european corporate governance institute

ecgi

Evidence-Based Policymaking: Promise, Challenges and Opportunities for Accounting and Financial Markets Research

Law Working Paper N° 397/2018
April 2018

Christian Leuz
University of Chicago, NBER and ECGI

© Christian Leuz 2018. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

This paper can be downloaded without charge from:
https://doi.org/10.67929/ecgi-law-397-2018

european corporate governance institute

ECGI Working Paper Series in Law

Evidence-Based Policymaking: Promise, Challenges and Opportunities for Accounting and Financial Markets Research

Working Paper N° 397/2018
April 2018

Christian Leuz

I acknowledge helpful comments from the editors, Luzi Hail, Christian Laux and, especially, Maximilian Muhn. I also thank Maximilian Muhn for his excellent research assistance.

© Christian Leuz 2018. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Abstract

The use of evidence and economic analysis in policymaking is on the rise, and accounting standard setting and financial regulation are no exception. This article discusses the promise of evidence-based policymaking in accounting and financial markets as well as the challenges and opportunities for research supporting this endeavor. In principle, using sound theory and robust empirical evidence should lead to better policies and regulations. But despite its obvious appeal and substantial promise, evidence-based policymaking is easier demanded than done. It faces many challenges related to the difficulty of providing relevant causal evidence, lack of data, the reliability of published research, and the transmission of research findings. Overcoming these challenges requires substantial infrastructure investments for generating and disseminating relevant research. To illustrate this point, I draw parallels to the rise of evidence-based medicine. The article provides several concrete suggestions for the research process and the aggregation of research findings if scientific evidence is to inform policymaking. I discuss how policymakers can foster and support policy-relevant research, chiefly by providing and generating data. The article also points to potential pitfalls when research becomes increasingly policy-oriented.

Keywords: Evidence-based policymaking, cost-benefit analysis, regulation, standard setting, accounting, finance, capital markets, causal inferences, political economy

JEL Classifications: A11, D61, D72, G18, G38, K22, L51, M48

Christian Leuz
Joseph Sondheimer Professor of International Economics, Finance and Accounting
University of Chicago, Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637, United States
phone: +1 773 834 1996
e-mail: cleuz@chicagobooth.edu

Evidence-Based Policymaking: Promise, Challenges and Opportunities for Accounting and Financial Markets Research*

Christian Leuz
University of Chicago, Booth School of Business & NBER

April 2018

Abstract

The use of evidence and economic analysis in policymaking is on the rise, and accounting standard setting and financial regulation are no exception. This article discusses the promise of evidence-based policymaking in accounting and financial markets as well as the challenges and opportunities for research supporting this endeavor. In principle, using sound theory and robust empirical evidence should lead to better policies and regulations. But despite its obvious appeal and substantial promise, evidence-based policymaking is easier demanded than done. It faces many challenges related to the difficulty of providing relevant causal evidence, lack of data, the reliability of published research, and the transmission of research findings. Overcoming these challenges requires substantial infrastructure investments for generating and disseminating relevant research. To illustrate this point, I draw parallels to the rise of evidence-based medicine. The article provides several concrete suggestions for the research process and the aggregation of research findings if scientific evidence is to inform policymaking. I discuss how policymakers can foster and support policy-relevant research, chiefly by providing and generating data. The article also points to potential pitfalls when research becomes increasingly policy-oriented.

JEL classification: A11, D61, D72, G18, G38, K22, L51, M48

Key Words: Evidence-based policymaking, cost-benefit analysis, regulation, standard setting, accounting, finance, capital markets, causal inferences, political economy

* This paper is based on my 2017 PD Leake Lecture at the ICAEW. I acknowledge helpful comments from the editors, Luzi Hail, Christian Laux and, especially, Maximilian Muhn. I also thank Maximilian Muhn for his excellent research assistance. I have worked as an economic advisor to the PCAOB and still serve as a consultant from time to time. I have also co-written independent research reports on policy questions for the FASB and the SASB. The views expressed in this paper are my own.

1. The Promise and Push for Evidence-Based Policymaking and Economic Analysis

Evidence-based policymaking has become increasingly popular in many areas, as can be seen by the proliferation of think tanks and initiatives supporting and advocating for evidence-based policymaking.[1] I define evidence-based policymaking as a rigorous attempt to base policy decisions (e.g., new regulation) on scientific and empirical evidence, including impact studies, cost-benefit analyses, program evaluation and academic research in general.[2]

The appeal of evidence-based policymaking is fairly obvious. Using science and empirical evidence as input to policy decisions makes imminent sense. Policymaking that is rooted in sound theory and empirical evidence should lead to better policies and regulations. Academic research can provide important empirical facts and advance our understanding of policy effects, both ex ante and ex post. Empirical facts and analysis, especially when rooted in theory, are said to impose more discipline on policymaking, which in turn should make it more resilient to political pressures, lobbying and capture (see also Zingales 2015). Besides, research receives substantial amounts of public funding and so it should make a contribution to society.

Given these arguments, policymakers, regulators and standard setters are increasingly under pressure to embrace this approach to policymaking and to justify their policies with research and empirical evidence. Accounting standard setters, auditing regulators, and financial regulators such as the SEC are no exception. Accounting standard setters (FASB and IASB) have long recognized, in their mission statements and in their conceptual frameworks, a responsibility for cost-benefit considerations. However, as noted by Schipper (2010), the standard setters do not

[1] Examples are the Coalition for Evidence-Based Policy as well as initiatives by think tanks such as the Pew Charitable Trusts or the Heritage Foundation.

[2] Cost-benefit analysis is a frequently used term for the economic analysis of regulation. In my mind, cost-benefit analysis is a tool of evidence-based policymaking. The latter term more broadly captures the idea that science and empirical evidence are used rigorously and comprehensively to inform policy decisions.

use conventional or formal cost-benefit analysis (see also Buijink 2006). Recently, the FASB and IASB have started conducting post-implementation reviews.[3] Moreover, the IASB is moving towards an "evidence-informed approach" to standard setting (Teixeira 2014). There are similar developments for financial market regulators. For example, the SEC has faced substantial pressures to perform cost-benefit analysis for its rule-making.[4] Currently, independent U.S. agencies like the SEC are not explicitly required by law to perform cost-benefit analysis (Coates 2015). However, several Congressional initiatives are under way that would require formal economic analysis.[5] In the UK, the financial agencies are required to perform and publish cost-benefit analyses for proposed rules (e.g., Financial Services and Markets Act 2000). In addition, several academics are pushing for formal economic or cost-benefit analysis in financial regulation (e.g., Posner and Weyl 2014).

Despite its appeal and promise, evidence-based policymaking is easier said (or demanded) than done. Research faces numerous challenges in generating evidence that informs and supports policymaking. Addressing these challenges requires substantial investments into the research infrastructure, from data generation to aggregation and transmission of research findings. In this article, I discuss challenges, but also opportunities for research as well as potential changes to the research infrastructure, so that research could better and more systematically support the use of evidence in policymaking. Throughout the article, I focus on research in the areas of accounting standard setting and financial markets regulation (e.g., disclosure or securities regulation), although many discussions probably apply to economic research more broadly.

---

[3] The IASB introduced post-implementation review (PIR) for major standards in 2007 and issued its first PIR in 2013 (conducted for IFRS 8). The FASB completed its first PIR in 2012 (pertaining to FIN 48). See also Ewert and Wagenhofer (2012) for a discussion of the PIR process for accounting standards.

[4] See, e.g., the 2011 *Business Roundtable v. SEC* decision, in which the court struck down a SEC voting rule for insufficient cost-benefit analysis.

[5] Examples are the SEC Regulatory Accountability Act, the Independent Agency Regulatory Analysis Act and the Regulatory Accountability Act.

Given this focus, I begin by asking to what extent accounting and financial markets research delivers insights and evidence that is useful to policymaking. Financial accounting research is directly related to policy and, in particular, accounting standard setting. More broadly, financial markets research should be able to contribute to financial market regulation, among other things, by studying relevant economic links and relations, such as the effect of information disclosure on market liquidity. Moreover, accounting research could inform debates about disclosure and transparency policies in many other areas outside the core accounting and financial reporting domain, such as environmental regulation, consumer protection, health care, and others. In addition, accountants perform many roles related to measurement, disclosure and compliance outside of financial reporting. While there is little doubt in my mind that accounting and financial markets research has important insights to offer, at present, we have little formal evidence on these contributions. It would be good to ask more systematically what we have learned from past research that is helpful to regulators and policymakers. Doing so would also point to opportunities for future research.

Next, I turn to the challenges that evidence-based policymaking poses for research. One of the key challenges is the ability to draw causal inferences. Causality plays a central role because without a causal relation, it is difficult to provide reliable advice to policymakers (or to the general public). Yet, accounting and financial market research, like many other social sciences, faces major limitations in its ability to generate causal evidence. Another and related challenge is the measurement of the regulatory "treatment." Many studies are based on regulatory or accounting changes, which alter the amount of public information, but do so in "unspecified" ways. As a result, accounting research rarely provides treatment effects in the form of elasticities, i.e., estimates that tell regulators by how much a market outcome improves if a

disclosure mandate increases information quality by X percent. Part of the problem is that we lack a standardized measure for the amount and quality of financial information. Making progress on this measurement problem is of first-order importance. In terms of generating such estimates, structural modeling and estimation could play an important role.[6] In addition, it could provide counterfactuals and what-if analyses. But I hasten to add that structural estimation is not a panacea and also has substantial limitations.

The biggest challenge for causal evidence and better policy-relevant estimates, however, is lack of relevant data that is sufficiently granular to identify and measure regulatory effects. This shortcoming is difficult to overcome without the help of policymakers and regulators. The same can be said for the lack of exogenous variation in regulatory changes. In addition, there are challenges to the reliability of research findings. For instance, discretion and incentives in the research and publication process likely play an important role in the relatively low reproducibility rates, which recently have been documented for several fields.[7] While I am not aware of such evidence on reproducibility rates of accounting and financial markets research, I am convinced that, in these areas, similar issues with respect to the reliability of research findings exist. Finally, there are a number of challenges related to the aggregation and transmission of research findings as well as the use of results by policymakers, especially when recognizing the political nature of policymaking.

Recognizing all these challenges, there are a number of steps that we could take if research is to inform and support policymaking. Specifically, I discuss ideas on how to organize and facilitate policy-relevant research, to increase the reliability of our research findings, and to

[6] Structural modeling and estimation refers to a technique for estimating "deep" structural parameters (which are typically policy invariant) using economic models. It is different from reduced-form estimation, which relies on (direct) statistical relations between observable variables.

[7] See, e.g., the Reproducibility Project in psychological science (Open Science Collaboration 2015).

systematically summarize, aggregate and communicate evidence. Importantly, if we are serious about evidence-based policymaking in accounting and financial regulation, then regulators and standard setters need to actively help with generating relevant data and fostering research, essentially building economic analysis into the process of rulemaking. I also conjecture that academic research needs to devote much more effort towards understanding the regulatory "plumbing," i.e., the details of regulation, its implementation and the interplay of rules, as it matters a great deal for regulatory outcomes and is often where things go wrong.

My main message is that evidence-based policymaking requires building an entire research infrastructure. We currently do not have such an infrastructure for accounting and financial markets research. To illustrate this point, I draw parallels to the creation and rise of evidence-based medicine, which highlights the necessary investments. We are still a long way from evidence-based standard setting or financial regulation. In my view, it is nevertheless worth starting a process towards a more systematic use of evidence in policymaking. Poorly designed or implemented policies can have major costs to market participants and society.[8]

I realize that there is a larger debate about whether evidence-based policymaking is even feasible (e.g., Hammersley 2013). The "evidence-based" notion is much more contentious in public policy than it is in medicine or safety regulation. At the heart of the criticism is that social science is quite different from medicine or more technical areas, such as automobile safety. Policymaking is inherently more political than medical practice guidelines (although I do not claim that they are apolitical). Judgment, political values and ideology play a much larger role in setting public policy. In addition, accounting standards and financial regulation tend to be quite

[8] See, e.g., survey results from the IGM Economic Experts Panels (2017) indicating that flawed financial sector regulation and supervision was the most important factor in the 2008 global financial crisis, which clearly was a very costly event.

different from rules in automobile safety, an area that has used cost-benefit analysis successfully. Among other things, the former offer much more discretion than the latter (see also Cochrane 2014). The general conclusion from this debate is that evidence-based policymaking in public policy or in financial regulation is likely harder and more fraught with problems than evidence-based medicine.[9] I agree and view the example of evidence-based medicine as more aspirational.

