What's your Identification Strategy? Innovation in Corporate Finance Research

Donald E. Bowen, Laurent Frésard, and Jérôme P. Taillard*

May 13, 2016

ABSTRACT

We study the diffusion of techniques designed to identify causal relationships in corporate finance research. We estimate the diffusion started in the mid-nineties, lags twenty years compared to economics, and is now used in the majority of corporate finance articles. Consistent with recent theories of technology diffusion, the adoption varies across researchers based on individuals' expected net benefits of adoption. Younger scholars, holders of PhDs in economics, and those working at top institutions adopt faster. Adoption is accelerated through networks of colleagues and alumnis and is also facilitated by straddlers who cross-over from economics to finance. Our findings highlight new forces that explain the diffusion of innovation and shape the norms of academic research.

*University of Maryland, University of Maryland, and Babson College respectively. Bowen can be reached at donald.bowen@rhsmith.umd.edu, Frésard can be reached at lfresard@rhsmith.umd.edu, and Taillard can be reached at jtaillard1@babson.edu. We are grateful to the Associate Editor, two anonymous Referees, Jean-Noel Barrot, Jess Cornaggia, Michel Dubois, Rudiger Falhenbrach, Xavier Giroud, Jerry Hoberg, Andrew Karolyi, Max Maksimovic, Toby Moskowitz, Nagpurnanand Prabhala, Michael Roberts, Alberto Rossi, Phil Strahan, René Stulz, and seminar participants at Boston College and the University of Maryland for helpful suggestions and comments.

Electronic copy available at: https://ssrn.com/abstract=2378460

I Introduction

How do individuals and organizations make decisions about the use of new technologies? This question is important as the diffusion of innovation is believed to be a key channel through which productivity growth is achieved (e.g. Aghion and Howitt (1992)). Simple intuition suggests that profitable innovations should be adopted almost instantaneously. Yet, various frictions can act as barriers that slow the diffusion of innovation. The rate of adoption of newly invented technologies can vary substantially across agents that differ in their knowledge of the costs and benefits of adopting, their access to the technologies, their learning ability, financial resources, previous experiences, human capital, and exposure to the technology through peers (e.g. Hall and Khan (2002) or Foster and Rosenzweig (2010)).

In this paper, we study the diffusion of a particular innovation: Econometric techniques designed to estimate causal relationships in empirical corporate finance research. The use of such techniques has recently become a widespread tool for corporate finance researchers, mirroring the trend observed in other areas of economics (e.g. labor or development).[1] Proponents claim that these econometric techniques offer a "credibility revolution" for researchers (see Angrist and Pischke (2010)), thus enhancing their ability to make causal statements, policy recommendations, and gain credibility in the marketplace for ideas.

Studying this specific innovation allows us to shed new light on the forces that drive the diffusion of innovation in financial research and the establishment of new academic norms. We construct a unique sample that comprises the *content* of all empirical corporate finance articles published in the leading ("top-three") finance academic journals between 1970 and 2012.[2] This sample is comprised of 1,796 articles written by 1,880 distinct authors for whom we hand-collect detailed biographical information. We consider five main identification techniques: Instrumental variables; difference-in-differences (and natural experiments); selection models; regression discontinuity design; and randomized experiments. We treat these identification techniques as innovations, and taken together, they form what we coin the "identification technology." We exploit the fine granularity of our dataset to examine in detail the patterns and determinants of technological adoption in academia.

[1] Attesting to the increasing focus on identification in corporate finance research, several recent methodology surveys outline current best practices to address issues related to endogeneity in corporate finance (e.g. Li and Prabhala (2007), Roberts and Whited (2013), Strebulaev and Whited (2012)), and several recent keynote speeches at major conferences have concentrated on identification (e.g. Toni Whited at the 2014 SFS Cavalcade or Wei Jiang at the 2015 SFS Cavalcade).

[2] These are the Journal of Finance (JF), the Journal of Financial Economics (JFE), and the Review of Financial Studies (RFS).

Electronic copy available at: https://ssrn.com/abstract=2378460

We first document a recent and fast adoption of the identification technology in corporate finance. The fraction of articles using the identification technology is close to zero until the end of the eighties and then steadily rises to reach more than 50% in 2012. The adoption pattern is consistent with an S-shaped diffusion curve, similar to diffusion patterns observed for "harder" types of technology (e.g. Griliches (1957), Mansfield (1961), and more recently, Comin and Hobijn (2010)). By fitting logistic diffusion curves to the data and using the conventional 5% adoption threshold, we estimate that the year 1998 marks the emergence of the identification technology in finance journals.

We then use various adoption thresholds to compare the emergence of the identification technology in top finance and top economics journals.[3] Focusing on the language of identification, we document that the identification technology emerges in finance journals about twenty years after its emergence in the top economics journals. Given the relatively blurred boundaries between economics and corporate finance research, the long adoption lag between the two fields suggests the presence of important frictions that impede the diffusion of the identification technology across research fields.

To shed light on the possible frictions at play, we examine the determinants of adoption and measure how adoption rates vary across researchers. We conjecture that a given researcher will adopt as soon as the expected net gains from using the identification technology are positive. Estimating such expected net gains is empirically challenging however, as we do not observe the relevant inputs in each researcher's utility function. We posit that researchers attempt to maximize their lifetime scientific impact and recognition, which we approximate by using citations and publications (see Hamermesh and Pfann (2012)).

We estimate that identification articles attract 22% more citations on average, after controlling for various author and paper-level characteristics. Such an identification "premium" started in 1995 for finance articles and as soon as 1982 for economics articles. In parallel, we document a substantial change in the composition of editorial boards at the leading finance journals, with a substantial increase in the fraction of board members who have adopted the technology in their own research. This shift can arguably increase the odds of publication for papers relying on the identification technology. The aggregate citation premium, combined with the change at the editorial board level, suggests that the benefits of adopting the identification technology increased over our sample period.

[3]Following the classification of Card and DellaVigna (2013), we focus on the top-three economics journals: the American Economic Review, the Journal of Political Economy, and the Quarterly Journal of Economics.

Electronic copy available at: https://ssrn.com/abstract=2378460

The costs and benefits of adopting the technology vary across researchers. Using a survival analysis to estimate the determinants of time-to-adoption across researchers, we show that knowledge about the net gains from adopting is related to the diffusion of the identification technology. In particular, our estimates indicate a faster adoption among researchers holding a PhD degree in economics compared to finance PhDs. The magnitude is large as researchers with PhDs in economics adopt almost 30% faster. This significant differential suggests that doctoral training in economics increases researchers' exposure to the identification technology, which in turn improves their inference about the net gains of adopting the innovation. Relatedly, we document that authors that have previously published articles in economics journals – whom we call "straddlers" – are among the early adopters in finance journals. These findings highlight that the diffusion of innovation can be accelerated when researchers with a given training migrate to a neighboring field of research (see Stoyanov and Zubanov (2012)).

We also find that the adoption speed is related to an author's research network. Specifically, we show that the adoption rate is positively and significantly related with the extent of prior adoption by a researcher's network of current and previous colleagues. Moreover, researchers adopt the identification technology significantly faster when a larger fraction of alumni researchers from their PhD-granting institutions have already adopted the technology. Both results are consistent with networks creating positive externalities that facilitate the diffusion of knowledge among researchers and reduce the informational constraints associated with learning about the new technology.[4]

We find no evidence that the school ranking of a researcher's PhD-granting institution is related to the likelihood of adopting the identification technology. In contrast, the ranking of a researcher's affiliation at the time of publication is significantly related to the speed of adoption. Researchers employed at top-tier institutions adopt the identification technology significantly sooner than their peers at lower-ranked institutions. The relationship is not necessarily a causal one: The faster adoption of researchers at top institutions could indicate that the larger financial and organizational resources allocated to research at top schools plays a role in explaining the observed "top-down" diffusion of identification techniques. But this finding could also suggest that top schools are simply better at selecting innovators and

[4]Learning through peers and similar networking effects have been documented in other settings. For instance, Bayer, Hjalmarsson, and Pozen (2009) finds strong peer learning effects among juveniles in correctional facilities. In academia, Azoulay, Graff Zivin, and Wang (2010) shows that superstar researchers generate significant positive spillover effects among their network of collaborators.

Electronic copy available at: https://ssrn.com/abstract=2378460

early adopters.

We estimate that seniority is significantly associated with the diffusion of the identification technology. Controlling for cohort and year effects, younger authors tend to adopt earlier. This age effect is consistent with models of vintage human capital implying that early-stage researchers are the primary adopters of new technology (e.g. Chari and Hoenhayn (1991) or MacDonald and Weisbach (2004)). The age effect could be due to the more flexible minds of young scholars (e.g. Darwin (1859)), their larger exposure to recent innovation, or higher incentives and greater amount of time to learn new techniques (e.g. Diamond (1980)). Alternatively, more senior scholars could have more vested interests that may render them less receptive to innovation (e.g. Cohen (1985)). Although they adopt later and less frequently on average, a significant number of senior researchers do adopt the identification technology over our sample period. We show that seniors are more likely to adopt by co-authoring papers and that their coauthors are younger on average for identification articles.

While we primarily focus on the diffusion of the identification technology, we also document changes *within* the technology itself. In particular, we show that best practices for the implementation of the technology – or simply the "norms" – are evolving. Specifically, we track the evolution of several technological refinements (e.g. the appearance of weak instrument tests for the implementation of instrumental variables) and confirm that the boundaries of the identification technology change significantly over time. This dynamic process is consistent with the notion that most innovations and scientific improvements are the result of experimentation and refinements (e.g. Popper (1959) or Kuhn (1962)). Moreover, the genuinely transient nature of best practices makes the judgment of the proper implementation of the identification technology at any given point in time a somewhat delicate exercise.

This study makes several contributions. First, it adds to the growing literature that studies the determinants of the diffusion of scientific advances and new technologies. Mokyr (2002) argues that while the stock of existing knowledge is important for creating new knowledge, the effective diffusion of knowledge across researchers requires them to be not only aware of existing knowledge, but also to be able to access it. Confirming this idea, existing research indicates that the reduction of informational, institutional, or legal barriers to knowledge spurs the creation and diffusion of new ideas. For instance, Agrawal and Goldfarb (2008), Ding, Levin, Stephan, and Winkler (2010), Furman and Stern (2011), and Kim, Morse, and Zingales (2009) show that improvements in communication technologies render researchers more productive, foster knowledge accumulation, and increase collaboration. Other studies

Electronic copy available at: https://ssrn.com/abstract=2378460

indicate that the cost of accessing existing ideas, as measured by geographical proximity (e.g. Azoulay, Graff Zivin, and Sampat (2012) and Head, Li, and Minondo (2015)) or protection of intellectual property (e.g. Williams (2013) and Teodoridis (2015)), is related to the flow of academic knowledge and inventive activities. Our study is complementary as we examine the forces that explain the diffusion of a *soft* innovation among researchers in a setting where legal and financial barriers are low. We find that human factors related to the cost of accessing the existing knowledge (e.g. PhD training and research networks) are empirically important and help explain the diffusion of new technologies across the boundaries of research fields.

Our work also relates to the "burden of knowledge" concept (e.g., Jones (2009) and Teodoridis (2015)). Jones (2009) shows that, as the stock of knowledge grows in the economy, researchers need to spend more time in training and narrow their field of study to reach the ever-expanding frontier of knowledge. As a consequence, researchers become more specialized and collaborate more. Our evidence is supportive of his theory. Researchers specialized in corporate finance adopt only with a significant lag the econometric techniques established outside their area. Further, the diffusion of the identification technology comes in part through early adopters in the field of economics who migrate and bring the technology to finance journals. We also highlight a mutually beneficial "cross-generational" collaboration, whereby seniors team up with juniors to overcome the extension of the frontier of knowledge.

Our findings further contribute to the literature studying the evolution of academic research in finance. This introspective and rhetorical literature is relatively thin in finance compared to economics and other fields in social sciences. This line of inquiry used to be more prevalent in the top-three finance journals.[5] The recent studies in this tradition concentrate on evaluating the adequacy of statistical tools used in corporate finance research.[6] Our focus is distinct, as we study the shift towards identification techniques in corporate finance research and the resulting changes in the profession.

Finally, we need to underline two limitations of our analysis. First, perhaps paradoxically, our paper is not an identification paper. Our study concentrates on analyzing patterns of

[5] Weber (1973) and Keenan (1991) study trends in finance publications. Kaufman (1984), Niemi (1987), and Alexander and Marby (1994) evaluate the performance of departments and journals. Zivney and Bertin (1992) and Borokhovich, Bricker, Brunarski, and Simkins (1995) examine the productivity of finance researchers. Corrado and Ferris (1997) analyze the influence of published articles on finance doctoral education. There are a few notable recent studies: Karolyi (Forthcoming) examines general trends in finance publications, Welch (2014) analyzes the refereeing process in finance and economics journals, and Brogaard, Engelberg, and Parsons (2014) examines how proximity with an editor influences research productivity.

