Financial Economists Roundtable

For release
October 2019

Statement on Evidence-Based Regulation and the Limits of Pilot Studies

The Financial Economist Roundtable (FER) is a group of senior financial economists who have made significant contributions to the finance literature and seek to apply their knowledge to current policy debates. The Roundtable focuses on microeconomic issues in investments, corporate finance, as well as financial institutions and markets, both in the U.S. and internationally. It aims to create a forum for intellectual interaction that promotes in-depth analyses of current policy issues to raise the level of public and private policy debate and improve the quality of policy decisions.

FER was founded in 1993 and meets annually. Members attending an FER meeting discuss specific policy issues on which the FER may adopt statements. When the FER issues a statement, it reflects a consensus among at least two-thirds of the attending members, and all the members who sign it support it. The statements are intended to increase the awareness and understanding of public policy makers, the financial economics profession, the communications media, and the public. FER distributes its statements to relevant policy makers and the media. This statement is the outcome of the FER’s discussion at its annual meeting, which took place on July 13-15, 2019, in Bolder, Colorado.

We signatories to this statement believe that financial regulators, when pursuing the goals of evidence-based decision making, should recognize both the benefits and limitations of the use of regulatory pilot studies. While pilot studies have potential to generate new knowledge about financial markets, regulators often can effectively evaluate the potential impact of a proposed regulation by analyzing archival data obtained from other markets or from similar situations in the past. They also can apply theory based on well-accepted economic principles.

We discuss why regulators and industry participants sometimes call for pilot studies that have little scientific value. We discuss some limitations of pilot studies that regulators should consider before proposing them. Finally, we conclude with a set of recommendations concerning the use of pilot studies in financial regulation.
For additional information, contact:

Larry Harris
University of Southern California
L.Harris@USC.edu

Robert McDonald
Northwestern University
r-mcdonald@kellogg.northwestern.edu

Charles Kahn
University of Illinois
cmkahn@illinois.edu

Chester Spatt
Carnegie Mellon University
cspatt@cmu.edu
Evidence-Based Regulation and the Limits of Pilot Studies

Statement of the Financial Economics Roundtable

Evidence-based regulation refers to regulatory decision-making based on accepted theory and analysis of data. If done well, evidence-based regulation can:

- Reduce regulatory uncertainty through the use of reliable knowledge,
- Reduce the potential for unintended consequences, and
- Counter the adoption of regulations serving special interests to the detriment of the public interest.

To enhance evidence-based regulation in the securities markets, financial regulators sometimes undertake “pilot studies” to learn about the consequences of regulatory alternatives. These studies usually involve applying a proposed rule (a treatment) to a group of securities, while maintaining the existing rule for another group (the control). After assigning securities to the two groups, researchers then consider what effects the contemplated rule change would have on the different groups over a study period.

The recent Tick Pilot Study (SEC, 2018) is an example. This study increased the minimum price increment (“the tick”) upon which traders can quote prices on U.S. equity exchanges from one cent to five cents for a set of stocks (the treatment group) for two years. Researchers then studied the effect of this change on the markets by examining differences between variables of interest for the treated five-cent stocks and comparable one-cent stocks. They found that a larger tick size did not lead to greater analyst coverage for the large (five-cent) tick stocks as study proponents had hoped. Instead, the increase in tick size substantially increased investor transaction costs.

Pilot studies are one of three broad methods that can bring evidence to the regulatory process. Regulators also can examine data collected from past events or from contemporaneous markets whose structures are comparable to those being considered. Finally, they may base their regulatory decisions on well-accepted theoretical principles that are useful for understanding the regulatory alternatives.

The members of the Financial Economists Roundtable met in July 2019 to discuss the benefits and potential concerns associated with regulatory pilot studies. This Statement summarizes our conclusions.

FER members strongly support evidence-based regulatory decision-making. Good evidence-based regulation need not require undertaking a pilot study. Such studies have potential to generate new knowledge about financial markets. However, they have drawbacks that can limit their usefulness and contribution to evidence-based regulation. In this statement, we discuss both the benefits and drawbacks of pilot studies, and make recommendations for how they can be used to facilitate evidence-based regulation. Regulators often can effectively evaluate the potential impact of a proposed regulation by analyzing archival data obtained from other markets or from similar situations in the past. They also can apply theory based on well-accepted economic principles.

