Field Experiments in Strategy Research

Aaron K. Chatterji
Michael Findley
Nathan M. Jensen
Stephan Meier
Daniel Nielson

I Z A

Field Experiments in Strategy Research

Aaron K. Chatterji
Duke University

Stephan Meier
*Columbia University
and IZA*

Michael Findley
University of Texas at Austin

Daniel Nielson
Brigham Young University

Nathan M. Jensen
Washington University in St. Louis

ABSTRACT

Field Experiments in Strategy Research[*]

Strategy research often aims to empirically establish a causal relationship between an independent variable and a dependent variable such as firm performance. For many important strategy research questions, however, traditional empirical techniques are not sufficient to establish causal effects with high confidence. We propose that field experiments have potential to be used more widely in strategy research, leveraging methodological innovations from other disciplines to address persistent puzzles in the literature. We first review the advantages and disadvantages of using field experiments to answer questions in strategy. We define two types of experiments, "strategy field experiments" and "process field experiments," and present an original example of each variety. The first study explores the liability of foreignness and the second study tests theories regarding corporate culture.

Corresponding author:

Stephan Meier
Graduate School of Business
Columbia University
3022 Broadway
New York, NY 10027
USA
E-mail: sm3087@columbia.edu

---

1. Introduction

Most empirical papers in the field of strategy aim to estimate a causal effect of an independent variable, for example, the number of acquisitions in a given period, on a dependent variable such as firm performance. Assessing the direction of causality correctly is crucial, not just for the advancement of the scholarly field, but for the guidance we offer to practitioners. However, empirically establishing causality is challenging in strategy and many other fields (e.g. Hamilton and Nickerson 2003). For example, some of the most interesting phenomena in strategy research that could be related to firm performance, such as the establishment of a strong corporate culture, the hiring of a new CEO, or the development of a new capability, do not occur randomly across firms. Other important attributes of organizations, such as founding conditions or country of origin, are likely time invariant and correlated with important unobserved factors.

Thus, although strategy scholars have made considerable progress on these topics, it remains difficult to definitively conclude that stronger culture drives greater profits, that insider CEOs do better than outsiders, or that capabilities are the source of heterogeneity in performance across firms. Similarly, while topics like imprinting and the liability of foreignness have been explored in many theoretical and empirical papers, it has been challenging to disentangle founding conditions or country of origin from other important factors. Whereas various econometric techniques, such as employing an instrumental variable (e.g. Bascle 2008), can sometimes be used to address these methodological concerns in strategy research, these hurdles persist. In this paper, we propose that field experiments are a promising methodology to address some of the limitations of prior work and shed light on many of the foundational questions in the field. We view this methodology as a complement, not a substitute, for existing strategy research

methods, since field experiments have strengths and limitations that are often different than our traditional techniques.

Experiments have long been used in disciplines such as social psychology and marketing to establish causality. In an experiment, the researcher randomly assigns research subjects to a treatment and a control group—in the same way that patients are assigned to treatment and placebo groups in randomized controlled trials in medical research. Building on this tradition, field experiments, the application of experimental methods outside a traditional lab, have recently led to a research method revolution in economics, a field that had been historically resistant to experimental methods.[1] This trend is especially prevalent in development economics (for overviews, see Duflo, Glennerster and Kremer (2006) and Banerjee and Duflo (2009)), behavioral economics (for overviews, e.g., Harrison and List (2004) and Levitt and List (2009)), public policy (Ludwig et al. 2011), and to a certain degree in organizational economics (see, Bandiera, Barankay, and Rasul 2011, for an overview). This revolution has not yet reached the field of strategy, however, for reasons we discuss below. Figure 1 shows that although the number of papers published based on "field experiments" or "randomized controlled trials" has exploded in economics journals, no corresponding trend has occurred in strategy and management journals. In Appendix A1 we provide a list of field experiments published and forthcoming at *SMJ* and provide a brief summary of how the results contribute to the field. The small number of field experiments in *SMJ* provides a window into the potential of this methodology in strategy research. Based on these papers, it appears that scholars can employ

---

[1] Psychology, a discipline that has traditionally used lab experiments, is also seeing a surge in field experiments. See Shadish and Cook (2009) for an overview. Similarly, there has been increasing interest in field experiments in marketing research. As one example, *The Journal of Marketing Research* issued a Call for Papers for a *Special Issue on Field Experiments* in Spring 2014.

field experiments to study topics that have been less explored due to lack of data. Further, since scholars can carefully control the research design, the method is amenable to testing competing theories, another important priority for the field. In sum, we view these four papers as promising examples of the use of field experiments to both adjudicate existing debates using new data and to test new theory.

We contend that strategy scholars are especially well-placed to utilize field experiments, since many important research questions are unaddressed by other fields using this method.[2] In particular, strategy research places significant emphasis, especially relative to economics, in explaining firm heterogeneity. For example, many questions of interest to strategy researchers but traditionally less so for organizational economists, are related to firm capabilities (Teece et al. 1997; Helfat et al. 2009). As we describe below, this gap provides an important opportunity for strategy researchers to formulate and test original theory using field experiments. However, the field experiment methodology has specific advantages and disadvantages for strategy research in particular, which we also review below.

To guide future scholarship, we define two general types of experiments, "strategy field experiments" and "process field experiments," to explicate the different approaches strategy scholars can employ to address key questions in the field. We conclude that although one might assert at first glance that field experiments have limited utility for strategy research, the opportunities to apply these methods are widespread.

Next, we detail two original field experiments, the first a 'strategy field experiment' and the second a 'process field experiment,' that illustrate the great potential but also the drawbacks

---

[2] We will focus on field experiments and not experiments conducted in the laboratory. Croson et al. (2007) argues that these laboratory experiments should also be part of the toolkit of strategy researchers. See Harrison and List (2004) for a taxonomy of the various kinds of field experiments.

of using field experiments for strategy research. Lastly, we reflect on the path forward for field experiments in strategy research and discuss what kind of research agenda could be formulated to guide future scholarship.

2. Methodological innovation: Answering "big" questions with "small" experiments

Advantages of field experiments in strategy research

Field experiments as a method allow the design and implementation of creative treatments to answer relevant questions that are otherwise very hard to address. While the ideal context to study the effect of interest might not be available in observational data, researchers using field experiments can create their own exogenous variation to cleanly identity causal relationships. This feature allows researchers to do question-driven research as opposed to being constrained by existing data, an important aspiration for the field. As one example, there might be little within-firm variation in national origin in the typical observational dataset, but in our first experiment we can manipulate perceptions of the audience about the origins of a company to test theory about the liability of foreignness.

