Economics Letters 235 (2024) 111541

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Abstract readability: Evidence from top-5 economics journals☆

Belicia Rodriguez [a], Kim P. Huynh [b,1], David T. Jacho-Chávez [a,*], Leonardo Sánchez-Aragón [c]

[a] Department of Economics, Emory University, Rich Building 306, 1602 Fishburne Dr., Atlanta, GA 30322-2240, USA
[b] Bank of Canada, 234 Wellington Ave, Ottawa ON, K1A 0G9, Canada
[c] Escuela Superior Politécnica del Litoral, ESPOL, Facultad de Ciencias Sociales y Humanísticas, Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador

ARTICLE INFO

JEL classification:
A1
C21
C51

Keywords:
Flesch–Kincaid
Readability
Lasso

ABSTRACT

Readability is a measure of how easy it is to read a text. Over time, general-interest journals have become more technical. This affects how accessible research is to a general audience. Our analysis looks at how readable abstracts are. We study the readability of abstracts of top five economics journals between 2000–2019. We collect the characteristics of the abstracts, papers, and authors of these papers. We find that abstracts with higher proportion of women co-authors are more readable. These results are robust to various readability measures and model specifications.

1. Introduction

Readability of a text is an important measure of peer-reviewed publications. In fact, research articles in Economics with higher readability scores are cited more often than papers with lower readability scores, see, e.g., McCannon (2019). Although scientific journals' main audience are often field experts, so authors' writing styles are aimed at them, the readability of scientific texts overall has been decreasing over time, affecting both the accessibility and reproducibility of research results, see Pontus et al. (2017). We look into the characteristics of the co-authors whose papers' abstracts have high readability scores. This research utilizes abstracts because of their understood word limits and structure (average of 6 sentences with a standard deviation of 2 in our sample), and because they are the most likely part of an article that is fully read and understood by potential readers, see, e.g., Anupriya and Karpagavalli (2015).

We find that abstracts of articles published between 2000 and 2019 in the top five (T5) general interest journals in Economics are more readable when either the majority or *all* of the co-authors in a paper are women, after controlling for paper characteristics and subfield of study. These results are robust to different readability measures, model specifications, and estimation routines while accounting for the co-authorship network structure of the data.

Additionally, we discuss the potential mechanisms of our findings, such as the likeliness for papers in macroeconomics and quantitative economics to be less readable and the lower rates of women in these subfields. Women may be less likely to co-author papers in macroeconomics and mathematical economics due to the gender disparities in treatment of economists in the profession. Dupas et al. (2021) show that female presenters are more likely to be asked questions that are neither valuable, nor constructive, nor collegiate during macroeconomic talks compared to talks in microeconomics or finance. Although our findings do not go into the causation for these results on readability, many studies that we discuss in the potential mechanism section dive into the possibilities of why we see a correlation between readability and gender demographics in our study.

2. Bibliometric dataset

Our bibliometric dataset contains 5077 peer-reviewed research articles published between January 2000 and December 2019 in T5 general interest economics journals: *American Economic Review* (AER), *Econometrica* (ECA), the *Journal of Political Economy* (JPE), the *Review of Economic Studies* (RES), and *The Quarterly Journal of Economics* (QJE). Using an Application Programming Interface (API) from IDEAS/RePEc and Scopus, we collected the abstract, names of co-authors, page

☆ We would like to thank the Bank of Canada for allowing us to use their Azure cloud resource, as well as Milca Valdez and Marcelo Ortiz for research assistance.

* Corresponding author.
E-mail addresses: belicia.rodriguez@outlook.com (B. Rodriguez), khuynh@bankofcanada.ca (K.P. Huynh), djachocha@emory.edu (D.T. Jacho-Chávez), lfsanche@espol.edu.ec (L. Sánchez-Aragón).

[1] The views expressed here do not necessarily represent the views of the Bank of Canada.

https://doi.org/10.1016/j.econlet.2024.111541
Received 12 September 2022; Received in revised form 26 December 2023; Accepted 10 January 2024
Available online 11 January 2024

Table 1
Summary statistics.