In this article, I sidestep this higher-level and more philosophical debate and instead focus on more specific challenges that arise if we take the desire for evidence-based policymaking in accounting and financial markets seriously. Nevertheless, it is important to have realistic expectations. Building the necessary research infrastructure takes time and even a more modest evidence-*informed* approach to policymaking requires a concerted effort by researchers and policymakers. But without these investments and efforts, and a commitment by policymakers to use scientific evidence, we will only pay lip service to the idea of evidence-based policy making.[10]

2. Contributions of Accounting and Financial Markets Research to Policymaking

In this section, I discuss the extent to which accounting and financial markets research delivers insights and evidence that is useful to policymaking. There is no question in my mind that accounting and financial markets research can contribute valuable insights to standard setters and policymakers. Financial accounting is intimately connected to the idea of setting standards for corporate disclosure and reporting. Normative considerations have a long tradition in early accounting research. Many accounting studies analyze the effects of accounting standards and disclosure requirements. Financial markets research often examines the effects of regulatory

[9] This view is not universally shared. See, e.g., Posner and Weyl (2015). See also my discussion in Section 6.

[10] A lack of such investments (and commitments) could also explain why Hahn and Tetlock (2008) conclude that the track record for economic analysis of regulatory decisions to date is sobering.

changes (such as the Securities and Exchange Acts or the Sarbanes-Oxley Act). Thus, accounting and financial markets research is often closely connected to policy debates and regulatory issues.

But while it is clear that this research has the potential to generate relevant insights to policymakers and regulators, it is less clear how this research has influenced policymaking. As far as I know, we do not have a systematic account of the contributions to policymaking or the impact of this research. There are examples where the work of economists has influenced policy in a major way, such as Friedman's analysis of an all-volunteer military (Singleton 2016). We also have subjective assessments. For instance, Heckman (2001) states in his Nobel lecture that "microeconometrics has contributed substantially to the evaluation of public policy." Zingales (2015) in turn has a more pessimistic view of finance's contributions to public policy in his presidential address. Hellwig (2015) is also more critical of the role of economic research and policy recommendations of economists.

In addition, there are impact studies but they tend to focus on citations and hence research impact in academe. Going beyond citations, Burton et al. (2017), analyze research impact using the Altmetric's Attention Score, which captures attention to published research from online sources such as news sites, policy documents, social media, etc. Such attention does not necessarily imply impact, but the metric captures broader interest in research. Burton et al. (2017) find that research in social science disciplines such as economics, finance, management, and psychology has far less impact (or interest) than research in the natural sciences.[11] The impact of accounting is the lowest among the business school disciplines; finance is the second lowest (but it still receives much more attention than accounting). One potential explanation is

[11] Bastow et al. (2014) study the impact of social science more broadly and come to similar conclusions. They discuss a number of structural reasons why the impact of social science is low compared to natural science, especially its external impact. They also note that there is little formal evidence on external impact.

that accounting and finance are smaller fields and much more specialized compared to management, economics or psychology.[12] Consistent with this view, accounting and finance have much lower attention scores than management, marketing and psychology in the Altmetric categories for news, blogs, and social media, but receive higher scores for policy documents. The latter underscores, at least relatively speaking, the policy relevance of accounting and finance research.[13] Bastow et al. (2014) make a similar observation with respect to the external impact of social science in general, noting that its links with policymakers and governments are far stronger and more developed than its other external links, consistent with its relevance to public policy (see also Abreu et al. 2009).

Even though we do not have much systematic evidence on the contributions to policymaking, it is still worth asking what we have learned from accounting and financial markets research that is helpful to the economic analysis of regulation and public policy. What are the insights that regulators and policymakers should know or could use? Have we established relations and facts that are both relevant and sufficiently reliable? These are important questions, and the answers are not forgone conclusions.[14]

One place to look for answers is in academic literature surveys that specifically focus on policy questions, standard setting or financial market regulation. Specific examples in accounting standard setting and financial regulation are Hail et al. (2010a, 2010b), Koch et al. (2013), Coates and Srinivasan (2014), Acharya and Ryan (2016), and Leuz and Wysocki (2016). These

[12] Other potential explanations are the quality of research, the relevance to practice or the relative appeal of the field. Despite the stereotypes, I do not believe that accounting is inherently less interesting. But the field would benefit from more innovation, fresh ideas and, in particular, from broader research topics. See also Burton et al. (2017) and the critiques by Demski (2007), Fellingham (2007), Hopwood (2007), Waymire (2012), and Wood (2016). I discuss ideas for broadening policy-relevant accounting research in Section 3.

[13] See also Bauguess et al. (2017), FMA presentation on "Role of Academic Research in Financial Market Policy" and Geoffrey and Lee (2018) for evidence on the role of research in academic research for SEC rulemaking.

[14] A related and much broader debate questions the ability of social science to provide cumulative insights in the same way as natural science does (e.g., Flyvbjerg 2001). I am sidestepping this debate here. See also footnote 11.

surveys provide an overview and critique of the literature. They highlight that the amount of relevant research is substantial. But in my read, it would be difficult to infer from these surveys whether an academic consensus on certain policy or regulatory issues has emerged.

An effort that is specifically geared towards eliciting whether such a consensus exists for certain public policy questions is the IGM Economic Experts Panel.[15] The panel comprises over 40 distinguished economists from the top-seven economics departments in the U.S. It covers scholars from the major areas of economics, different age cohorts, and from different political persuasions.[16] The panel members receive a policy question every two weeks, with which they can agree or disagree; they can also vote that the answer is uncertain. Members indicate the level of confidence they have in their expertise in the specific matter or question. Gordon and Dahl (2013) use responses from this panel to explore the extent to which economists agree or disagree on major public policy issues. The underlying question is whether the views of economists are based on accumulated academic evidence or, alternatively, reflect different camps and political leanings. Interestingly, Gordon and Dahl (2013) conclude that there is "close to full consensus" among panel members when the underlying economic literature is large. There is little evidence of different camps; political leanings seem to be of little importance. Thus, their main finding supports a broad consensus among top U.S. economists for many public policy questions.[17] This evidence suggests that research in economics is cumulative, leading to an academic consensus for certain policy issues. Hence, there should be a body of research that is useful to the policymakers. While we lack such surveys and evidence for accounting and finance, I presume that a similar consensus for many broad policy issues exists or at least could emerge as the

[15] Recently, the IGM added a second expert panel with European economists. In the interest of full disclosure, I am a co-director of the IGM and a member of the European IGM Economic Experts Panel (http://www.igmchicago.org/european-economic-experts-panel).
[16] For details on the construction of the panel, see www.igmchicago.org/igm-economic-experts-panel.
[17] For qualitatively similar conclusions see Alston et al. (1992) and Fuller and Geide-Stevenson (2014).

literature accumulates.[18]

Reflecting on the literature on standard setting and financial markets regulation, I would say there are policy-relevant economic links and also conceptual insights for which a consensus has likely emerged. An example for a link that immediately comes to mind is the positive relation between corporate disclosure and market liquidity. This link is well supported by economic theory and numerous studies with very different research designs (see surveys by Verrecchia 2001, Leuz and Wysocki 2016). An example for a conceptual insight from the accounting literature is the tradeoff between relevance and reliability of financial information, which plays a central role for standard setting and for which we have theory and evidence (e.g., Watts and Zimmerman 1986, Dechow 1994, Kanodia et al. 2004, Bratten et al. 2013). Another example is the notion of reporting incentives shaping reporting practices. The idea is that accounting standards, for good reason, give substantial discretion to firms. This discretion implies that managerial reporting incentives, which are shaped by markets, contracts and institutions, heavily influence reporting practices, leading to substantial heterogeneity even when operating under the same standards. This insight is of fundamental importance in standard setting and there is a good amount of evidence supporting it (e.g., Watts and Zimmerman 1986, Ball et al. 2003, Leuz et al. 2003, Burgstahler et al. 2006, Cascino and Gassen 2015). There are many other examples; I simply listed a few that are close to my own work and to the topic of this article.

However, the aforementioned insights are general economic links and tradeoffs. While they are surely relevant to policymakers, they still need to be interpreted in a specific context and do not provide (quantitative) answers to specific policy questions. In fairness, we also have many studies examining the economic consequences of particular changes in the accounting standards

[18] See, for example, the IGM panel question on stock prices and market efficiency, which supports this conjecture.

and in financial market regulation (see survey by Leuz and Wysocki 2016). This work is often closer in spirit of program evaluation, providing specific estimates for effects of certain regulatory changes. But as I discuss in more detail in Section 4, this research also faces many challenges and the general insights for new policy decisions can be quite limited.

The bottom line from this discussion is that accounting and financial markets research has something to offer to policymakers and regulators, but that we need a much more systematic account of the literature and its impact as well as better ways to aggregate the policy implications of extant research findings. I come back to this point in Sections 4.4 and 5.2.

3. A Brief Digression: Opportunities for Future Accounting Research

Given my discussion of the contributions of extant accounting research to policymaking in the previous section, it is clear that many opportunities for future research remain. We still need much more research before we can move to evidence-based standard setting or financial market regulation. Before I turn to the challenges of such research and offer some thoughts on what it would take to move towards evidence-based policymaking, I briefly digress to discuss a few specific opportunities for accounting research. The discussion focuses on opportunities outside the traditional or core domain of accounting with the hope that by broadening accounting research, we increase its external impact, be it on practice, other social sciences or society more generally.

There are many important issues that society and businesses face that are connected to accounting, including concepts like transparency, accountability, trust, verifiability, governance, communication, goalsetting, budgeting, measurement, control, security, accuracy, taxation, sustainability, and corporate social responsibility (see also Burton et al. 2017). In my mind,

accounting research should be concerned with and should have something to say about these and many other important concepts in business and society; after all accountants are often involved in the underlying processes. Of course, these concepts are not only relevant to accounting. But by considering them as part of the accounting domain and by studying them from an accounting angle, we increase our links with and contributions to other social sciences.

A specific arena in which such broadening of the accounting domain would be both natural and fruitful are disclosure and transparency mandates. Such mandates are increasingly used as a public policy instrument in lieu of more conventional regulation in many areas outside accounting's core domain, including corporate governance, environmental, food safety, restaurant hygiene, consumer protection, mine safety, health care, and conflicts of interest disclosures to name but a few. The idea is to compel disclosures, rather than to restrict or mandate certain behaviors or business practices, with the expectation that transparency incentivizes desirable (or discourages undesirable) behavior. Thus, the ultimate regulatory motivation goes beyond informing consumer, investors, or other recipients and aims to induce real effects, i.e., behavioral changes by the sender and/or those involved in producing and disclosing the information (e.g., auditors). Given the prevalence of many transparency and disclosure mandates, understanding their real effects is of first-order importance to policymaking and society.

Nascent research in this area shows that the effects of disclosure mandates are not always positive and can be quite pernicious (see the overviews in Fung et al. 2007, Dranove and Jin 2010, Leuz and Wysocki 2016). But the key point here is not to argue in favour or against transparency regimes but that accounting research could make important contributions to our understanding of these regimes, and that these contributions would be germane to accounting.