This Statement identifies how pilot studies used in financial regulation differ from those used in other contexts. We discuss why regulators and industry participants sometimes call for pilot studies that have little scientific value. We discuss some of the limitations of pilot studies that regulators should consider before proposing them. Finally, we conclude with a set of recommendations concerning the use of pilot studies in financial regulation. The Appendix considers additional examples of past and proposed pilot projects.
Pilot Studies in Financial Regulation

Financial regulators use the term “pilot study” to describe a research procedure in which they introduce a proposed regulation to a limited set of securities or institutions and then assess the results with the intent of determining whether the rule change covering the broader market is desirable. In contrast, pilot studies in most other contexts are preliminary studies conducted to determine whether a larger study would be feasible. For example, pharmaceutical companies conduct pilot studies to determine the feasibility of large-scale drug trials and whether they would effectively produce actionable information. This difference in terminology is one source of some misconceptions about the role and usefulness of pilot studies in financial regulation.

Pilot studies are broadly attractive because they are small-scale trials of proposed changes. The operation of the markets during the study potentially produces evidence that can specifically inform their decision-making. Moreover, unlike economic laboratory experiments (frequently run with college students as subjects), pilot studies involve real actors in real environments risking real stakes.

Pilot studies in financial regulation often include random assignment of controls, although such studies may sometimes adopt non-random assignments instead. When using random assignment, the prospect for a pilot is even more attractive, as it resembles a randomized controlled trial (“RCT”), which is commonly regarded as the “gold standard” for experimentation. RCTs are highly regarded because they facilitate causal inference. However, as we point out below, financial regulation pilot studies, even with random assignment, are subject to significant limitations that do not normally arise in randomized trials conducted in fields such as medicine.

Calls for pilot studies generally reference the need to better understand an issue before acting. However, other interests also may motivate those who press for undertaking pilot studies. For example, a regulator already may have a well-developed understanding of the consequences of a particular regulation and its value to the public. But if skepticism is widespread, the regulator may choose to undertake a pilot study to increase public confidence in the regulatory process. Regulators also may initiate pilot studies to benefit special interests that want to promote their favored policies, or to block or defer the adoption of policies they oppose. Finally, individual academics or researchers working for regulators, consulting firms, or think tanks also may propose or support pilot studies to further their careers and secure their employment.

---

1 Pilot studies are part of a continuum in the regulatory arsenal including demonstration projects, beta tests, and phase-ins. Uncertainty about whether the regulation will be adopted distinguishes pilot studies from the other approaches. In circumstances where a regulation has already been adopted, regulators use beta tests and phase-ins to uncover and adjust for any problems in its implementation. Finally, in some circumstances, regulators may use retrospective impact studies to determine the effectiveness of regulations after implementation.

2 For a discussion of pilot studies in medicine, see for example https://nccih.nih.gov/grants/whatnccihfunds/pilot_studies.

3 A related technique that has been recommended for use in evaluating financial regulation is investor testing of disclosure requirements. See Prowar (2018).

4 For a discussion of the advantages and disadvantages of RCTs in medicine and development economics, see Banerjee et al (2016) and Deaton and Cartwright (2018).

5 Legislators frequently call for the initiation of pilot studies. A typical recent example is Congresswoman Carolyn Maloney (D-NY) asking the head of Libra (the Facebook currency initiative) in a House Financial Services Committee hearing in July whether he would commit to a “pilot” given all the complications for the broader society of such a currency (see https://www.washingtonpost.com/business/technology/the-latest-facebook-wont-commit-to-starting-small-on-libra/2019/07/17/2def05fa-a8b3-11e9-8733-48c87235f396_story.html). She prefaced her question by indicating her view that Facebook should just drop the project, arguing that currency is a core governmental function.
Since pilot studies are expensive and subject to misuse, regulators must understand the potential benefits that pilot studies can bring to the regulatory process as well as their limitations relative to other research methods.