Variables	Mean	SD	Min	Max	N
F-K grade	15.52	2.69	6.84	55.30	4988
American Economic Review	15.34	2.61	7.50	29.87	1500
Econometrica	15.96	2.59	9.09	30.86	892
Journal of Political Economy	15.31	2.81	7.30	43.70	801
Review Economic Studies	15.65	2.48	9.30	30.88	976
The Quarterly Journal of Economics	15.41	2.98	6.84	55.30	819
Abstract-related variables					
Number of sentences	5.50	1.88	1	24	4988
Number words	122.55	40.15	23	470	4988
Number of syllables	228.42	73.01	52	835	4988
Paper-related variables					
Proportion of women	0.14	0.27	0	1	4988
Only females	0.05	0.22	0	1	4988
Only males	0.74	0.44	0	1	4988
Both gender	0.21	0.41	0	1	4988
Number of authors	2.19	0.94	1	8	4988
Number of pages	34.04	11.80	3	103	4988
Journals					
American Economic Review	0.30	0.46	0	1	4988
Econometrica	0.18	0.38	0	1	4988
Journal of Political Economy	0.16	0.37	0	1	4988
Review Economic Studies	0.20	0.40	0	1	4988
The Quarterly Journal of Economics	0.16	0.37	0	1	4988
Other variables					
Papers not connected	0.11	0.31	0	10	4988
Papers with observed JEL codes	0.63	0.48	0	1	4988

Note: Descriptive statistics such as sample mean (Mean), standard deviation (SD), minimum (Min), maximum (Max), and sample size (n) for all variables are presented here.

numbers, and Journal of Economic Literature (JEL) codes. In addition, we cross-verified this information with the individual T5 journal websites. After excluding non-full length articles, we had 4988 abstracts to use for our analysis. Our analysis does not include short papers, comments, replies, erratum, special issues, and Nobel prize lectures. We used a programming package to surmise the genders of each author based on their first name. In addition, we manually searched the genders of the names the package classified as mostly female, mostly male, or unknown. The proportion of female authors (Share of women), number of pages (Number pages), and number of authors (Number authors) were also extracted from the data.

2.1. Readability measures

We calculate the readability score for each abstract using the Kincaid et al.'s (1975) grade level formula:

$$R_{\text{F-K grade}} = 0.39 \times \frac{W}{S} + 11.8 \times \frac{W_s}{\text{W}} - 15.59,$$

where W, S, and W_s are the total number of words, sentences, and syllables respectively. Since words and sentences are part of the formula, we did not include them as independent variables in our regression analysis below. The formula can be interpreted as either the U.S. grade level, or the number of years of formal education needed to understand the text. Our abstract has a Flesch–Kincaid grade level (F–K grade) of 9.8 for example, i.e., U.S. 9–10th graders are expected to understand it. Table 1 shows that the average abstract in the T5 journals can be read by an individual with some college education. We chose to use this readability measure since the formula is well-known and easy to interpret. In the supplemental section, we have the results that use (Dale and Chall's, 1948) readability scores instead, and the results are qualitatively the same as those here.

2.2. Control variables

Table 1 shows the descriptive statistics. Of the 4884 total authors, 851 are women. Of the 4988 total papers, 1053 were published by gender-diverse co-authors (Both genders), 3689 were published by male-only co-authors (Only males), and 246 were published by female-only co-authors (Only females). On average, women make up 14% of the co-authors in a paper. Among the 63% of papers that self-reported at least one JEL code, 48.7% identified with Microeconomics (D), 21.4% identified with Labor and Demographics Economics (J), and 18.5% identified with Industrial Organization (L). Machine learning techniques (see the supplemental materials) were used to impute at least one JEL codes for articles without self-reported ones (Imputed JEL code). AER and RES combined make up 50% of the papers with abstracts. JPE, QJE, and RES each make up 16%–17% of papers. Using the authors' names and articles, we constructed a co-authorship network, i.e., a paper is classified to be connected to another if they have at least one co-author in common. A total of 214 connected subgraphs (disjoint sets of papers and connections) were identified and used as clusters in our analysis. About 11% of articles do not share co-authorship connections with these clusters (Clusters).

