INNOVATION OVERLOAD?

Byoung-Hyoun Hwang, Hugh Hoikwang Kim and Kai Wu*

This Draft: January 2020

We conjecture that some individuals are increasingly overwhelmed by the myriad of new ideas proposed to them each year. As a result, even minor obstacles to understanding an innovation causes such innovation to not be adopted. To test this conjecture, we turn to ideas proposed in scientific journal articles. We find that something as seemingly frivolous as how difficult to read a journal article is strongly negatively predicts an article's subsequent number of citations. This negative predictability spills over to patent descriptions and patent forward citations.

JEL Classification: A10, M30.

Keywords: Innovation Diffusion, Bounded Rationality.

* Hwang is affiliated with the Cornell SC Johnson College of Business, Dyson School of Applied Economics and Management, Cornell University, Warren Hall, Ithaca, NY 14853. Kim is affiliated with the Darla Moore School of Business, University of South Carolina, 1014 Greene Street, Columbia, SC 29208. Wu is affiliated with the Central University of Finance and Economics, Beijing, China. E-mail: bhwang@cornell.edu, hugh.kim@moore.sc.edu and wukai8759@cufe.edu.cn. The authors would like to thank Richard Wu, who provided excellent research assistance and Alex Butler, Clifton Green, David Hirshleifer, Craig Holden, Soo Kim and Siew Hong Teoh for helpful comments.

1. Introduction

Innovations are an important source of a country's economic growth and, by some accounts, even the most important driver (Rosenberg 2004, OECD 2015). Factors that have been argued to stimulate the creation of innovations include, among others, product market competition, regulatory environment, managerial incentive structure, behavioral traits of managers, and sources of financing available (Hall and Lerner 2010; Ederer and Manson 2011; Kerr and Nanda 2015; He and Tian 2018).

Stimulating innovation is only part of the equation however as any innovation – no matter how useful – can only have economic impact when widely adopted. A long line of work rooted in history, sociology and economics examines the factors that potentially hinder the diffusion of seemingly useful products and technologies (Rogers 1995; Hall 2005).

In this paper, we point to a possibly critical conundrum. The number of new ideas proposed each year has reached an unprecedented level. The STM report notes that there are "*about 33,100 active scholarly peer-reviewed English-language journals in 2018, collectively publishing some 3 million articles a year*" (page 25). The report notes further that the annual number of journal articles has been growing more than exponentially. Similarly, the US Patent and Trademark Office (USPTO) notes that it granted 339,992 patents in 2018. As scientific journal articles, the annual number of patents granted has been rising more than exponentially.

People have limited capacity for gathering and processing information (Simon 1955). With the number of innovations to be considered each year so high, it appears plausible that at least some potential adopters are feeling overwhelmed. In this study, we speculate that, in such a world of "innovation overload," at least some individuals stop learning about an innovation if there are even minor inconveniences in understanding an innovation. This in turn hampers the diffusion of potentially important ideas. We hereafter refer to this possibility as the "innovation overload hypothesis."

To test the innovation overload hypothesis, we turn to innovations proposed in scientific journal articles. Scientific journal articles represent the primary forum through which advances in the sciences are reported and discussed. As we describe in Section 3, the construction of many of our variables is labor-intensive. We therefore cannot consider all articles published in all scientific journals. Instead, we focus our analysis on journal articles published in the field of financial economics (or, simply, finance).[1] We do not believe the observations we make in this paper are particular to finance, however, and we later also provide descriptive statistics for articles published in the fields of economics and other areas of business (accounting, management, marketing, operations and information systems).

We download all papers that were published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. As our measure of innovation diffusion we use the number of Google Scholar citations that the relevant article garners as of September 2016.

As a potential source of minor inconveniences in understanding an innovation we take the readability of the article in which the innovation is couched. We apply a copy-editing software to each journal article and search for the presence of writing faults that prior literature argues lowers the readability of a text. We then construct our readability measure as the number of writing faults per one hundred words. Writing faults include the use of passive verbs, hidden verbs, complex words, abstract words, overused words and clichés, legal words, wordy phrases, overwriting, foreign words and long sentences. The main body of the text describes and presents examples of each writing fault; it also discusses the relevant literature.

Consistent with the innovation overload hypothesis, when relating the readability of scientific journal articles to their subsequent number of citations within a regression framework, we find that papers of lower readability receive substantially fewer citations than papers of higher readability. The economic significance of our predictability is substantial: our estimates imply that a one standard deviation increase in the number of writing faults per one hundred words comes with 7% fewer citations. Our results easily survive the inclusion of various controls, such as years-since-publication, author affiliation, number of conference and seminar presentations, whether the paper won a best-paper award, number of authors,

[1] We choose finance because we, the authors of this study, all work in finance departments and, as such, are familiar with the publication process in the finance area.

whether the paper is a theory paper, title length, subfield of the paper (e.g., “Portfolio choice and investment decisions,” “Bankruptcy and litigation”) and journal- and Journal-of-Economic-Literature (JEL) fixed effects.