Let me illustrate this point with a disclosure study in health care. Dranove et al. (2003) study mandated cardiac surgery report cards in New York and Pennsylvania. They find evidence of improved matching of patients with hospitals, which presumably is an intended and expected outcome. But they also document selection behavior by the providers, in essence, suggesting that doctors, worried about bad report cards, start screening patients, which in turn makes it harder for sicker patients to obtain treatment. On net, the disclosure regime led to worse health outcomes, particularly for sicker patients, which is obviously very problematic.

However, this evidence does not necessarily imply that using disclosure regimes in health care is a bad idea. Instead it points to a measurement and communication problem. The selection effect likely occurs because doctors felt that the users of the report cards (e.g., patients) would not be able to appropriately adjust for patient characteristics (e.g., how sick the patient was) and hence might infer low-quality treatment when the outcomes for the most serious cases are less favourable. In my view, accounting research should work on such measurement and communication problems helping us better understand how to reap the benefits and avoid the pitfalls of disclosure regimes.

I will give two more examples for opportunities to broaden accounting research. The first example is the link in banking between financial disclosure, accounting measurement and financial stability. Does more disclosure increase or hurt financial stability (e.g., by enhancing market discipline or encouraging runs, respectively)? Does the expected loss model for loan accounting lead to more or less procyclicality in bank lending? These are first-order policy questions (see also Acharya and Ryan 2016). Again, there is nascent research on these topics (e.g., Bushman and Williams 2015, Domikowsky et al. 2017, Granja 2018), but currently we are not able to answer these important policy questions.

Second, we still need more work on externalities and market-wide effects of disclosure regulation. Such effects are central to the economic justification of disclosure mandates. Accounting research has studied such effects in capital markets (e.g., Bushee and Leuz 2005, Badertscher et al. 2013), although the number of studies is low and the evidence still needs to be corroborated and extended. We also need more work on the effects of disclosure regulation on competition, firm productivity, and the allocation of resources (e.g., labor and capital), which has just begun (e.g., Breuer 2017, Choi 2018).

In sum, there are encouraging signs that accounting research is branching out, becoming broader, and building bridges to other fields. I value this trend as it would increase the external impact of accounting research and, at the same time, make it more relevant and helpful to policymakers and regulators.

4. Challenges for Research Supporting Evidenced-Based Policymaking

Having discussed the promise of evidence-based policymaking and opportunities for new research, I now turn to a number of specific challenges for research if it is to inform and support evidence-based policymaking. I discuss four major challenges using the context of accounting and financial markets research; similar challenges arise in many other areas of economic research.

First, evidence-based policymaking requires a discussion about the "quality standard" for research evidence that is to be used to inform policymakers. For this, causal inferences play a critical role, yet they are very hard to obtain. The second challenge is the measurement of the treatment itself, e.g., the changes induced by past policy changes. Making progress on this challenge is central to estimating policy effects in a form that is more useful to policymakers,

e.g., when evaluating potential future policy changes. Third, I point out that the biggest challenge to providing more policy-relevant research is lack of data. This issue is closely connected to the scarcity of causal evidence. Lastly, I discuss challenges related to the reliability of research findings and their transmission to policymakers, essentially recognizing incentive problems and biases in conducting and publishing research as well as the political nature of policymaking.

4.1 Importance of Causal Inferences and the Tradeoff between Internal and External Validity

Causality is clearly of central importance for evidence-based policymaking. First, policymakers care about the magnitudes of potential effects, e.g., when conducting a cost-benefit analysis. But estimated magnitudes really only matter once we have causal inferences. Estimates based on mere associations contain or reflect other factors and hence have to be used very cautiously in a cost-benefit analysis. Second, without a causal relation and an understanding of the mechanism, it is difficult (and in some cases even unethical) to provide policy recommendations.

I illustrate this point with an example outside of accounting and financial markets, namely, the question of whether drinking alcohol in moderate amounts confers health benefits. Being a fan of red wine, I like to think that having a glass at night is not only enjoyable but also good for my health. Obviously, a policymaker or doctor would want to be sure that there is sufficient evidence that drinking moderate doses of alcohol has (net) health benefits before providing a recommendation to the public or a patient. This question has been heavily studied and some even argue that the evidence is sufficiently compelling to advise abstainers to drink. Stockwell et al. (2016) identify over 2,600 studies that are potentially relevant for the question. Many studies as well as several meta-analyses suggest a J-shaped relation between alcohol consumption and

mortality risk, indicating reduced risk for occasional and low-volume drinkers but higher risk for higher volume drinkers (e.g., Brien et al. 2011, Ronksley et al. 2011).

However, the existing evidence should not be taken without concern. There are large life-style and other differences between abstainers, moderate drinkers and those that drink more. These differences could confound the findings given the underlying studies are not randomized-control trials. In particular, there are selection concerns about the abstainer group as it often contains former drinkers, including people who ceased drinking alcohol for health reasons. Stockwell et al. (2016) report that 65 out of 87 studies included in their formal meta-analysis had former drinkers in the abstainer group, potentially creating substantial bias. Once Stockwell et al. (2016) adjust for abstainer bias and various study characteristics, the relation between alcohol consumption and mortality looks essentially linear, with mortality risk increasing as alcohol consumption increases.

While this evidence sadly implies that my occasional glass of wine does not provide health benefits, it nevertheless holds two important lessons. The first lesson is that having many studies with similar results may not be sufficient proof to support a particular policy or conclusion. Studies often share similar identification challenges or selection problems. If so, simply "piling up" studies does not help or address the fundamental challenges for causal inferences. Aggregating studies is more useful when the individual studies have fairly orthogonal research-design challenges or when all studies provide causal estimates but the magnitudes depend on the context or exhibit measurement error.

The second lesson is that selection problems are pervasive, especially in social science, and in many contexts we cannot address the ensuing inference problems by performing randomized-control trials (RCTs). Such trials are considered the gold standard in medical research. But in my

particular example a long-run RCT would be infeasible (if not unethical). The same is likely true for many other policy issues. For instance, we could (or would) not randomize monetary policy even though understanding the effects of interest rate changes is of great interest to central bankers. While we could and should conduct field experiments (or randomized pilot studies) much more often than we currently do,[19] for many questions in accounting standard setting and financial market regulation, RCTs or field experiments are simply not feasible. Moreover, field experiments are probably less well suited to study long-run and general-equilibrium effects, which are nevertheless very important. Similarly, it is difficult to study spillover effects and externalities based on small-scale experiments. Therefore, I do not think RCTs will play the same central role in accounting and financial regulation that they play in medicine and have played for the rise of evidence-based medicine (see Section 6).[20]

There are of course also situations, in which nature provides random assignment to treatment (e.g., due to weather) or settings, in which lotteries were used for fairness reasons (e.g., drafts, program or school admissions). However, these natural experiments are rare and there are not enough of these situations to answer all the policy questions we would like to answer. Thus, evidence-based policymaking in accounting and financial regulation by and large has to rely on empirical studies using archival data that is generated without explicit or naturally occurring randomization. Of course, we can draw causal inferences from such data, at least under certain conditions or when making certain assumptions, e.g., using regression-discontinuity designs or difference-in-differences analyses. There is a huge literature on the identification of treatment effects (for overviews, see Heckman 2001, Angrist and Pischke 2009, 2014).

[19] See also Floyd and List (2016) and Section 5.3 in this article.

[20] Chemla and Hennessy (2018a) show that subjects' rational expectations about future health benefits from participating in a trial could contaminate RCTs even in medicine. Expectations about the future are even more important (and hence a bigger issue) in policy experiments. See also Chemla and Hennessy (2018b).

In accounting and financial regulation, treatment effect studies often exploit regulatory changes or new mandates. Such settings have the advantage that a certain disclosure or a particular accounting treatment is imposed on firms, which in turn mitigates selection concerns that typically arise in voluntary disclosure or accounting choice settings. Nevertheless, these regulatory settings pose several major identification challenges when estimating causal effects. Leuz and Wysocki (2016, Section 2) discuss a number of these challenges in more detail, but let me mention a few.

First, even though regulation is imposed on firms, selection problems arise when firms can opt out or have ways to avoid the regulation (Heckman 1997). Second, new regulation or changes in regulation do not occur in a vacuum. They are often a response to financial or political crises or other major events (e.g., a corporate scandal). Financial markets also respond to these events, making it difficult to isolate the regulatory effects (Ball 1980). Third, and related to the previous challenge, regulatory changes tend to apply to a larger group of firms from a (single) point in time onwards. As a result, the empirical analysis is susceptible to other institutional changes, general time trends as well as market-wide shocks that are concurrent with but unrelated to the regulatory change. Fourth, firms and investors often anticipate regulatory changes, even before the first firms adopt the new rules. Furthermore, a regulatory change can signal future regulatory actions, for instance, a tougher stance when it comes to enforcement.

As a result of all these challenges, causal evidence from empirical studies on standard stetting and financial market regulation is still rare (Leuz and Wysocki 2016). Moreover, these challenges highlight that regulatory studies require careful research design and deep institutional knowledge, including of the process by which the regulatory change came about. Such institutional knowledge is particularly important when articulating why a particular setting

allows us to identify and estimate the economic effects and when discussing what the potential threats to identification are. It is often the specific features of the institutional setting that afford us identification or allow us to rule out alternative explanations.

However, it is precisely this reliance on the specifics of the institutional settings that brings us back to the well-known tradeoff between internal and external validity. Put differently, there is a price that we pay for identification. Many studies on standard setting and financial market regulation that provide causal estimates do so in very specific settings and as a result, their estimates (or the magnitudes of their estimates) have limited generalizability (see also Leuz and Wysocki 2016, Glaeser and Guay 2017). This limitation also arises with field experiments.

To illustrate this point, I refer to two studies that I admire very much. First, Duflo et al. (2013) perform a field experiment in environmental auditing to study the conflicts of interest that arise when firms choose and pay their auditor. The study captures many of the key economic tradeoffs in auditing; it is in the field and uses randomization. Based on this setup, the study shows that conflicts of interests related to auditor choice and payment firm can corrupt audit outcomes. While this and other insights from the study are of general importance, we would not use the results from this experiment on auditing plant emissions in India, where the auditors were paid less than a $1,000 per audit, and apply them directly to policy issues in corporate auditing of financial reports in the U.S. The markets, legal institutions, audit processes and also fee arrangements and magnitudes are too different.

My second example is the study by Iliev (2010) examining the effects of the Sarbanes-Oxley Act (SOX) on firm value, audit fees and corporate reporting. The study is a good example for a setting in which a size threshold allows for a regression-discontinuity design, which can give us quasi-random assignment for firms that are close to the $75 million public-float threshold. The

study provides convincing evidence that, among other things, SOX caused an increase in audit fees (by $700,000). But the estimated increase is very local in that it applies to firms just around the threshold, and it cannot simply be extrapolated to infra-marginal firms that are much larger.

In sum, the message of this section is that causal estimates are difficult to obtain and when we have causal inferences, they often come with limited generalizability, posing significant challenges to evidence-based policymaking.

4.2 The Measurement Challenge and Structural Modeling

Another challenge for the use of accounting and financial markets research in policymaking is the measurement of the treatment. In medicine, the measurement of the treatment (or dosage) is often straightforward (e.g., 100 milligrams of a particular active ingredient). Knowing the precise treatment is obviously very important when estimating treatment effects and also when comparing effects across studies or conducting meta-analyses. Thus, in medical studies, one can compute the effect on mortality as well as the side effects per dose of treatment (e.g., 100 milligrams). The equivalent in economics is expressing treatment effects in the form of elasticities. For instance, it would be very helpful to securities regulators to have estimates along the following lines: increasing the amount of public information by 1% increases market liquidity by Y% and decreases the cost of capital by Z%.[21] However, such estimates are rarely available in public policy. Economic elasticities are very hard to obtain (e.g., Cochrane 2014). Part of issue is again identification. But the other issue is measurement of the treatment. We generally do not know by how much a regulatory change increased information, i.e., the dosage.