3. Results

Our empirical strategy is based on the following specification for article a:

$$\begin{aligned}\log(\text{F-K grade})_a = {} & \beta_1 \times \log(\text{Number authors})_a + \beta_2 \\ & \times \log(\text{Number pages})_a \\ & + \beta_3 \times \text{Both genders}_a + \beta_4 \times \text{Imputed JEL code}_a \\ & + \theta \times \text{Gender variable}_a \\ & + \text{Journals} + \text{JEL codes} + \text{Cluster} + \text{Years} + \varepsilon_a,\end{aligned} \tag{1}$$

where Gender variable defines a specification, i.e., Female, Share of women, or Male, and Journals, JEL codes, Cluster, and Years are sets of corresponding fixed effects that are estimated along the coefficients multiplying the baseline set of regressors that include the natural logarithm (log) of the number of authors, the log of the total number of article pages, an indicator of whether the co-author team is gender-diverse, and an indicator of whether the paper does not report a JEL code and therefore was imputed. The AER, year 2000, JEL code D (Microeconomics), and the disconnected articles are set as the reference group in all our analysis below.

3.1. Kitagawa–Blinder–Oaxaca Decomposition

shows the Kitagawa–Blinder–Oaxaca Decomposition, see, i.e., Kitagawa (1955), Blinder (1973), and Oaxaca (1973) excluding papers written by mixed-gender authors and using the Female indicator as the Gender variable in (1). The decomposition output reports the mean predictions by gender and their difference in the first panel for both readability measures. The mean of the $\log(\text{F-K grade})$ is 2.731 for males and 2.711 for females, yielding a readability gap of 0.0197. In the second panel of the decomposition output, the readability gap is further divided into three parts. The first part reflects the mean decrease in female-papers' readability if they had the same characteristics as male-papers. The decrease of 0.0125 indicates that differences in endowments account for 63.4% of the readability gap, statistically significant at the 5% level. The second term quantifies the change in female-papers' readability when applying the male-papers' coefficients to the female-papers' characteristics, also significant at the 5% level. The third part is the interaction term that measures the simultaneous effect of differences in endowments and coefficients. The second column of shows the decomposition for $\log(\text{Dale–Chall})$. In this case, both endowments and coefficients decomposition are significant at 5%. In terms of magnitudes, these results show that the differences in readability between genders are driven by the unexplained effect (difference in coefficients) and by differences between endowments (differences in measured mean covariates for men and women) when using both readability measures (F–K grade and Dale–Chall).

Table 2
Summary statistics for JEL codes.

Variables	Observed JEL codes						With JEL codes imputed					
	Mean	SD	n	Females	Males	Both genders	Mean	SD	n	Females	Males	Both genders
A	0.008	0.09	26	0.00	88.46	11.54	0.005	0.07	26	0.00	88.46	11.54
B	0.004	0.07	14	0.00	100.00	0.00	0.003	0.05	14	0.00	100.00	0.00
C	0.212	0.41	662	2.57	76.89	20.54	0.260	0.44	1298	3.24	78.97	17.80
D	0.487	0.50	1521	4.14	73.70	22.16	0.506	0.50	2522	3.93	76.01	20.06
E	0.221	0.41	690	3.33	76.09	20.58	0.208	0.41	1038	3.37	77.46	19.17
F	0.122	0.33	380	5.26	69.21	25.53	0.107	0.31	535	4.67	71.96	23.36
G	0.169	0.37	527	3.98	74.38	21.63	0.162	0.37	810	3.70	74.69	21.60
H	0.133	0.34	416	6.49	66.11	27.40	0.135	0.34	671	5.96	68.11	25.93
I	0.120	0.32	375	9.60	55.47	34.93	0.123	0.33	616	9.90	60.23	29.87
J	0.214	0.41	668	6.14	66.02	27.84	0.221	0.41	1102	7.89	67.60	24.50
K	0.040	0.20	126	9.52	67.46	23.02	0.044	0.20	217	10.60	68.66	20.74
L	0.185	0.39	577	6.07	70.02	23.92	0.186	0.39	929	5.71	71.47	22.82
M	0.038	0.19	118	7.63	61.02	31.36	0.062	0.24	309	7.12	66.34	26.54
N	0.045	0.21	142	9.86	69.72	20.42	0.054	0.23	269	10.78	70.63	18.59
O	0.157	0.36	491	9.78	62.32	27.90	0.164	0.37	816	9.44	66.42	24.14
P	0.024	0.15	74	6.76	62.16	31.08	0.029	0.17	147	6.12	70.75	23.13
Q	0.037	0.19	115	3.48	68.70	27.83	0.038	0.19	192	5.73	70.83	23.44
R	0.057	0.23	177	6.78	75.14	18.08	0.055	0.23	274	8.39	75.91	15.69
Y	0.000	0.02	1	0.00	100.00	0.00	0.000	0.01	1	0.00	100.00	0.00
Z	0.029	0.17	92	4.35	57.61	38.04	0.041	0.20	205	8.29	63.90	27.80