While we believe our results are most naturally explained by the innovation overload hypothesis, there are other possible interpretations of our results. In particular, it might be that “high-quality innovators” not only make discoveries of greater impact, but they also communicate them more clearly (“quality hypothesis”). Relatedly, innovations that are inherently more complex may simply have to be couched in language that is more difficult to process (“complexity hypothesis”).

We cannot rule out the quality- or the complexity hypothesis. But one piece of evidence against the quality- and the complexity hypotheses is that the correlations between our measure of readability and two indicators, which likely capture the quality and complexity of an innovation: winning a prestigious research award and being a theory paper, are virtually zero. That is, it is simply not true in the data that award-winning papers have fewer writing faults than non-award-winning papers or that theory papers are couched in more-difficult-to-process language than their empirical counterparts.

Another piece of evidence against the quality- and the complexity hypotheses is that our readability effect becomes much stronger in the second half of our sample period. The annual number of articles published in finance journals is noticeably higher in the second half of our sample period. The innovation overload hypothesis thus predicts that our readability effect become stronger. Neither the quality- nor the complexity hypothesis shares this prediction.

To assess the generalizability of our finding, we extend our analysis to patents. Patents are described mostly in sketches, but still consider a non-material amount of text. We construct a representative sample of patents and relate the readability of patent descriptions to their number of forward citations within a regression framework. Similar to our findings based on innovations proposed in scientific journal articles, we find that the readability of a patent description strongly predicts a patent’s number of forward citations. As with scientific journal articles, this predictability strengthens substantially as the number of patents to be considered each year increases.

2. Background and Hypothesis

The question of what prevents a seemingly useful innovation from being adopted widely has been examined from two perspectives: (1) a historical, sociological perspective, and (2) an economic perspective. The historical, sociological perspective argues that the level of adoption depends on what an innovation's perceived benefit is, how compatible an innovation is with a potential adopter's social norm, how complex an innovation is, and whether an innovation can easily be tested and evaluated before full adoption (Rogers 1995). As few potential adopters operate in a vacuum, the diffusion of innovation also depends on certain social conditions, such as whether an innovation was first heard about via mass media or word-of-mouth and what the social structure is among potential adopters. Rogers uses the historical, sociological framework to explain why certain innovations, such as the use of contraceptives or the boiling of water prior to consumption, diffused in some groups but not in others.

In comparison, the economic perspective views the diffusion of innovation as the sum of "individual calculations that weigh the incremental benefits of adopting a new technology against the costs of change" in the presence of uncertainty (Hall 2005). Regarding the costs of adopting a new technology, the literature emphasizes that most of the cost comes not from acquiring the technology per se, but from the difficulty of learning how to use the new technology and any complementary investment required (Bresnahan, Brynjolfsson and Hitt 2002). Other factors that the economics literature notes affect the level of adoption are how uncertain the benefits and costs of change are, the regulatory environment, the prevailing culture, and the overall market structure (Hall).

Our hypothesis is most closely related to what the historical, sociological perspective refers to as "complexity" and what the economic perspective refers to as the difficulty of learning how to use a new technology.

Because scientists produce hundreds of journal articles each year, staying current, even in one's "narrow" area of expertise, represents a non-trivial task. Given the limited amount of attention and time the intended audience can spend on each journal article, we speculate that at least some in the audience stop reading about a new idea if the corresponding article is filled with writing faults and, as a result, becomes difficult to process. Any new ideas described in such articles thus do not become adopted and cited widely.

Hypothesis: *Innovations proposed in scientific journal articles that are filled with writing faults are adopted and cited less widely.*

The null hypothesis, which, ex ante, appears equally plausible to us, is that new ideas proposed in scientific journal articles are evaluated by experts, who presumably are highly trained in processing journal articles. Writing faults, which, at most, cause only minor inconveniences in learning about an innovation thus are too frivolous to meaningfully affect the degree of innovation diffusion. On this view, we should observe no link between our measure of readability and our measure of innovation diffusion.