[21] Another issue is that arises in prospective analyses is that regulators do not know by how much an intended regulation would increase public information in markets. But if we could compute the information effects of past regulations, we could develop estimates for future regulation.

Let me explain this issue in the context of standard setting or financial regulation. In this area, we have many studies that exploit regulatory acts and accounting changes to estimate treatment effects. But even when these studies provide causal estimates, they do not deliver elasticities along the lines suggested above. For instance, we do not know by how much the EU's Transparency Directive, the SEC's rule mandating the Compensation Discussion & Analysis, or IFRS adoption changed the amount of information available to investors. Put differently, we do not know how strong the regulatory treatment was. Often, the focus is on first determining whether there was a treatment as a result of the regulatory change (e.g., Christensen et al. 2013 for IFRS). Given this issue, it is difficult to compare across treatment-effect studies and to conduct meta-analyses.

Evidence-based policymaking requires more than (causal) evidence on for directional relations. Ideally, we would have quantitative estimates for the effect in the form of elasticities, which regulators can use in cost-benefit analyses or prospective analyses. Towards this end, we need substantial progress in measuring the amount and quality of information in financial reports and disclosure documents, which would then enable us to measure or quantify the treatment.

There are a few approaches that start taking us in this direction. What makes these approaches different is not that they provide quantitative estimates but that they start from theory and explicitly show how (or under which conditions) we can *identify* the relevant constructs from observables.[22] An example is Nikolaev (2017). His approach uses accounting relations between accruals and cash flows as well as the fact that accruals reverse to identify the quality of the accounting information (or system) in a GMM estimation framework. The key construct is a

[22] These new approaches are different from conventional earnings quality measures (see survey by Dechow et al. 2010), which cannot separate aspects of the business process (e.g., economic volatility) from the quality of the reported information. See also discussion in Nikolaev (2017).

quantitative estimate for information quality at the firm or industry level (see also Choi 2018 for embedding this measure in a macroeconomic comparison of accounting systems). Another example is Smith (2018). He models the link between corporate disclosure and option prices. Based on this model, he can extract the precision of reported earnings using market reactions at the earnings announcements and option prices. His approach would even enable firm-specific, point-in-time estimates. These two examples are very promising. We need to push them further, so that we can deliver the kind of estimates evidence-based policymaking requires.

In this regard, structural estimation (as defined in footnote 6) holds much promise to move the literature on standard setting and financial regulation forward. I say this for two reasons. First, the accounting system naturally provides structure that can be exploited for identification (Nikolaev 2017). The same holds true for financial regulation or taxation (e.g., McClure 2018). The structure makes it clear what it takes to estimate the parameters of interest and what data are required.[23] As Heckman (2001, Table 4) points out, structural estimation is particularly suited for policy analyses. It provides, as he puts it, ingredients for extrapolation to new environments. The underlying idea is that (policy-invariant) structural parameters have better "transportability" to other settings. Moreover, structural modeling enables us to compute counterfactuals (i.e., perform what-if analysis), which is particularly useful for prospective policy assessments.

While the structural approach holds promise for evidence-based policymaking, it is not a panacea. Like any other approach, it is not without its own problems and limitations (see also discussion in Angrist and Pischke 2010, Nevo and Whinston 2010, Gow et al. 2016). Structural modeling should therefore be one of many approaches that we pursue.

[23] Even though the model structure provides another source of identification, identification should not solely rely on the model. As Nevo and Whinston (2010, p. 71) put it, "structural analysis is not a substitute for credible inference." It is also important to ask what additional information or data a structural approach could bring to bear that could not be used in reduced form.

4.3 The Biggest Problem: Lack of Data

The challenges discussed in the previous two sections are major obstacles for research used in evidence-based policymaking. But in my mind the biggest problem is lack of data. We generally do not have the relevant and sufficiently granular data and we also lack exogenously generated data. Surely, accounting and finance researchers could further improve their research designs and embrace new econometric methods. But insufficient data are at the heart of the aforementioned challenges to causal inference and treatment measurement.

For instance, Posner and Weyl (2015) argue strongly in favor of cost-benefit analysis for financial regulation. They concede: "If the data do not exist, or are noisy, or if no plausible identification strategy has been developed, then regulators will not be able to determine valuations with any confidence. This creates a dilemma." They also note that lack of data is not an argument against performing economic analysis of regulation. I concur. But I believe that the data situation they describe is the norm, and we should acknowledge this challenge when considering evidence-based policymaking (see also Coates 2015, Leuz and Wysocki 2016).

At some level, it is well known that identification and endogeneity are fundamentally data problems. For instance, econometrics treats the selection problem as a missing data problem. My point here is more specific to evidence-based policymaking. We need to recognize that, despite the recent explosion in the availability of data, we are missing the relevant data for many regulatory and policy issues that we would like to answer. In many cases, the relevant data do not yet exist. In other cases, the relevant data are proprietary or not observable to researchers. As a result, studies have to rely on relatively crude proxies. A good example is audit research, as much of the audit process is unobservable to researchers and hence audit studies rely heavily on (discretionary) accruals or restatements as proxies for audit quality. However, it is well known

that both proxies also reflect reporting choices by the firm that is being audited. Thus, we need measures that more specifically reflect audit quality, given a client's reporting choices. Similarly, our studies generally use highly aggregated numbers. For most firms, accounting numbers from consolidated financial statements reflect hundreds, if not, thousands or millions of transactions, possibly from many subsidiaries. Thus, when we study changes in the accounting numbers before and after a new accounting rule came into effect, we do not observe the change in the numbers solely due to the new standard, which in turn creates the challenge of separating other changes in the economics of the firm from the changes induced by the new standard.

To make matters worse, new regulation or accounting standards generally come into effect roughly at the same time for all firms in a particular market or country, creating concerns about concurrent events (e.g., Christensen et al. 2013). In addition, standards are not implemented at random times, which means that the data are generated in ways that create many of the identification problems that I have already discussed in Section 4.1. Thus, if we are serious about evidence-based policymaking, addressing the data problem is a key place to start. I come back to this issue in Section 5.3, where I provide several suggestions on how to generate new data and to mitigate the problems discussed here.

4.4 Reliability of Research Findings and Political Influences on the Research Process

In this section, I draw attention to the reliability of research findings. As researchers, we obviously care about the reliability of our findings, irrespective of their use in policymaking. But when research findings are used to inform policymakers or to support policies, the reliability of the findings is an important dimension to consider explicitly.

In recent years, many concerns about the reliability of scientific publications have been

raised (Begley 2013, Begley and Ioannidis 2015). The Reproducibility Project in Psychology (Open Science Collaboration 2015) shows that the rates are surprisingly low: Only 39% of the studies were judged to have replicated the original results. Moreover, the replicated effects were half the magnitude of original effects (but see also Gilbert et al. 2016). A related study (but with a much smaller sample) in experimental economics indicate reproducibility for 60% of the studies, well below what would be implied by reported p-values, and again they find smaller effect sizes (Camerer et al. 2016).[24] Brodeur et al. (2016) conduct an analysis of 50,000 p-values reported in three widely cited general economics journals between 2005 and 2011. They find that (borderline insignificant) p-values between 0.10 and 0.25 are less common than one would expect. To my knowledge, we do not yet have a similar reproducibility project in accounting, but the *Critical Review of Finance* has created a Replication Network. One could argue that the heavy reliance of accounting and finance research on databases like CRSP and Compustat should increase the reproducibility, and it might. However, a recent study by Basu and Park (2016) shows that accounting research exhibits similar p-value patterns consistent with selective reporting of statistically significant results.

There is a growing recognition that discretion in empirical analyses is at the heart of the matter. With discretion, researchers might engage in "p-hacking" or "fishing" by estimating many (reasonable) specifications and selecting those that deliver significant results (e.g., Simmons et al. 2011). However, as Gelman and Loken (2013) point out, researchers' degrees of freedom can also lead to multiple comparisons, even when researchers do not actively engage in p-hacking or fishing. Their "garden of forking path" argument is that it is sufficient that researchers make design choices and perform analyses contingent on the data and the results for

[24] McCullough et al. (2006) report that only 22% of 62 macroeconomic studies could be successfully replicated, despite data and code archives. See also Chang and Li (2018) for similar conclusions. Both studies also show that the rates are higher for journals with data and code policies.

inferences to be biased. The issue could arise subconsciously. To be clear, discretion in research can be very useful, e.g., allowing researchers to explore and better understand the data. But it is important to recognize that discretion is a double-edged sword (Gelman and Loken 2013, Bloomfield et al. 2018). One way to counter the negative effects of discretion and to boost the credibility of the findings is to pre-register the analysis.[25]

Aside from discretion, biases in research findings can come from incentives in the publication process, e.g., a tendency by editors to favor surprising results. Similarly, it is generally hard to publish null results.[26] Again, pre-registration of studies can help mitigate these biases. Other useful ideas are code and data sharing policies (e.g., Höffler 2017; see also polices at the Journal of Accounting Research or the Journal of Finance) as well as creating platforms for post-publication review, allowing other researchers to comment on published studies, engage in online discussions and publish more formal reviews (e.g., Swoger 2014). We also need much more replication of published studies in accounting and finance to gauge and ensure the reliability of our findings.[27] Thus, if we want evidence-based policymaking for standard setting and financial regulation, we need to explicitly discuss the reliability of our research findings and find ways to counter the shortcomings in the research and publication process.

In addition, evidence-based policymaking would likely increase the political influence on the research process, another important challenge that needs to be recognized. Policymaking is

[25] E.g., Monogan (2013) and Chambers (2014). The Journal of Accounting Research recently experimented with a registration-based publication process for its 2017 conference (https://research.chicagobooth.edu/arc/journal-of-accounting-research/2017-registered-reports). Interestingly, the frequency of null results in the registered reports of the conference issue is much higher than what is typical in published accounting studies. See Bloomfield et al. (2018) for more discussion.

[26] Open Science Collaboration (2015) reports that 97 out of 100 studies in the Reproducibility Project had positive results. As Abadie (2018) points out, insignificant results or non-rejections could be quite informative.

[27] Berry et al. (2017) find that less than 30% of the studies in AER's centenary volume were replicated by other studies. In contrast, Hamermesh (2017) argues that the most important studies in labor economics are "replicated" in that their ideas are tested in other settings and contexts. While this is an important argument, it applies more to general ideas and relations, which is less helpful to policymakers relying on specific results in specific studies.

inherently political. One concern is that the use of evidence in policymaking could lead to researchers seeking particular results and to so-called "policy-based" research. There is considerable debate on the politics of evidence-based policy (e.g., Byrne 2011, Cairney 2016).

A related concern is that cost-benefit analysis is not "neutral" and used politically, e.g., to prevent regulation (e.g., Driesen 2006). The very process of cost-benefit analysis could facilitate regulatory capture (Cochrane 2014), which brings me to the influence of those having an interest in the policies (e.g., industry or lobby organizations). Research funding, advisory relations, and access to proprietary data can create conflicts of interest for researchers (see my own disclosures). For instance, there is evidence that research funded by drug companies is associated with more positive drug effects (e.g., Bekelman et al. 2003, Sismondo 2008). Mechanisms range from implicit bias to drug companies' control over research design or their suppression of unfavorable research results (Collier and Iheanacho 2002, Moore and Loewenstein 2004, Sage 2006). These conflicts of interest exist not only in medical research.

All these arguments highlight that the various influences on the research process need serious consideration, especially if research findings are used for evidence-based policymaking. I shall note that having academics involved in the policy process is not necessarily bad. As Zingales (2015, p. 1329) puts it: Researchers "should get more involved in policy (while not in politics)." But we need to recognize the political nature of policymaking, the potential political influences on research, and put in place appropriate safeguards.