**Pilot Studies and Other Sources of Evidence**

Regulators rely on various approaches to gather evidence to inform decision-making. Sources of valuable information include theoretical arguments, archival studies, qualitative studies, laboratory experiments, and field experiments. Each source has its strengths and limitations that depend on the question to be resolved. No clear hierarchy ranks the value of these sources of evidence for all questions. Moreover, for many questions, no single study or approach alone will produce compelling evidence. Regulators must develop a package of evidence, with each source contributing additional information to support the decision. Collecting evidence is costly, so regulators must consider how much and what types of evidence to collect when considering regulatory changes.

The Tick Pilot Study illustrates the potential value of low-cost archival data. Researchers can estimate the effect of an increase in the tick size from one cent to five cents for $50 stocks by examining trade data for similar $10 stocks because a five-cent tick on $50 is economically similar to a one-cent tick on $10. Moreover, the tick size decreases implemented in 1997 and 2001 in the U.S. stock markets produced considerable information about how tick size affects the markets. With this prior evidence, critics of the proposed pilot did not expect that the Tick Pilot Study would contribute much to our understanding of the effects of a change in tick size.

When considering pilot studies, regulators must ensure that the proposed studies have the statistical power necessary to reliably identify the effects that interest them. The study period must be sufficiently long, and other factors must not overly influence the metrics that they hope to examine. A poorly designed pilot study could easily lend credence to results that are misleading or unreliable, and thus lead to poor regulatory decisions.

The statistical power problem is particularly important if the expected effects are material but difficult to measure. For example, a study of whether to allow firms to report their financials semi-annually instead of quarterly would require many years of data to determine how less-frequent financial reporting affects investors, analysts, and issuers. A randomized study of this question would have the further difficulty that information from companies that report quarterly would spill over to the assessment of those that report less frequently, a point we discuss in detail below.

Pilot studies are costly. Planning, implementation, and subsequent analysis are expensive, and the costs that firms and market participants incur can greatly exceed the costs borne by regulators. Moreover, pilot studies that change prices and other incentives potentially distort asset allocations and price informativeness, which can produce economic costs far beyond the markets. Pilot studies also may lead to extended litigation, as participants placed in one or the other samples protest the unfairness of the allocation or aggrieved parties try to protect structures that allow them to profit.

---

6 Harris (1994) used this research strategy to predict the effects of decreasing tick sizes. The actual effects that were observed when the markets moved to smaller tick sizes were close to the predicted effects. Albuquerque et al. (2019) examine possible indirect effects of changes in tick size.

7 The study’s main new result was the identification of an unanticipated effect associated with an issue completely unrelated to the analyst coverage issue: HFTs regularly exploited wide spreads by trading ahead of other traders at “inverted exchanges,” with significant costs to public investors. See Angel et al. (2015).
Pilot Studies in Finance Are Not the Gold Standard

Randomized control trials can produce reliable information about causal relations when they are well designed. For example, medical researchers use double-blind RCTs to test the efficacy and risks of drugs in real people rather than the lab.

Pilot studies often resemble the RCTs that researchers widely use in other contexts to obtain evidence about whether a treatment caused an outcome. For this reason, one might think evidence from pilot studies is superior to that from archival studies. But, even the best-designed pilot studies in financial markets significantly differ from most RCTs used in science in at least three important respects: they are not double-blind, they suffer spillover effects, and they often cannot continue long enough to obtain reliable results.

**Double-Blind**

Trials in financial markets using random assignment cannot be double-blind. In a double-blind design, neither the researchers nor the subjects know which group is the treatment group. This ensures that differences in effects between treated and untreated subjects are due to the treatment, and not due to patients or researchers altering their behavior in response to the subject’s assignment. Double-blind RCTs are sometimes described as the gold standard in drug trials.

In contrast, participants in market experiments always can know the regime in which they are participating, and many may have strong views about the desirable outcome of the experiment and any resulting regulatory decisions. Incentives for gaming the experiment and the potential for collusion among market participants can result in misleading conclusions from randomized trials.