Second, field experiments are ideally suited to assess specific processes and activities inside firms, where our traditional data sources are unable to provide much detail. As we discuss below, the outcome variables of interest in these field experiments need not be firm performance, but could be antecedents based on prior theory. Using this approach, the results of field experiments have great potential to reveal actionable insights for managers seeking to improve firm performance through the development of capabilities or, as we discuss in our second experiment, establishing a particular corporate culture.

More broadly, experiments make it possible to vary one factor at a time and therefore provide "internally" valid estimates. In the spirit of Shadish, Cook and Campbell (2002), field

experiments allow the researcher to be more confident that any difference between the treatment and control means is due to the intervention. This feature creates the ability to use multiple conditions to disentangle forces that covary in the real world. Importantly however, this method does not automatically confer external validity (Lynch, 1982, 1983), which we discuss in detail below.

Challenges for field experiments in strategy research

Despite these positive attributes, there are several drawbacks to the use of field experiments.

External validity. First, since this methodology promises strong internal validity through the random assignment of treatment and control conditions, it might be intuitive to think that there are advantages related to external validity as well. However, Lynch (1982) points out that no research methodology in and of itself can generate externally valid results. External validity can be thought of as whether the results from the experiment can be generalized to some population/context and/or can be generalized across populations/contexts (Cook and Campbell, 1979). Most importantly, "generalizability can only be judged on the basis of replication across settings, subjects, stimuli, and responses." (Dipboye, 1990:26).

One of the fundamental challenges that prevents generalization across settings from a field experiment is that various background factors can substantially interact with the treatment, unbeknownst to the researcher (Lynch 1983). This kind of oversight can create challenges to external validity via two channels. First, researchers might not test for the background factor and treatment interactions, committing an aggregation fallacy because of unobserved heterogeneity. As a result, researchers report the main effect of the treatment but the "simple effects" vary according to unobserved background factors (Hutchinson, Kamakura, and Lynch, 2000). In the

second case, researchers may hold background factors fixed when they actually interact with the treatment, leading to treatment effects that would be different if the background variables were held constant at another level (Lynch 1982).

While it is impossible to generate perfect external validity from field experiments, there are a few strategies to mitigate these concerns. First, the treatment should be replicated in different contexts. Second, process field experiments (see below) that test the underlying mechanisms more directly should be conducted, since insights about processes advance our understanding of firms more generally. Indeed, while typical field experiments excel at generating causal effects in which researchers can place high confidence, they are limited in their ability to indicate the causal mechanisms that link interventions to outcomes. Process field experiments, by randomly assigning the mechanisms themselves, provide a possible way forward (see also Ludwig et al. (2011) for detailed discussion with regards to policy field experiments). Third, theoretical development can help anticipate the specific ways that an observed treatment effect might change if extrapolated to a specific, but different situation (see Lynch 1999).[3]

Feasibility. There are also additional challenges to field experiments for strategy scholars in particular. Strategy scholars have traditionally sought to explain firm performance by variation in industry structure and firm capabilities. Very few field experiments could ever be large enough or sustained for long enough to generate appreciable differences in firm performance. Next, it is quite difficult to manipulate attributes of an entire industry or firm positions experimentally. Finally, while working inside a single firm may provide a promising vantage

---

[3] Closely related to the problem of external validity is a concern expressed by Heckman (1992) that the organizations that agree to participate in an experiment may be different from those that do not (see Allcott and Mullainathan (2012) for an example). While there are some limited strategies to mitigate this challenge, it is still difficult for researchers to assess the extent of this bias in interpreting the results of their experiments.

point to observe firm capabilities, prior work has found that competitive advantage relies on bundles of resources and capabilities that are endogenously linked to one another and possibly impossible to separate even in a well-designed study (e.g. Siggelkow, 2001).

As a result, many hypothetical research designs to answer important questions in the field are clearly impractical. For example, randomly assigning firms to different competitive positions in the market and then measuring their return on invested capital (ROIC) five years later is almost impossible. And randomly changing the culture of half of the business units in a firm and then measuring their subsequent performance compared to the other half of divisions where the culture was presumably constant is also likely not feasible.[4] While there are unique opportunities to manipulate core attributes of firms experimentally (as we do in our first experiment), such instances may be rare. Further, as field experiments in strategy are usually conducted inside firms, they can be difficult to implement since enlisting employees to help manage the experiment and securing top management buy-in are not always easy.

The examples above constitute a reasonable critique for why field experiments cannot be applied universally in strategy research. However, although running experiments to directly estimate the effect of certain strategic choices on firm performance may be too impractical, running experiments that teach us about the activities that conceptually underlie how those strategic choices relate to firm performance is far more feasible. In other words, methodological innovations are unlikely to occur that will allow us to randomly assign corporate culture to firms and observe their subsequent performance. However, carefully designed experiments could manipulate one of the key processes that underpin culture and measure its effect on a process

---

[4] More generally, if individual-level randomization is not feasible and randomization has to be done across larger units, like teams or business units, statistical power (or the lack thereof) might be a limitation.

outcome that affects performance. This approach can help us answer the "big" questions in strategy with "small" experiments (as has been argued in development economics, e.g. Banerjee and Duflo (2009)).

We believe that field experiments can make headway in strategy in a similar spirit, through targeting processes identified by theory and assessing shorter-term outcomes we believe matter for firm performance and other longer-term outcomes of interest. To make this approach explicit, we categorize experiments into two broad categories (inspired by a discussion about field experiments in public policy in Ludwig et al. (2011)): "strategy field experiments" and "process field experiments".

Two types of field experiments in strategy research

The two types of field experiments shed light on a strategic question using different approaches. In general, managers are interested in understanding how a choice X, i.e. the effect of span of control (e.g. Rajan and Wulf 2006), affects an outcome Y, i.e. firm performance. However, X affects Y through a mechanism M, i.e. span of control affects communication channels between managers and subordinates (Bloom et al. 2009). The two types of experiments either study the effect of X on Y directly or focus on the mechanism M in order to learn about X $\rightarrow Y$.[5]

Strategy field experiments: We call experiments "strategy field experiments" if they test the effect of significant strategic choices, X, on firm performance, Y. For example, in studying the effect of span of control, a strategy field experiment would reduce the span of

---

[5] The two types of field experiments are not always mutually exclusive. It is possible to design a field experiment using different treatment arms that can test all aspects of the chain $X \rightarrow M \rightarrow Y$. For an interesting example on pricing, see Karlan and Zinman (2009). However, we believe that the distinction we draw between different kinds of experiments is still useful in determining appropriate research designs and choosing which questions to address.

control for a given firm's middle managers in a randomly assigned subset of divisions. Then the researcher would compare the performance of treated and control business units. The advantages of a strategy field experiment are twofold: First, it directly tests an important strategic choice and measures the effect on a key performance metric. Second, because it tests a strategic choice in its entirety, this research design can take into account the *total effect* of the strategic choice. For example, reducing span of control might not only affect the manager directly, but also influence the selection of potential managers, allowing the firm to hire higher-quality managers. A strategy field experiment can capture such broader effects.