3. Data and Key Variables

3.1. Measure of Innovation Diffusion

Our sample comprises 2,618 papers from 2005 through 2014, 716 of which are from the *Journal of Finance*, 1,048 of which are from the *Journal of Financial Economics*, and 854 of which are from *The Review of Financial Studies*. For each of these papers, we assess the degree of innovation diffusion by manually searching the number of citations in Google Scholar (https://scholar.google.com/intl/en/scholar/about.html) via title, authors, and year of publication. We collect the number of citations as of September 16–20 2016, which should ensure that differences in citations do not reflect differences in points of data collection. We only consider articles published from 2005 through 2014 (as opposed to articles published from 2005 through 2016) to give each article some time to diffuse. As shown in Table 1, the average paper in our sample generates 204 citations. The median number of citations is 115; the 10^{th} and 90^{th} percentiles are 25 and 465, respectively.

3.2. Measure of Readability

The development of our readability measure is couched within a growing body of work in accounting and finance that examines how financial market participants respond to disclosure documents when the text is difficult to process. Lawrence (2013) and Elliott, Rennekamp and White (2015) provide evidence that investors shun firms whose disclosure documents are difficult-to-read. Hwang and Kim (2017) go one step further and argue that the associated reduction in investor demand causes such firms to trade at substantial discounts relative to their fundamentals.

Our measure is most similar to that adopted in Hwang and Kim (2017). We save each article as a separate Microsoft Word document. We then use a program called StyleWriter, a manuscript editor that, once installed on a computer, searches Word documents for "writing faults."

The writing faults are: the use of passive verbs, hidden verbs, complex words, abstract words, overused words and clichés, legal words, wordy phrases, overwriting, foreign words and long sentences. Appendix Table 1 provides examples of each writing fault; Appendix Table 1 also provides possible corrections to each writing fault.

Our readability measure, *Readability*, is the number of occurrences of the above writing faults, scaled by the number of words and multiplied by (100) and (-1).

$$Readability = \frac{\sum_{i=1}^{10} WritingFaults_i}{\#Words} \times (100) \times (-1). \qquad (1)$$

Multiplying by one hundred later helps us interpret the coefficient estimates. We multiply by negative one so that higher readability scores imply more easily readable documents.

To assess the validity of our measure, we randomly assign finance PhD students the introduction sections of articles that, as per our measure, earn "low readability" scores. We repeat this procedure with introduction sections of articles that earn "high readability" scores. As discussed further in Online Appendix Figure 1 and Online Appendix Table 1, we find that students largely agree with the output generated by our readability measure, as they perceive introductions with low readability scores to be significantly more difficult to read than those with high readability scores.

Table 1 shows that there is great variation in our readability measure in our sample. The 10th and 90th percentiles for *Readability* are -7.5 and -4.8; the mean is -6.13. The mean of -6.13 implies that, on average, there are 6.1 writing faults for every one hundred words. For reference, the *Readability* of this paper is -5.6, which puts this paper in the 75th percentile.

In additional analyses, we compare the readability of articles published in finance journals to the readability of articles published in (general) economics as well as other areas in business. As (general) economics "A-level" journals, we include the *American Economic Review*, the *Journal of Political Economy* and the *Quarterly Journal of Economics*. Our list of A-level business journals is that of the *Business School Research Rankings™* compiled by the University of Texas (Dallas) (http://jindal.utdallas.edu/the-utd-top-100-business-school-research-rankings). Since the computation of *Readability* is labor-intensive, we consider only the 1,752 articles in the above journals that were published in 2014, the last year of our sample period.

Online Appendix Table 2 reveals strong differences across journals and fields. Economics journals fare well with the average *Readability* ranging from -6.02 for the *Journal of Political Economy* to -6.12 for *The Quarterly Journal of Economics*. Accounting and marketing journals also fare relatively well with the quantitative journals having slightly fewer writing faults than the more psychology-based journal. Management journals tend to have lower readability scores. The journals with the lowest readability scores however are in the fields of Operations and Information Systems (*Information Systems Research* = -8.33; *Journal of Operations Management* = -8.71).

3.3. Other Article and Author Characteristics

While innovations proposed in finance journal articles represent a relatively homogeneous pool, there remain important differences in quality as well as differences in article and author characteristics that likely affect citation counts outside of the readability channel.

In an attempt to control for these differences, we include the following variables in our regression analysis:

(1) The number of years since publication as of 2016, *Years since Publication*. One of the most established patterns in the innovation diffusion literature is that when the number of adoptions is plotted against the number of years the corresponding innovations has been present, the plot follows an S-shape (Hall 2005). That is, the rise in the number of adoptions proceeds slowly at first, accelerates as it spreads throughout the population, and then slows down again as the innovation reaches its saturation point. To account for the presence of such an S-shaped pattern in our setting, we also include a squared term of *Years since Publication*.