**Spillovers**

Spillovers, which occur when treatments also affect control subjects, generally occur in financial trials. The goal of pilot studies in financial markets is typically to determine global effects if the regulation were fully adopted. However, treated and control subjects all operate in the same markets, so that investors or traders affected by the treatment may behave differently with respect to the controls. If investors regard treated and untreated securities as substitutes or complements, spillovers will occur. The results then may overstate or understate the global effects so that generalizing the pilot study results to a fully treated or fully untreated market may be misleading. In contrast, in most medical studies, the treatment only affects the treated subject.

For example, in the Tick Pilot Study, the increase in transaction costs for five-cent tick stocks made arbitrage more expensive between five-cent tick stocks and all other stocks, including one-cent tick stocks. The increased cost of arbitrage thus likely affected the pricing of correlated one-cent tick stocks. Price efficiency measures separately computed for the treated and control stocks thus both would be affected so that their difference would misestimate the effect of the rule on price efficiency if the rule were implemented for all stocks.

**Study Length**

A third issue for financial market pilot studies concerns the length of time necessary to observe significant treatment effects. Studies must be long enough to capture the effects in question. These effects might not

---

8 Boehmer et al. (2019) discuss spillovers that affected results from the uptick pilot that studied the effects of restrictions on short sales. Al-Ubaydli et al. (2019) emphasize spillovers as one of the threats to the scalability of the results of pilot programs.

9 Herd immunity—the protection that unvaccinated people obtain when many others are vaccinated—is a notable exception in which spillovers can affect even a medical RCT.
emerge quickly if market participants are unaware of the treatment or its implications for their decision-making, or if the response to the treatment has substantial fixed implementation costs. For example, if a five-cent tick does increase incentives for financial analysts to cover a stock, the effect might not have been observed within the two years of the study because analysts may not have recognized or responded to the new incentives.\footnote{Simple economic theory suggests that the hoped-for increase in analyst support would not be forthcoming. The only reasonable mechanism whereby increased analyst support would flow from an increase in the tick size is if dealers expected that they could obtain greater order flows by providing more analyst support. Even assuming that this relation is strong enough to cover the costs of analyses, dealers face the free-rider problem: Any analyses that they provide to their clients, and certainly all that they provide to the entire market, will also accrue to the benefit of other dealers, which decreases their incentives to provide analyses. The tick pilot might have obtained its desired effects had it been applied to dealers who have substantial market power, but most dealers compete in highly competitive dealing markets. Furthermore, the hiring of additional analysts is risky and costly if the pilot program does not become permanent.}

Financial market pilot studies differ from medical RCTs because people might choose not to respond to new rules if they do not believe that they will be permanent, especially if the costs of responding are high. A pilot study, therefore, might fail to identify the full effect of the proposed regulation if it were made permanent.

In summary, pilot studies, even if using random assignment, should not be viewed as the gold standard for collecting evidence about regulatory alternatives in financial markets. Financial market pilot studies face significant problems due to the incentives and opportunities for interaction and influence among the subjects. Potential spillovers ensure that many pilot studies may not be well suited for questions in which the policy is intended to change the overall environment. Given these shortcomings, as well as the high costs to firms and market participants in many cases, regulators always should weigh the relative value of pilot tests against the possibility of obtaining better and cheaper evidence from other sources.

**Recommendations**

Financial regulators should conduct pilot studies when their expected benefit exceeds their cost. Net benefits are likely positive for a potential policy proposal when

- Significant uncertainty exists about the effects of the proposal. Uncertainty arises if the relevant theory is ambiguous, underdeveloped or controversial, or no existing empirical work can credibly identify the effects.
- Regulators can design and implement a pilot study that reduces uncertainty. Spillovers and strategic responses should not have a significant impact on the results, and the results must be relevant to the issues in question.
- The study is not too costly relative to other methods of obtaining evidence. The costs of pilot studies include the writing of the regulatory proposals, monitoring, reporting, the education of market participants, and the costs associated with data collection and analysis. They also include market disruption during the study and the costs borne by firms and investors in dealing with those disruptions.
- The study will be persuasive to a broad range of interested parties. Although no evidence may ever change the views of some interested parties, enough people must be potentially persuaded by the results to make the study worth undertaking.
When writing cost-benefit analyses for the regulatory pilot, regulators already attempt to address many of these costs. Inasmuch as possible, regulators should present the proposed regulatory change and the associated pilot study in the same proposing document with clear discussions about the objectives and the linkages between potential empirical results and regulatory outcomes as well as why other methods are not feasible to study the proposed change. Regulators should articulate in advance their plans for analyzing data from the proposed studies to avoid subsequent cherry-picking of the empirical findings. Finally, regulators should share data with practitioner and academic researchers to ensure that the results are robust to alternative explanations and can be replicated.