[6] Examples include Atanasov and Black (Forthcoming), Campello, Galvao, and Juhl (2013), Gormley and Matsa (2014), Karolyi (2011), Petersen (2009), Roberts and Whited (2013), and Thompson (2011).

Electronic copy available at: https://ssrn.com/abstract=2378460

adoption and highlighting factors that are significantly related to these adoption patterns, but not necessarily in a *causal* manner. Second, while we believe that our framework is useful to shed new light on some of the forces that drive technological adoption, we recognize that our academic setting is atypical and hence, extrapolation of our results to other contexts deserves careful thought.

The paper proceeds as follows. In Section II, we provide the conceptual framework for our study. In Section III, we describe the construction of our sample. In Section IV, we show general diffusion patterns of the identification technology. Section V explores the determinants of diffusion at the researcher-level. Section VI highlights technological refinements over time. Section VII concludes.

II Conceptual Framework

How do finance researchers decide to adopt or not the identification technology? To help organize the discussion and guide our empirical analysis, we follow the theoretical literature on technology adoption. The backbone of this literature is an empirical regularity: When the number of users of a new innovation is plotted versus time, the resulting curve is typically an S-curve (or ogive distribution). This general pattern appears natural as one imagines adoption proceeding slowly at first, accelerating as it spreads, and slowing down as the relevant population becomes saturated (e.g. Comin and Hobijn (2010)).

In an attempt to explain this pattern, existing models of technology diffusion typically consider the decision of an agent (e.g. a firm, an organization, or an individual) to adopt a new technology at a given date. The agent chooses to adopt when the expected net benefits of adopting are positive. The central prediction of existing models is that potential adopters will adopt at different dates (or not adopt at all). The asynchronous adoption pattern originates from differences in agents' inherent characteristics (e.g. skills, experience, or beliefs) and/or information about the expected net benefits associated with the technology (see Rogers (2003)).

Early work typically models the diffusion of a new technology using epidemic theories according to which, perfectly homogeneous agents decide to adopt a new technology when they (exogenously) receive information relating to its existence and its associated net benefits (e.g. Griliches (1957) or Mansfield (1961)). In these models, the pattern and speed of adoption is entirely driven by the awareness of the new technology (or lack thereof) and the velocity

Electronic copy available at: https://ssrn.com/abstract=2378460

of information spreading. In contrast, more recent theoretical work considers heterogeneous agents that can differ on various dimensions and focuses on the decision to adopt a new technology (e.g. Hall and Khan (2002) or Foster and Rosenzweig (2010)). Potential adopters have different characteristics and differential access to information about the technology and, as a result, form distinct expectations about the costs and benefits from adopting the new technology.[7] In this framework, the pattern of adoption is dictated by intrinsic differences among potential adopters that can include, for instance, differences in education levels, access to information, financial capital, personal experiences, social connections, or risk aversion.

Estimating the expected return to adoption of a new technology is notoriously challenging. Our specific academic setting is no exception. To make our empirical analysis tractable, we assume that finance researchers are economic agents who make decisions that maximize the present value of their expected lifetime utility.[8] Based on this assumption, they adopt a new technology when doing so increases their expected utility. Arguably, the utility function of researchers is complex and features many different inputs (and associated weights) that can vary significantly across individuals. For instance, financial gains associated with additional grant funding, salary increases, or tenure decisions can all increase the utility of researchers. But, there can also be non-pecuniary elements, such as titles, honorary memberships (e.g. editorial board positions, Distinguished Fellowships), solicitations to give talks or serve as experts in the media and courts, or simply the pursuit of true scientific endeavors.

In this context, a scholar's decision to adopt a technological innovation in his research is made by trading off the expected adoption costs and the increased expected utility associated with the use of the innovation. While it is a priori unclear what elements increase a researcher's utility, we posit that a researcher's utility is directly linked to her reputation and impact in the scientific community.[9] Based on existing studies that examine the determinants of researchers' career paths (e.g. Hamermesh, Johnson, and Weisbrod (1982), Moore, Newman, and Turnbull (2001), or Hamermesh and Pfann (2012)), we assume that researchers' scientific impact and recognition is determined by the number of peer-reviewed publications obtained and citations they garner. Therefore, our analysis builds on the assumption that a

[7] See Kremer and Miguel (2007) for a specific example of heterogeneous agents modeling.

[8] As emphasized by Foster and Rosenzweig (2010), the estimation of return to adoption is a general problem in the literature. In the case of technology used by profit maximizing entities, it is clear that technology profitability is a good measure of return to adoption. For technologies that improve an agent utility (e.g. those that improve health, happiness, or scientific impact) measurement of returns is much less straightforward. Agents choose to use a technology based on gains in terms of welfare, which cannot be easily observable.

[9] Hamermesh and Pfann (2012) show that reputation is positively related to both pecuniary and non-pecuniary rewards that provide researchers with utility.

Electronic copy available at: https://ssrn.com/abstract=2378460

researcher adopts the identification technology in her research when she expects that the use of the new technology will increase her overall research impact, net of adoption costs.

Based on this simple conceptual framework, our empirical analysis first describes the general pattern and speed of adoption of the identification technology in empirical corporate finance research. In a second step, we concentrate on the determinants of adoption to understand what factors drive – or hinder – the diffusion of the identification technology. We focus on characteristics of researchers and their research networks that could affect their awareness of the identification technology and/or its potential benefits, their incentives to adopt, and their adoption cost. In doing so, we explicitly recognize the heterogeneous nature of the pool of researchers we study.

III Sample Construction

This section details the sample. We first explain the construction of the main sample of empirical corporate finance articles and authors. Second, we outline the procedure we followed to define and identify articles that rely on the identification technology.

A Empirical Corporate Finance Research

We start by constructing a dataset containing detailed information on all published articles in the Journal of Finance (JF), Journal of Financial Economics (JFE), and Review of Financial Studies (RFS) for the period 1970 to 2012. Using the JSTOR database for JF and RFS articles and ScienceDirect for JFE articles, we collect the following variables for each article: Year of publication, title, authors, journal, volume, issue, and page range. We screen for errata, comment, correction, reply, or discussion articles. We apply further screens to filter common issue fillers (e.g. "Back Matter"), results related to annual meetings, and abstracts of doctoral dissertations. The initial sample contains 3,579 JF articles, 1,969 JFE articles, and 1,249 RFS articles for a total top-three sample of 6,797 articles.[10]

We then identify empirical corporate articles as follows. First, we define an article as empirical if it performs statistical tests on real data (see Hamermesh (2013)). With this broad definition, the only excluded articles are: (1) Pure theory articles; (2) articles that use only simulated data; and (3) articles that calibrate a model based on prior publication results and not self-generated inference.

[10]The Internet Appendix presents some general trends in publication between 1970 and 2012. Overall the patterns in our sample are consistent with the findings of other articles focusing on publication trends in economics journals (e.g. Hamermesh (2013), Card and DellaVigna (2013), and Ellison (2002)).

Electronic copy available at: https://ssrn.com/abstract=2378460

Second, we define corporate finance articles using the Journal of Economic Literature (JEL) codes and resort to a manual classification for the subset of articles for which JEL codes are not available. All JFE articles since 1993 and most RFS articles since 2006 directly contain JEL codes (1,994 articles). Of the remaining articles (4,803), we search the EconLit database and find JEL codes for 2,402 articles. We assign articles into the corporate finance group if they contain at least one JEL code starting with G3 ("Corporate Finance and Governance").[11] For articles without JEL codes (2,401 articles), we manually define them as corporate finance or non-corporate finance. To enhance integrity and accuracy, each article was independently evaluated by two of the authors.[12] We identify a total of 1,796 empirical corporate finance articles in the top-three finance journals between 1970 and 2012.

For each article we gather citation data from Thomson Reuters' Web of Knowledge (WOK). We collect the "Times Cited" variable, which tracks the total number of citations received by the article across all databases indexed by WOK, as of November 2013. Additionally, we collect the year, title, authors, and journal of each citation. From this, we are able to track the citations a paper receives every year. 99.6% of our sample contains citation information from WOK.[13]

Finally, we collect detailed biographical information on authors from their personal or professional websites. Between 1970 and 2012, 1,880 distinct authors published at least one empirical corporate article in a top finance journal. In particular, we identify the field of PhD (e.g. finance or economics) for 1,610 authors (85.6%), year of PhD graduation for 1,620 authors (86.2%), PhD-granting institution for 1,718 authors (91.4%), and the employment history (i.e. career path) for 1,554 authors (82.7%).

B The Identification Technology

We devise a simple classification procedure to determine whether an article uses an econometric technique designed to identify causal relationships. We consider five widely-used techniques: (1) Instrumental variables (IV), (2) difference-in-differences and (quasi-)natural

[11] See http://www.aeaweb.org/jel/guide/jel.php for a description of JEL classifications.

[12] Note that we assign "banking" articles into the corporate finance group if the main variable of interest is a corporate decision (e.g. capital structure or the size of the board) made by non-financial firms. To illustrate the outcome of our classification, Table IA.1 in the Internet Appendix presents the assignment of the two most cited articles every year into empirical (or not) and corporate (or not), as well as the source for the classification (JFE, RFS, or EconLit JEL codes or Manual). A visual inspection of this table reveals that our classification performs well. Out of 90 cases, the only potential misclassification is the article by Bansal and Yaron (2004) that is classified as corporate finance based on the JEL codes from EconLit.

[13] Excluding JFE articles in 1974 and 1975 and RFS articles in 1988 and 1989, for which WOK did not have records as of the time of our data collection.

Electronic copy available at: https://ssrn.com/abstract=2378460

experiments (DD), (3) selection models, (4) randomized experiments (field or lab), and (5) regression discontinuity designs (RDD). Henceforth, we use DD as shorthand that includes quasi-natural experiments. We label these five techniques as identification techniques. Together they form the "identification technology."[14]

We start by flagging articles that *may* use each technique by searching the full text of every article for a list of keywords (a dictionary) associated with each technique. As an example, we use the following terms to form the IV-specific dictionary: "instrumental variable", "two stage least square", "three stage least square", "2SLS", and "3SLS" (allowing for all common permutations, hyphenation, and plural form). The dictionary for all five techniques is reported in Table A1 of Appendix A. The dictionaries are designed to catch all articles that refer to the technique—that is, we purposefully err on the side of encouraging false positives to reduce the number of false negatives. In a second step, we verify manually each flagged article and remove all false positives. We then classify each corporate finance articles as "identification" if the paper uses the identification technology, and "non-identification" if it does not.[15] We provide more details about the classification in Appendix A.

C Descriptive Statistics

Table I provides a description of our sample of published papers in the top-three finance journals between 1970 and 2012. Empirical corporate finance articles represent 26% of all published articles in the top-three finance journals over the entire sample period. Panel A of Table I highlights that this fraction has largely increased over time; specifically, the share of empirical corporate finance articles has grown from 8% in the 1970s to 39% during 2010-2012.

[Insert Table I about here]

Panel B of Table I provides details on the use of the identification technology within the subset of empirical corporate finance papers over the sample period. We uncover 408 articles that adopt the identification technology. Taken together, identification (ID) articles represent 22.7% of all published empirical corporate finance articles over the 1970-2012 period. These articles are written by 632 distinct authors. IV is the most established technique, hence it is

[14] In unreported analyses, we also considered matching techniques and structural estimations as part of the identification technology. The results presented throughout the paper are virtually identical with the addition of these two techniques.

[15] Note that while we classify articles based on the econometric technique they use, we do not attempt to evaluate whether these techniques are applied "correctly". For instance, we remain agnostic as to whether a given natural experiment provides a truly exogenous source of variation in the data or whether an instrument satisfies the exclusion restriction. We believe this exercise is outside of the scope of this paper (see Section VI for further discussion). We manually checked a random sample of 100 non-flagged articles, and found only four false negatives. We conclude that our empirical approach minimized the potential impact of false negatives.

Electronic copy available at: https://ssrn.com/abstract=2378460

perhaps not surprising to find that it is the most represented technique with more than half of all ID articles using an IV approach.[16] The difference-in-differences category (which includes all natural experiments) comes in second, followed by selection models, regression discontinuity designs, and, lastly experiments (with only five articles). Regression discontinuity is the most recent identification technology to emerge in top finance journals.

IV Diffusion Patterns

We start the analysis by examining the overall rate of adoption of the identification technology in corporate finance. We then measure the adoption lag between finance and economics journals.

A The Evolution of Identification

Based on our classification, we simply define p_t as the fraction of the population (i.e., empirical corporate finance articles) that has adopted the identification technology in year t. We plot p_t over the period 1980-2012 in Figure I.[17] We observe a substantial increase in the adoption rate over the sample period (the solid line) from virtually zero in the eighties to more than 50% in 2012.