However, strategy field experiments have at least two major drawbacks: First, they are often difficult to conduct. Second, strategy field experiments cannot typically provide insights into the mechanism for why a certain strategic choice has an effect on performance. There is a complicated intervention *X*, producing a treatment effect on *Y*, but the researcher cannot always identify the "active ingredient." As such, any results would be especially vulnerable to criticisms about generalizability or external validity discussed above.[6] Our perspective is that there are some contexts where the merits of strategy field experiments will outweigh the drawbacks. In other cases, process field experiments (discussed below) and other methodologies such as lab experiments, regression analysis and qualitative work can work in concert to shed light on important research questions.

We see our first field experiment as an illustration of a 'strategy field experiment', which manipulates a fundamental characteristic of the firm. We manipulate the implied country of origin of a confederate company and test the effects on investment incentives offered by US cities as a test of the literature on the liability of foreignness. While we are not directly testing a

---

[6] For a similar critique about randomized controlled trials in development economics, see, Deaton (2010).

firm's market entry strategy, we are testing for differences in the cost of a foreign firm's choice to enter via direct investment relative to that of a domestic firm. While a strategy field experiment is especially appealing for this research question, our second study highlights a line of inquiry where process field experiments are more appropriate.

Process field experiments: "Process field experiments" do not test a strategic choice or firm attributes directly, but rather test the activity or mechanism that is the theoretical underpinning of the causal relationship of interest. For example, in the case of span of control, before designing a process field experiment, the researcher would need to reflect on prior theory for why reducing span of control would affect performance. Perhaps a smaller span of control facilitates better communication across hierarchy levels. A researcher interested in this domain could randomize access to a particular communication channel used by managers and subordinates and measure the impact on employee productivity in tackling a short-term task such as the completion of a single project. This approach would be a "small" but tractable experiment that relies on the theoretical connection between communication and span of control and the correlation between performance on a particular project and long-term performance.[7]

The clear disadvantage of process field experiments is that they only test the effect of the strategic lever indirectly and, as such, might miss more general effects of the strategic choice. However, process field experiments have significant advantages as well. First, process field experiments are generally more feasible to implement due to their smaller scale. Moreover, as Ludwig et al., (2011) point out, even small experiments on testing mechanisms can inform larger policy or strategic choices. Second, they provide evidence on the process underlying why a

---

[7] Process field experiments should not be confused with 'process evaluations'. Process evaluations are used in field experiments to study the degree to which the intervention was implemented correctly and can thus help in the interpretation of the results (for an example, see Oakley et al., 2006).

certain strategic choice is effective, which can enhance opportunities to apply the lessons learned into different contexts and inform theory. To discriminate between different theories, it is often not enough to know whether a certain strategy works, but to understand why it worked. Understanding the process more deeply will also improve our advice for managers, who may devise better ways to affect the outcome of interest (i.e. working on opening communication channels rather than simply changing the span of control in the example above). Our categorization has clear connections to the concept of mediating variables, where the relationship between two variables X and Y might be mediated by M (see Zhao et al. (2010)). Process field experiments are ideal when the relationship of interest is M $\rightarrow$ Y. In such a case, manipulating M directly might be preferable to exploring just X $\rightarrow$ Y (see Spencer et al. (2005); Bullock et al. (2010); and Green et al. (2010)).

We see our second field experiment as an illustration of a 'process field experiment'. We cannot manipulate corporate culture in a large sample of firms and explore performance patterns over time. Instead, we test an important link in the causal chain between identity and cooperation that speaks to the broader importance of culture in organizations. To extrapolate from this style of experiment, one needs to tie variation in employee identity to culture and cooperative behavior among employees to firm performance. As discussed below, these two sets of linkages are well established in the literature.

Our distinction between strategy and process field experiments is similar in spirit to Calder et al.'s (1981) distinction in consumer behavior research. On one hand, they present "effects application" as a test of a general theory, similar to our strategy field experiments. Alternatively, they argue that "theory application," similar to our process field experiments, is grounded in different philosophical foundations and presents a different set of methodological

challenges. They advise that researchers clearly identify the primary goal of a given study and design it accordingly. Following Calder et al.'s (1981) logic, we recommend that strategy field experiments prioritize realism, while process field experiments evaluate mechanisms with the cleanest possible test. As Ludwig et al. (2011) argue, "mechanism" experiments can still be very valuable even if they do not mirror policy interventions that would actually be implemented.

In the next section, we describe both of our field experiments. Due to space constraints, we only briefly review the prior literature on these well-studied topics and place many of the details about how we designed and implemented these two studies in an Online Appendix, which also includes suggestions for implementing experiments and managing ethical concerns.

3. Field experiment 1: Liability of foreignness and investment promotion

A core argument in international business scholarship is that firms are often confronted with substantial barriers to operating abroad, and these barriers shape a firm's strategic choices on when, how, and if to enter a foreign market (Hymer 1976). This "liability of foreignness" is a broad definition that encompasses a number of factors that can disadvantage firms operating in overseas markets, ranging from geographic distance, to cultural distance, to economic nationalism (Zaheer 1995).[8] Some evidence suggests that this liability of foreignness has been shown to have a substantial impact on the performance of foreign firms operating in host countries (Zaheer 1995; Zaheer 2002 and Zaheer and Mosakowski 1997).[9] This topic has important implications for the field of strategy. One of the most cited works on the topic notes in the first paragraph that, "This liability of foreignness has been the fundamental assumption driving theories of the multinational enterprise…" (Zaheer 1995: 341).

---

[8] A related literature is the work on home-bias in financial asset markets (established by French and Poterba 1991).
[9] See Miller and Eden (2006) for a discussion on this topic.