(2) A relative ranking of the authors' affiliations based on publication numbers, *Affiliation Ranking*. We turn to the *Finance Research Ranking* compiled by the Arizona State University (http://apps.wpcarey.asu.edu/fin-rankings/rankings). The *Finance Research Ranking* counts for each institution the number of publications in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies*; it runs from 1990 through the present. To avoid look-ahead bias, we save the Top 50 ranking as of 2004, the year before our sample begins. To facilitate interpretation, the number one institution receives a score of 50, the number two institution receives a score of 49, and so on. Any institution that is not represented in the Top 50 receives a score of zero. We compute the average score across the institutions with which the authors of the paper in question are primarily affiliated with.[2]

(3) The number of conferences and seminars the paper has been invited for presentation prior to publication, *Number of Presentations*.

(4) Whether the article won a "best-paper award," *Award Paper*. Each year, the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies* award best-paper prizes to the best published articles in that year. Our *Award Paper* variable equals one if the relevant article won such a prize, and zero otherwise.[3]

---

[2] Taking the score of the highest ranking institution instead of the average score does not noticeably alter any of our results (results available upon request).

[3] The awards are the Amundi Smith Breeden Prize and the Brattle Group Prize for the *Journal of Finance*; the Jensen Prize and the Fama-DFA Prize for the *Journal of Financial Economics*; and the Michael J. Brennan Best Paper Award for *The Review of Financial Studies*.

(5) The number of authors, *Number of Authors*.

(6) Whether the article is a theory paper, *Theory Paper*. *Theory Paper* equals one if the corresponding text contains the term "proof" along with either one of the following terms: "proposition," "theorem," "lemma," "corollary".

(7) The length of the title in words, *Length of Title*.

(8) The number of JEL codes, *Number of JEL Codes*. For papers published in the *Journal of Finance*, which do not have JEL codes, the *Number of JEL Codes* variable is set at zero. In our regression analysis, we include journal-fixed effects along with fixed effects based on the article's first JEL code.

Table 2 presents a correlation matrix. Table 2 shows that, as expected, *Affiliation Ranking*, *Number of Presentations* and *Award Paper* are all significantly positively correlated with *Citations*. *Theory Paper* is significantly negatively correlated with *Citations*. Importantly, none of these measures reliably associate with *Readability*. That is, there are no systematic differences in the occurrence of writing faults between "high-affiliation-rank" papers and "less-high-affiliation-rank" papers, between papers that have been presented many times prior to publication and papers that have been presented fewer times, between award-winning papers and non-award winning papers, and between theory papers and non-theory papers. This finding does not lend support to the notion that innovations of greater quality or greater inherent complexity systematically come with fewer or more writing faults.

4. Readability and the Diffusion of Innovation through Scientific Journal Articles

To quantify the effect of readability on innovation diffusion, we estimate the following regression equation:

$$Y_i = \alpha + \beta\, Readability_i + \delta' X_i + \varepsilon_i. \tag{2}$$

where *i* indexes an article. The dependent variable is the natural logarithm of the number of citations. We take the natural logarithm since *Citations* is highly right-skewed. *X* includes our explanatory variables described above. As alluded to earlier, we also include journal fixed effects and fixed effects based on the paper's first JEL code. *T*-statistics are based on standard errors adjusted for heteroscedasticity and clustered by the first JEL code.

We present our regression results in Table 3. Depending on the set of controls, the coefficient estimate for *Readability* reported in Columns 1 through 3 ranges from 0.063 (*t*-statistic = 3.83) to 0.083 (*t*-statistic = 5.40). Our regression analysis thus indicates that, holding all else equal, one less writing fault per one hundred words is followed by ~7% more citations. This result is consistent with the notion that making it easier to learn about an innovation facilitates an innovation from being adopted, or, put in analogue form, making it more difficult to learn about an innovation hinders an innovation's level of diffusion.

In further tests, we assess whether readability affects the *rate* or the *degree* of innovation diffusion. That is, do more-inconvenient-to-learn innovation simply take longer to become adopted, but, in the end, achieve the same level of adoption as their less-inconvenient-to-learn counterparts, or, is the number of adoptions permanently lower for the more-inconvenient-to-learn innovations?

To speak to this question, we include an interaction term between *Readability* and *Years since Publication*. If the effect of readability on innovation diffusion is temporary and does not influence the long-term number of adoptions, the negative impact of *Readability* on the number of citations should weaken over time. In other words, as values for *Years since Publication* become larger, the magnitude of the estimated coefficient for *Readability* should become smaller. The coefficient estimate for the interaction term should therefore be negative and statistically significant.