[Insert Figure I about here]

Similar to many hard innovations, the adoption path of the identification technology resembles a logistic distribution (S-shape curve), which is marked by a relatively long period of slow diffusion that is followed by a period of rapid diffusion. To formally estimate the speed of adoption and date the emergence of the identification technology in corporate finance, we fit a logistic diffusion model to the data. Specifically, we estimate the following equation:

$$p_t = \frac{k}{1 + e^{-(\alpha+\beta\times t+\varepsilon_t)}}, \tag{1}$$

where α and β are parameters that determine the position and slope of the S-curve, and ε_t is a normally distributed error term. The variable k represents the "ceiling" value of the S-curve. It captures the equilibrium fraction of the population that will eventually adopt the identification technology. Because the adoption is still under way (it reaches 52% in 2012) and will likely not achieve 100%, we cannot easily estimate k. In our baseline estimates, we

[16] Note that an identification article can be associated with several techniques, and as such, the techniques listed in this section are not necessarily mutually exclusive.

[17] Although we identify one IV article prior to 1980, we decided to be consistent throughout and present results, tables, and figures using the sample years 1980-2012.

Electronic copy available at: https://ssrn.com/abstract=2378460

set $k = 1$, but we also show results for $k = 0.8$ and $k = 0.6$ and conclude that the choice of k has little bearing on our conclusions. Moreover, we obtain a close fit when superimposing the fitted value $\tilde{p}_t$ on top of p_t in Figure I.

[Insert Table II about here]

Following the convention in the technology diffusion literature, we determine that a technology has "emerged" when it becomes adopted by more than 5% of the population ($p_t > 5\%$). Using this definition, the date of emergence is given by the implied year $\tilde{t}_{5\%}$ that corresponds to the year when the fitted S-curve crosses the 5% threshold ($\tilde{p}_t = 5\%$).[18] We report the estimated value for $\tilde{t}_{5\%}$ in the first column of Table II (Panel A). Our baseline estimation indicates that 1998 marks the emergence of the identification technology in finance journals. As a robustness check, panels B and C reveal that this estimate is fairly insensitive to our choice of k. The 5% adoption threshold is crossed in 1996 if we set $k = 0.8$ and in 1995 if we set $k = 0.6$.

Our estimates reveal a fast diffusion rate. Although we do not have a good benchmark (i.e., estimates of the diffusion speed of related innovation), the first column of Table II (Panel A) indicates that the diffusion rate doubles in only two years after its emergence. Starting from the baseline emergence level of 5% in 1998, the diffusion rate increases to 10% in 2000 ($\tilde{t}_{10\%} = 2000$). It doubles again to 20% over the following three years ($\tilde{t}_{20\%} = 2003$). Overall, the diffusion rate increases from 5% to 50% in only 11 years ($\tilde{t}_{50\%} = 2009$).

The remaining columns in Table II further report the date of emergence of several of the ID techniques, namely instrumental variables, difference-in-differences, and selection models. We concentrate on these three techniques because we cannot estimate similar dates of emergence for regression discontinuity design and experiments since their fitted adoption curves do not cross the 5% adoption threshold during the sample period. Panel A indicates that the implied emergence year for the IV approach is 2001; it is 2006 for difference-in-differences, and 2011 for selection models. (Results from Panel B and C are similar to those in Panel A and hence are not discussed for brevity.) Remarkably, the speed of adoption is roughly similar across the three techniques going from 5% to 10% in the space of about two years, and reaching 20% approximately six years after reaching the 5% threshold.

[18] This number is computed using the 5% threshold of the cumulative distribution of our estimated normalized logistic function. Mathematically, $\tilde{t}_{5\%} = 1980 + \frac{-2.94-\hat{\alpha}}{\hat{\beta}}$. The year 1980 corresponds to the first year of the sample period over which we estimate the model.

Electronic copy available at: https://ssrn.com/abstract=2378460

B Origins and Lags

The identification technology was not originally developed to tackle corporate finance questions but rather originated as a set of solutions to endogeneity problems in other subfields of economics. To better characterize the diffusion of the identification technology in corporate finance, we contrast its emergence with that observed in its original terrain: Top economics journals.

We complement our sample by including *all* articles published in the top-three economics journals between 1970 and 2012 (i.e., the American Economic Review (AER), the Journal of Political Economy (JPE), and the Quarterly Journal of Economics (QJE)). For this exercise, our original sample of 1,796 empirical corporate finance articles is augmented by 8,156 articles published in the top-three economics journals. To simplify our task of contrasting the emergence of the identification technology in both sets of journals, we focus only on the *language* associated with the identification technology (instead of the actual use of identification techniques). We rely on a dictionary of distinct keywords that are unambiguously associated with the concept of identification (presented in Table A2 in Appendix A). The dictionary comprises (1) general terms (e.g. exogenous variation) as well as (2) the same technique-specific terms we use to flag potential identification article in the first step of our classification scheme. For each article, we create a dummy variable that equals one if an article uses at least one of the keywords and aggregate these dummies by year for both economics and corporate finance articles.

[Insert Table III about here]

The results from estimating diffusion-curve models on each sample separately are reported in Table III. We observe a clear upward trend in the frequency of the vocabulary associated with identification within both economics and empirical corporate finance articles. In the first column, we estimate that the 5% and 10% adoption thresholds are reached in economics journals in 1963 and 1974, respectively. These estimates confirm that the identification technology has been used in economics journals for a very long time. For instance, Stock and Trebbi (2003) indicate that the instrumental variable regression technique was first introduced by Wright (1928) to estimate the slopes of supply and demand curves for animal and vegetable oils. Relatedly, our textual search identifies seven articles referring to instrumental variables in economics journals in 1970 (the first year of our sample). In contrast, the 5% adoption threshold is reached in finance only in 1992, while the 10% threshold is crossed in 1996. This

Electronic copy available at: https://ssrn.com/abstract=2378460

implies adoption lags of approximately 29 and 22 years for corporate finance (each difference is statistically significant). The average lag taken across the four adoption thresholds (5% to 20%) is approximately 20 years.[19]

Taken together, these estimates provide an empirical validation to the widely held belief that finance "lags" economics when it comes to adopting new methodological tools. While the identification vocabulary emerges with a significant lag in corporate finance, we observe that the adoption rate is significantly faster within finance journals. Table III reveals that the adoption rate progressed from 5% to 20% in 23 years in economics journals (1963-1986). In contrast, it took only eight years to reach the 20% adoption threshold in corporate finance (1992-2000).

C Empirical Corporate Finance in Economics Journals

Our main focus is on empirical corporate finance articles published in the top-three finance journals. Yet, to provide a complete picture of the adoption of the identification technology in corporate finance, we extend our definition of "corporate finance" research to include empirical corporate finance articles published in the top-three economics journals. We identify these articles based on their JEL codes, and classify an article as corporate finance if it contains at least one JEL code starting with G3. We identify 242 corporate finance articles published in the economics journals between 1980 and 2012. We read each article to determine if it is (a) empirical, and (2) uses one of the five identification techniques. We find 160 empirical corporate finance articles, 45 of which use the identification technology.

[Insert Table IV about here]

Table IV replicates Table II but expands the set of empirical corporate finance articles (i.e., the population in equation (1)) to include both finance and economics journals (1,878 articles). The estimated adoption of the identification technology is notably earlier when we include articles from economics journals. The 5% adoption threshold is crossed in 1990, and the 10% threshold crossed in 1996, as opposed to 1998 and 2000 when we concentrate only on finance journals. With this more inclusive definition of published work in empirical corporate finance, the adoption lag with economics journals is now shorter. This confirms the overall later adoption of the identification technology in finance journals compared to economics journals. Moreover, the estimates suggest that the identification technology in empirical

[19] We obtain similar results if we assume lower saturation levels; namely we estimate the lag at 20.5 years with $k = 0.8$ and at 12.3 years with $k = 0.6$.

Electronic copy available at: https://ssrn.com/abstract=2378460

corporate finance first appears through articles published in the top economics journals. We confirm this result in Section V.C when we look in greater details at the adoption patterns of a subset of authors that "straddle" the boundary between finance and economics journals.

V The Determinants of Adoption

As highlighted in Section II, an agent's decision to adopt a new technology at a given point in time depends on the expected net gains of adopting. We examine in this section factors that could be associated with such gains. We start by approximating the aggregate expected net gains from adopting using citations and publication odds and then assess how various researcher characteristics influence the speed with which they adopt the identification technology.

A Aggregate Benefits

Theory predicts that an agent will adopt a new technology when the (present value of the) expected net gains from adopting are positive. So, to understand agents' decisions and the observed diffusion pattern more broadly, we need to be able to estimate expected net gains for each agent. This is notoriously difficult because the factors (and weight) that enter each agent's utility function are not observed by the econometrician (see Foster and Rosenzweig (2010) for more details about measurement issues). Although we are not able to fully overcome this challenge, we can examine whether the use of the identification technology is associated with tangible aggregate benefits measured across all researchers. To do so, we assume that researchers seek to maximize their lifetime scientific impact and recognition in the profession. Following Hamermesh and Pfann (2012), we thus focus on the number of peer-reviewed publications and citations as metrics for scientific impact. We offer a two-pronged approach to approximate the aggregate net gains from adoption and their evolution over time.

A.1 Citations

We first concentrate on citations and examine whether articles using the identification technology display a different citation pattern compared to non-identification articles. We want to assess whether researchers in a given year T could have been aware of the difference in citations attracted by identification papers (i.e. the potential benefits).

Electronic copy available at: https://ssrn.com/abstract=2378460

To do so, we track the citations that identification papers obtain before T, and compare them to that of matched articles that do *not* use the identification technology. We follow Azoulay, Stuart, and Wang (2014) and select a single matched article for each identification article. We use a non-parametric approach designed to guarantee balance on crucial covariates. Specifically, for each identification article, we define the set of potential matches as all the non-identification articles published in the same journal, year, and quarter to ensure that we compare articles of the same vintage. We further restrict to articles where the most prolific author, measured using the number of previous publications, is in the same quantile as the focal identification article. If there is more than one article in this set, we select the non-identification article whose authors have the closest number of accrued citations. Any remaining ties are broken randomly. This procedure generates matches for 75% of identification articles.[20]

[Insert Figure II about here]

Panel A of Figure II plots the average difference in (log of) citations between pairs of identification and matched articles for each period starting in 1980 and ending in year T. The shaded area on the graph corresponds to the 95% confidence bounds, where the standard errors are clustered by match-pair. The expanding window nature of our procedure means that the furthest point to the right covers the entire sample from 1980 to 2012. Over the full sample, identification articles attract 22% more citations than matched non-identification articles. This identification "premium" is statistically significant with a t-statistic of 3.40. The premium was first significantly positive in finance in 1995, suggesting the use of the technology translates into some benefits for researchers and that such benefits were apparent as early as the mid-1990s.[21]

We repeat a similar analysis in Panel B of Figure II for articles published in the leading economics journals (AER, JPE, QJE) but define identification articles using the presence of the identification "language" (see Section IV). Interestingly, we observe an identification

[20] This non-perfect matching is expected because our non-parametric matching is prone to a "curse of dimensionality", where the proportion of matched articles decreases with the number of strata that are imposed. In an unreported analysis, we find similar results if we only match on journals and semester. In addition, we alternatively consider a series of expanding regressions where we regress, for each article, the (logarithm of the) number of citations obtained in a year on a dummy variable indicating whether an article uses the identification technology as well as control variables (log of the maximum of the log of accrued citations across the article's authors, the number of authors, the log of the number of pages, as well as fixed effects corresponding to the interaction terms of the year of publication and age of the paper fixed effects). These regressions use all of the articles in the sample. The results and conclusions are qualitatively similar.

[21] In an unreported analysis, we obtain a very similar pattern if we focus on differences in citations (in levels) instead of differences in the logarithm of citations.

Electronic copy available at: https://ssrn.com/abstract=2378460

premium in economics that started as early as 1982. This result is consistent with the earlier adoption patterns of the identification technology observed in the economics journals (see Sections IV.B and IV.C). Moreover, the magnitude of the identification premium in economics articles is roughly similar to that estimated for empirical corporate finance articles.[22]

A.2 Editorial Boards

To measure the net gains of adopting the identification technology, we would ideally want to gauge whether identification articles are more likely to get published. However, because we only observe published articles (and not all submitted articles), we cannot estimate the likelihood of getting published in finance as a function of whether or not a researcher adopts the technology. Instead, we provide indirect evidence on the likelihood of getting published by focusing on the characteristics of editorial board members at the leading finance journals. We conjecture that the odds of publication for articles using the identification are higher when a larger fraction of editors have already adopted the technology in their own research. The motivation behind this hypothesis is that adopting editorial board members are more likely to be familiar with the identification technology and view more positively studies that also use it.