Prior work has built an empirical foundation for the plausibility of the liability of foreignness, yet these studies employ indirect tests using observational data. In these instances, unobserved variables may confound results, self-selection may introduce bias, and thus causality is difficult to demonstrate. For example, two different studies published in *SMJ* on the liability of foreignness in the financial services industry found different results with regard to domestic and foreign bank performance (Miller and Parkhe 2002; Nachum 2003). These differences could arise from using indirect measures of the liability of foreignness. Other studies have taken a more direct approach to measuring the liability of foreignness, but have not settled the debate on the underlying mechanisms driving patterns in the data. For example, Mezias's (2002) article tested the liability of foreignness using data on labor lawsuit judgments in the United States, finding that foreign firms were more likely to be the subject of labor lawsuits than a matched set of domestic firms.

Moreover, empirical studies of the liability of foreignness face problems of endogeneity, where the firms that are most likely to overcome the liability of foreignness are the firms we are most likely to study and compare against domestic firms. For this reason, it is quite possible that endogeneity works against finding liability of foreignness, for example if firms that can mitigate this liability are more likely to be active abroad. For this and other reasons, it is very difficult to estimate the exact magnitude of liability of foreignness based on observational studies. We can overcome this issue and related challenges by focusing on the narrower question and using a field experiment.

In this field experiment we examine how the firm's country of origin affects whether government officials are willing to provide firm-specific incentives to a potential investor.[10] The use of business incentives, ranging from grants, tax holidays, or low-cost loans, is quite common across countries. In the United States, states and municipalities offer these incentives to firms, often in bidding wars to attract new investment. We explore how country of origin affects a firm's propensity to receive these incentives from government officials.

Sample

Our subjects of study are the 3,117 United States municipalities with populations above 10,000 that list a mayor, city manager, or other executive online. We refer to these municipalities, ranging from small towns and counties to large cities, as "cities" in this paper. We focus on municipalities because U.S. cities and states are increasingly active in the promotion of investment, providing as much as $80 billion per year in financial incentives to firms.[11] There are also enough US cities to ensure a sufficiently large sample size.

Experimental Protocol

Our first step was to identify a "confederate" firm that was considering an investment in a new manufacturing facility. The firm was willing to provide concrete details on its future investment, including projected numbers on job creation and capital investment based on the operations of two existing plants. We signed a confidentiality agreement with the firm, assuring that the name of the company would not be used in the experiment. In return for collaborating as

---

[10] While perhaps counterintuitive, there are examples where firms strategically choose to change their home market, and thus their "foreignness." Lenovo's acquisition of IBM's PC division or the recent spate of "tax inversions" can be viewed this way, though other less dramatic corporate decisions are also germane to our study. The narrowest interpretation of our experiment is a relatively clean estimate of how much national origin matters in a particularly significant corporate decision, investing in a new facility.

[11] Please see "The United States of Subsidies." http://www.nytimes.com/interactive/2012/12/01/us/government-incentives.html (Last accessed on June 23rd, 2014)

a confederate in the study, the firm was offered our analysis of incentive pledges. In Appendix A2 we briefly outline practical challenges as well as ethical considerations in this field experiment.

Next, we legally incorporated a consulting company that mimicked existing U.S.-based investment-promotion and incentive-management companies that often act as brokers between city economic development officials and firms exploring investments. These companies are generally small operations and often do not publish their client lists on their websites. We incorporated our company, Globeus Consulting, as an LLC in Delaware in 2013. We created a company website and a board of consultants (all academics willing to lend their names for the purposes of the experiment) and listed a company president. Three research assistants, using their actual names, served as "Associates" that directly contacted cities through email addresses registered through our website.[12] Our consulting company was built for the purpose of academic research and thus the authors received no compensation nor were the authors directly engaged in any further negotiations with the investor and city governments.[13] We installed web-tracking software that aided the analysis of our experiment.[14]

---

[12] We were careful to minimize any personal interaction between our RAs and city representatives. Our interactions were largely limited to email communications with a clear protocol. Individual inquires were handed by our lead research assistant or Professor AUTHOR. We highlighted that our confidentiality agreement between our consulting firm and client barred us from providing any additional information on the potential investor.
[13] We purposely designed our experiment to minimize the amount of time public officials would dedicate to our inquiry. While a back and forth negotiation would increase the realism of our experiment, we believe that this would lead to a more extensive time investment by public officials.
[14] We included two forms of tracking. For each experimental condition we provided a unique link that routed the respondent to our website. Thus while all respondents viewed the same websites, we could track subjects from each group based on the routing link. Second, our website allows for IP tracking, enabling us to both track unique users and the location of the visitor. This allows us to compare the web activity (pages viewed, time on the site, etc.) by treatment group and individual.

Next, we built a data set of the mayor/executive, and any listed economic development official for all of the 3,117 municipalities.[15] Then we emailed all of the cities with details on the proposed investment, which would include $2 million in capital investment and 19 full-time jobs. These numbers came directly from our client and are near the lower bound of the type of investments that would be of interest to most cities. It is thus possible that our study is under-reporting the general amount of interest in attracting investment.

In our email (see Appendix A3.1) we directed the client to a web form (via Qualtrics) to fill out specific details on the types of incentives offered. This impersonal interaction was designed to ensure that all cities were treated with the same message and to help limit the amount of time city officials dedicated to our inquiry. We include all of the Qualtrics questions in Appendix A3.2. The main question we examine in this paper is whether or not a city offers grant dollars or loans to investors, although we were careful to highlight that these were non-binding offers. Our approach contained two types of experimental interventions, only one of which we will present in detail in this article.[16]

While we were not authorized by our client to change any of the details of the investment, we were authorized by our client to vary "the pitch" of our consulting company. We did this in two ways. First, we began each email with the following paragraph randomizing across the three conditions (treatment conditions highlighted below but not in the original email):

---

[15] Our universe of cities is all cities with a population of 10,000 listed by ICMA. As part of our data collection efforts we also coded the date of the mayoral/council elections. Due to the lack of information on the web we directly called a number of municipalities.

[16] In our first treatment, we varied the timing of when the investor was willing to announce their intent to invest, randomizing between 2 months before or 1 month after the mayor or city councilors' next reelection month (without mentioning the reelection). Note that our full-factorial randomization strategy ensures that our first treatment will not produce any bias in our analysis of our liability of foreignness treatment because the two treatments were randomly assigned independently of one another.