**Conclusion**

Good regulation requires good information. When faced with genuine uncertainty about the effectiveness of proposed regulations or genuine divergence of beliefs among informed opinions, regulators must consider using a variety of research tools to gather evidence.

Pilot studies are potentially powerful methods to inform regulatory decision making but are by no means the silver bullet that they are sometimes portrayed to be. While a well-designed pilot study often will prove informative for policymakers, the inherent costs and limitations of pilot studies suggest that regulators should be mindful of the usefulness of other less expensive evidence-based methods to inform their decision making.

Regulators sometimes may run a pilot study as a demonstration project to convince a skeptical public of the soundness of proposed regulatory reform. But the use of pilot studies in regulation is a double-edged sword, as rent-seeking constituencies can use trials as a delaying tactic or to obfuscate an issue. Worse, a poorly conceived study may generate misleading results which could lead to poor regulatory decisions.

Pilot studies are just one tool in the regulator’s research toolbox. We support efforts by financial regulators to consider using them as part of an overall evidence-based strategy. Pilot studies can provide complementary evidence to findings from qualitative studies, empirical archival studies (including natural experiments), and laboratory experiments. All evidence should be interpreted in the light of well-accepted theory.

**References**

Al-Ubaydli, Omar, John List, and Dana Suskind, 2019. “The science of using science: Towards an understanding of the threats to scaling experiments,” NBER working paper 25848


---

11 This discipline is sometimes called the “registry approach.”


APPENDIX

In the body of this Statement we illustrated many of our main points with references to the tick size pilot carried out by the SEC in 2015-2018. In this appendix we briefly examine two other past pilot studies: The Regulation SHO pilot study of the removal of up-tick restrictions on short sales (the “Up-Tick Pilot”), and a phased-in increase in transparency for corporate bond prices, using a treatment group and matched sample (the “TRACE Rollout”). Although both studies produced data upon which valuable results were obtained, both studies were primarily undertaken as demonstration projects that confirmed strong theoretical expectations. We also consider a current proposal to conduct a pilot study of delayed reporting for large corporate bond transactions in TRACE.12

1. Up-Tick Pilot

The Up-Tick Pilot Study commenced on May 2, 2005 as part of the SEC’s implementation of Regulation SHO governing short-selling of securities. This regulation mandated that a sample of stocks consisting of every third stock in the Russell 3000 ordered by volume would not be subject to short-sale price tests for the duration of the program. The New York Stock Exchange up-tick rule required that a short sale had to occur on a price higher than the last different price. This restriction, which was imposed in 1937, prevented short-sellers from arranging trades that would actively push prices down. A similar restriction in NASDAQ prevented short-sellers from selling on the bid.

The pilot study results showed that the repeal of the up-tick rule did not affect volatility or other significant market characteristics for the treated stocks. Short-sale restrictions in conjunction with the one-penny tick imposed only slight restrictions on short-sellers.13 Accordingly, the results of the up-tick pilot were unsurprising. However, the results helped alleviate societal concerns about shorting, and thereby helped facilitate repeal of the short-sale price tests.14

2. TRACE Rollout

On January 23, 2001, the SEC approved proposed NASD rules to collect reports of US OTC bond trades though its TRACE (Trade Reporting and Compliance Engine) system. Dissemination of trade reports occurred in stages, starting with a limited set consisting mostly of large value corporate grade bonds and extending in stratified steps, until by February 1, 2005 all TRACE-eligible securities were included.15 Although the roll-out schedule for TRACE dissemination was not labeled a pilot study, it was constructed to obtain information in stages about whether public dissemination of trade prices for various bond qualities would be beneficial or detrimental to the markets. Trade reports for the largest high-quality bonds were disseminated first because any potential dangers of data dissemination were thought to be smallest for these bonds.