[Insert Figure III about here]

We gather information on editors and associate editors directly from the journals' websites (for the RFS and JFE) and annual activity reports (for the JF).[23] The data is available starting in 1990 for the RFS and 1997 for the JFE and JF and includes information on 269 individuals. We define both editors and associate editors as editorial board members and (manually) link each member to our sample of all authors. Then, for each member-year observation, we create a dummy variable that equals one if the board member has published at least one identification article by that year and zero otherwise.[24] Figure III displays a substantial change in the composition of journals' editorial boards over our sample period. We find no adopting editors before 1997. The fraction of adopting editors climbs above 20% in 2007 and 40% in 2011 and reaches 52% in 2012. This compositional change is sizable and has occurred across all three top finance journals. This shift in board composition towards

[22]Note however that the larger number of articles in economics journals provide more precise estimates of the identification premium.

[23]The JFE website only reports the current editorial board, so we use http://archive.org/web/ to retrieve annual snapshots of the website.

[24]Note that this definition of adopting editors provides a lower bound estimate of adopting editors, since we only use ECF articles in the top-three finance journals to detect the use of identification techniques in their research.

Electronic copy available at: https://ssrn.com/abstract=2378460

identification techniques confirms the widespread diffusion in the field and is consistent with increased benefits for researcher adopting the identification technology over time.

B Characteristics of Researchers

Next, we exploit the fine granularity of our sample to examine researcher-level determinants of adoption. We specifically focus on factors that can influence a researcher's awareness about the technology and its expected net benefits, as well as factors related to the cost of adopting it: (1) the doctoral training of researchers, (2) their research network, (3) the ranking of their doctoral and current institution, and (4) their seniority in the profession.

We rely on a survival analysis to shed light on the "time-to-adoption" among finance researchers. We use a linear model to estimate the factors that drive the *timing* of adoption for each individual researcher and identifies whether a given factor induces faster adoption. For each author-year, the dependent variable (adoption) is equal to zero before the author adopts the identification technology (i.e., publishes her first ECF identification article in a finance journal), and is equal to one if the author adopts the identification technology in a given year.[25] Once an author adopts, we exclude her from the panel as is standard in survival analysis.

The panel of author-years comprises the career years of all authors who publish at least one empirical corporate finance article in the JF, JFE, or RFS between 1980 to 2012. To capture the publishing career of each author, we rely on the detailed biographical information we collect on PhD year and employment. Each author enters the panel the year when their publishing career begins as indicated by the first of the following three events: (1) PhD graduation year; (2) First academic position; (3) First publication in top-three finance journal. We use this starting date for each author to compute their "professional" age. An author leaves the sample when their publishing career ends as indicated by the last of the following two events: (1) Last academic position ends (retirement); (2) Last publication in top-three finance journal.[26] We drop years before 1980 and the sample ends in 2012, so there is left and right censoring in the data.

Because many authors' characteristics mechanically increase over time (e.g. seniority or the set of former colleagues), we saturate our linear model with age, cohort, and calendar

[25] In unreported test, we define the year of adoption based on publications in ECF articles *as well as* articles in the top-three economics journals (AER, JPE, and QJE). The inferences are unchanged.

[26] The results are robust to altering the conditions based on the year of first and last publication. For example, the results are unchanged if we define the entry year as the first publication year minus three years.

Electronic copy available at: https://ssrn.com/abstract=2378460

year fixed effects. This specification guarantees that our estimates truly isolate the effect of researcher-specific factors on adoption time from any age, cohort, or calendar year effects (see Hall, Mairesse, and Turner (2015) for a discussion of age-cohort-time effects).

B.1 Doctoral Training

We first consider the type of doctoral degree of each researcher. Based on detailed (hand-collected) biographical information for 1,610 scholars, we create a dummy variable that equals one if a researcher holds a PhD degree in economics. The fraction of economic PhDs in the sample is 31.5%. Results reported in the first column of Table V reveal a positive association between adoption and a PhD degree in economics.[27] The estimated coefficient is 0.007, indicating that the adoption of economics PhDs is almost 30% stronger than the average researcher (as the unconditional adoption rate in the panel is 0.024). In terms of actual time-to-adoption, our estimates imply that a median researcher with an economics PhD adopts about 2.3 years faster than the same researcher with any other PhD degree. In particular, the average cohort of economics PhDs reaches the 20% adoption threshold in 12.8 years, compared to 10.5 years for researchers with non-economics PhDs.

[Insert Table V about here]

As documented earlier, the identification technology emerged in economics roughly 20 years prior to its emergence in finance. Hence, doctoral training in economics is likely to have exposed and informed researchers about the identification technology both earlier and more prominently. The quicker adoption of economics PhD holders is thus consistent with the idea that increased exposure improves researchers' awareness and inference about the new technology. In addition, estimates from Table V indicate that the diffusion of innovation across research fields accelerates when researchers with a given training migrate to a neighboring field of research. This result is in line with Stoyanov and Zubanov (2012) which shows that labor mobility produces knowledge spillovers across firms and Azoulay, Graff Zivin, and Manso (2011) which documents similar knowledge spillovers across scientists in the life-sciences.

[27] Non-economics PhDs are coined Finance PhDs in the text for simplicity although we recognize that researchers publishing in empirical corporate finance might have non-economics PhDs from other fields (e.g. accounting, business administration, psychology, mathematics, physics, etc.).

Electronic copy available at: https://ssrn.com/abstract=2378460

B.2 Academic Networks

Next, we consider researchers' academic network. Arguably, knowledge about the expected net gains associated with using the identification technology could be facilitated by interactions with other researchers. As proxies for such interactions, we define three types of social networks. First, we define a network of colleagues (*Adopting Colleagues*) by tracking employment history so as to link a researcher to all other researchers located at the same institution during his tenure at the institution. The network is time-varying and is composed of all past and current colleagues as of year t. We also define two other networks based on alumni ties. In particular, we define a second network that comprises all researchers that receive a PhD from the same PhD granting institution (*Adopting Alumni*) and a third network that links each researcher to all researchers working at his PhD granting institution in the two years preceding his PhD graduation year (*Adopting PhD Faculty*).[28] Because we want to examine the influence of a given researcher's network on his decision to adopt in a given year t, we compute the fraction of a researcher's network members that have adopted the identification technology by the prior year $(t-1)$.

The estimations displayed in column (2) of Table V indicate that the adoption of the identification technology is related to social interactions among academics. We estimate a faster adoption when a larger fraction of a researcher's academic network has already adopted. This result holds for networks of current and former colleagues, as well as network of alumni. The estimated coefficients for the network variables range between 0.003 and 0.011. Given that each network variable is normalized by its standard deviation, these estimates indicate adoptions of the identification technology that are 30% to 45% faster in response to a one standard deviation increase in the adoption rate of the members of their academic networks.

The positive link between a researcher's adoption and that of its peers is consistent with researchers' networks creating positive externalities that facilitate the diffusion of knowledge among individuals. For instance, Azoulay, Graff Zivin, and Wang (2010) finds that co-authorship ties play an important role in researchers productivity, and Head, Li, and Minondo (2015) shows that academic linkages facilitate knowledge flows (measured using theorem citations) among mathematicians. Our results suggest that academic networks significantly reduce the informational constraints associated with learning about the net gains of the new

[28] We manually obtain the career path and institutions of employment for 1,554 authors. For the remaining 246 authors, we infer the institutions from the author information on each publication. We fill years in which an author does not publish by carrying forward all institutions from the last year the author published. We carry back the institution of the first publication.

Electronic copy available at: https://ssrn.com/abstract=2378460

technology and foster the diffusion of innovation.

B.3 Institutions' Rankings

We next examine whether adoption varies across school rankings. We follow the classification of Kim, Morse, and Zingales (2009) to define the top-25 academic institutions. When we focus on the ranking of a researcher's PhD-granting institution, the results reported in the second row of Table V reveal no evidence that the school ranking of a researcher's PhD-granting institution is related to his time-to-adoption. While we find a strong link between a researcher's alumni-based network and his adoption, adoption appear similar for researchers holding PhDs from top-tier institutions than for researchers with PhDs from lower-ranked institutions.

In sharp contrast, estimates in the third row reveal that researchers adopt the identification technology significantly faster when they are employed at top-tier institutions. The estimated coefficients indicate a difference in adoption speed between 11% (column (2)) and 21% (column (1)) when we compare researchers at the top institutions to researchers in lower-ranked institutions. Kim, Morse, and Zingales (2009) show that since the nineties, researchers at top institutions are not associated with more productive research. Yet, we find that these researchers have adopted the identification technology significantly faster.

This positive effect of institution ranking is strong and could be consistent with several explanations. Although we focus on a soft innovation that requires limited costly research equipment (e.g. a lab or a supercomputer), the larger financial or organizational resources allocated to research in top schools could facilitate the faster adoption of new techniques. Alternatively, differences in research incentives and tenure requirements could induce researchers at top schools to be more innovative and adopt new techniques faster. On the other hand, the relationship between ranking and adoption may go the other way; our estimate could simply indicate that top schools are better at selecting future innovators and early adopters.

B.4 Seniority

We find evidence that researchers' seniority and status are significantly related to the diffusion of the identification technology. In columns (3) and (4) of Table V, we include authors' career age (and its squared term) as explanatory determinants of adoption speed and limit the set of fixed effects to cohort and calendar years. We observe a negative association between age and adoption time; more senior authors adopt the technology later. Moreover, results in column

Electronic copy available at: https://ssrn.com/abstract=2378460

(4) indicate that the age-adoption relation is concave, with an implied maximum adoption rate approximately nine years into a researcher's career.

This seniority effect is consistent with models of vintage human capital implying that early-stage researchers are the primary adopters of new technology (e.g. Chari and Hoenhayn (1991) or MacDonald and Weisbach (2004)). Several theories could lead to lower adoption costs for younger, less-established researchers. Lower costs could arise from more flexible minds of young scholars (e.g. Darwin (1859)), their larger exposure to recent innovation, higher incentives, and more time to learn new techniques (Diamond (1980)). Alternatively, more senior scholars could have more vested interests that may render them less receptive to innovation (e.g. Cohen (1985)).

Remarkably, the association between age and adoption mirrors the negative link between adoption time and authors' number of citations observed in every specification of Table V. Insofar as citations take time to accumulate, citation count is likely higher for more senior researchers. Yet, citations is also a commonly-used metric to identify "high-status" researchers. In an attempt to disentangle the effects of seniority and status on adoption, we define a subset of researchers as exhibiting high-status if they have received a best-paper award from one of the top journals, or are fellows of the American Finance Association.[29] Results in column (5) confirm that seniority and status are distinct, as we estimate a faster adoption for the subset of high status researchers.

[Insert Table VI about here]

To better understand the role of seniority in the diffusion of the identification technology, Table VI provides an analysis of coauthorship teams across different career age groups. In Panel A, we examine how seniority relates to solo-authored identification articles. Columns under "Authorship" report the number of authorships by age group for all published articles ("All"), identification articles ("ID"), all solo-authored articles ("Solo"), and solo-authored identification articles ("Solo ID").[30] Columns under "Percent by Career Age" report the likelihood of a solo-authored identification article among three different subsets of publications. The fraction of solo-authored identification articles among all authorships ("Solo ID/All") is 0.2% for authors with more than 20 years seniority and 3.3% for authors in the first 10 years

[29] We consider the Smith-Breeden (1989-) and Brattle Group (1999-) Prizes for the *Journal of Finance*, the Jensen (1997-) and Fama-DFA (1997-) Prizes for the *Journal of Financial Economics*, and the Michael J. Brennan (2000-) and Young Researcher (2006-) Prizes for the *Review of Financial Studies*. There are 81 researchers in the high-status category.

[30] Authorship is credited per author per paper. For instance, two authors on one paper corresponds to two authorships.

Electronic copy available at: https://ssrn.com/abstract=2378460

of their career. The difference in proportions of -3.0% is statistically significant. Furthermore, the difference in proportions of identification articles being solo-authored ("Solo ID/ID") between the senior (row 3) and junior (row 1) authors is -12.2% and statistically significant. The magnitude in the last column, which reports the fraction of identification articles within solo-authored articles, is similar but not significant. These results add to the findings in Table V by showing that senior researchers are more likely to adopt by co-authoring papers.

Panel B of Table VI investigates the co-authorship structure across researchers at different stages of their careers. Each row computes the average career age difference (between an author and their coauthors) for a given career age category. The first row of Panel B focuses on authors with a career age of less than 10 years. Given the positive age gap we find, we conclude, perhaps not surprisingly, that junior authors tend to collaborate with more senior authors. More interestingly, the average age gap relative to coauthors is greater for published work using the identification technology (7.3 years vs. 6 years) and the difference is statistically significant. Corroborating those results, we find that senior researchers (as measured by a career age exceeding 20 years) have younger coauthors on average. Further, we also find a significantly greater average age gap for identification articles compared to non-identification articles (-12.2 years vs. -10.7). While senior researchers typically co-author with more junior researchers, they team up with even younger researchers when they adopt the identification technology.