4.5 Transmission of Research Findings and Political Incentives in Using Evidence

The next issue is the transmission of research findings to policymakers as well as political incentives in using evidence. It is of course critical that research findings are easily accessible,

appropriately synthesized and communicated. This process is an important part of evidence-based policymaking, which cannot be left to its own devices and needs serious consideration, especially when considering the political incentives for policymakers when using evidence.

A common claim is that policymakers ignore, do not understand or do not act on evidence (e.g., Cairney 2016). Consistent with this notion, Blinder (1987) formulates the economic policy version of Murphy's Law as: "Economists have the least influence on policy where they know the most and are most agreed; they have the most influence on policy where they know the least and disagree most vehemently." He also offers O'Connor's Corollary: "When conflicting economic advice is offered, only the worst will be taken." Hahn and Tetlock (2009) review the track record of economic analysis for regulatory decisions and find little evidence that economic analysis has substantially improved regulatory decisions.

One reason for these pessimistic views is the political process by which public policies are generally chosen (see also Acemoglu and Robinson 2013). A good example is the outsized influence of special interests when the costs and benefits of a policy are asymmetrically distributed. But beyond reasons related to the political process, the accessibility of evidence and the use of research findings by policymakers can play a role. Policymakers may not be aware of the results for a particular policy issue or not have the training to understand them. Moreover, political factors and incentives can lead to the selective use of evidence (e.g., Bastow et al. 2014, p. 144). Watts and Zimmerman (1979) have pointed out a long time ago that the existence of different interest groups creates a demand for prescriptive research (or theory) and a "market for excuses" (see also Horngren 1973, Zeff 1974). Related to this argument is the concern that regulators and policymakers have a tendency to "cherry-pick" evidence to legitimize or support policy (Byrne 2011). One could argue that this concern is precisely why we need evidence-based

policymaking, i.e., a more rigorous and systematic (rather than selective) approach to supporting policy by research and evidence. At the same time, the debate alerts us to the concern that, given the political nature of public policy, research is unlikely to be used in a neutral fashion.[28]

But even setting politics aside, the synthesis and transmission of research findings is a major undertaking. Policymakers typically lack the time to effectively search and synthesize the relevant research literature. Also, there is usually a major knowledge gap between researchers and policymakers. As not all research findings are created equal, policymakers not only need to understand the findings themselves, but should also have a solid understanding of the underlying research methods and their limitations. My earlier discussion of the reliability of research findings underscores this point. Moreover, many research findings need to be interpreted and applied to the specific policy question at hand. In my casual assessment, it is rare that accounting and financial markets research directly speaks to or prescribes a particular policy choice.

Policymakers probably do not have the time, training or the necessary institutional support to overcome these challenges. Thus, researchers need to take a more active role in the synthesis, transmission and communication of research findings. But we need to do this in an unbiased and systematic fashion that is helpful to policymakers.[29] Stanley and Jarrell (1989) argue that the typical literature reviews in economic journals contribute very little to research synthesis. Thus, we need to go beyond them. I come back to this issue in Section 5.2.

[28] In addition, Taber and Lodge (2006) provide evidence that citizens are biased-information processors when given evidence in policy debates. Moreover, so-called "false balance" in media articles can distort perceptions of expert opinion even when participants have all the information needed to correct for its influence (Koehler 2016). These studies illustrate the complexity of the transmission process for academic evidence.

[29] Evidence in Sumner et al. (2014) provides a cautionary tale, showing that the exaggeration in health-related news in the media is positively correlated with such exaggeration in university press releases. Thus, we also need to think about the incentives of researchers and their universities communicating research findings.

5. Looking Forward: How to Better Support Policymaking

The previous section highlights numerous and significant challenges for evidence-based policymaking. Thus, it is important to have realistic expectations as to what academic research can deliver at present. At the same time, using evidence in policymaking holds significant promise. Moreover, policymakers and standard setters face significant pressures to perform economic analyses and to practice evidence-based policymaking.

Therefore, in this section, I discuss a number of potential routes we could take such that accounting and financial markets research could better support policymaking. Specifically, I discuss suggestions (not necessarily new ones) for how to organize and facilitate policy-oriented research, how to better aggregate findings, and how regulators could help and enable more and better policy-relevant research. These ideas tie directly into the challenges that I discussed in Section 4. The key message that I am building towards is that we need to make significant changes to the research process and investments into an infrastructure if evidence is to inform and systematically support policymaking.

5.1 Ideas for Policy-Oriented Research on Accounting Standards and Financial Regulation

Currently, research in accounting and finance is conducted in silos that are generally defined by methods or field and methods. Examples of such silos are "empirical financial accounting" or "empirical corporate finance" (see also Bloomfield et al. 2016). I suspect that these silos also exist when we cite studies and when we conduct literature reviews. However, regulators, standard setters and policymakers care less about the methods (or the field) and instead focus on the topic or the answer to a policy question. Thus, we should find ways to organize research and conferences around topics and policy questions, bringing together scholars from different fields

using different methods. For example, disclosure is a topic that is being studied by many fields (e.g., accounting, economics, finance, sociology, etc.). Cross-fertilization across fields may stimulate new and also more robust research. Along the same lines, I argue for a plurality of methods when tackling policy questions, either in the same or in multiple studies. For instance, studies could combine descriptive and quantitative approaches (e.g., provide regression analyses and interview or survey evidence). Similarly, there are studies in industrial organization comparing structural and reduced-form estimates for the same setting (Ashenfelter and Hosken 2010). All methods have something to contribute, and it is important to keep an open mind about approaches and methods.

To organize research more around policy questions, we should consider creating academic journals that specialize in publishing research on certain policy issues or program evaluations. If these journals have policy impact, they would be attractive research outlets. From an evidence-based policymaking perspective, there is value in studies that (re-)examine similar policies in different countries, states, cities or time periods. However, top-tier journals in economics, finance or accounting are likely to publish only the first of these studies. Policy-oriented journals could publish the others. These journals could also be the outlet for replications. As noted in Section 4.4, we need more replications, but also more stress testing of extant results, to boost the reliability of our research evidence. In this regard, it is important for all journals to require code and data sharing, as these policies lower the costs of replications and extensions (see evidence in footnote 24). In addition to academic journals, we could create for fora that allow standard setters and policymakers to openly solicit new research that would inform their decisions as well as to engage with the academic community. The recently formed EFRAG Academic Panel is an example that goes in this direction.

The final suggestion for research that would better support policymaking is specific to accounting and financial regulation. As discussed in Section 4.2, measuring the strength of accounting or regulatory treatment is a major obstacle in the estimation of policy effects that are more generalizable or at least easier to aggregate across studies. Thus, we should invest in new research that allows us to better quantify the amount of information in accounting numbers, financial statements, annual reports, or disclosures, which in turn would allow us to compute changes in the amount of information around regulatory changes.

5.2 Ideas on How to Aggregate Policy-Relevant Research Findings

The transmission and communication of research findings to policymakers is an important part of evidence-based policy. As discussed in Section 4.5, this process faces several major challenges, including political influences. In this section, I discuss several ideas related to the aggregation and communication of existing evidence that would be more independent from the policymakers and reduce selective use of evidence or political "cherry-picking."

In my view, accounting and financial market research needs to start by developing a "canon" of economic relations, tradeoffs, and effects that are both relevant to standard setters and regulators but also well-understood in the literature and reliably estimated. One step towards such a canon could be a sequence of surveys asking researchers: What is the set of policy-relevant results or relations in this area that are well understood and reliably estimated? Which results in this literature do we expect to be able to replicate? Such surveys, which are similar to the IGM Economic Expert polls mentioned earlier, would give us a better sense for which results we have a shared academic consensus and provide a starting point. As discussed in Section 2, the consensus tends to reflect the depth of the literature. Establishing such a consensus also makes it harder for interested parties to deny such a consensus exists and, simultaneously, easier for the

general public to see when interested parties refer to economic effects or results that are not supported by the academic consensus.

We also need to collect and aggregate findings by policy issue or question. Towards this end, several "research clearinghouses" across a range of policy areas have been established in recent years. Examples are the "Clearinghouse for Labor Evaluation and Research" (CLEAR) and the U.S. Department of Education's "What Works Clearinghouse." The Pew-MacArthur Results First Initiative (2014) shows that such clearinghouses exist for studies on adult and juvenile justice, child welfare, mental health, pre-K to higher education, and substance abuse. Among other things, these clearinghouses conduct systematic literature reviews to identify effective public programs, most of them using explicit criteria for evaluating the strength of evidence as well as structured summaries that allows policymakers to easily compare the relative effectiveness of programs (Pew-MacArthur Results First Initiative 2014). At present, I am not aware of formal assessments of these clearinghouses and their systematic reviews. However, in my view, the idea of such clearinghouses is appealing and a step in the right direction, especially if they are operated independently and their systematic reviews follow scientific guidelines for evaluating evidence. We could create such clearinghouses for research on standard setting and financial regulation, including (independent) post-implementation reviews.

The idea of systematic reviews stems originally from the Cochrane Collaboration in medicine. Systematic Cochrane reviews are very different from the typical literature reviews in accounting or finance. They are comprehensive in that they review all available primary research and summarize the best available evidence. The reviews follow extensive guidelines for every step in the process. These guidelines are established in the Cochrane Handbook for Systematic Reviews (Higgins and Green 2011), which cover the selection of studies, assessments of bias and

criteria for classifying evidence by its epistemological strength (Chapter 12). Cochrane reviews often conduct meta-analyses or include statistical analyses to compare effects in different studies. They provide conclusions divided into implications for practice and implications for research. The reports are available in two versions, one being a plain-language summary to help with the transmission and communication of the findings and the conclusions to wider audiences.

To produce a systematic review on a particular medical intervention, a group of researchers registers a protocol. The review is a massive collaborative effort involving many researchers, often a worldwide network. Most of the Cochrane reviews are concerned with particular medical interventions and practice guidelines, although some focus on policy issues, such as incentives for smoking cessation, electronic cigarettes for smoking cessation, restricting or banning alcohol advertising to reduce alcohol consumption in adults and adolescents.

The Cochrane Collaboration was an integral part of the development towards evidence-based medicine (see Section 6), suggesting that a similar effort would likely be central to evidence-based policymaking. An important question here is whether we even could perform such reviews for accounting and financial markets research. In thinking about this question, it is important to realize how closely connected Cochrane reviews are to the rise of RCTs.[30] Having RCTs facilitates meta-analyses, which is presumably one reason why meta-analyses are much more common in medicine than they are in accounting, economics or finance. Meta-analyses are trickier and less powerful for non-randomized studies and, at a minimum, they must consider potential confounders in the underlying studies (Higgins and Green 2011).

Generally speaking, aggregation of findings is much harder and more qualitative for public

[30] Using RCTs is the norm for Cochrane reviews. However, the Cochrane Collaboration realizes that there are questions of interest that cannot be answered by randomized trials and hence provide guidance on when it might be appropriate to include non-randomized studies (Chapter 13).

policy research. Interestingly, systematic reviews in medicine seem to focus primarily on internal validity, i.e., on whether the evidence is sound. However, in evidence-based policymaking, it is just as important to guide policymakers with respect to the external validity and applicability of the findings (Avellar et al. 2017). The closest analog to Cochrane reviews in public policy are the Campbell reviews, which are formally connected to the Cochrane Collaboration. They provide systematic reviews for public policy issues, including criminal justice, education, and social welfare policy. Campbell reviews also follow guidelines, although the guidance is less explicit and standardized than for Cochrane reviews, consistent with the view that systematic reviews in public policy are harder to conduct. Furthermore, simple comparisons in Google Scholar suggest that Campbell reviews have had less impact than Cochrane reviews. Nevertheless, Campbell reviews demonstrate that a form of systematic review for public policy is feasible, but that it requires a massive effort and faces substantial challenges. At present, Cochrane reviews as in medicine are more aspirational for accounting and financial markets research.