12 Another proposed pilot study on market access fees is currently under litigation. See Osipovich (2019).
13 See Diether et al. (2009).
14 This repeal subsequently became more contentious during the financial crisis. However, a Wall Street Journal editorial (2008) argued that the pilot study made this rule the most economically vetted SEC rule-making in many years.
Several studies examined the data from the first dissemination stages. Goldstein et al. (2007), Edwards et al (2007), and Bessembinder et al (2006) all found that the increased trade transparency decreased customer transaction costs. The decision to report trade prices for all TRACE-eligible bonds was based on these results.

3. TRACE Large Trade Delayed Reporting

OTC bond traders must report their trades to FINRA’s TRACE system within 15 minutes after they are arranged. Presently, TRACE disseminates prices and quantities of all trades as it receives the report. TRACE initially disseminates quantities for large trades—those over $5M in investment grade bonds and those over $1M in speculative grade bonds—only as “5MM+” or “1MM+” with full quantity dissemination occurring months later.

Large bond dealers claim that immediate dissemination of their trades allows others to front run them. In response, the SEC Fixed Income Market Structure Committee (“FIMSAC”) recommended that FINRA delay trade reporting of large bonds (those over $10M in investment grade bonds and those over $5M in speculative grade bonds) by two trading days, after which TRACE would disseminate prices and full quantities. FIMSAC also recommended that TRACE immediately disseminate full sizes for all trades smaller than these thresholds. In other words, transparency would decrease for the largest trades and increase for trades at an intermediate level.

FINRA proposes to conduct a one-year pilot study to identify the effects of the proposed changes. To isolate the effects of the delayed trade reporting from the additional size reporting, FINRA proposed a two-way study of effects which will require the comparison of three sets of treated bonds to each other and to a control sample. However, the initial TRACE rollout studies already identified the value to investors of trade transparency, and even those studies arguably were not necessary because the underlying economic principle that information is valuable is well-tested. Thus, additional insights are unlikely from the proposed pilot study.

---

16 See https://www.finra.org/industry/notices/19-12.
FER Members Signing the Statement, “Evidence-Based Regulation and the Limits of Pilot Studies”

Rashad Abdel-Khalik  
University of Illinois

Reena Aggarwal  
Georgetown University

Edward I. Altman  
New York University

Brad Barber  
UC Davis

Arnoud Boot  
University of Amsterdam

Charles W. Calomiris  
Columbia University

Jennifer Conrad  
University of North Carolina

Elroy Dimson  
Cambridge Judge Business School

Franklin R. Edwards  
Columbia University

Robert A. Eisenbeis  
Cumberland Advisors

Mark J. Flannery  
University of Florida

Lawrence R. Glosten  
Columbia University

Martin J. Gruber  
New York University

Kathleen Hanley  
Lehigh University

Larry Harris  
University of Southern California

Richard J. Herring  
University of Pennsylvania

Charles Kahn  
University of Illinois

Andrew Karolyi  
Cornell University

Jan Pieter Krahnen  
Goethe University Frankfurt

Ravi Jagannathan  
Northwestern University

Kose John  
New York University

Jonathan Karpoff  
University of Washington

Craig Lewis  
Vanderbilt University

Robert Litan  
Brookings Institution

Dennis Logue  
Dartmouth College

Deborah Lucas  
Massachusetts Institute of Technology

Robert McDonald  
Northwestern University

Olivia S. Mitchell  
University of Pennsylvania

Terrance Odean  
UC Berkeley

Frank Partnoy  
UC Berkeley

Stephen Penman  
Columbia University

Rafael Repullo  
CEMFI

Jay Ritter  
University of Florida

Stephen Schaefer  
London Business School

Catherine Schrand  
Ramaswamy University of Pennsylvania

Lemma W. Senbet  
University of Maryland

Erik Sirri  
Babson College

Chester S. Spatt  
Carnegie Mellon University

Laura T. Starks  
University of Texas

Siew Hong Teoh  
UC Irvine

Ingo Walter  
New York University

Ingrid Werner  
The Ohio State University