I am an associate with GLOBEUS Consulting (see our website here [insert hyperlink]). GLOBEUS is a new consulting firm that specializes in matching cities with prospective firms. I work in the GLOBEUS group focusing on investors based in **[TREATMENT 1: the United States / TREATMENT 2: Japan / TREATMENT 3: China]** and am contacting you to see if your city would be a good match for a client I am representing.

Second, we signed each email with the name of one of our three associates (research assistants) with a group email address and country team. Below is the email signature.

Associate Name
[us / japan / china]_client_team@globeusconsulting.com
Selection & Incentives Associate Globeus Consulting—**[U.S. / Japan / China]** Client Team www.globeusconsulting.com

We specifically focus on differences between U.S., Japan, and China due to previous work documenting public perceptions of Japan and China. In a public opinion study by the Pew foundation, only 42% of Americans had a favorable view of China as compared to a 77% favorability of Japan (Pew 2007). AUTHOR (2013) find that the country of origin has a major impact on FDI preferences. Using survey experiments in the United States and the United Kingdom they found very little difference in support for Japanese investment relative to the control condition of "foreign investment" without a country name. In contrast they found considerable skepticism towards Chinese investment.

With this background in mind, we first test whether the city respondents will favor U.S. firms over foreign firms in general and then test whether city respondents are less likely to respond to Chinese investors compared to Japanese investors.[17] Prior to random assignment we stratified the subject pool on a set of covariates that might affect the outcome and thus implemented the experiment using block randomization on the criteria of state, population size

---

[17] The term "investor" is widely used in this context to refer to the management team of the operating company seeking to make an investment. While this word could potentially be misinterpreted by some participants, we believe that this term of art is well-understood by city officials involved in economic development projects.

(above or below median), government type (council-manager or other), elected vs. appointed executives, and the quarter of the year for the next election. Block randomization ensures that the blocking covariates are balanced across experimental conditions and provides a stronger methodological basis for simple difference-in-means analysis.[18] Also prior to fielding our experiment, we pre-registered both our research design and analysis plan as a means of pre-committing the study to experimental conditions and analysis strategies. This step is used to limit the temptation to "fish" for statistical associations after the fact.

Results

For the purpose of this study we focus on three main indicators: (1) response rates to our inquiry, (2) cities' offering some form of financial incentive, and (3) the natural log of the mean dollars offered in grants per new job for this investment.[19] We supplement the analysis of the main outcomes with a description of the website traffic by treatment group (see Appendix A3.3).

[Insert Table 1 here]

In Table 1 Panel A, we document the response rate to our inquiries on behalf of our client. Our response rates range from 14.7% (Japan Treatment) to 15.8% (China) to 17.6% (US treatment). These relatively small differences in response rates are significant at the 90% level for U.S. vs. foreign (Japan and China pooled) and U.S. vs. Japan. This provides modest evidence that cities tend to respond more to inquiries on behalf of U.S. firms over foreign firms in general and Japanese firms (though not Chinese firms) in particular. This rather weak evidence of treatment effects may undermine the conventional wisdom that foreign firms are treated

---

[18] As a robustness check, blocking covariates should also be used in regression analysis to control for the stratification variables alongside the treatment conditions. In this experiment, regression analysis with the blocking covariates as controls produces substantively similar findings to those reported in the difference-in-means analysis.
[19] We phrased this per-job since many incentives programs allocate dollars on a per-job basis. We included the total number of jobs in the email proposal (25).

differently from domestic firms in the allocation of discretionary benefits. The liability of foreignness argument would expect that cities would respond much more frequently to U.S. firms, especially over Chinese companies. This was not the case in our study. In the case of response rates to U.S. vs. Chinese firms, our results suggest that any differences that might be detected with greater statistical power would be modest.

More telling still was the lack of differences in the number of respondents offering incentives to our client firm, as shown in Table 1 Panel B. The client firm in the U.S. condition received incentive offers from 9.84 percent of cities. The firm in the Japan condition received offers from a nearly identical 9.77 percent. The firm in the China condition received a slightly lower proportion of offers at 8.30 percent, but this difference was not significant statistically. The relatively tight 95% confidence intervals around the difference in treatment means suggests that any statistical differences undetected by this test are likely quite small substantively, especially in the U.S. vs. Japan comparison. The null hypothesis that there is zero difference between incentives for U.S. vs foreign firms cannot be rejected at any conventional level of confidence. Likewise, there is no statistically significant difference between incentive offers for the firm in the Japan vs. China conditions, providing no meaningful evidence for our second prediction. These results are robust to regression analysis with the blocking criteria employed as control variables. See Table A3.4.1 in the Online Appendix.

This pattern persists when we examine the size of the incentive offers measured in the natural log of grant dollars per job. We employed the natural log because the dollar values of the offers are highly skewed, but the results are substantively similar if unlogged dollar amounts are employed. As shown in Table 1 Panel C, the logged dollar values of offered grants are very closely matched across conditions when we code non-responses as zero, again providing no

significant evidence of discrimination against foreign firms. In Table 1 Panel D, we dropped all non-responses. We find a similar pattern.

Our empirical analysis thus far provides modest evidence for treatment effects on response rates but no significant evidence for treatment effects on actual incentive offers. We find no significant differences between the offers made to the U.S. firm compared to the putative Japanese or Chinese firm.[20][21][22]

In summary, our results provide limited evidence of discrimination against foreign firms in the allocation of incentives. City officials were just as likely to respond to emails when they believed they were interacting with Japanese and Chinese investors, though there was some modest evidence that U.S. firms were more likely to receive a response. However, this slight response-rate bias did not extend to actual incentives offers. Rather, the rate of offers to the U.S.-based client was statistically indistinguishable from rates for the Japan- and China-based clients.[23]

We suggest caution in interpreting our results, given that our field experiment was designed to precisely test just one mechanism underpinning the liability of foreignness. As a result, it is not straightforward to extrapolate this result to other theorized and practical barriers to operating abroad, including cultural distance, linguistic challenges and other issues. For

---

[20] All of the results reported above are substantively similar when robustness is checked using regression analysis and covariates.

[21] As a supplemental analysis we provide documentation of activity on our website, from August 1, 2013 (the start of our experiment) to August 29, 2013. We installed tracking software on our website that allows for the tracing of location and individual IP addresses. More important for this project, in emails to the different groups we provided a different routing address that all forwarded to our main website. Our tracking software thus allows us to examine the activities on our website by treatment group (U.S. vs. Japan vs. China). We document our results in Appendix A3.3, but note here that similar to our findings above, we find very weak evidence for the liability of foreignness.

