies for professional success in the economics labor market. The effects seem to hold even when controlling for many of the attributes that could have potentially created trends in name distributions of economists' surnames (country of origin, ethnicity, religion). Furthermore, the effect pertains to an assortment of success measures that occur in different stages in an economist's career (tenure, Econometric Society fellowship, Nobel Prize and Clark Medal winnings), providing yet another indication that the effects are not spurious.

We suspect that the channel by which the effects are created is the accepted norm in economics of alphabetical ordering of credits in collaborative work. It is essentially the only institutional structure creating asymmetries between market participants with different surname initials. Indeed, we also document a significant relationship between alphabetical placement and participation in multi-authored projects and willingness to deviate from the accepted norm and list authors non-alphabetically. Such patterns are consistent with this channel of "alphabetical discrimination," and suggest that market participants are aware of it and respond. Furthermore, alphabetical placement

seems to have no significant consequences on academic success in psychology, in which publications specify authors predominantly according to their intellectual contribution.

Some possible mechanisms by which the alphabetical ordering norm can produce alphabetical discrimination are:

- Memory: commonly, only the first author is mentioned when referring to a paper with more than two authors (by way of the "et al." addition). The work of first authors, with lower average surname initials, may be easier to remember.
- Attention: the fact that first authors appear first on every mention of their collaborative work, as well as the fact that reference lists are normally ordered alphabetically, may draw attention to authors with lower average surnames. In fact, these sort of influences on attention appear to be heavily exploited in the realm of advertising (the 2003-2004 Los Angeles Westside Yellow Pages reveal more than 450 listed businesses with names containing a seemingly redundant initial A, as in "A-Approved Chimney Services," "A Any Way Bail Bonds," "A Budget Moves," etc.).
- Social Science Citation Index: the book format of the social science citation index references work according to first authors only, creating potential biases in citation counts favoring authors with lower initials. While the online version of the citation index corrects for this by accounting for all authors of the referenced work, this is so only for published work in journals covered by the citation index. For other types of research, such as working papers or books, only first authors are accounted for, so some bias may still exist. Unfortunately, we are not aware of any rigorous study of these possible effects.

We remain agnostic as to which of these (or other) mechanisms are at work. Nonetheless, we maintain that some policy implications may be drawn from the observed effects of alphabetical placements. For example, the economics profession could require the termination of the use of "et al,"[11] the order of citations can be determined by their importance or order of appearance in the text, and credits can be randomized or ordered by contribution (as in most other academic disciplines). At the individual level, economists entering the labor market could change their names.[12]

[11] The current style guidelines of the *AER* require the use of "et al." when referring to a paper by three or more authors.

[12] Indeed, one of us is currently contemplating dropping the first letter of her surname.

References

[1] Bertrand, Marianne and Mullainathan, Sendhil (2004), "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, **94(4)**, 991-1013.

[2] Ellison, Glen (2002), "The Slowdown of the Economic Publishing Process," *The Journal of Political Economy*, **110(5)**, 947-993.

[3] Engers, Maxim, Gans, Joshua S., Grant, Simon, and King, Stephen P. (1999), "First-Author Conditions," *The Journal of Political Economy*, **107(4)**, 859-883.

[4] Gans, Joshua S., editor (2001), *Publishing Economics: Analyses of the Academic Journal Market in Economics*, Edward Elgar Publishing.

[5] Hudson, John (1996), "Trends in Multi-Authored Papers in Economics," *The Journal of Economic Perspectives*, **10(3)**, 153-158.

[6] Laband, David N. and Tollison, Robert D. (2000), "Intellectual Collaboration," *The Journal of Political Economy*, **108(3)**, 632-662.

[7] National Research Council Report (1996), http://www.socialpsychology.org/ranking.htm

[8] Over, Ray and Smallman, Susan (1970), "Citation Idiosyncrasies," *Nature*, **228**, 1357.

[9] Persico, Nicola, Andrew Postlewaite, and Dan Silverman (2004), "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height," *The Journal of Political Economy*, 11**2(5)**, 1019-1053.

[10] Rosenblat, Tanya S. and Mobius, Markus M. (2004), "Getting Closer or Drifting Apart?," *Quarterly Journal of Economics*, **119(3)**, 971-1009.

[11] Shevlin, Mark and Davies, Mark N. O. (1997), "Alphabetical Listing and Citation Rates," *Nature*, **388**, 14.

[12] Thursby, Jerry G. (2000), "What do We Say about Ourselves and What Does It Mean? Yet Another Look at Economics Department Research," *Journal of Economic Literature*, **38(2)**, 383-404.

[13] Zuckerman, Harriet A. (1968), "Patterns of Name Ordering Among Authors of Scientific Papers: A Study of Social Symbolism and Its ambiguity," *American Journal of Sociology*, **74(3)**, 276-291.

Table A1: Departmental Ranking

Ranking	Economics[a]	Psychology[b]
1	Harvard	Stanford
2	Stanford	Michigan
3	UChicago	Yale
4	MIT	UCLA
5	Princeton	Illinois
6	Yale	Harvard
7	Berkeley	Minnesota
8	Penn	Penn
9	Northwestern	Berkeley
10	Minnesota	UCSD
11	UCLA	CMU
12	Columbia	Washington
13	Rochester	Princeton
14	Michigan	Cornell
15	Wisconsin	Wisconsin
16	UCSD	Texas
17	NYU	Columbia
18	Cornell	U Chicago
19	CalTech	Virginia
20	Maryland	Indiana
21	BU	Ohio State
22	Duke	Oregon
23	Brown	Colorado
24	Virginia	Northwestern
25	UNC	UNC
26	U of Washington	UC Irvine
27	MSU	UMass
28	Illinois	Rutgers
29	Washington U (St. Louis)	USC
30	Iowa	Purdue
31	Texas	Rochester
32	Ohio State	Penn State
33	Johns Hopkins	Duke
34	Pittsburgh	NYU
35	Texas A&M	Johns Hopkins

[a] Source: Thursby (2000), Table 1.

[b] Source: National Research Council Report (1996).