C Straddling Authors

The aim of this section is to provide a more complete picture of adoption patterns by identifying authors who publish at least one empirical corporate finance article in a finance journal and one article in a economics journal during our sample period. We define these authors as "straddlers". Given the general pattern of diffusion from the field of economics to the field of finance, we expect these straddlers to play a vital role in the transfer of the technology across the two fields. Further, because our estimates of authors' time-to-adoption are based on articles published in finance journals, our interpretation could potentially be biased if a large number of authors in our sample adopt the identification technology earlier, but do so in articles published in the top-three economics journals. To assess the robustness of our findings, we investigate these straddlers and their impact in the diffusion process below.

Table VII presents the characteristics of straddlers. Overall, we identify 291 authors with publications in both disciplines. They represent 16.3% of our sample of authors. For

Electronic copy available at: https://ssrn.com/abstract=2378460

each straddler, we further check whether the article(s) published in the economics journals (a) uses the identification technology, and (b) is in empirical corporate finance (using the same classification as in Section III.A). Panel A reveals that straddlers are more likely to adopt. While 36.4% of all authors adopt the technology in our sample, the proportion rises to 56.5% among straddlers. Importantly, with regard to potential biases, only a small number of straddlers adopt the technology in an economics article prior to adopting it in a finance article (32 authors). An even smaller number of straddlers (8 authors) adopt the technology in an empirical corporate finance article published in a top economics journal before adopting it in a finance journal. Ultimately, among all authors who adopt the identification technology in finance journals, only a handful adopted the identification technology in published work outside of finance journals beforehand. This result dispels the bias concerns raised above.

[Insert Table VII about here]

Panel B reveals that straddlers are significantly more likely to have PhD degrees in economics and degrees from top schools. The former result is to be expected given that we condition our sample of straddlers on having at least one publication in the top-three economics journals (in addition to having a publication in the top-three finance journals).

In Panel C, we show that straddlers adopt sooner the identification technology than non-straddlers. Specifically, when considering publications in both sets of journals, the average year of adoption among straddlers is 2003 compared to 2007 for non-straddlers. Even considering adoption *only* within finance journals, straddlers adopt earlier (2005 vs. 2007). Both differences are statistically significant. The average year of adoption for straddlers is slightly earlier in economics journals (2002) relative to finance journals (2005). Results from Panel C confirm that straddlers are early adopters of the identification technology and hence play a potentially important role in transferring the identification technology to the field of finance.

We test further this conjecture in a multivariate regression of time-to-adoption that controls for age, cohort, and year fixed effects. The last column of Table V replicates the baseline time-to-adoption regression reported in column (1) but adds binary variables to capture the notion of straddlers. Including directly a straddler dummy variable, however, induces substantial look ahead bias. Thus, we define a binary variable, *Prev Econ*, which is equal to one if the author has an article published in one of the top economics journals in previous years. We also define *Prev Econ-ID*, which is equal to one if any of the previous economics articles also used the identification technology. Consistent with the idea that straddling authors represent an important vehicle of technology diffusion, we observe a significantly faster time

Electronic copy available at: https://ssrn.com/abstract=2378460

to adoption for authors with previous identification articles published in economics journals. Importantly, we also see that the estimates of the other determinants are unaltered.

VI Technological Refinements

Our analysis so far has focused on a "static" definition of innovation. However, innovation is by nature a dynamic concept. As with other technologies, we expect the identification technology to evolve over time as more researchers learn its benefits and limitations. Epistemologists have long argued that scientific progress evolves through experimentation or trial-and-error. As a consequence, scientific norms (or what could be considered "best practices") also evolve over time as more scientists use and criticize the new techniques (see Popper (1959) and Kuhn (1962)). To buttress the dynamic nature of the identification technology, we provide evidence of the changing nature of "standard" tests, diagnostics, and procedures performed when implementing identification techniques.

In Figure IV, we focus on the tests and diagnostics that are performed when applying the two most used identification techniques: Instrumental Variables (IV) and Difference-in-differences (DD). We track IV and DD articles over time and verify whether they contain keywords related to diagnostic tests recommended in survey articles such as Roberts and Whited (2013). For IV articles, we count the number of articles that mention the following terms: "weak instrument","Stock and Yogo", "overidentification", "exclusion restriction", or "Hausman test". For DD articles, we count the number of articles that mention "falsification test", "parallel trend", "placebo test", or "Bertrand, Duflo, and Mullainathan (2004)." Common permutations of each keyword are allowed.

[Insert Figure IV about here]

Figure IV reveals a slow and gradual implementation (or discussion) of these diagnostic tests. For instance, whereas IV techniques emerge in the mid-nineties, discussion of "weak instrument" or "overidentification" come about 10 years later. Similarly, whereas difference-in-differences appear around 2000, the discussion of "parallel trend" within ECF papers first appears in 2007. Overall, this gradual emergence of different diagnostic tools over our sample period suggests that researchers are learning-by-doing and refine the implementation of these identification techniques over time (see the surveys by Hall and Khan (2002) or Foster and Rosenzweig (2010) for a detailed discussion of learning-by-doing). More generally, the patterns in Figure IV are consistent with changing scientific "norms" over time.

Electronic copy available at: https://ssrn.com/abstract=2378460

As put forth by Kuhn (1962), such a changing nature constitutes a "healthy" scientific process, whereby new techniques and discoveries are first introduced, then scrutinized and criticized, and as a result improved. The identification technology is no exception. For instance, several recent papers cast doubt on the merits and appropriateness of the identification technology (e.g. Hennessy and Strebulaev (2015)), while others evaluate whether the implementation of the technology is performed "correctly" in practice (e.g. Atanasov and Black (Forthcoming)). This line of research is highly constructive as it provides a description of *current* best practices. Yet, as we document in this section, the evolving nature of scientific discoveries make it difficult to pass judgment on the merits of a particular application at any given point in time.

VII Conclusion

This study documents a secular rise of the identification technology in corporate finance research as of the mid-nineties. Although this rise lags the emergence of these techniques in the top economics journals by approximatively 20 years, the adoption rate is steeper in finance, with more than 50% of all published articles in empirical corporate finance having adopted the technology by 2012. We find significant heterogeneity in the speed of adoption across finance researchers. Consistent with this choice being driven by each individual's awareness of the technology and its net benefits, we find that scholars that are less senior, hold a PhD in economics, and work for the top institutions display faster adoption. In addition, we find evidence that a researcher's adoption is related to that of his research networks composed of colleagues and alumni sharing the same alma mater. Our findings highlight new forces that facilitate the diffusion of innovation in academic research.

Our results suggest several interesting avenues for future research. In particular, it would be interesting to examine how the identification technology disseminates into other related research fields, such as accounting, strategy, or management, and estimate general adoption lags across disciplines. More generally, it would be interesting to study the diffusion of other research tools (e.g. textual analysis or computing power) and assess whether similar economic forces are at play.

From a practical perspective, our findings point to important changes in the publication norms in the field of corporate finance. These changes can have potentially important implications, particularly for the career path of young scholars. Competition for space in the

Electronic copy available at: https://ssrn.com/abstract=2378460

top journals has grown fiercer over time, making it increasingly difficult for young scholars to achieve a given set of publication benchmarks.[31] At the same time, editorial boards are now composed of researchers that have adopted the identification technology. Young authors, in particular those with training in economics and those favored by top schools, seem to have largely responded to the incentives created by this new environment by adopting the identification technology much faster in recent years.

[31] Based on information provided by the journals, the overall acceptance rates for submission at the top-three journals is about one third as high today as it was in the early nineties.

Electronic copy available at: https://ssrn.com/abstract=2378460

References

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica 60 (2), 323–351.

Agrawal, A., 2013. The impact of investor protection law on corporate policy and performance: Evidence from the blue sky laws. Journal of Financial Economics 107 (2), 417–435.

Agrawal, A., Goldfarb, A., 2008. Restructuring research: Communication costs and the democratization of university innovation. American Economic Review 98 (4), 1578–1590.

Alevy, J. E., Haigh, M. S., List, J. A., 2007. Information cascades: Evidence from a field experiment with financial market professionals. The Journal of Finance 62 (1), 151–180.

Alexander, J., Marby, R., 1994. Relative significance of journals, authors, and articles cited in financial research. Journal of Finance 49 (2), 697–712.

Angrist, J., Pischke, J.-S., 2010. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Journal of Economic Perspectives 24 (2), 3–30.

Atanasov, V., Black, B., Forthcoming. Shock-based causal inference in corporate finance and accounting research. Critical Finance Review.

Azoulay, P., Graff Zivin, J. S., Manso, G., 2011. Incentives and creativity: Evidence from the academic life science. Rand Journal of Economics 42 (3), 527–554.

Azoulay, P., Graff Zivin, J. S., Sampat, B., 2012. The diffusion and scientific knowledge across time and space: Evidence from professional transitions for the superstars of medicine. In: Lerner, J., Stern, S. (Eds.), The Rate and Direction of Inventive Activity Revisited. University of Chicago Press, Ch. 2, pp. 107–155.

Azoulay, P., Graff Zivin, J. S., Wang, J., 2010. Superstar extinction. The Quarterly Journal of Economics 125 (2), 549–589.

Azoulay, P., Stuart, T., Wang, Y., 2014. Matthew: Effect or fable? Management Science 60 (1), 92–109.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance 59 (4), 1481–1509.

Bayer, P., Hjalmarsson, R., Pozen, D., 2009. Building criminal capital behind bars: Peer effects in juvenile corrections. Quarterly Journal of Economics 124 (1), 105–147.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust difference-in-differences estimates. Quarterly Journal of Economics 119 (1), 249–275.

Borokhovich, K., Bricker, R., Brunarski, K., Simkins, B., 1995. Finance research productivity and influence. Journal of Finance 50 (5), 1691–1717.

Bris, A., Welch, I., Zhu, N., 2006. The costs of bankruptcy: Chapter 7 liquidation versus chapter 11 reorganization. Journal of Finance 61 (3), 1253–1303.

Brogaard, J., Engelberg, J., Parsons, C., 2014. Networks and productivity: Causal evidence from editor rotations. Journal of Financial Economics 111 (1), 251–270.

Campa, J. M., Kedia, S., 2002. Explaining the diversification discount. Journal of Finance 57 (4), 1731–1762.

Campello, M., Galvao, A., Juhl, T., 2013. Policy heterogeneity in empirical corporate finance. Working Paper. Cornell University.

Electronic copy available at: https://ssrn.com/abstract=2378460

Card, D., DellaVigna, S., 2013. Nine facts about top journals in economics. Journal of Economic Literature 51 (1), 144–161.

Chari, V., Hoenhayn, H., 1991. Vintage human capital, growth, and the diffusion of new technology. Journal of Political Economy 99 (6), 1142–1165.

Chava, S., Roberts, M., 2008. How does financing impact investment? the role of debt covenant violations. Journal of Finance 63 (5), 2085–2121.

Cohen, B., 1985. Revolution in Science. Harvard University Press: Cambridge, MA.

Comin, D., Hobijn, B., 2010. An exploration of technology diffusion. American Economic Review 100 (5), 2031–2059.

Corrado, C., Ferris, S., 1997. Journal influence on the design of finance doctoral education. Journal of Finance 52 (5), 2091–2102.

Cox, J. C., Smith, V. L., Walker, J. M., 1984. Theory and behavior of multiple unit discriminative auctions. The Journal of Finance 39 (4), 983–1010.

Darwin, C., 1859. On the origin of species. John Murray: London.

Diamond, A., 1980. Age and the acceptance of cliometrics. Journal of Economic History 40 (4), 838–841.

Ding, W., Levin, S., Stephan, P., Winkler, A., 2010. The impact of informaton technology on academic scientists' productivity and collaboration patterns. Management Science 56 (9), 1439–1461.

Ellison, G., 2002. Evolving standards for academic publishing: A q-r theory. The Journal of Political Economy 110 (5), 994–1034.

Foster, A., Rosenzweig, M., 2010. Microeconomics of technology adoption. Annual Review of Economics 2, 395–424.

Furman, J., Stern, S., 2011. Climbing atop the shoulders of giants: The impact of institutions on cumulative research. aer 101 (5), 1933–1963.

Giroud, X., Mueller, H., Stomper, A., Westerkamp, A., 2012. Snow and leverage. Review of Financial Studies 25 (3), 680–710.

Gormley, T., Matsa, D., 2011. Growing out of trouble? corporate responses to liability risk. Review of Financial Studies 24 (8), 2781–2821.