Our results reported in Column 4 of Table 3 show that the coefficient estimate for the interaction term is not reliably negative, implying that the negative effect of low readability is permanent. In fact, since the impact of *Readability* on the percentage change in the number of citations is similar across *Years since Publication*, we can infer that the difference in raw citation numbers between high- and low-readability papers widens over time due to compounding.[4] Online Appendix Figure 2, which plots the predicted

[4] Put differently, let *X* be the citation number of articles with high readability and *Y* be the citation number of articles with low readability. Given the insignificant estimate for the interaction term, we can infer that log (X) – log(Y) = constant (>0) over time, which implies that (X)/(Y) = constant (>1) over time. This, in turn, implies that (X) – (Y) increase over time as *X* and *Y* rise.

number of citations of high-readability papers versus low-readability papers over time, further illustrates this point.

In Columns 5 and 6 of Table 3, we explore whether our predictability is stronger in the second half of our sample period, i.e., from 2010 through 2014, than in the first half of our sample period, i.e., from 2005 through 2009. The annual number of articles published in our three finance journals is ~20% higher in the second half of our sample period. To potential adopters, it has thus become more challenging to carefully consider each article. In line with this argument, we find that the estimate for *Readability* increases from 0.053 (*t*-statistic = 2.88) in the first half to 0.096 (*t*-statistic = 4.55) in the second half. We do not think that either of the alternative hypotheses, i.e., the quality hypothesis and the complexity hypothesis, can account for the observed strengthening of the predictability.

We conclude this section with a few notes on the coefficient estimates for the control variables. The estimates are similar across our five columns. Here, we discuss the estimates reported in Column 3. The coefficient estimate for squared *Years since Publication* is negative and statistically significant suggesting that a plot of number of citations against number of years since publication, indeed, follows an S-shaped pattern. The number of authors on a paper strongly positively associates with the number of citations (coefficient estimate = 0.115, *t*-statistic = 5.98). One possible explanation for this finding is that co-authors help improve the quality of the work. An alternative perspective is that co-authors help raise awareness of a paper through their own personal networks (Kerr, 2008). The coefficient estimate for *Affiliation Ranking* is 0.009 (*t*-statistic = 8.55), suggesting that a ten-rank difference in the average author's *Affiliation Ranking* comes with 9% higher subsequent citation counts. The estimate for *Number of Presentations* is 0.024 (*t*-statistic = 7.58), suggesting that presenting the paper one more time at a conference or a university prior to publication increases subsequent citation counts by 2.4%. Not surprisingly, receiving an award strongly and positively contributes to citation counts. Our estimate implies that, holding all else equal, award-winning papers, on average, receive 46.4% more citations than non-award papers. Theory papers receive 45.4% fewer citations (*t*-statistic = -10.78) perhaps as innovations in theory papers are inherently more complex and more complex innovations diffuse less widely. Interestingly, the length of an article's title negatively associates with the number of citations. One possible explanation is that innovations described with short titles are broader.

5. Readability and the Diffusion of Innovation through Patents

Scientific journals represent only one channel through which innovation are disseminated. Patents represent another vehicle and in this section we explore whether the observations made for scientific journal articles carry over to patents.

We crawl all utility-patent-filing documents in html format from the USPTO website (http://patft.uspto.gov/netahtml/PTO/srchnum.htm). We remove all patents granted prior to 1976 because filing documents are available as high-quality text files only starting in 1976 and we require high-quality text files to construct our readability measure. We merge our data with the patents data provided by Kogan, Papanikolaou, Seru, Stoffman (2017), which runs from 1926 through 2010.[5] As the construction of our readability measure is labor-intensive, we randomly select 1% of these patents. In the end, we have 12,851 patents granted between 1976 and 2010 with data available for all our dependent and independent variables.

Our regressions are similar to the ones presented in Table 3 for scientific journal articles:

$$Y_i = \alpha + \beta\, Readability_i + \delta' X_i + \varepsilon_i. \tag{3}$$

where *i* indexes a patent. The dependent variable is the natural logarithm of the number of forward citations received by a patent as described in other patents' filing documents through 2010. We take the natural logarithm since *Patent Citations* is highly right-skewed. *Readability* is the number of writing faults in a patent description per 100 words, multiplied by (-1). *X* includes our explanatory variables: *Years since Granting*, which is the number of years since a patent has been granted (as of 2010); *Economic Value of Patent*, which is the estimated value of a patent based on the stock market reaction to the corresponding patent's granting, scaled by 100; *Firm-Level Innovation Value*, which is the *Economic Value of Patent* aggregated to the firm-level, over the corresponding firm's book value, scaled by 1,000; and *Firm-Level*

[5] https://iu.app.box.com/v/patents

Number of Patents, which is the number of patents granted to the relevant firm as of 2010, scaled by 100. *Patent Citations*, *Economic Value of Patent* and *Firm-Level Number of Patents* are all from Kogan, Papanikolaou, Seru, Stoffman's (2017) dataset.[6] We include USPTO three-digit technology class fixed effects to account for potentially confounding effects that technology-field specific characteristics have on patent citations. *T*-statistics are based on standard errors adjusted for heteroscedasticity and clustered at the technology class-level.