5.3 Ideas on How Policymakers Can Help Generate Data and Facilitate Analyses

In Section 4.3, I argue that the biggest obstacle to more and better policy-relevant research as well as causal inferences in accounting and financial markets research is lack of data. If the goal is evidence-based policy in these areas, then we need to overcome this problem and find ways to increase data availability or to even actively generate data. On both dimension, firms and regulators can help.[31]

First, regulatory agencies often have relevant data from their regulatory or supervisory activities that they could make available for economic analysis. Many of them already do so (e.g., PCAOB and SEC), but they could go much further. Firms can also provide much more data

[31] This section draws on ideas that I have also expressed in Leuz and Wysocki (2016).

for regulatory analyses (and not just analyses of their business practices that more directly benefit them). After all, if certain regulations are (net) costly to firms, they should have something to gain from better economic analysis. Of course, there are issues related to data security and confidentiality, but we can and should find ways to address these issues.

Second, whenever regulators and standard setters create new rules that mandate certain disclosure, reporting or auditing practices, they should contemplate including mandates for firms (or their auditors) to collect and keep relevant data around these regulatory changes. Such data could be shared with researchers for ex-post economic analyses or post-implementation reviews. As I discussed in Section 4.3, we need much more granular data about the specific changes that occur at the firm level as a result of a new accounting standard or a new financial regulation in order to perform better economic analyses.

Let me illustrate the basic idea with a new accounting rule for asset impairments. Once the rule has been promulgated (and ideally even before it is effective to generate pre-period benchmark data), firms would be required to keep specific data on their impairment testing, in particular, what the impairment would have been under the old rule in the post-period as well as details about the changes made under the new rule. We also would need data on impairments that management considered and information on why impairments were not taken or not necessary. The latter information is usually not available to researchers. Moreover, these data would be much more granular and specific to the new accounting rule, which would allow us to measure the treatment and information consequences much better. In addition, firms should have relevant data on the implementation costs and other effects from the new rule. Including data-collection requirements into new accounting standards or financial regulation would substantially improve the data situation for economic analysis.

There are two obvious counterarguments to this idea. First, firms' regulatory burden is already high and such requirements would further increase it. Second, the data are proprietary. My response to the first concern is that we cannot have it both ways. If we are serious about evidence-based policymaking precisely because we are concerned that new rules and regulations could be harmful or *net* costly to firms, then studying the effects of new rules is absolutely worthwhile and could bring substantial cost savings to firms and society. We simply need to invest into smarter policymaking. However, we could reduce the data collection costs and spread the burden around by randomly selecting a sufficient number of firms, rather than asking the entire population to collect the relevant data.

To alleviate the second concern, the data would be made available to researchers on a confidential basis following, for instance, the model of the U.S. Census Bureau. The Census Bureau has very confidential data and their process for making this data available has worked well, as far as I can tell. It would be important to ensure that the data are made available for research in a way that minimizes political influences. In addition, I would advocate for making the data available to researchers only, and not to the respective regulator or the supervisor. Doing so would alleviate firms' concerns that the regulator or the supervisor could use the data for enforcement actions against the firm, which would likely increase the quality of the data.[32]

This discussion more generally illustrates that accounting rules or financial market regulation need to be written with ex-post evaluation in mind. Regulators and standard setters should ask during the rulemaking phase: What data will we need and how would we able to tell that the rule works as intended?

[32] If such data would *also* be useful for supervisory or enforcement purposes, then making them available for this purpose is a regulatory decision that is separate from the initial economic analysis of mandate.

In addition to making data available and generating data through mandates, standard setters and regulators can implement new standards or rules in ways that are more conducive to economic analysis. For instance, having thresholds above or below which new rules apply would facilitate regression-discontinuity designs, which under certain conditions are reasonably close to random assignment. Furthermore, new rules or standards could become effective in a staggered fashion, i.e., phased-in over time.[33] Doing so would help the identification of causal estimates, as it allows studies to better control for concurrent events with appropriate time-fixed effects. Such a research design can be further enhanced by exploiting differences in firms' fiscal year ends (e.g., Daske et al. 2008, Christensen et al. 2013). However, using a staggered implementation along with firms' fiscal year ends for identification purposes is not a panacea. The effective dates need to be plausibly exogenous, e.g., pre-determined or tied to arbitrary characteristics like the ticker symbol (see, e.g., the phase-in of the Eligibility Rule for the OTCBB). If firms can choose when to adopt or implement the new rules, then we are back to standard selection concerns. In addition, we need to consider if the staggered implementation leads to spillover (or anticipation) effects, which in turn could contaminate the estimation of the regulatory effects.

Thresholds and implementation staggering often raise concerns about fairness. But again, there is a tradeoff. We essentially need to weigh the potential societal gains from better regulation due to economic analysis and evidence-based policymaking against the fairness concerns that arise with arbitrary implementation schemes.[34]

[33] The staggering could be along at least two dimensions. First, components of a "regulatory package" might be implemented in a staggered fashion to facilitate the evaluation of the components. Second, the entire regulatory package could be applied to cohorts of firms in a time-staggered fashion to allow for better identification of regulatory effects using regulated and yet unrelated firms.

[34] See also Abramowicz et al. (2011). They push even further and argue for randomly assigning individuals, firms, or jurisdictions to different legal rules.

A final way in which regulators could contribute to evidence-based policymaking is by explicitly conducting experiments and pilot studies. In contrast to the suggestions above, pilot studies can also help with prospective analyses. Such studies (with randomization) have been conducted. A good example is the Regulation SHO pilot program that the SEC did on short sale restrictions (e.g., Li and Zhang 2015). Another example is the FINRA tick size pilot program. I would encourage regulators to perform such pilot studies (with randomization) more often.

In sum, regulators and standard setters can contribute substantially to better economic analysis and in doing so lead the way towards evidence-based policymaking. While some of the above suggestions would likely entail major changes in the way we set accounting standards and write financial regulation, these changes are worth considering. There are potentially large gains from better and smarter regulation, especially when we are convinced that current regulation is burdensome. Moreover, I am convinced that we will not make headway towards evidence-based policymaking if we do not address the data and related causal inference issues.

6. Creating an Infrastructure for Evidence-Informed Policymaking

In my view, we are still a long way from evidence-based policymaking in accounting and financial markets. We need to temper our expectations, starting with the objectives, which is one reason why I prefer the more modest term "evidence-*informed* policymaking." The systematic use of academic evidence to inform standard setting, regulation and policy requires substantial investments into the research infrastructure, including the synthesis and transmission of findings. It is not something that policymakers can simply decide to do on their own. Building the necessary research infrastructure takes time and, if taken seriously, evidence-based policymaking requires a concerted and long-term effort by researchers and policymakers.

Let me illustrate this general point with the rise of evidence-based medicine. The core idea of evidence-based medicine is to apply the best available research to clinical decision making and practice guidelines.[35] It places less emphasis on expert intuition and unsystematic clinical experience and instead emphasizes the systematic examination and accumulation of clinical research. This approach is now the leading paradigm for clinical practice. The emergence of evidence-based medicine is viewed as one of the 15 most important milestones in medicine, along with the discovery of antibiotics (BMJ 2007), illustrating the tremendous upside potential for the systematic use of academic evidence.

But the development of evidence-based medicine was a massive and long-term effort. It started in the 1970s with David Sackett, Alvan Feinstein and other researchers at McMaster University in Canada, at the time a new medical school, which established the world's first department of clinical epidemiology and biostatistics (Zimerman 2013). McMaster University created a new curriculum ("problem-based learning method") and Sackett developed courses (and criteria) for the "critical appraisal" of the literature.[36] The latter illustrates that methodological changes were central to the development of evidence-based medicine. Consistent with this notion, its development is closely connected to the rise of randomized-control trials and meta-analyses. Iain Chalmers and the creation of the Cochrane Collaboration in 1993, named in honor of Archie Cochrane's earlier contributions, played a central role, facilitating the review, aggregation and dissemination of relevant clinical research through a worldwide, network-based approach (Chandler and Hopewell 2013, Smith and Rennie 2014). The development of evidence-

[35] For more elaborate definitions see Guyatt (1991), the Evidence-Based Medicine Working Group (1992), and Sackett et al. (1996). It was first coined as the "systematic approach to analyze published research as a basis for clinical decision making." (Claridge and Fabian 2005).

[36] For more detailed accounts of the history, see Zimerman (2013) and Smith and Rennie (2014). Evidence-based medicine is closely related to clinical epidemiology and in this sense the methods were not new and it has much older roots (Claridge and Fabian 2005). For instance, "the father of evidence-based medicine" David Sackett acknowledges that he was inspired by the work of Thomas Chalmers (Sackett 2010).

based medicine further involved academic courses for medical students, researchers and physicians, a series of journal articles and editorials, workshops and conferences, and the creation of electronic databases for clinical trials. The movement also had substantial institutional support by one of the leading medical journals (JAMA) and its editor Drummond Rennie (Zimerman 2013). Overall, it took more than 20 years for evidence-based medicine to become widespread and widely accepted.

This historical account on evidence-based medicine illustrates why I believe that evidence-informed policymaking requires major effort and substantial investments. The article delineates a number of suggestions on how accounting and financial market research could better support evidence-informed policymaking. Among other things, I discuss that such policymaking requires that we systematically aggregate findings and evidence, regularly perform replications, and conduct many more closely related studies in different settings. We also need to obtain help from regulators (and firms) in generating relevant data and exogenous regulatory variation. But just as in medicine, developing this approach will take time.

Importantly, however, I view evidence-based medicine more as an aspirational example. I do not argue that evidence-informed policymaking is the same as evidence-based medicine or that accounting and financial markets research could deliver what medical research has delivered for evidence-based medicine. As discussed earlier, there are substantial differences between natural science and social science as well as medicine and public policy (e.g., financial market regulation). These differences are another reason why we should use the term of evidence-informed policymaking. Research can inform policymakers but it cannot fully determine policy. The notion that, once we have enough rigorous research, it can tell us the right or optimal policy

without much further judgement is naïve. Professional judgment by policymakers will continue to play a significant role in economic and public policymaking (see also Coates 2015).[37]

But even then, research can deliver important quantitative but also qualitative insights and improve policymaking, and therefore the efforts are – at least in my mind – not in vain. As Blinder (1987, p. 10) says, the "fact that economists do not know everything does not mean that they do not know anything." Also, in many instances, just knowing certain empirical facts can be helpful to policymakers or policy debates. Let me illustrate this point with an example. The financial crisis led to a major policy debate over the role of fair-value (or mark-to-market) accounting, as it was viewed by many as an important factor contributing to or exacerbating the crisis. This debate led to pressures on the accounting standard setters and banking regulators to change the accounting rules (e.g., Financial Stability Forum 2009). However, basic descriptive evidence on banks' balance sheets and activities could have informed standard setters and banking regulators that it was unlikely that fair-value accounting played a significant role. Examples for relevant facts include the fraction of financial assets that banks reported at fair value, what fraction of these assets were actually marked to market as opposed to a matrix or a model, and information on when banks started taking write-downs on their financial assets, relative to when the crisis unfolded (see discussions by Laux and Leuz 2009, 2010). This is not to say that the question about potential negative effects of fair-value accounting during a crisis could have been answered without a more extensive empirical analysis. This example merely serves to illustrate that basic empirical facts can be very useful to policy debates.

To me, the glass is half full. There are many challenges and we have to tread carefully. But considering the potential costs of poorly designed or implemented regulation, it is worth going

[37] Even in evidence-based medicine, expert knowledge and context are relevant and still needed (Sackett et al. 1996).

down the path towards a more systematic use of evidence in policy making, in general and also when it comes to accounting standards and financial markets regulation specifically. This path requires the cooperation of researchers and policymakers. We need a concerted and systematic effort as well as substantial infrastructure investments, if we do not want to pay lip service to the term and the idea.

References

Abadie, A. 2018. Statistical Non-Significance in Empirical Economics. *Working Paper, MIT*.

