Economics Letters 173 (2018) 100–103

Contents lists available at ScienceDirect

Economics Letters

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

Easy to read, easy to cite?

Michael Dowling, Helmi Hammami, Ousayna Zreik *

Rennes School of Business, France

HIGHLIGHTS

- Readability of Economics Letters articles is positively related to future citations.
- This finding is robust to testing across subcategories.
- Methods and macroeconomic papers particularly benefit from good readability.

ARTICLE INFO

Article history:
Received 18 June 2018
Received in revised form 21 September 2018
Accepted 25 September 2018
Available online xxxx

JEL classification:
A1
B4

Keywords:
Readability
Research impact
Bibliometrics
Economics letters

ABSTRACT

Ease of readability of *Economics Letters* abstracts, and number of works cited in an article, is positively related to future citations. Readability appears to particularly matter for mathematical and quantitative methods and macroeconomics papers, while number of works cited is generally important across all articles.

1. Introduction

McCloskey (1985, 1999) eloquently and persuasively argues for the benefits of a clear and concise writing style in the expression of economic ideas. While the public benefits of clear economic communication are self-evident, there has not been a large scale investigation of whether there are benefits from clear writing in terms of creating impact from economic research publications. We address this question with an analysis of the relationship between readability and citations for articles published in *Economics Letters* (EL).

Our analysis is based primarily on a range of formal readability measures. For each abstract published in EL from 2003–2012 (3229 regular articles) we calculate the Flesch Reading Ease Score (FRES), Gunning Fog Index (FOG), and SMOG Index (SMOG) (Bailin and Grafstein, 2016). These measures are widely-applied approaches to measuring how complex a piece of writing is to read.

To these readability measures we also add a count of works cited (references) per article, excluding self-citations and reciprocal citations.[1] Prior estimates suggest that up to 40% of references in a paper are merely perfunctory, with the remaining references used mainly for constructive argumentative functions such as justifying the research question and building theory (Swales, 1986). As EL has a policy of 2000-word maximum article length including references this provides a strong disincentive towards perfunctory referencing. This therefore suggests that the use of more references in an article is for constructive argumentation reasons and therefore to enhance readability and understandability. We thus incorporate the count of works cited as an alternative whole-article perspective on readability, while acknowledging that this is a noisy measure and keeping our main focus on the pure readability measures.

Our work is the first since Laband and Taylor (1992) to analyze the relationship between readability and impact in economic research. This prior study did not use formal readability measures for their core tests,[2] just investigated 311 articles, and analyzed

* Correspondence to: Rennes School of Business, 2 rue Robert d'Arbrissel, 35065 Rennes, France.
E-mail address: ousayna.zreik@rennes-sb.com (O. Zreik).

https://doi.org/10.1016/j.econlet.2018.09.023
0165-1765/

[1] Reciprocal citation is defined as subsequent citation by an author cited in an article's references (Sivadas and Johnson, 2005).

[2] Although Footnote 9 of their paper notes that they ran unreported tests on the relationship between a formal readability score and subsequent citations. The

Table 1
Descriptive statistics.

Panel A: Citations	
Mean	3.02
Median	1.00
Standard deviation	6.07
Kurtosis	147.40
Skewness	9.18
No citations %	31.99%
Citations top 10%	50.15%
Total citations	9,743
Number articles	3,229
Panel B: Readability	
Abstract: FRES	42.46 (15.24)
Abstract: FOG	16.08 (3.97)
Abstract: SMOG	14.20 (2.71)
Article: References	10.65 (5.25)
Panel C: JEL categories	
JEL_C	40.42%
JEL_D	32.58%
JEL_E	20.00%
JEL_F	10.81%
JEL_J	13.01%
Other	44.26%

Notes: Panel A: Descriptive statistics for citations to all regular articles published in the *Economics Letters* journal during 2003–2012. Citation data covers citations for the five years following year of publication. Panel B: Means and standard deviations (in brackets) for the three readability scores as defined in Section 2. References variable is a count of the number of references per article excluding self-citations and reciprocal citations. Panel C: Division of articles published by JEL general category as defined in Section 2. Articles with JEL codes in more than one category are counted for every category.

Table 2
Overall tests of readability and citations.