Our regression results presented in Table 4 suggest that patents filed with more readable descriptions receive more citations than similarly valued counterparts with less readable descriptions. Depending on the set of controls, the coefficient estimate for *Readability* reported in Columns 1 through 3 ranges from 0.011 (*t*-statistic = 1.96) to 0.034 (*t*-statistic = 4.99). Our regression analysis thus indicates that, holding all else equal, one less writing fault per one hundred words is followed by 1% to 3% more citations.

To examine whether the impact of readability on patent citation numbers has increased over time, we split our sample period in half and estimate our regression with the full set of controls separately for each subsample. The results presented in Columns 4 and 5 show that our previously found predictability is coming entirely from the second half of our sample period during which the total number of patents granted per year is 169,210 on average (compared with 80,914 in the first half of our sample period).

In sum, we find that our results based on scientific journal articles carry over to patents and are consistent with our proposition that even minor inconveniences in learning about an innovation can hinder the diffusion of an innovation.

6. Conclusion

Our study notes that on the "supply side," there is considerable variation in readability. Some papers and patent descriptions read rather well; others suffer from numerous writing faults. Some scientists and lawyers

[6] We present summary statistics for the above variables in Online Appendix Table 3. Compared with scientific journal articles, the number of citations for patents is lower (an average of 11.62 compared with an average of 203.56 for scientific journal articles) and the number of writing faults in the patent description is higher (an average of 11.31 per 100 words compared with an average of 6.13 for scientific journal articles).

may be overconfident and erroneously believe their writing to be superb. Others may not care about their writing. Still others may take pride in their ability to construct complex phrases and use terms such as “inter alia” and “lacuna.”

On the “demand side,” we propose that at least some potential adopters are increasingly overwhelmed by the myriad of new ideas proposed to them such that even minor obstacles to understanding an innovation causes an innovation to not be adopted.

If it is true that at least some individuals and institutions are beginning to “feel some innovation overload,” we face a possibly critical conundrum: The more we innovate, the harder it becomes to evaluate and fully absorb each innovation, thereby hampering any positive effect coming out of greater innovative activities. Our evidence, while not causal, corroborates the innovation overload hypothesis and also suggests that the economic magnitude of the effect is substantial. Given the importance of innovations to economic growth, future research may consider subjecting the innovation overload hypothesis to additional tests and exploring possible solutions for how to overcome problems arising from such overload.

References

Bresnahan, T.F., Brynjolfsson, E., Hitt, L.M., 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economic* 117(1), 339–376.

Cutts, M., 2013. Oxford guide to Plain English. Oxford University Press.

Ederer, F., Manso, G., 2011. Incentives for innovation: Bankruptcy, corporate governance, and compensation systems. Handbook of law, innovation, and growth (Cheltenham: Edward Elgar Publishing).

Elliot, W.B., Rennekamp, K., White, B.J., 2015. Does concrete language in disclosures increase willingness to invest? *Review of Accounting Studies* 20(1), 839-865

Hall, B., 2005. Innovation and Diffusion. The Oxford handbook of innovation (Oxford and New York: Oxford University Press).

Hall, B., Lerner, J., 2010. The financing of R&D and innovation. Handbook of the economics of innovation (North Holland: Elsevier).

He, J., Tian, X., 2018. Finance and corporate innovation: A survey. *Asia-Pacific Journal of Financial Studies* 47, 165–212.

Hwang, B.H., Kim, H., 2017. It pays to write well. *Journal of Financial Economics* 124(2), 373–394.

Johnson, R., Watkinson, A., Mabe, M., 2018. The STM report: An overview of scientific and scholarly publishing.

Kerr, W., 2008. Ethnic scientific communities and international technology diffusion. *Review of Economics and Statistics* 90(3), 518-537.

Kerr, W., Nanda, R., 2015. Financing innovation, *Annual Review of Financial Economics* 7, 445–462.

Kogan, L. Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics* 132(2), 665–712.

Lawrence, A., 2013. Individual investors and financial disclosure. *Journal of Accounting and Economics* 56(1), 130-147.

OECD, 2015. The innovation imperative: Contributing to productivity, growth and well-being. OECD Publishing.

Rogers, E., 1995. Diffusion of innovations. New York, Free Press of Glencoe.

Rosenberg, N., 2004. Innovation and economic growth, *OECD Report*.

Simon, H.A., 1955. A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99-118.

Simon, H.A., 1959. Theories of decision-making in economics and behavioral science. *The American Economic Review*, *49*(3), 253-283.