[22] We also tested city responses, incentive offers, and log of dollars using logit regressions in Appendix A3.4. We include a model with our blocking covariates as control variables and a model that interacts our treatments with each blocking covariate. Our results remain unchanged.

[23] The largest contrast between the treatment of US and foreign firms were found in the interest in our "US Clients" versus "International Clients" tabs on our webpage.

ethical reasons, we only made initial contact with the cities, and it is plausible that further

negotiations between a given city and investors could have different potential outcomes based on

the country of origin. However, our design was informed by discussions with economic

development professionals to closely mimic this common interaction between firms and

governments. In sum, we thus see our study as a complement to existing work on the liability of

foreignness. Our study has the advantage of using the narrow lens of experimentation to

compare how identical proposals are treated prior to investment. But the clear limitation is that

field experiment methodology does limit what we can claim about how cities would react to

different types of investments or how foreign firms are treated relative to domestic firms after the

investment has been made.

4. Field experiment 2: Culture, identity and prosocial behavior

Our second field experiment can be categorized as a process field experiment and

explores the canonical question of whether organizational culture influences firm performance.

Prior strategy literature has argued and shown empirically that a key driver of firm performance

is organizational culture (Barney 1986; Sørensen, 2002), where strong organizational cultures

are often described as having widely shared norms and values that motivate employees (OReilly

and Chatman 1996; Sørensen 2002; Goldberg 2011; Gibbons and Henderson 2012). Using

traditional methods, it is challenging to identify a causal effect, because, for example, strong

performance could in turn drive widespread agreement on norms and values instead of the other

way around. Resolving this empirical issue is not only difficult, but also has important lessons

for practice.

In our parlance, a process field experiment aimed at this question would take the

following approach. Such an experiment tests the process underlying why we think culture has

an effect on performance. For example, prior work has found that individuals strive to be connected to one another, organizations, or ideas in an effort to construct identity (Ashforth and Mael, 1989; Shih, Pittinsky and Ambady 1999; Chen et al. 2010). Strong cultures are built by articulating norms and values with which employees identify. Not only can stronger identity influence individual performance (e.g Shih et al. 1999), but it can also facilitate cooperation and has been posited to be a crucial factor in explaining firm performance (Kogut and Zander 1996). Thus one potential process underlying how culture affects firm performance is through the influence of employee identity on the likelihood of cooperation. A process field experiment would aim to answer the broader question of how culture influences performance by manipulating employee identity and exploring the impact on cooperation on a particular task (see Cable et al. 2013 for another field experiment investigating employee identity and behavior). This is the approach we will use in our experiment.

In addition to investigating the effect of firm identity, we will also analyze which identity (firm or business unit) is more important.[24] A trade-off could exist in strengthening identities at different organizational levels. Identity at the lowest level of the organization could have the most significant impact on prosocial behavior and cooperation between members of that small group (e.g. Bohnet and Frey 1999; Goette et al. 2012). However, such an intervention might reduce cross-group cooperation, which could harm the firm as a whole. Competition between business-units could potentially pronounce the possible negative effect of strengthening business-unit identity on cross-business-unit cooperation. We consider these various trade-offs explicitly in our design.

---

[24] Prior research has shown that individuals have different identities. For example, one individual could identify simultaneously as a female, an Italian, a researcher, a strategist, and so forth. A professor at a business school might identify most strongly with the university, the business school, or her own department.

Experimental Design

The field site for this experiment is a Fortune 500 medical device company with more than 45,000 employees dispersed across the world. Through a former student, we established a relationship with the company's "innovation" team.[25] One of the key responsibilities of this team is to manage the company's internal knowledge-sharing platform. This platform allows employees to participate in "Facebook-like" social networks across the firm and ask and answer questions of their fellow employees in a similar way to the popular site Yahoo! Answers.[26] Corporate management and the "innovation team" were very interested in increasing engagement on this platform to foster increased collaboration and strengthen corporate culture more generally. In our study we focus on 4,185 engineers at the firm across 6 business units. Our experiment involved manipulating employee identity using three different kinds of email prompts and observing subsequent cooperative behavior using data from the internal ideation platform.

We measure cooperation in the context of knowledge transfer. Like many companies, the firm uses their intranet platform to facilitate knowledge transfer across employees and encourage novel ideas by "breaking down silos" within and across business units and geographies. Employees can ask "the company" questions related to work. These questions range from requests to help solve a particular technical problem to queries about possible interactions between various materials to broader and more abstract scientific and technical concerns. The company sends out a bi-monthly email from the same email account to all employees with a list of featured questions selected by the innovation team after culling through a larger list provided

---

[25] To protect the anonymity of the firm, we will sometimes use more general descriptions of business units, products, and corporate attributes.
[26] See answers.yahoo.com as an example.

by employees across the firm. All employees receive the same featured questions. We measure whether employees open this email, whether they click on a link in the email, and whether they answer any of the questions. This approach provides us with an objective measure of whether the employee engages in cooperative behavior within the firm. We can observe 3 mailings (1.5 months) before and 3 mailings (1.5 months) after our 2 intervention mailings (1 month).

We opted for a subtle intervention that only marginally manipulated the text of the emails the innovation team was sending out.[27] In doing so, we rely on standard research methods in social psychology to manipulate identity. In the email that went out to employees of our firm, we changed the text in ways to make particular identities salient. We compare this firm-identity treatment to a control in which the identity of the firm is not directly made salient.

To experimentally test the effect of identity on different organizational levels, we not only made the firm identity salient, but also implemented three treatments (in addition to having a control group ("T0: Control Group"): (1) making firm identity salient ("T1: Firm Identity"), (2) making business unit (BU) identity salient (T2: BU Identity"), and (3) making business-unit identity salient plus introducing competition between the business units ("T3: BU Competition"). Please see Appendix A4 for the text of all the treatments, summary statistics and confirmation that our randomization (we stratified according to business unit) succeeded in creating groups with similar observable characteristics.

Comparing the different treatments allows us to test 1) whether making identity salient (at either the firm or the BU level) increases cooperation (comparing T0 to either T1, T2, or T3);

---

[27] It is important to think about how subtle a given treatment intervention will be. The trade-off in this decision is that if the treatment is not very subtle or even implausible, subjects in the experiment may become aware of the intervention, and their behavioral reactions might be influenced, comprising internal validity. One advantage of this specific subtle intervention was that it was extremely cheap to execute, which would positively influence any cost-benefit analysis.

2) whether firm or BU identity is more powerful in increasing cooperation (comparing T1 to T2);

3) whether competition between BU increases cooperation even further and/or has detrimental

effects (comparing T2 to T3).