Abramowicz, M., Ayres, I., and Listokin, Y. 2011. Randomizing Law. *University of Pennsylvania Law Review*, 159 (4), 929-1005.

Abreu, M., Grinevich, V., Hughes, A., and Kitson, M. 2009. Knowledge Exchange between Academics and the Business, Public and Third Sectors. Available from: https://eprints.soton.ac.uk/357117/ [Accessed 1 March 2018].

Acemoglu, D., and Robinson, J.A. 2013. Economics Versus Politics: Pitfalls of Policy Advice. *Journal of Economic Perspectives*, 27 (2), 173-192.

Acharya, V.V., and Ryan, S.G. 2016. Banks' Financial Reporting and Financial System Stability. *Journal of Accounting Research*, 54 (2), 277-340.

Alston, R.M., Kearl, J.R., and Vaughan, M.B. 1992. Is There a Consensus among Economists in the 1990's? *The American Economic Review*, 82 (2), 203-209.

Angrist, J.D., and Pischke, J.-S. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton, NJ: *Princeton University Press*.

Angrist, J.D., and 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.

Angrist, J.D., and Pischke, J.-S. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton, NJ: *Princeton University Press*.

Ashenfelter, O., and Hosken, D. 2010. The Effect of Mergers on Consumer Prices: Evidence from Five Mergers on the Enforcement Margin. *The Journal of Law and Economics*, 53 (3), 417-466.

Avellar, S.A., Thomas, J., Kleinman, R., Sama-Miller, E., Woodruff, S.E., Coughlin, R., and Westbrook, T.P.R. 2017. External Validity: The Next Step for Systematic Reviews? *Evaluation Review*, 41 (4), 283-325.

Badertscher, B., Shroff, N., and White, H.D. 2013. Externalities of Public Firm Presence: Evidence from Private Firms' Investment Decisions. *Journal of Financial Economics*, 109 (3), 682-706.

Ball, R. 1980. Discussion of Accounting for Research and Development Costs: The Impact on Research and Development Expenditures. *Journal of Accounting Research*, 18, 27-37.

Ball, R., Robin, A., and Wu, J.S. 2003. Incentives Versus Standards: Properties of Accounting Income in Four East Asian Countries. *Journal of Accounting and Economics*, 36 (1-3), 235-270.

Bastow, S., Dunleavy, P., and Tinkler, J. 2014. The Impact of the Social Sciences: How Academics and Their Research Make a Difference. London, UK: *Sage*.

Basu, S., and Park, H.-U. 2016. Publication Bias in Recent Empirical Accounting Research. *Working Paper, Temple University*.

Bauguess, S., Countryman, V., and Phatak, N. 2017. The Role of Academic Research in Financial Market Policy. Division of Economic and Risk Analysis, U.S. SEC: *FMA Annual Meeting - Boston 2017*.

Begley, C.G. 2013. Six Red Flags for Suspect Work. *Nature*, 497 (7450), 433-434.

Begley, C.G., and Ioannidis, J.P.A. 2015. Reproducibility in Science: Improving the Standard for Basic and Preclinical Research. *Circulation Research*, 116 (1), 116-126.

Bekelman, J.E., Li, Y., and Gross, C.P. 2003. Scope and Impact of Financial Conflicts of Interest in Biomedical Research: A Systematic Review. *JAMA*, 289 (4), 454-465.

Berry, J., Coffman, L.C., Hanley, D., Gihleb, R., and Wilson, A.J. 2017. Assessing the Rate of Replication in Economics. *American Economic Review*, 107 (5), 27-31.

Blinder, A.S. 1987. Hard Heads, Soft Hearts: Tough-Minded Economics for a Just Society. Boston, MA: *Addison Wesley Publishing Company*.

Bloomfield, R., Nelson, M.W., and Soltes, E. 2016. Gathering Data for Archival, Field, Survey, and Experimental Accounting Research. *Journal of Accounting Research*, 54 (2), 341-395.

Bloomfield, R., Rennekamp, K., and Steenhoven, B. 2018. No System Is Perfect: Understanding How Registration-Based Editorial Processes Affect Reproducibility and Investment in Research Quality. *Journal of Accounting Research*, Forthcoming.

BMJ. 2007. Milestones on the Long Road to Knowledge. *BMJ*, 334 (suppl 1), s2-s3.

Bratten, B., Choudhary, P., and Schipper, K. 2013. Evidence That Market Participants Assess Recognized and Disclosed Items Similarly When Reliability Is Not an Issue. *The Accounting Review*, 88 (4), 1179-1210.

Breuer, M. 2017. How Does Financial-Reporting Regulation Affect Market-Wide Resource Allocation? . *Working Paper, University of Chicago*.

Brien, S.E., Ronksley, P.E., Turner, B.J., Mukamal, K.J., and Ghali, W.A. 2011. Effect of Alcohol Consumption on Biological Markers Associated with Risk of Coronary Heart Disease: Systematic Review and Meta-Analysis of Interventional Studies. *BMJ*, 342 (7795), 1-15.

Brodeur, A., Lé, M., Sangnier, M., and Zylberberg, Y. 2016. Star Wars: The Empirics Strike Back. *American Economic Journal: Applied Economics*, 8 (1), 1-32.

Buijink, W. 2006. Evidence-Based Financial Reporting Regulation. *Abacus*, 42 (3-4), 296 -301.

Burgstahler, D.C., Hail, L., and Leuz, C. 2006. The Importance of Reporting Incentives: Earnings Management in European Private and Public Firms. *Accounting Review*, 81 (5), 983-1016.

Burton, F.G., Summers, S.L., Wilks, T.J., and Wood, D.A. 2017. An Evaluation of Research Impact in Accounting, Economics, Finance, Management, Marketing, Psychology, and the Natural Sciences. *Working Paper, Brigham Young University*.

Bushee, B.J., and Leuz, C. 2005. Economic Consequences of SEC Disclosure Regulation: Evidence from the OTC Bulletin Board. *Journal of Accounting and Economics*, 39 (2), 233-264.

Bushman, R.M., and Williams, C.D. 2015. Delayed Expected Loss Recognition and the Risk Profile of Banks. *Journal of Accounting Research*, 53 (3), 511-553.

Byrne, D.S. 2011. Applying Social Science: The Role of Social Research in Politics, Policy and Practice. Bristol, UK: *Policy Press*.

Cairney, P. 2016. The Politics of Evidence-Based Policy Making. London, UK: *Springer Nature*.

Camerer, C.F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., and Wu, H. 2016. Evaluating Replicability of Laboratory Experiments in Economics. *Science*, 351 (6280), 1433-1436.

Cascino, S. and Gassen, J. 2015. What drives the comparability effect of mandatory IFRS adoption? *Review of Accounting Studies*, 20 (1), 242-282.

Chambers, C. 2014. Registered Reports: A Step Change in Scientific Publishing. Available from: https://www.elsevier.com/reviewers-update/story/innovation-in-publishing/registered-reports-a-step-change-in-scientific-publishing [Accessed 1 March 2018].

Chandler, J., and Hopewell, S. 2013. Cochrane Methods - Twenty Years Experience in Developing Systematic Review Methods. *Systematic Reviews*, 2 (1), 76-81.

Chang, A.C., and Li, P. 2018. Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Often Not". *Critical Finance Review*, Forthcoming.

Chemla, G., and Hennessy, C. 2018a. Subject Rational Expectations Will Contaminate Randomized Controlled Medical Trials *Working Paper, Imperial College*.

Chemla, G., and Hennessy, C. 2018b. Rational Expectations and the Paradox of Policy-Relevant Natural Experiments. *Working Paper, Imperial College*.

Choi, J.H. 2018. Accrual Accounting and Resource Allocation: A General Equilibrium Analysis. *Working Paper, Stanford University*.

Christensen, H.B., Hail, L., and Leuz, C. 2013. Mandatory IFRS Reporting and Changes in Enforcement. *Journal of Accounting and Economics*, 56 (2-3), 147-177.

Claridge, J.A., and Fabian, T.C. 2005. History and Development of Evidence-Based Medicine. *World Journal of Surgery*, 29 (5), 547-553.

Coates, J.C. 2015. Cost-Benefit Analysis of Financial Regulation: Case Studies and Implications. *Yale Law Journal*, 124 (4), 882-1011.
Coates, J.C., and Srinivasan, S. 2014. SOX after Ten Years: A Multidisciplinary Review. *Accounting Horizons*, 28 (3), 627-671.
Cochrane, J.H. 2014. Challenges for Cost-Benefit Analysis of Financial Regulation. *The Journal of Legal Studies*, 43 (S2), S63-S105.
Collier, J., and Iheanacho, I. 2002. The Pharmaceutical Industry as an Informant. *The Lancet*, 360 (9343), 1405-1409.
Daske, H., Hail, L., Leuz, C., and Verdi, R. 2008. Mandatory IFRS Reporting around the World: Early Evidence on the Economic Consequences. *Journal of Accounting Research*, 46 (5), 1085-1142.
Dechow, P., Ge, W.L., and Schrand, C. 2010. Understanding Earnings Quality: A Review of the Proxies, Their Determinants and Their Consequences. *Journal of Accounting and Economics*, 50 (2-3), 344-401.
Dechow, P.M. 1994. Accounting Earnings and Cash Flows as Measures of Firm Performance: The Role of Accounting Accruals. *Journal of Accounting and Economics*, 18 (1), 3-42.
Demski, J.S. 2007. Is Accounting an Academic Discipline? *Accounting Horizons*, 21 (2), 153-157.
Domikowsky, C., Foos, D., and Pramor, M. 2017. Loan Loss Accounting Rules and Bank Lending over the Cycle: Evidence from a Global Sample. *Working Paper, Deutsche Bundesbank*.
Dranove, D., and Jin, G.Z. 2010. Quality Disclosure and Certification: Theory and Practice. *Journal of Economic Literature*, 48 (4), 935-963.
Dranove, D., Kessler, D., McClellan, M., and Satterthwaite, M. 2003. Is More Information Better? The Effects of "Report Cards" on Health Care Providers. *Journal of Political Economy*, 111 (3), 555-588.
Driesen, D.M. 2006. Is Cost-Benefit Analysis Neutral. *University of Colorado Law Review*, 77 (2), 335-404.
Duflo, E., Greenstone, M., Pande, R., and Ryan, N. 2013. Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India. *The Quarterly Journal of Economics*, 128 (4), 1499-1545.
Evidence-Based Medicine Working Group. 1992. Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine. *JAMA*, 268 (17), 2420-2425.
Ewert, R., and Wagenhofer, A. 2012. Using Academic Research for the Post-Implementation Review of Accounting Standards: A Note. *Abacus*, 48 (2), 278-291.
Fellingham, J.C. 2007. Is Accounting an Academic Discipline? *Accounting Horizons*, 21 (2), 159-163.
Financial Stability Forum. 2009. Report of the Financial Stability Forum on Addressing Procyclicality in the Financial System. Available from: http://www.financialstability-board.org/publications/r_0904a.pdf [Accessed 1 March 2018].
Floyd, E., and List, J.A. 2016. Using Field Experiments in Accounting and Finance. *Journal of Accounting Research*, 54 (2), 437-475.
Flyvbjerg, B. 2001. Making Social Science Matter: Why Social Inquiry Fails and How It Can Succeed Again. Cambridge, UK: *Cambridge University Press*.
Fuller, D., and Geide-Stevenson, D. 2014. Consensus among Economists—an Update. *The Journal of Economic Education*, 45 (2), 131-146.
Fung, A., Graham, M., and Weil, D. 2007. Full Disclosure: The Perils and Promise of Transparency. Cambridge, UK: *Cambridge University Press*.
Gelman, A., and Loken, E. 2013. The Garden of Forking Paths: Why Multiple Comparisons Can Be a Problem, Even When There Is No "Fishing Expedition" or "P-Hacking" and the Research Hypothesis Was Posited Ahead of Time. *Working Paper, Columbia University*.
Geoffrey, R., and Lee, H. 2018. Academic Research and the Security and Exchange Commission's Rulemaking Process. *Working Paper, University of Chicago*.
Gilbert, D.T., King, G., Pettigrew, S., and Wilson, T.D. 2016. Comment on "Estimating the Reproducibility of Psychological Science". *Science*, 351 (6277), 1037-1037.