Note: JEL codes stand for the *Journal of Economic Literature* classification codes. Descriptive statistics such as sample mean (Mean) and standard deviation (SD), for all variables are presented here. Columns n, Females, Males and Both genders show the number of/proportion of papers classified by each JEL Code in these categories.

Table 3
Kitagawa–Blinder–Oaxaca Decomposition.

	log(F–K grade)	log(Dale–Chall)
Differential		
Prediction_1	2.7310**	2.4265**
	(0.0047)	(0.0023)
Prediction_2	2.7113**	2.4187**
	(0.0072)	(0.0051)
Difference	0.0197**	0.0077**
	(0.0065)	(0.0035)
Decomposition		
Endowments	−0.0125*	−0.0281**
	(0.0075)	(0.0067)
Coefficients	0.0298**	0.0134**
	(0.0091)	(0.0034)
Interaction	0.0024	0.0225**
	(0.0082)	(0.0045)
N	3935	3935

Note: Clustered standard errors at the network disconnected subgraph level are in parentheses. * p-value < 0.10, ** p-value < 0.05. The reference group is male in baseline specification (1) excluding papers written by both genders. Clustered standard errors at the network disconnected subgraph level are in parentheses.
Standard errors in parentheses.
* $p < 0.10$.
** $p < 0.05$.

Table 4
Double-selection Lasso linear estimation results, JEL codes imputed.

	(1) log(F–K grade)	(2) log(F–K grade)	(3) log(F–K grade)
log(Number authors)	−0.0071	−0.0072	−0.0071
	(0.0044)	(0.0044)	(0.0045)
log(Number pages)	0.0185**	0.0184**	0.0184**
	(0.0033)	(0.0033)	(0.0033)
Both genders	−0.0077*	0.0004	0.0151**
	(0.0045)	(0.0038)	(0.0062)
Female	−0.0209**		
	(0.0058)		
Share of women		−0.0187**	
		(0.0058)	
Male			0.0233**
			(0.0068)
Observations	4988	4988	4988
Number potential controls	257	257	257
Number controls selected	22	22	20
$\chi^2(4)$	38.988	36.726	38.433

Note: Clustered standard errors at the network disconnected subgraph level are in parentheses. * p-value < 0.10, ** p-value < 0.05. The Chi-squared test, $\chi^2(4)$, is a Wald test of the coefficients of the 4 variables of interest shown in each specification jointly equal to zero in each specification.
Standard errors in parentheses.
* $p < 0.10$.
** $p < 0.05$.

3.2. Specification curves

Fig. 1 displays the specification curves (Gao et al., 2021) for the regression analysis with standard errors clustered at the co-authorship sub-graph level. There are a total of 262 parameters including the constant. Panel A displays the point least squares estimates of the coefficient that multiplies a dummy if the paper was written by an only-female research team, and the 95% cluster-robust confidence intervals are presented for 15 different model specifications. In each specification, readability is negatively related with the only-female dummy variable. Since higher scores correspond to more unreadable writing, this indicates that only-female research teams have written easier-to-read abstracts than only-men or mixed-gender co-author teams. Panel B shows that higher proportions of females in a research team are associated with easier-to-read abstracts. Panel C shows that only-male research teams have written more difficult-to-read abstracts than only-female or mixed-gender counterparts. These findings are robust for different readability measures as the dependent variable (see the supplemental materials for further details).