We will interpret differences between the treatments as being driven by identity.[28]

However, as with all priming experiments, it is not entirely clear whether the prime actually

affects identity or not. Consider our comparison of T0 to T1/T2: In this instance, our intervention

could make the identity of the firm or the BU salient or alternatively highlight the culture of

prosocial behavior in general, as T1/T2 contain prosocial language more explicitly than T0. To

partially remedy this concern, we compare T1 and T2, where the only difference is the mention

of the firm or the business unit and all of the other language remains constant.

In addition to testing whether the treatments have a main effect on cooperation, we also

want to test whether heterogeneous treatment effects occur, that is, whether certain groups react

differently to the treatment (see Appendix A2 for a discussion about the pitfalls of conducting

subgroup analyses). In our field experiment, we will test whether employees who are actively

engaged with the firm prior to the experiment are affected by the treatments differently than

employees who are not engaged. Increasing the engagement of employees who were not as

active in providing public goods within the organization was a key objective of our internal team.

Interestingly, this is an important consideration for future academic work on this topic since it

could be that typical initiatives to strengthen corporate culture only impact a subset of

employees.

---

[28] The different messages also differ in terms of length. While T1 and T2 have the same length, the control email is shorter and T3 is longer. It is possible that message length affects behavior independent of the content. If so, comparing T1 vs. T2 should help mitigate this concern.

We use prior activity on the internal platform as an objective measure of engagement by creating an "Engagement Index." The index takes values from 0, "not engaged," to 4, "most engaged."[29] A second possible method for thinking about engagement is the distance from headquarters (HQ). The medical device company has multiple business units spread out across the United States and the world, so we also tested for heterogeneous treatment effects on employees who work in HQ as opposed to other locations.

Results

In principle, the analysis of the results of a field experiment is straightforward, because the randomization allows a simple comparison of means between the control group and the treatment groups.[30] However, individual data before and/or after the intervention period allows us to take into account slight differences between the treatment groups to estimate the treatment effects more precisely. Figure A4.3.1 in the Appendix which presents the raw data (Figure A4.3.2 for above or below median engaged) illustrates that it can be important to look at changes in behavior in the intervention period while taking pre-treatment period differences into account.

We estimate panel regressions with a proxy for cooperation (opening of the email in Table 2 and clicking on a link or answering a question from a colleague in Appendix A4) as the dependent variable. For independent variables, the dummy "Treatment Periods (=1)" capture the effect on the two treatment periods for the control group. The dummies for the different treatments are 1 in the two periods in which we implemented our intervention and 0 otherwise,

---

[29] For each possible activity on the platform, (i) whether the employee has created a user account on the internal platform, (ii) ever posted a comment, (iii) joined more than two employee groups, or (iv) ever asked a question to the firm, we assigned 1 point. We then summed them up to create the Engagement index.

[30] This is only true if we look at the intent-to-treat effect, i.e. at behavior of all individuals that get assigned to control and treatment. If we condition on behavior, i.e. whether an individual opened the email, or look at subgroups, we are working with non-random samples and the empirical approach must be adjusted accordingly.

i.e. "T1/T2/T3 x Treatment Periods" show the difference of the treatments compared to the control group for the treatment periods. As all the regressions include individual fixed-effects, we estimate difference-in-differences. We also added a variable that captures any time trend.[31] All the regressions include individual fixed effects to capture time-invariant individual effects. We also cluster the standard errors at the individual level, to adjust the standard error for the fact that we have eight observations per individual. The regressions either include or exclude the post-treatment periods as it is unclear whether behavior after the treatment intervention is over, (i.e. email text reverts to normal), should be seen as a continuation of the level before the intervention or was affected by the treatment intervention. In most of the models we include the post-treatment period and look at whether the interventions 'break' the time trend.[32]

Table 2 reveals a number of results about the effect of our identity manipulation on opening of the email. The results mainly support the observations from the Figures in the Appendix. Whether we exclude the post-intervention periods (column (1)) or not (column (2)), making firm identity salient increases the proportion of opening by between 2.2 percentage points ($p<0.1$) and 2.9 percentage points ($p<0.05$). Making BU identity salient has a small effect when excluding the post-intervention period. The effect is about half as big as making firm identity salient. However, the difference between the two treatments is not statistically significant. When taking the post-intervention pattern into account, both firm and BU identity increases opening 2.9 or 2.6 percentage points, respectively. Adding competition has no

---

[31] While there is no pre-treatment time trend differences between the treatments (analysis is available on request), there is a general decreasing time trend – even for the control group. Our discussions with corporate management indicate that this is likely attributable to seasonality coinciding with the end of the fiscal quarter, during which employees are often busier than usual.

[32] Further analyzing employee behavior in the post-treatment periods might reveal interesting insights. In Appendix A4.3., we provide some analysis on post-treatment behavior and find differences between the different treatments. Exploring these differences more systematically is an interesting topic for future work.

statistically significant effect on opening the email either way. None of the differences between the treatments are statistically significant at the 10-percent level (except for the difference between firm identity and BU competition in the model shown in column (3)).

Comparing the reaction of employees who were engaged or not engaged on the platform shows that our effects almost entirely come from individuals who were not previously engaged at the firm. Firm identity increases opening by 4.3 percentage points for individuals less engaged prior to the intervention (p<0.01). BU identity increases opening by less, but still by 3.1 percentage points (p<0.05). By contrast, competition has no statistically significant effect on opening and is half as strong as BU identity alone based on the point estimate. However, the difference between individuals who exhibited low or high previous engagement is not statistically significant at any conventional level.

Comparing the effect of the intervention for employees working in the HQ or not also reveals interesting patterns. Judging from both the size of the effect and the significance level, the results suggest that for employees who don't work in the HQ, making BU identity salient is more powerful than appealing to firm identity. BU identity increases the proportion opening the email by 2.6 percentage points (p<0.1). This effect is reversed for employees in the HQ: The firm identity intervention increases opening by 4.4 percentage points (p<0.05) versus 2.7 percentage points (BU identity) (p=0.152). Again, the coefficients are not different between the subgroups "Not HQ" and "HQ" at any conventional significance level.