Gormley, T., Matsa, D., 2014. Common errors: How to (and not to) control for unobserved heterogeneity. Review of Financial Studies 27 (2), 617–661.

Griliches, Z., 1957. Hybrid corn: An exploration in the economics of technological change. Econometrica 25 (4), 501–522.

Hall, B., Khan, B., 2002. Adoption of new technology. In: Jones, D. (Ed.), New Economy Handbook. Elsevier Science, Amsterdam, The Netherlands, pp. 5–20.

Hall, B., Mairesse, J., Turner, L., 2015. Identifying age, cohort, and period effects in scientific research productivity: Discussion and illustration using simulated and actual data on french physicists. Economics of Innovation and New Technology 16 (2), 159–177.

Hamermesh, D., Pfann, G., 2012. Reputation and earnings: The roles of quality and quantity in academe. Economic Inquiry 50 (1), 1–16.

Hamermesh, D. S., 2013. Six decades of top economics publishing: Who and how? Journal of Economic Literature 51 (1), 1–11.

Electronic copy available at: https://ssrn.com/abstract=2378460

Hamermesh, D. S., Johnson, G. E., Weisbrod, B. A., 1982. Scholarship, citations and salaries: Economic rewards in economics. Southern Economic Journal 49 (2), 472–481.

Head, K., Li, Y. A., Minondo, A., 2015. Geography, ties, and knowledge flows: Evidence from citations in mathematics. Working Paper. University of British Columbia.

Hennessy, C., Strebulaev, I., 2015. Beyond random assignment: Credible inference of causal effects in dynamic economies. NBER Working Paper No. 20978.

Jones, B., 2009. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? Review of Economic Studies 76 (1), 283–317.

Karolyi, A., 2011. The ultimate irrelevance proposition in finance? Financial Review 46 (4), 485–512.

Karolyi, A., Forthcoming. Home bias, an academic puzzle? Review of Finance.

Kaufman, G., 1984. Rankings of finance departments by faculty representation on editorial boards of professional journals: A note. Journal of Finance 39 (4), 1189–1197.

Keenan, M., 1991. Fifty years of the American Finance Association. Journal of Finance 46 (3), 1125–1156.

Kim, E, H., Morse, A., Zingales, L., 2009. Are elite universities losing their competitive edge? Journal of Financial Economics 93 (3), 353–381.

Kremer, M., Miguel, E., 2007. The illusion of sustainability. Quarterly Journal of Economics 122 (3), 1007–1065.

Kuhn, T., 1962. The Structure of Scientific Revolutions. The University of Chicago Press.

Li, K., Prabhala, N., 2007. Self-selection models in corporate finance. In: Eckbo, B. E. (Ed.), Handbook of Empirical Corporate Finance. Elsevier, Amsterdam, Netherlands, Ch. 2, p. 37–83.

MacDonald, G., Weisbach, M., 2004. The economics of has-beens. Journal of Political Economy 112 (1 pt. 2), 289–310.

Mansfield, E., 1961. Technical change and the rate of imitation. Econometrica 29 (4), 741–766.

Mokyr, J., 2002. The gifts of Athena: Historical origins of the knowledge economy. Princeton University Press.

Moore, W., Newman, R., Turnbull, G., 2001. Reputational capital and academic pay. Economic Inquiry 39 (4), 663–671.

Niemi, A., 1987. Institutional contributions to the leading finance journals, 1975 through 1986: A note. Journal of Finance 42 (5), 1389–1397.

Petersen, M. A., 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. Review of Financial Studies 22 (1), 435–480.

Popper, K., 1959. The Logic of Scientific Discovery. London: Hutchinson.

Roberts, M., Sufi, A., 2009. Control rights and capital structure: An empirical investigation. Journal of Finance 64 (4), 1657–1695.

Roberts, M. R., Whited, T. M., 2013. Endogeneity in empirical corporate finance, . In: Constantinides, G., Stulz, R., Harris, M. (Eds.), Handbook of the Economics of Finance. Vol. 2, Part A. Elsevier, Amsterdam, Netherlands, Ch. 7, pp. 493–572.

Rogers, E., 2003. Diffusion of Innovation, 5th Edition. Free Press: New York.

Electronic copy available at: https://ssrn.com/abstract=2378460

Stock, J., Trebbi, F., 2003. Who invented instrumental variable regression? Journal of Economic Perspective 17 (3), 177–194.

Stoyanov, A., Zubanov, N., 2012. Productivity spillovers across firms through worker mobility. American Economic Journal: Applied Economics 4 (2), 168–198.

Strebulaev, I. A., Whited, T., 2012. Dynamic models and structural estimation in corporate finance. Foundations and Trends in Finance 6 (1-2), 1–163.

Teodoridis, F., 2015. Generalists, specialists, and the direction of inventive activity. Working Paper. University of Southern California.

Thompson, S. B., 2011. Simple formulas for standard errors that cluster by both firm and time. Journal of Financial Economics 99 (1), 1–10.

Weber, J., 1973. A note on keeping abreast if developments in the field of finance. Journal of Finance 28 (1), 161–165.

Welch, I., 2014. Referee recommendations. Review of Financial Studies 27 (9), 2773–2804.

Williams, H., 2013. Intellectual property rights and innovation: Evidence from the human genome. Journal of Political Economy 121 (1), 1–27.

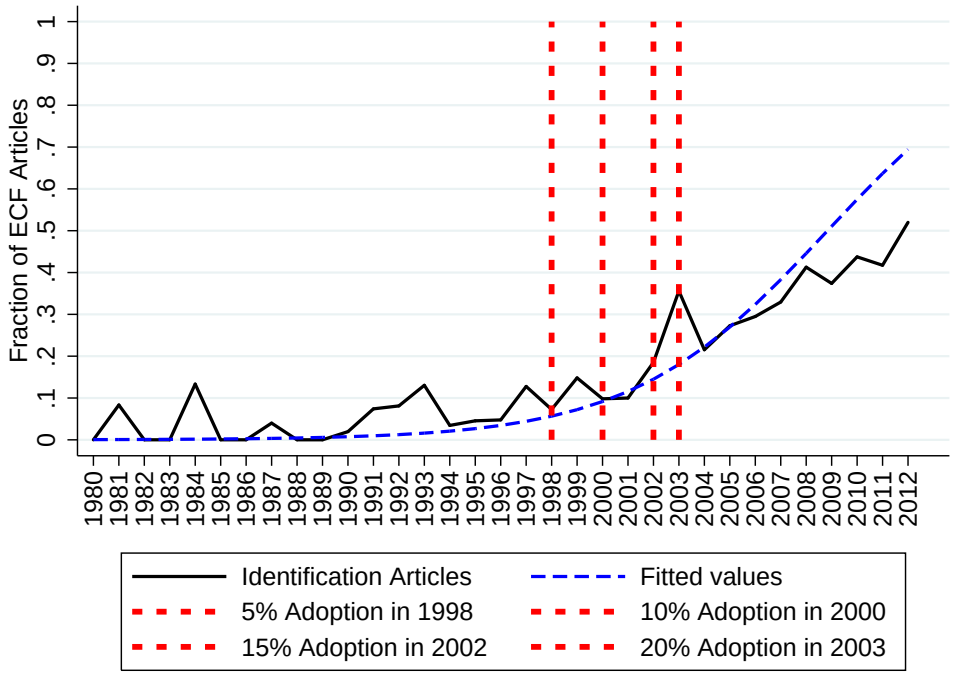
Wright, P., 1928. The Tariff on Animal and Vegetable Oils. MacMillan, New York.

Zivney, T., Bertin, W., 1992. Publish or perish: What competition is really doing. Journal of Finance 47 (1), 295–329.

Electronic copy available at: https://ssrn.com/abstract=2378460

Figure I: Adoption of the Identification Technology

This figure presents the evolution over time of the fraction of empirical corporate finance (ECF) articles published in the JF, JFE, or RFS that use the identification technology as well as the fitted values obtained from estimating the baseline diffusion model specified in Equation (1) and estimated in column (1) of Table II. The sample period is from 1980 to 2012. The five techniques that make up the identification technology are: (1) Instrumental variables; (2) Difference-in-differences; (3) Selection models; (4) Regression discontinuity design; and (5) Randomized experiments. The classification is detailed in Section III.B. The solid line represents the actual fraction, while the dashed line represents the fitted fraction. We include four vertical lines to highlight the years where the fitted fraction crosses the 5%, 10%, 15%, and 20% adoption thresholds.

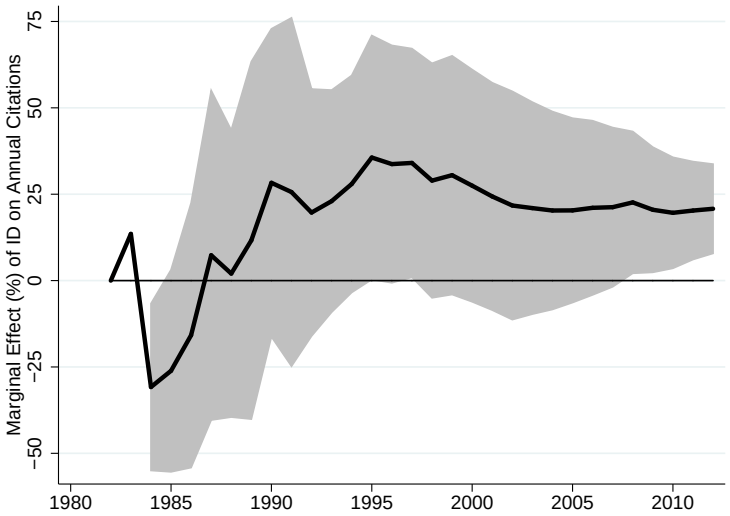


Electronic copy available at: https://ssrn.com/abstract=2378460

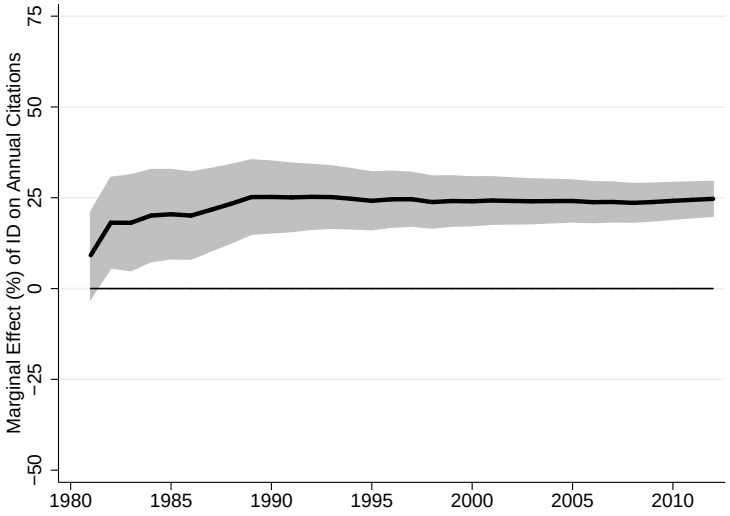
Figure II: Evolving Relationship Between Citations and Adoption

This table presents the results of a citation analysis on two distinct samples. The sample in Panel A consists of all ECF identification articles published in the JF, JFE, and RFS over the period 1980-2012. The sample in Panel B consists of all identification articles published in the top-three economics journals (AER, JPE, and QJE) over the same period. Following Azoulay, Stuart, and Wang (2014), we select one matched non-identification article for each identification article. Details on the matching procedure can be found in Section V.B. For each year t after an identification article i is published, we compute the difference, $\Delta_{i,t}$, in log of citations between the identification article and its match. For each year T between 1980 and 2012, we compute the average of $\Delta_{i,t}$, where $t \in [1980, T]$, and report this as the solid line below. The gray area represents the 95% confidence interval based on standard errors clustered by matched pair. In Panel A, an article is classified as an identification article as detailed in Section III.B. In Panel B, an article is classified as an identification article if the article adopts the language of identification by using any of the identification technology keywords displayed in Table A2. Citation data comes from Web of Knowledge.