3.3. Sparsity

Given the large number of fixed effects in our baseline specification in (1), we implement (Belloni et al.'s, 2014) Double Lasso regression to conduct robust model selection for our regression. These methods' hyperparameters are chosen by plug-in methods, where uncertainty about what controls to include can be accommodated. Table 4 shows the results. On average the number of years of formal education needed to read an only-female-author article's abstracts is 2.09% lower than mixed-gender or only-male authors, i.e., a 0.209 change in the

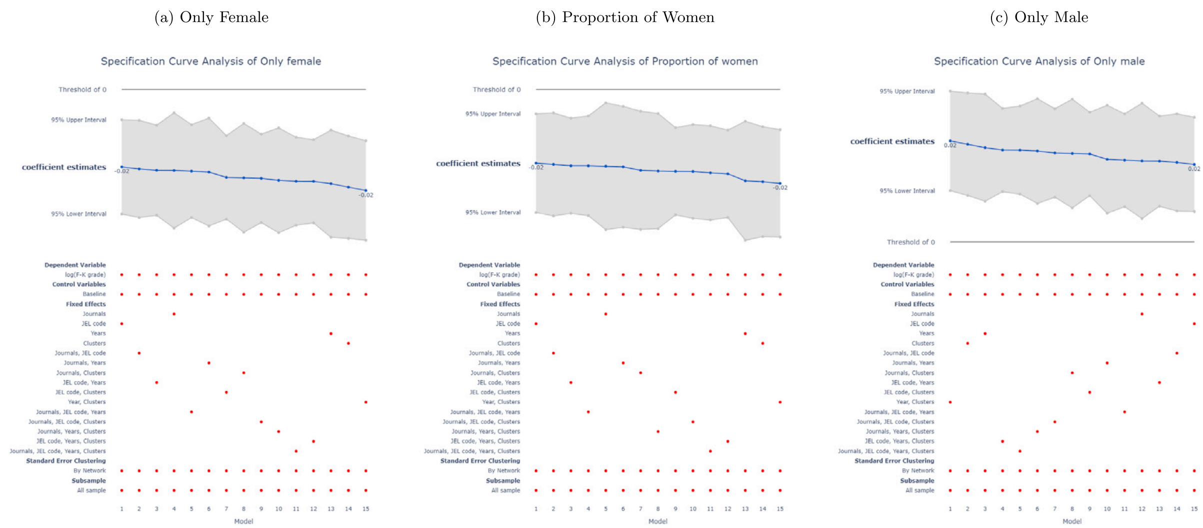

Fig. 1. Specification curves – Flesch–Kincaid reading grade. Note: Baseline specification includes an intercept as well as the natural logarithm of the number of authors and number of pages. It also includes dummies for papers that do not belong to a network subgraph, or do not report any JEL code and therefore imputed. A dummy for a mixed-gender co-author team is also included. Fixed effects include groups of dummy variables for 4 journals (using the AER as its reference), 19 years (using the year 2000 as its reference), 19 JEL codes (using Microeconomics – D as its reference), and 214 network disconnected subgraphs (cluster) membership.

average grade level needed to understand the article's abstract. Similarly, the average number of years of formal education needed to read an only-male-author article's abstract is 2.33% higher otherwise. Adding a female co-author to a solo-authored paper written by a man will decrease the number of years needed to understand its abstract by 1.87%. These results are qualitatively robust to different choices of hyperparameters (cross-validation and adaptive Lasso) and estimators such as double machine learning (Chernozhukov et al., 2018) – see the supplemental materials.

4. Discussion

We have shown that the readability of an abstract improves if there is at least one female co-author on the academic paper. One possible reason for this is that female authors may be less likely to co-author papers from subfields that tend to have abstracts containing mathematical equations, which can affect the readability score. If female co-authors are less likely to co-author papers in quantitative economics or macroeconomics, which both rely heavily on abstract mathematical theories compared to other subfields, then they may have more readable abstracts. Table 2 in supplemental materials provide evidence that abstracts in mathematical and quantitative methods (C) and macroeconomics (E) are two of the five hardest JEL codes to read.