Appendix Table 1
List of Writing Faults, Examples and Possible Corrections

(1) Writing fault	(2) Example	(3) Example how to avoid the corresponding writing fault
Passive verbs	*We must re-think how our resources will be best used to provide world-class customer service*.	*We must re-think how to best use our resources to provide world-class customer service*.
Hidden verbs	*. . . to make an application for employment*.	*. . . to apply for employment*.
Complex words	*While third parties sometimes endeavor to ameliorate relationships . . .*	*While third parties sometimes try to improve relationships . . .*
Abstract words	*We need to install more output devices*.	By avoiding abstract words, writers can clarify the message they are trying to convey: *We need to install more printers*.
Overused words and Cliches	*The patient was then informed about the parameters of treatment available . . . we must more carefully study the parameters of our health care system*. and *open a can of worms, we beg to differ, wakeup call*	The former are popular terms used in a variety of settings; they can essentially mean anything a writer wants them to mean; the latter are phrases that have become devalued through overuse: *The patient was then informed about the types of treatment available . . . we must more carefully study the limitations of our health care system*.
Legal words	*forthwith*	*immediately*
Wordy phrases	*an appreciable number of, has a requirement for*	*many, requires*
Overwriting	*It is completely unnecessary.*	*It is unnecessary.*
Foreign words	*The results show a high urban crime rate, inter alia . . . our paper helps fill a lacuna in the literature . . .*	*The results show a high urban crime rate, among others . . . our paper helps fill a gap in the literature . .*
Long sentences	Regarding "long sentences," there is no objective criterion as to what constitutes a long sentence. Cutts (2013) in the Oxford Guide to Plain English, for instance, recommends an average sentence length of 15–20 words. In our study, we follow our software's definition of a long sentence, which is a sentence with more than 35 words.	

Table 1
Descriptive Statistics for Scientific Journal Articles Sample

This table presents summary statistics for our main variables in the scientific journal articles sample. The sample includes 2,618 scientific journal articles published in the *Journal of Finance*, the *Journal of Financial Economics*, and *The Review of Financial Studies* from 2005 through 2014. The *Journal of Finance* has 716 papers, the *Journal of Financial Economics* has 1,048 papers, and *The Review of Financial Studies* has 854 papers. *Citations* is the Google citations the paper receives as of September 16–20, 2016. *Readability* is the number of writing faults per one hundred words multiplied by negative one. *Years since Publication* is the number of years since publication (as of year 2016). *Number of Authors* is the number of authors listed in the paper. *Affiliation Ranking* is the average "ranking score" across the schools the authors are primarily affiliated with as detailed in Section 3.3. *Number of Presentations* is the number of conferences and seminars the paper was presented at prior to publication. Each year, the *Journal of Finance*, the *Journal of Financial Economics* and *The Review of Financial Studies* award best-paper prizes to the best published papers in that year. *Award Paper* equals one if the paper won such a prize. *Theory Paper* equals one if a paper contains the term "proof" along with either one of the following terms: "proposition," "theorem," "lemma," "corollary". *Length of Title* is the number of words in the title. *Number of JEL Codes* is the number of JEL codes listed in the paper.

	N	Mean	StDev	10th Percentile	Median	90th Percentile
Citations	2,618	203.56	293.14	25.00	115.00	465.00
Readability	2,618	-6.13	1.12	-7.50	-6.10	-4.80
Years since Publication	2,618	6.19	2.83	2.00	6.00	10.00
Number of Authors	2,618	2.39	0.87	1.00	2.00	3.00
Affiliation Ranking	2,618	18.00	16.40	0.00	15.67	44.50
Number of Presentations	2,618	8.68	6.50	1.00	8.00	17.00
Award Paper	2,618	0.05	0.21	0.00	0.00	0.00
Theory Paper	2,618	0.14	0.35	0.00	0.00	1.00
Length of Title	2,618	8.80	3.48	5.00	8.00	13.00
Number of JEL Codes	2,618	2.28	1.84	0.00	2.00	5.00

Table 2
Correlation Matrix for Scientific Journal Articles Sample

This table presents Pearson correlation coefficients across our variables. Correlations that are significant at the 5% level are in bold.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) *Citations*	1.000									
(2) *Readability*	**0.056**	1.000								
(3) *Years since Publication*	**0.353**	**-0.045**	1.000							
(4) *Number of Authors*	0.019	**-0.041**	**-0.047**	1.000						
(5) *Affiliation Ranking*	**0.151**	0.031	**0.039**	**-0.100**	1.000					
(6) *Number of Presentations*	**0.059**	0.004	**-0.177**	**0.080**	**0.132**	1.000				
(7) *Award Paper*	**0.149**	0.039	0.023	-0.023	**0.135**	**0.072**	1.000			
(8) *Theory Paper*	**-0.083**	-0.034	0.035	**-0.114**	**0.062**	**0.086**	-0.015	1.000		
(9) *Length of Title*	-0.024	**-0.070**	**0.127**	**0.062**	**-0.049**	**-0.083**	**-0.050**	**-0.141**	1.000	
(10) *Number of JEL Codes*	**-0.071**	0.001	-0.007	**0.099**	-0.023	-0.036	**-0.079**	**-0.044**	**0.150**	1.000