Panel A: Empirical Corporate Finance



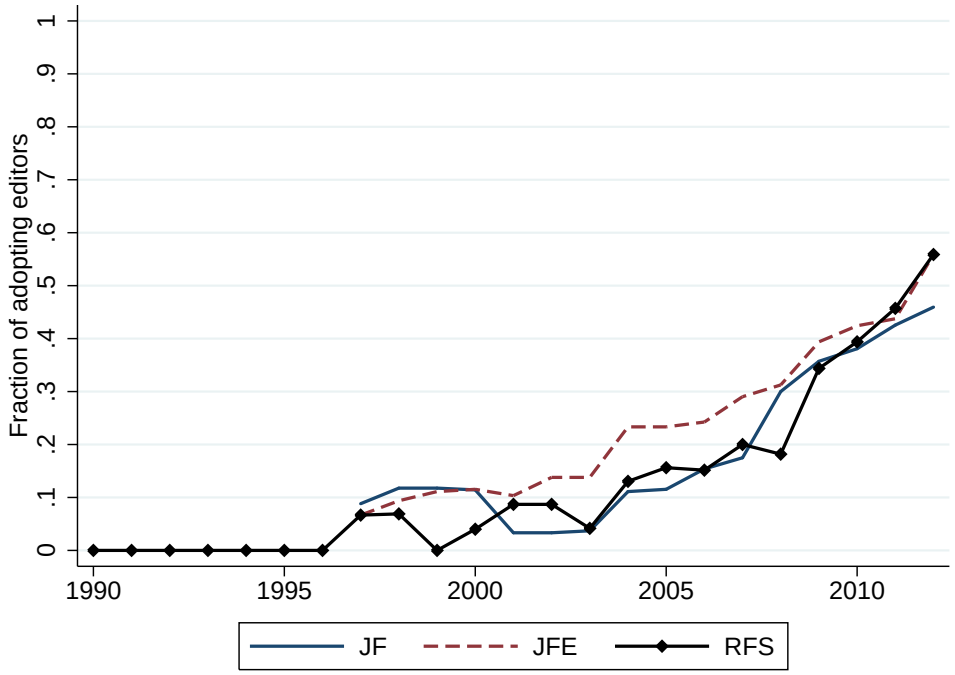
Panel B: Economics



Electronic copy available at: https://ssrn.com/abstract=2378460

Figure III: Evolution of Adopting Editors

This figure presents the evolution over time of the fraction of editorial board members (editors and associate editors) at the JF, JFE, or RFS that have adopted the identification technology ("Adopting Editors"). The sample period is from 1990 to 2012. We define a board member in a given year as an Adopting Editor if he or she has written at least one paper that uses the technology by that year. The classification of articles into identification and non-identification is detailed in Section III.B.

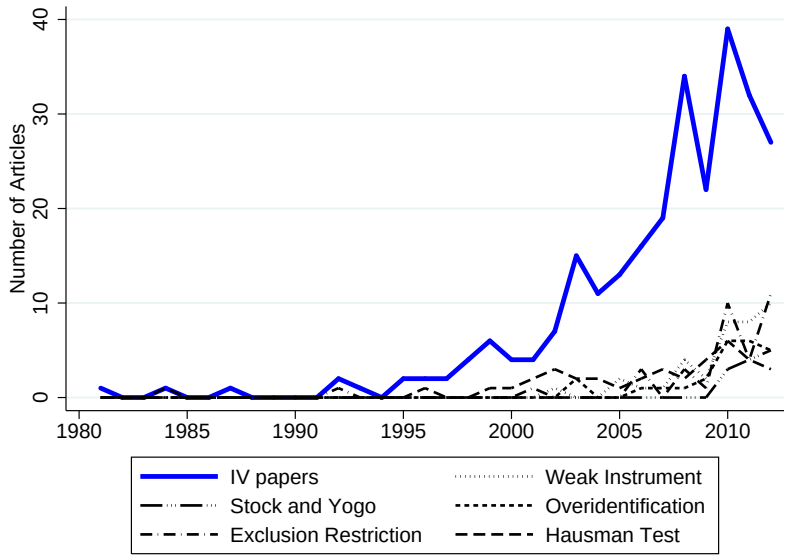


Electronic copy available at: https://ssrn.com/abstract=2378460

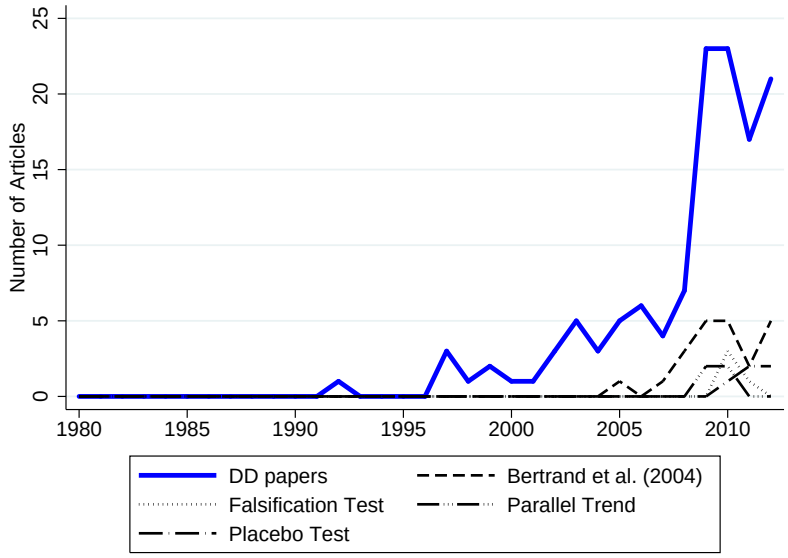
Figure IV: Technological Refinements

This figure presents the evolution over time of the number of ECF articles published in the JF, JFE, or RFS that adopt the instrumental variables ("IV") approach or difference-in-differences ("DD') approach, as well as the implementation of diagnostic tests associated with these techniques. The sample period is from 1980 to 2012. Each ECF article is classified into an IV or DD category, depending on whether it adopts the technique (see Section III.B). We further search within these classified articles for keywords associated with diagnostic tests ("weak instruments","Stock and Yogo", "overidentification", "exclusion restriction", or "Hausman test" for IV; and "falsification test", "parallel trend", "placebo test", or "Bertrand, Duflo, and Mullainathan (2004)" for DD). Common permutations of each keyword are allowed. Panel A displays the prevalence of the IV technique (solid line) and its associated diagnostic terms (dashed lines). Panel B displays the prevalence of the DD technique (solid line) and its associated diagnostic terms (dashed lines).

Panel A: IV



Panel B: DD



Electronic copy available at: https://ssrn.com/abstract=2378460

Table I: Descriptive Statistics

The sample comprises all articles published in the Journal of Finance (JF), Journal of Financial Economics (JFE), and Review of Financial Studies (RFS) between 1970 and 2012. Panel A reports the number and fraction of articles that are classified as Empirical, Corporate Finance (CF), and Empirical Corporate Finance (ECF) by decade, as well as for the whole sample period (Section III.A provides more details). Panel B presents the use of the five techniques that compose the identification technology in ECF articles. Each ECF article is classified as having adopted or not these techniques based on a three-step procedure as detailed in Section III.B. The columns report the number and fraction of ECF papers using these different identification techniques along with the year the technique was first used in ECF.

Panel A: Classification of Sample Articles

Classification:		Empirical		CF		ECF	
	Total	N	(%)	N	(%)	N	(%)
Total	6,797	4,889	72%	2,438	36%	1,796	26%
1970s	976	513	53%	220	23%	78	8%
1980s	1,228	732	60%	356	29%	226	18%
1990s	1,519	1,141	75%	534	35%	409	27%
2000s	2,196	1,766	80%	921	42%	740	34%
2010-2012	878	737	84%	407	46%	343	39%

Panel B: Use of Identification Technologies

	N	(% of ECF)	First Use
Any identification technology (ID)	408	22.7%	1977
Instrumental variables (IV)	266	14.8%	1977
Difference-in-differences (DD)	126	7.0%	1992
Selection models (Selection)	72	4.0%	1990
Regression discontinuity design (RDD)	10	0.6%	2006
Randomized experiments (Experiments)	5	0.3%	1984

Electronic copy available at: https://ssrn.com/abstract=2378460

Table II: Adoption of the Identification Technology

This table presents the estimates and fitted values of the empirical model of technology diffusion (S-curve) presented in Equation (1). The sample corresponds to a time series of 33 annual observations computed from all ECF articles published in the JF, JFE, and RFS over the period 1980-2012. In column (1), the dependent variable (p_t) is the fraction of ECF articles classified as using the identification technology in year t. In columns (2) to (4), we examine separately the three most prominent identification techniques found in the sample (respectively IV, DD, Selection). The classification procedure is detailed in Section III.B. Panels A, B, and C report the estimates and implied adoption thresholds for three different ceiling levels, $k = 1$, $k = 0.8$, and $k = 0.6$ respectively. An implied adoption threshold year is defined as the year in which the fitted values ($\tilde{p}$) cross a given fraction of the sample population (respectively 5%, 10%, 15%, 20%, 50%). For brevity, we report the estimates only for the specification in Panel A. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	ID	IV	DD	Selection
Panel A: Ceiling Parameter k = 1				
β	0.260***	0.263***	0.306***	0.229***
	(5.91)	(5.99)	(18.65)	(10.43)
α	-7.487***	-8.416***	-10.77***	-10.01***
	(-7.21)	(-7.96)	(-24.07)	(-18.12)
R^2	0.598	0.553	0.782	0.473
$\hat{p}$ =5%	1998	2001	2006	2011
$\hat{p}$ =10%	2000	2004	2008	—
$\hat{p}$ =15%	2002	2005	2010	—
$\hat{p}$ =20%	2003	2007	2011	—
$\hat{p}$ =50%	2009	2012	—	—
Panel B: Ceiling Parameter k = 0.8				
$\hat{p}$ =5%	1996	2000	2005	2010
$\hat{p}$ =10%	1999	2003	2007	—
$\hat{p}$ =15%	2001	2004	2009	—
$\hat{p}$ =20%	2002	2006	2010	—
$\hat{p}$ =50%	2007	2011	—	—
Panel C: Ceiling Parameter k = 0.6				
$\hat{p}$ =5%	1995	1998	2004	2008
$\hat{p}$ =10%	1997	2001	2006	2012
$\hat{p}$ =15%	1999	2003	2008	—
$\hat{p}$ =20%	2000	2004	2009	—
$\hat{p}$ =50%	2005	2009	—	—

Electronic copy available at: https://ssrn.com/abstract=2378460

Electronic copy available at: https://ssrn.com/abstract=2378460

Table III: Adoption in Economics and Finance Journals

This table presents the estimates and fitted values of the empirical model of technology diffusion (S-curve) presented in Equation (1) applied to two distinct samples. The first sample ("Econ.") consists of all articles published in the top-three economics journals (AER, JPE, and QJE). The second sample ("ECF") consists of all empirical corporate finance articles published in the JF, JFE, and RFS over the period 1980-2012. In all estimations, the dependent variable (p_t) is the fraction of articles that adopt the *language* associated with the identification technology in year t. We define that an article adopts the language of identification if it mentions at least one of the keywords associated with the identification technology (see Table A2 for details). Panel A reports the position (α) and slope (β) estimates of the logistic S-curve. Panel B reports the implied adoption thresholds, which are the years in which the fitted values ($\tilde{p}$) cross several adoption thresholds. We report tests of the difference between the slope coefficient and fitted threshold years across samples (Econ. vs ECF). The last row displays the adoption lag computed as the average difference in number of years for the crossing of the four adoption thresholds. We report estimates and fitted values for three different ceiling levels, $k = 1$, $k = 0.8$, and $k = 0.6$. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Ceiling (k):	$k = 1$			$k = 0.8$			$k = 0.6$		
	Econ.	ECF	Difference	Econ.	ECF	Difference	Econ.	ECF	Difference
Panel A: Estimation of S-curves									
β	0.068***	0.209***	-0.141***	0.080***	0.224***	-0.144***	0.115***	0.255***	-0.140***
	(18.63)	(5.59)		(17.99)	(5.99)		(11.62)	(5.55)	
α	-1.808***	-5.520***		-1.602***	-5.406***		-1.454***	-5.327***	
	(-26.55)	(-7.94)		(-19.35)	(-7.76)		(-8.35)	(-6.86)	
R^2	0.918	0.502		0.913	0.537		0.823	0.524	
Panel B: Implied Adoption Thresholds									
$\hat{p}$ =5%	1963	1992	-29***	1963	1991	-28***	1967	1989	-22***
$\hat{p}$ =10%	1974	1996	-22***	1973	1994	-21***	1974	1992	-18***
$\hat{p}$ =15%	1981	1998	-17***	1978	1996	-18***	1978	1994	-16***
$\hat{p}$ =20%	1986	2000	-14***	1983	1998	-15***	1981	1995	-14***
Average:			-20.5			- 20.5			-17.5