Probit: No Cite v Cite	FRES	FOG	SMOG
FRES	0.002		
	(0.002)		
FOG		−0.013**	
		(0.006)	
SMOG			−0.015*
			(0.009)
References	0.024***	0.024***	0.024***
	(0.005)	(0.005)	(0.005)
Constant	0.138	0.432***	0.438***
	(0.119)	(0.140)	(0.161)
LR chi-square	51.47***	54.77***	53.14***
Negative binomial: Count of cite			
FRES	0.003**		
	(0.002)		
FOG		−0.017***	
		(0.006)	
SMOG			−0.021**
			(0.010)
References	0.029***	0.029***	0.039***
	(0.005)	(0.005)	(0.007)
Constant	0.805***	1.230***	1.248***
	(0.123)	(0.154)	(0.178)
LR chi-square	68.03***	70.86***	68.75***
OLS: asinh of count of cite			
FRES	0.001		
	(0.001)		
FOG		−0.010**	
		(0.005)	
SMOG			−0.012*
			(0.007)
References	0.029***	0.029***	0.029***
	(0.004)	(0.004)	(0.004)
Constant	0.871***	1.108***	1.106***
	(0.096)	(0.114)	(0.131)
F	8.10***	8.40***	8.22***

Notes: Table reports probit tests of whether or not an article is cited over the subsequent five years, negative binomial tests of count of citations, and OLS tests of inverse hyperbolic sine (asinh) of count of citations. Number of articles is 3229 and year fixed effects are included. Dataset and variables as defined in Section 2. Standard errors in parentheses.
*$p < 0.10$.
**$p < 0.05$.
***$p < 0.01$.

papers from a prior generation of economics writing. Our approach though has been recently successfully applied in other disciplines (Dolnicar and Chapple, 2015; Lei and Yan, 2016). There has also been some limited analysis of the readability of economic discourse (Diamond and Levy, 1994; Jansen, 2011) and to analyze the differences between male and female economists' writing styles in abstracts of top-ranked economics journals (Hengel, 2017).[3] This literature on which we build suggests that, for academic publications, the relationship between readability and impact is mixed at best, and has at times shown that less readable articles can receive more citations.

Indeed an early critique of McCloskey (1985) by High (1987) argued that "writing well misdirects time and labor" and that "time spent on writing is time taken away from reading another article or running another regression" (p. 544). A readability analysis of the management literature by Armstrong (1980) even argued that greater writing complexity (or 'unintelligible management' as Armstrong termed it) bestowed greater prestige on the academic within the academic community. The remainder of this article thus investigates if the unfortunate Armstrong perspective or the positive McCloskey message holds more weight in determining impact for economic research.

2. Methodology

We select abstracts through *Scopus* from all 3229 regular publications in EL for the years 2003–2012. While abstracts are just a small portion of an article, they tend to be the most widely read, and there is generally consistency between abstract writing style and the writing style of other sections of a paper (Hartley et al., 2003).

Citation information on articles is obtained from *Scopus* for the following five years after the year of publication (thus, up to and including 2017 for 2012 articles). The articles in our dataset attracted 9,743 citations excluding self-citations and reciprocal citations. The average five-year citation rate per article is quite high at 3.02, but the median number of citations is much lower, and 32% of articles received no citations. This skewness is also evident at the other end with the top 10% of articles receiving 50% of all citations. See Table 1 (Panel A) for further citation descriptive statistics.

Three readability measures are calculated.[4] FRES is calculated as: 206.835–1.015 x (average words per sentence) - 84.6 x (average syllables per word). FOG is calculated as:

$$FOG = 0.4\left[\left(\frac{words}{sentences}\right) + 100\left(\frac{complex\_words}{words}\right)\right],$$

sample was the first page of 58 articles from *Review of Economics and Statistics* published in 1978, and no relationship is found. While this attests to the comprehensive insight of Laband and Taylor (1992), it is perhaps not too surprising that a very limited sample of quite uniform articles would struggle to show significance.

[3] Hengel (2017) incorporates citations, but not as a dependent variable, as it is not the focus of her research.

[4] We use the package *Textatistic* in Python (https://pypi.org/project/textatistic/) to construct these measures, and manually cross-check a sample of output to ensure accuracy.