Table 3
Readability and Innovation Diffusion: Evidence based on Scientific Journal Articles

This table presents coefficient estimates from regressions of the natural logarithm of *Citations* on various article characteristics. *T*-statistics are reported in parentheses and are based on standard errors adjusted for heteroscedasticity and clustered by the first JEL code. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

	(1) Published 2005-2014	(2) Published 2005-2014	(3) Published 2005-2014	(4) Published 2005-2014	(5) Published 2005-2009	(6) Published 2010-2014
Readability	0.063***	0.083***	0.072***	0.133***	0.053***	0.096***
	(3.83)	(5.40)	(4.98)	(3.40)	(2.88)	(4.55)
Years since Publication		0.322***	0.336***	0.276***	-0.311	0.498***
		(4.48)	(5.20)	(5.56)	(-1.17)	(3.07)
Sqr. Years since Publication		-0.011**	-0.010**	-0.010***	0.026*	-0.031*
		(-2.49)	(-2.53)	(-2.59)	(1.77)	(-1.74)
Readability × *Years since publication*				-0.101		
				(-1.52)		
Number of Authors			0.115***	0.116***	0.094***	0.132***
			(5.98)	(6.02)	(3.34)	(5.74)
Affiliation Ranking			0.009***	0.009***	0.009***	0.008***
			(8.55)	(8.54)	(6.47)	(5.39)
Number of Presentations			0.024***	0.024***	0.031***	0.018***
			(7.58)	(7.71)	(4.78)	(5.94)
Award Paper			0.464***	0.466***	0.482***	0.385*
			(2.65)	(2.62)	(3.17)	(1.75)
Theory Paper			-0.454***	-0.457***	-0.582***	-0.317***
			(-10.78)	(-10.94)	(-10.13)	(-4.59)
Length of Title			-0.025***	-0.025***	-0.026***	-0.023***
			(-5.13)	(-4.96)	(-4.03)	(-2.87)
Number of JEL Codes			0.013	0.014	0.014	0.034
			(0.80)	(0.82)	(0.46)	(1.40)
Constant	5.307***	4.519***	4.344***	4.731***	7.082***	4.040***
	(43.33)	(16.36)	(15.44)	(18.92)	(5.78)	(10.79)
Journal FE	Yes	Yes	Yes	Yes	Yes	Yes
1st JEL Code FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.148	0.343	0.416	0.417	0.293	0.386
Number of observations	2,618	2,618	2,618	2,618	1,206	1,412

Table 4
Readability and Innovation Diffusion: Evidence based on Patents

This table presents coefficient estimates from regressions of the natural logarithm of *Patent Citations* on various patent characteristics. *T*-statistics are reported in parentheses and are based on standard errors adjusted for heteroscedasticity and clustered by the USPTO three-digit technology class. Statistical significance at the 10%, 5%, and 1% levels is denoted by *, **, and ***, respectively.

	(1) Granted 1976-2010	(2) Granted 1976-2010	(3) Granted 1976-2010	(4) Granted 1976-1993	(5) Granted 1994-2010
Patent Readability	0.034***	0.016***	0.011*	-0.002	0.017***
	(4.99)	(2.89)	(1.96)	(-0.21)	(2.91)
Years since Granting		0.242***	0.236***	0.016	0.312***
		(34.69)	(34.74)	(0.42)	(25.61)
Sqr. Years since Granting		-0.006***	-0.006***	-0.000	-0.009***
		(-28.74)	(-28.78)	(-0.62)	(-13.48)
Economic Value of Patent			0.082**	0.230	0.038
			(2.30)	(1.58)	(1.14)
Firm-level Innovation Value			0.006***	0.042***	0.004***
			(4.54)	(2.74)	(3.17)
Firm-level Number of Patents			-0.006***	-0.017*	-0.006***
			(-5.47)	(-1.70)	(-5.15)
Constant	2.330***	1.035***	1.003***	3.720***	0.871***
	(27.05)	(15.51)	(14.09)	(7.12)	(11.55)
USPTO Technology Class FE	Yes	Yes	Yes	Yes	Yes
R^2	0.035	0.399	0.403	0.160	0.445
Number of observations	12,851	12,851	12,851	3,867	8,984