The question stands of why female co-authors are less likely to co-author papers in quantitative economics or macroeconomics. A potential explanation for this phenomena could be seen in the preferences of economic subfields amongst female authors. Beneito et al. (2021) show that undergraduate female students are more likely to outperform undergraduate males in microeconomics, and undergraduate females were more likely to rate microeconomics as more interesting and less technically complex. It may be possible that female economic students retain these preferences for microeconomics in their careers, but we do not explore this possibility in this study.

While we have been discussing the women who have already co-written papers in economics, there are also external factors that influence the participation of women in co-authored papers. Voena et al. (2019) discuss how women and men should strive to spend the same amount of time between work and child care. If women are responsible for more of the child care compared to men, then they may not have enough time to dedicate to research and co-authoring academic papers. In addition, if the field of economics is already suffering from a lower percentage of women, then women may not want to enter the field, potentially seeing the low proportion of women as uninviting.

Note that the examples we give touch on a few of the potential mechanisms for the results of our paper and are not an exhaustive list. There is a growing literature on gender dynamics in economics that dives into other reasons as to why women are less represented in economics and could further explain the results of our paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econlet.2024.111541.

References

Anupriya, P., Karpagavalli, S., 2015. LDA based topic modeling of journal abstracts. In: 2015 International Conference on Advanced Computing and Communication Systems. pp. 1–5.

Belloni, Alexandre, Chernozhukov, Victor, Hansen, Christian, 2014. Inference on treatment effects after selection among high-dimensional controls. Rev. Econom. Stud. 81 (2), 608–650.

Beneito, Pilar, Boscá, José E., Ferri, Javier, García, Manu, 2021. Gender imbalance across subfields in economics: When does it start? J. Hum. Cap. 15 (3).

Blinder, Alan S., 1973. Wage discrimination: Reduced form and structural estimates. J. Hum. Resour. 8 (4), 436–455, URL http://www.jstor.org/stable/144855.

Chernozhukov, Victor, Chetverikov, Denis, Demirer, Mert, Duflo, Esther, Hansen, Christian, Newey, Whitney, Robins, James, 2018. Double/debiased machine learning for treatment and structural parameters. Econom. J. 21 (1), C1–C68.

Dale, Edgar, Chall, Jeanne Sternlicht, 1948. A formula for predicting readability. Educ. Res. Bull. 27 (1), 11–20.

Dupas, Pascaline, Sasser Modestino, Alicia, Niederle, Muriel, Wolfers, Justin, Collective, The Seminar Dynamics, 2021. Gender and the Dynamics of Economics Seminars. Working Paper 28494, National Bureau of Economic Research.

Gao, Mingze, Leung, Henry, Qiu, Buhui, 2021. Organization capital and executive performance incentives. J. Bank. Financ. 123, 106017.

Kincaid, J., Fishburne, R.P., Rogers, R.L., Chissom, B.S., 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Tech. Rep. Research Branch Report 8–75, Chief of Naval Technical Training: Naval Air Station Memphis.

Kitagawa, Evelyn M., 1955. Components of a difference between two rates. J. Amer. Statist. Assoc. 50 (272), 1168–1194, URL http://www.jstor.org/stable/2281213.

McCannon, Bryan C., 2019. Readability and research impact. Econom. Lett. 180, 76–79.

Oaxaca, Ronald, 1973. Male-female wage differentials in urban labor markets. Internat. Econom. Rev. 14 (3), 693–709, URL http://www.jstor.org/stable/2525981.

Pontus, Plavén-Sigray, Matheson, Granville James, Schiffler, Björn Christian, Thompson, William Hedley, 2017. The readability of scientific texts is decreasing over time. eLife 6.

Voena, Alessandra, Goldin, Claudia, Guerrieri, Veronica, 2019. URL https://www.chicagobooth.edu/review/why-are-there-so-few-women-economists.