Glaeser, S., and Guay, W.R. 2017. Identification and Generalizability in Accounting Research: A Discussion of Christensen, Floyd, Liu, and Maffett (2017). *Journal of Accounting and Economics*, 64 (2), 305-312.
Gordon, R., and Dahl, G.B. 2013. Views among Economists: Professional Consensus or Point-Counterpoint? *American Economic Review*, 103 (3), 629-635.
Gow, I.D., Larcker, D.F., and Reiss, P.C. 2016. Causal Inference in Accounting Research. *Journal of Accounting Research*, 54 (2), 477-523.
Granja, J. 2018. Disclosure Regulation in the Commercial Banking Industry: Lessons from the National Banking Era. *Journal of Accounting Research*, 56 (1), 173-216.
Guyatt, G.H. 1991. Evidence-Based Medicine. *ACP Journal Club*, 114 (2), A16.
Hahn, R. and Tetlock, P. 2008. Has Economic Analysis Improved Regulatory Decisions? *Journal of Economic Perspectives*, 22 (1), 67–84.
Hail, L., Leuz, C., and Wysocki, P. 2010a. Global Accounting Convergence and the Potential Adoption of IFRS by the US (Part I): Conceptual Underpinnings and Economic Analysis. *Accounting Horizons*, 24 (3), 355-394.
Hail, L., Leuz, C., and Wysocki, P. 2010b. Global Accounting Convergence and the Potential Adoption of IFRS by the US (Part II): Political Factors and Future Scenarios for US Accounting Standards. *Accounting Horizons*, 24 (4), 567-588.
Hamermesh, D.S. 2017. Replication in Labor Economics: Evidence from Data, and What It Suggests. *American Economic Review*, 107 (5), 37-40.
Hammersley, M. 2013. The Myth of Research-Based Policy and Practice. London, UK: *Sage*.
Heckman, J.J. 1997. Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *The Journal of Human Resources*, 32 (3), 441-462.
Heckman, J.J. 2001. Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture. *Journal of Political Economy*, 109 (4), 673-748.
Hellwig, M. 2015. Neoliberal Religious Sect or Science? On the Relation between Academic Research and Real-World Policy Advice in Economics, Working paper, Max Planck Institute for Research on Collectie Goods 2015/17.
Higgins, J.P., and Green, S. 2011. Cochrane Handbook for Systematic Reviews of Interventions. Available from: http://handbook.cochrane.org [Accessed 1 March 2018].
Höffler, J.H. 2017. Replication and Economics Journal Policies. *American Economic Review*, 107 (5), 52-55.
Hopwood, A.G. 2007. Whither Accounting Research? *The Accounting Review*, 82 (5), 1365-1374.
Horngren, C.T. 1973. The Marketing of Accounting Standards. *Journal of Accountancy*, 136 (4), 61-66.
IGM Economic Experts Panels. 2017. Factors Contributing to the 2008 Global Financial Crisis. Available from: http://www.igmchicago.org/surveys-special/factors-contributing-to-the-2008-global-financial-crisis [Accessed 1 March 2018].
Iliev, P. 2010. The Effect of SOX Section 404: Costs, Earnings Quality, and Stock Prices. *The Journal of Finance*, 65 (3), 1163-1196.
Kanodia, C., Sapra, H., and Venugopalan, R. 2004. Should Intangibles Be Measured: What Are the Economic Trade-Offs? *Journal of Accounting Research*, 42 (1), 89-120.
Koch, A.S., Lefanowicz, C.E., and Robinson, J.R. 2013. Regulation FD: A Review and Synthesis of the Academic Literature. *Accounting Horizons*, 27 (3), 619-646.
Koehler, D.J. 2016. Can Journalistic "False Balance" Distort Public Perception of Consensus in Expert Opinion? *Journal of Experimental Psychology: Applied*, 22 (1), 24-38.
Laux, C., and Leuz, C. 2009. The Crisis of Fair-Value Accounting: Making Sense of the Recent Debate. *Accounting, Organizations and Society*, 34 (6), 826-834.
Laux, C., and Leuz, C. 2010. Did Fair-Value Accounting Contribute to the Financial Crisis? *Journal of Economic Perspectives*, 24 (1), 93-118.
Leuz, C., Nanda, D., and Wysocki, P.D. 2003. Earnings Management and Investor Protection: An International Comparison. *Journal of Financial Economics*, 69 (3), 505-527.

Leuz, C., and Wysocki, P.D. 2016. The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of Accounting Research*, 54 (2), 525-622.
Li, Y., and Zhang, L. 2015. Short Selling Pressure, Stock Price Behavior, and Management Forecast Precision: Evidence from a Natural Experiment. *Journal of Accounting Research*, 53 (1), 79-117.
McClure, C. 2018. Determinants of Tax Avoidance. *Working Paper, Stanford University*.
McCullough, B.D., McGeary, K.A., and Harrison, T.D. 2006. Lessons from the JMCB Archive. *Journal of Money, Credit and Banking*, 38 (4), 1093-1107.
Monogan, J.E. 2013. A Case for Registering Studies of Political Outcomes: An Application in the 2010 House Elections. *Political Analysis*, 21 (1), 21-37.
Moore, D.A., and Loewenstein, G. 2004. Self-Interest, Automaticity, and the Psychology of Conflict of Interest. *Social Justice Research*, 17 (2), 189-202.
Nevo, A., and Whinston, M.D. 2010. Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference. *Journal of Economic Perspectives*, 24 (2), 69-82.
Nikolaev, V.V. 2017. Identifying Accounting Quality. *Working Paper, University of Chicago*.
Open Science Collaboration. 2015. Estimating the Reproducibility of Psychological Science. *Science*, 349 (6251), 943.
Pew-MacArthur Results First Initiative. 2014. Evidence-Based Policymaking: A Guide for Effective Government. Available from: http://www.pewtrusts.org/~/media/assets/2014/11/evidencebasedpolicymakingaguideforeffectivegovernment.pdf [Accessed 1 March 2018].
Posner, E.A., and Weyl, E.G. 2014. Benefit-Cost Paradigms in Financial Regulation. *The Journal of Legal Studies*, 43 (S2), S1-S34.
Posner, E.A., and Weyl, E.G. 2015. Cost-Benefit Analysis of Financial Regulations: A Response to Criticisms. *The Yale Law Journal Forum*, 124, 246-262.
Ronksley, P.E., Brien, S.E., Turner, B.J., Mukamal, K.J., and Ghali, W.A. 2011. Association of Alcohol Consumption with Selected Cardiovascular Disease Outcomes: A Systematic Review and Meta-Analysis. *BMJ*, 342 (7795), 1-13.
Sackett, D.L. 2010. A 1955 Clinical Trial Report That Changed My Career. *Journal of the Royal Society of Medicine*, 103 (6), 254-255.
Sackett, D.L., Rosenberg, W.M.C., Gray, J.A.M., Haynes, R.B., and Richardson, W.S. 1996. Evidence Based Medicine: What It Is and What It Isn't. *BMJ*, 312 (7023), 71-72.
Sage, W.M. 2006. Some Principles Require Principals: Why Banning Conflicts of Interest Won't Solve Incentive Problems in Biomedical Research. *Texas Law Review*, 85 (6), 1413-1463.
Schipper, K. 2010. How Can We Measure the Costs and Benefits of Changes in Financial Reporting Standards? *Accounting and Business Research*, 40 (3), 309-327.
Simmons, J.P., Nelson, L.D., and Simonsohn, U. 2011. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22 (11), 1359-1366.
Singleton, J.D. 2016. Slaves or Mercenaries? Milton Friedman and the Institution of the All-Volunteer Military. In: *Milton Friedman: Contributions to Economics and Public Policy*, edited by R. Cord and J. D. Hammond. Oxford, UK: Oxford University Press, 499-522.
Sismondo, S. 2008. Pharmaceutical Company Funding and Its Consequences: A Qualitative Systematic Review. *Contemporary Clinical Trials*, 29 (2), 109-113.
Smith, K. 2018. Option Prices and Disclosure. *Working Paper, Stanford University*.
Smith, R., and Rennie, D. 2014. Evidence-Based Medicine—an Oral History. *JAMA*, 311 (4), 365-367.
Stanley, T.D., and Jarrell, S.B. 1989. Meta-Regression Analysis: A Quantitati ve Method of Literature Surveys. *Journal of Economic Surveys*, 3 (2), 161-170.
Stockwell, T., Zhao, J., Panwar, S., Roemer, A., Naimi, T., and Chikritzhs, T. 2016. Do "Moderate" Drinkers Have Reduced Mortality Risk? A Systematic Review and Meta-Analysis of Alcohol Consumption and All-Cause Mortality. *Journal of Studies on Alcohol and Drugs*, 77 (2), 185-198.

Sumner, P., et al., 2014. The association between exaggeration in health related science news and academic press releases: retrospective observational study, *British Medical Journal*, 349:g7015.
Swoger, B. 2014. Post Publication Peer-Review: Everything Changes, and Everything Stays the Same. *Scientific American*, March 26, 2014.
Taber, C.S., and Lodge, M. 2006. Motivated Skepticism in the Evaluation of Political Beliefs. *American Journal of Political Science*, 50 (3), 755-769.
Teixeira, A. 2014. The International Accounting Standards Board and Evidence-Informed Standard-Setting. *Accounting in Europe*, 11 (1), 5-12.
Verrecchia, R.E. 2001. Essays on Disclosure. *Journal of Accounting and Economics*, 32 (1), 97-180.
Watts, R.L., and Zimmerman, J.L. 1979. The Demand for and Supply of Accounting Theories: The Market for Excuses. *The Accounting Review*, 54 (2), 273-305.
Watts, R.L., and Zimmerman, J.L. 1986. Positive Theory of Accounting. Englewood Cliffs, NY: *Prentice-Hall*.
Waymire, G.B. 2012. Seeds of Innovation in Accounting Scholarship. *Issues in Accounting Education*, 27 (4), 1077-1093.
Wood, D.A. 2016. Comparing the Publication Process in Accounting, Economics, Finance, Management, Marketing, Psychology, and the Natural Sciences. *Accounting Horizons*, 30 (3), 341-361.
Zeff, S.A. 1974. Comments on 'Accounting Principles – How They Are Developed'. In: *Instituional Issues in Public Accounting*, edited by R. R. Sterling. London, UK: Routledge, 172-178.
Zimerman, A.L. 2013. Evidence-Based Medicine: A Short History of a Modern Medical Movement. *American Medical Association Journal of Ethics*, 15 (1), 71.
Zingales, L. 2015. Presidential Address: Does Finance Benefit Society? *The Journal of Finance*, 70 (4), 1327-1363.

european corporate governance institute

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.org

european corporate governance institute

ECGI Working Paper Series in Law

Editorial Board

Editor	Luca Enriques, Allen & Overy Professor of Corporate Law, Faculty of Law, University of Oxford
Consulting Editors	John Coates, John F. Cogan, Jr. Professor of Law and Economics, Harvard Law School
	Paul Davies, Senior Research Fellow, Centre for Commercial Law, Harris Manchester College, University of Oxford
	Horst Eidenmüller, Freshfields Professor of Commercial Law, University of Oxford
	Amir Licht, Professor of Law, Radzyner Law School, Interdisciplinary Center Herzliya
	Roberta Romano, Sterling Professor of Law and Director, Yale Law School Center for the Study of Corporate Law, Yale Law School
Editorial Assistants	Tamas Barko , University of Mannheim
	Sven Vahlpahl, University of Mannheim
	Vanessa Wang, University of Mannheim