Table IV: Broader Definition of Empirical Corporate Finance

This table presents results for the same adoption model (S-curve) as in Table II estimated on a broader sample of empirical corporate finance (ECF) articles, which also includes ECF articles published in the top economics journals (AER, JPE, and QJE). Classification of the articles in these journals follows the same procedure used for finance articles, which is detailed in Section III.B. For brevity, only the adoption threshold years are reported in the different panels. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	ID	IV	DD	Selection
Panel A: Ceiling Parameter k = 1				
$\hat{p}$ =5%	1990	1996	2004	2010
$\hat{p}$ =10%	1996	2001	2006	—
$\hat{p}$ =15%	1999	2004	2008	—
$\hat{p}$ =20%	2002	2007	2009	—
$\hat{p}$ =50%	—	—	—	—
Panel B: Ceiling Parameter k = 0.8				
$\hat{p}$ =5%	1988	1994	2003	2009
$\hat{p}$ =10%	1993	1999	2005	—
$\hat{p}$ =15%	1997	2003	2007	—
$\hat{p}$ =20%	1999	2005	2008	—
$\hat{p}$ =50%	2010	—	—	—
Panel C: Ceiling Parameter k = 0.6				
$\hat{p}$ =5%	1986	1992	2002	2008
$\hat{p}$ =10%	1991	1997	2004	2011
$\hat{p}$ =15%	1994	2000	2006	—
$\hat{p}$ =20%	1996	2002	2007	—
$\hat{p}$ =50%	2006	2012	2012	—

Electronic copy available at: https://ssrn.com/abstract=2378460

Table V: Determinants of Adoption

This table presents the results of a time-to-adoption analysis using OLS. The sample comprises the author-years between 1980 and 2012 of the careers of authors who publish at least one ECF article in the JF, JFE, or RFS. See Section V.B for more details about the panel data construction. The dependent variable is an indicator variable that equals one once an author adopts the identification technology in any article in the ECF sample, and zero before. Observations for an author after adoption are dropped as is standard. Each ECF article is classified as using identification techniques based on a three-step procedure as detailed in Section III.B. *Econ PhD* equals one if the author obtained a PhD in Economics. *TopPhD* equals one if the author holds a PhD from a top-25 institution, as defined by Kim, Morse, and Zingales (2009). *Prev Papers* is the number of articles the author published in the JF, JFE, or RFS by the prior year. *Prev Cites* is the number of citations (via Web of Knowledge) received by previously published articles up to the prior year. The *Colleagues* network comprises authors that were previously employed at the same institution in the same year. The *Alumni* network comprises authors that previously received a PhD from the same institution. The *PhD Faculty* network comprises authors that were on staff at the PhD-granting institution in the two years preceding graduation with a PhD. *Adopting Colleagues, Adopting Alumni, and Adopting PhD Faculty* are the fraction of each network that adopted the identification technology by the prior year. *Career Age* is the number of years since an author's career began (see Section V.B). *High Status* is a binary variable equal to one if the author has received a best-paper award or is a fellow of the American Finance Association. *Prev Econ* is a binary variable equal to one if the author has an article published in one of the top economics journals in previous years. *Prev Econ-ID* is a binary variable equal to one if the author has an article using the identification technology published in one of the top economics journals in previous years. To facilitate interpretation, *Log(1+Prev Papers), Log(1+Prev Cites)*, and the adopting network variables are standardized. Standard errors are clustered by author. We report t-statistics in parentheses. The symbols ***,**, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Econ PhD	0.007***	0.005*	0.007***	0.007***	0.007***	0.005**
	(2.96)	(1.73)	(3.19)	(2.95)	(3.03)	(2.32)
Top PhD	-0.002	-0.001	-0.002	-0.002	-0.002	-0.002
	(-0.67)	(-0.20)	(-0.73)	(-0.64)	(-0.84)	(-0.74)
Top School (Current)	0.005**	0.003	0.005**	0.005**	0.004*	0.004*
	(2.12)	(1.06)	(2.03)	(2.27)	(1.86)	(1.80)
Log(1+Prev Papers)	0.001	0.001	0.005**	0.002	0.005**	0.002
	(0.43)	(0.37)	(2.20)	(0.76)	(2.05)	(0.70)
Log(1+Prev Cites)	-0.005*	-0.005	-0.009***	-0.007**	-0.010***	-0.006**
	(-1.91)	(-1.59)	(-3.38)	(-2.49)	(-3.65)	(-2.10)
Adopting Colleague		0.011***				
		(4.02)				
Adopting Alumni		0.011***				
		2.84)				
Adopting PhD Faculty		0.003				
		(1.36)				
Career Age			-0.001**	0.003***	-0.001**	
			(-2.13)	(4.54)	(-2.11)	
Career Age^2				-0.000***		
				(-8.22)		
High Status					0.025**	
					(2.23)	
Prev Econ						-0.001
						(-0.22)
Prev Econ-ID						0.023**
						(2.45)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Career Age FE	Yes	Yes	No	No	No	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	21,383	16,899	21,383	21,383	21,383	21,383
Authors	1,473	1,190	1,473	1,473	1,473	1,473
Adj. R^2	0.051	0.056	0.048	0.050	0.049	0.52
Avg. Adoption Rate	0.024	0.027	0.024	0.024	0.024	0.24

Electronic copy available at: https://ssrn.com/abstract=2378460

Table VI: Seniority

This table presents publication patterns in ECF and the subset of identification articles by career age groups. Panel A focuses on characteristics of solo-authored work, while Panel B focuses on the nature of collaborative work. The sample comprises the author-years between 1980 and 2012 of the careers of authors who publish at least one ECF article in the JF, JFE, or RFS. Each ECF article is included in the "identification" or "non-identification" category based on a three-step procedure as detailed in Section III.B. Each row corresponds to a 10 year career age bin. We set year zero of a career to the minimum of the year of PhD, first publication, or first academic position. We omit observations 30 years after a career begins and any author whose career began before 1970. In columns (1) through (4) of Panel A, *All* and *ID* count the number of authorships among all ECF articles (respectively all ECF identification articles) for each age bin, where every author on an article receives full authorship credit. *Solo* and *Solo ID* count the number of authorships among all solo-authored articles, and respectively all identification solo-authored articles. Columns (5) through (7) show proportions of solo-authored identification authorships relative to all authorships, all authorships of identification papers, and all solo-authorships respectively. The bottom rows *Diff* and *P-value* compute the difference (and statistical significance) in the proportions of solo-authored identification articles between the first age group (Career Age 0-9) and the last age group (Career Age 20-29). Panel B reports the average age difference between an author and coauthors. We compute the average coauthor age for each author-year, and then average the age gap within the given age group. Higher values indicate more senior coauthors. For each group, we compute the average age gap for the subset of coauthors of identification articles (column (1)) and non-identification articles (column (2)). The difference in the average age gap across the two types of articles is computed along with the t-statistic from a test of equality in means in columns (3) and (4), respectively.

Panel A: Solo-Authored Identification Articles

	Authorships				Percent by Career Age		
Career Age	All	ID	Solo	Solo ID	Solo ID/All	Solo ID/ID	Solo ID/Solo
0-9	1,960	485	256	64	3.3%	13.2%	25.0%
10-19	1,208	281	59	15	1.2%	5.3%	25.4%
20-29	460	105	7	1	0.2%	1.0%	14.3%
Diff					-3.0%	-12.2%	-10.7%
P-value					<0.01	<0.01	0.52

Panel B: Avg. Age Difference of Coauthors by Article Type

Career Age	ID	Non-ID	Difference	T-Stat
0-9	7.3	6.0	1.3	3.04
10-19	-2.3	-2.1	-0.2	-0.37
20-29	-12.2	-10.7	-1.5	-1.70

Electronic copy available at: https://ssrn.com/abstract=2378460

Table VII: Straddlers

This table explores the characteristics of authors straddling between economics and finance journals. We define authors as "straddlers" if they have at least one publication in a top-three finance journal and one publication in a top-three economics journal at any point in their career. The sample contains all authors who publish at least one ECF article in the JF, JFE, or RFS during 1980 - 2012. Panel A provides a comparison between straddlers and non-straddlers in terms of adoption rate. *Adopting Author* is a binary variable equal to one if an author adopts the identification technology in at least one published paper over the sample period. Panel B offers descriptive statistics on the academic background (training) of straddlers vs. non-straddlers in terms of field of PhD studies and PhD institution. *Econ PhD* equals one if the author obtained a PhD in Economics. *TopPhD* equals one if the author holds a PhD from a top-25 institution, as defined by Kim, Morse, and Zingales (2009). In Panel C, we compute in the first row the average first year of adoption for all authors (column (1)), non-straddlers (column (2)) and straddlers (column (3)). The second and third row provide the average first year of adoption only within the subset of the top-three finance journals and top-three economics journals respectively. In column (4), we compute a Z-statistic for the test of equality in proportions across straddlers and non-straddlers in Panel A and B, and a t-statistic from a test of equality in means in Panel C.

Author sample:	All	Non-straddlers	Straddlers	Z-stat/T-stat
Panel A: Authors				
All	1,780	1,489	291	
Adopting Authors	648	484	164	
Adopting Authors	36.4%	32.5%	56.4%	-7.86
Panel B: Author Training				
Econ PhD	30.9%	23.7%	65.8%	-14.41
Top PhD	57.7%	53.4%	79.8%	-8.37
Panel C: Year of Author's First ID Article by Field				
Overall	2006	2007	2003	9.48
Fin	2007	2007	2005	4.11
Econ	2002	N/A	2002	

Electronic copy available at: https://ssrn.com/abstract=2378460

Appendix A Detailed Classification

This Appendix details our procedure to identify articles that use the identification technology. Table A1 presents the list of keywords we use for each of the five identification techniques we use to flag articles. Below, we describe the criteria we use to assign flagged articles into each category, and remove all false positives (i.e. flagged articles that do *not* use identification techniques). Table A2 lists the keywords that we use to define the language associated with the identification technology.

- **Instrumental Variables**. The IV designation requires an article to apply an IV, 2SLS, 3SLS, or GMM estimation and show the results in the article. 215 articles satisfy this requirement. Roberts and Whited (2013) cite Giroud, Mueller, Stomper, and Westerkamp (2012) as a successful implementation of IV in corporate finance. An additional 51 articles are also assigned to the IV category for the following reasons: (1) Untabulated IV results, (2) reduced form IV approach, (3) IV approach without disclosing the instruments used, (4) unusual procedure claimed to be equivalent to IV by the authors, and (5) non-valid instruments as stated by the authors themselves. We find a total of 266 IV articles.

- **Difference-in-Differences**. The DD category includes articles that perform a difference-in-differences estimation as defined in Roberts and Whited (2013). Examples include Agrawal (2013) and Gormley and Matsa (2011). The DD classification further includes articles that make use of an exogenous shock as a natural experiment or quasi-experimental framework outside of a difference-in-differences framework. We find a total of 126 DD articles.

- **Selection Models**. The Selection category requires that an article implements a selection model following the traditional two-stage approach of the Roy model or the Heckman model as presented in Li and Prabhala (2007) and report the estimates. Examples include Campa and Kedia (2002) and Bris, Welch, and Zhu (2006). The Selection category further contains articles that either discuss untabulated results or use a non-standard technique to perform an estimation that corrects for self-selection biases. We find a total of 72 Selection articles.

- **Regression Discontinuity Design**. The RDD category requires that an article explicitly performs a RDD estimation as defined in Roberts and Whited (2013). Examples

Electronic copy available at: https://ssrn.com/abstract=2378460

include Chava and Roberts (2008) and Roberts and Sufi (2009). We find a total of 25 RDD articles.

- **Randomized Experiments**. The Randomized Experiment category requires that an article explicitly perform a field or laboratory experiment with a treatment and control condition. Examples include Cox, Smith, and Walker (1984) and Alevy, Haigh, and List (2007). We find a total of 5 Randomized Experiment articles.

Electronic copy available at: https://ssrn.com/abstract=2378460

Table A1: Classifying Identification Articles

We manually identify articles that adopt the identification technology in Section III A. This table presents the search terms that we use to flag articles for manual inspection. We explicitly provide plural terms and hyphenation during the search process. For brevity, we list here permutations within parentheses.

Technique	Key words used in search
IV	instrumental variable(s); 2sls; two(-)stage(s) least square(s); 3sls; three(-)stage(s) least square(s)

DD	difference(s)(-)in(-)difference(s); diff-in-diff;
	quasi(-)natural; natural experiment(s);
	exogen(e)ous(-)shock(s)

Selection	self(-)selection; switching(-)regression(s)
RDD	regression discontinuity; regression discontinuities

Experiments	random(ized) experiment(ation)(s);
	lab(oratory) experiment(ation)(s);
	field experiment(ation)(s);

Electronic copy available at: https://ssrn.com/abstract=2378460

Table A2: Identification Language Terms

In Section IV, we compare the adoption rate of the identification *language* in the top-three finance journals (JF, JFE, RFS) to that of the top-three economics journals (AER, JPE, QJE). We define an article as having adopted the identification language if at least one of the terms in the dictionary below or at least one of the terms in the dictionaries listed in Table A1 is found in the text. We explicitly provide plural terms and hyphenation during the search process. For brevity, we list here permutations within parentheses.

Key words used in search
causal affect(s); causal effect(s); causal impact(s)
exogen(e)ous variation
reverse causality; reverse causation
identification strategy; identification strategies

Electronic copy available at: https://ssrn.com/abstract=2378460