In sum, we sought to test an important link in the chain between strong corporate culture and performance by manipulating employee identity and testing the impact on cooperative behavior. Our results show some effects of our intervention on the opening of emails containing questions from colleagues, the first step in cooperation. However, there is no equivalent effect of

the intervention on whether an employee clicked on any of the links in the email (Table A4.3.1 in the Appendix) or whether they actually answered any of the questions from their colleagues (Table A4.3.2 in the Appendix). We did find interesting differences in cooperative behavior according to whether the employee was previously engaged and whether they worked at HQ, pointing to interesting opportunities for future analysis. These results are in the spirit of Kogut and Zander (1996), who discussed the role of identity inside organizations. For many firms that are highly acquisitive, integrating new employees and building collective identities are a significant challenge in the post-acquisition process. Future process experiments might further inform the best strategies to navigate these issues.

5. Discussion

One of the most unique features of strategy research is the potential to make prescriptions that guide behavior in the field, often managerial actions with significant economic value at stake. These questions are invariably the most difficult to answer, in large part because the underlying causal models are not well understood. Although rigor and relevance are often posed as being at odds in our research, we think otherwise. Using field experiments is one way to provide rigorous answers to relevant questions, building on a long tradition of influential strategy research. In fact, given that so many strategy scholars work closely with firms already, our community has a particular advantage in formulating and executing useful field experiments.

The two experiments presented in this paper apply a new research methodology to canonical questions in strategy research and reveal useful empirical implications. More broadly, we use these studies to illustrate the broader applicability of this method to the field. In doing so, we identify both important drawbacks and advantages of using field experiments in strategy research. Neither kind of field experiment we present allow researchers to account for all of the

impact of mediating and background variables. External validity cannot be guaranteed. However, field experiments are especially useful to establish internal validity – a challenge in much of the literature that links various firm characteristics to performance. Field experiments also may provide a high level of "naturalism" that has appealing properties for researchers. Conducting research in an actual business environment also ought to endow the results with greater relevance for managers, which may increase the chance that the findings have an impact on managerial decision-making.

As a result, we argue that using "strategy field experiments" and "process field experiments" in concert can yield valuable insights for strategy researchers and practice. In general, the opportunities to do strategy field experiments will be few but the results are more likely to be provocative and significant for future research and practice. This approach will work best when the firm attribute under study operates through the perceptions of outside stakeholders, government officials, the media or competitors. For their part, process field experiments rely on the development and refinement of strong theory to direct researchers to the relationships that matter. This attribute will drive continued interest in theoretical development alongside the iterative implementation of new field experiments. Especially as the field moves toward studying activities inside firms, the process field experiment approach has considerable merit, especially since it allows us to create new data rather relying on existing datasets. Taken together, we believe that the field experiment method, encompassing the two approaches discussed above, can be a key driver in the development of strategy research going forward.

References

Allcott, H. and S. Mullainathan (2012). "External Validity and Partner Selection Bias", Working Paper.

Ashforth, B. E. and F. Mael (1989). "Social identity theory and the organization." <u>Academy of Management Review</u> **14**(1): 20-39.

Banerjee, Abhijit V., and Esther Duflo (2009). "The experimental approach to development economics." <u>Annual Review of Economics</u> **1**:151-178.

Bandiera, O., I. Barankay and I. Rasul. (2011). "Field experiments with firms." <u>The Journal of Economic Perspectives</u> **25**(3): 63-82.

Barney, J. B. (1986). "Organizational Culture: Can It Be a Source of Sustained Competitive Advantage?" <u>The Academy of Management Review</u> **11**(3): 656-665.

Bascle, G. (2008). "Controlling for endogeneity with instrumental variables in strategic management research." <u>Strategic Organization</u> **6**(3): 285-327.

Bertrand, Marianne, and Sendhil Mullainathan (2004). "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." <u>American Economic Review</u> 94 (4): 991-1013.

Bloom, Nicholas, et al (2009). "The distinct effects of information technology and communication technology on firm organization." <u>NBER working paper</u> 14975.

Bohnet, I. and B. S. Frey (1999). "Social distance and other-regarding behavior in dictator games: Comment." <u>The American Economic Review</u> **89**(1): 335-339.

Bullock, John G., Donald P. Green, and Shang E. Ha. (2010) "Yes, but what's the mechanism?(don't expect an easy answer)." <u>Journal of Personality and Social Psychology</u> 98.4:550.

Butler, Daniel M. and David E. Broockman (2011). "Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators." <u>American Journal of Political Science</u> 55 (3): 463-477.

Cable, D., Gino, F., and Staats, B. (2013). "Breaking Them In or Eliciting Their Best? Reframing Socialization around Newcomers' Self-expression." <u>Administrative Science Quarterly</u> **58** (1): 1-36.

Calder, Bobby J., Lynn W. Phillips, and Alice M. Tybout. "The concept of external validity." <u>Journal of Consumer Research</u> (1982): 240-244.

Carpusor, Adrian G., and William E. Loges (2006). "Rental discrimination and ethnicity in names." <u>Journal of Applied Social Psychology</u> 36 (4): 934-952.

Chen, Y. et al. (2010). "Which Hat to Wear? Impact of Natural Identities on Coordination and Cooperation." <u>Working paper</u>.

Cook, Thomas D., and Donald T. Campbell. "Quasi-experimentation: Design and analysis for field setting." MA: Houghton Mifflin (1979).

Croson, R., et al. (2007). "Using experiments in corporate strategy research." <u>European Management Review</u> **4**(3): 173-181.

Deaton, A. (2010). "Instruments, Randomization, and Learning about Development." <u>Journal of Economic Literature</u> **48**(2): 424-455.

Dipboye, Robert L. "Laboratory vs. field research in industrial and organizational psychology." International review of industrial and organizational psychology 5 (1990): 1-34.

Duflo, E. , Glennerster, R., and Kremer, M. (2006), "Using Randomization in Development Economics: A Toolkit," NBER working paper 333.

French, K. and J. Poterba (1991). "Investor Diversification and International Equity Markets". American Economic Review **81** (2): 222–226.

Gibbons, R. and R. Henderson (2012). "Relational contracts and organizational capabilities." Organization Science **23**(5): 1350-1364.

Glennerster, Rachel, and Kudzai Takavarasha (2013). Running Randomized Evaluations: A Practical Guide. Princeton University Press.

Goette, L., et al. (2012). "Competition between organizational groups: its impact on altruistic and antisocial motivations." Management Science **58**(5): 948-960.

Goldberg, A. (2011). "Mapping Shared Understandings Using Relational Class Analysis: The Case of the Cultural Omnivore Reexamined1." American Journal of Sociology **116**(5): 1397-1436.

Green, Donald P., Shang E. Ha, and John G. Bullock. "Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose." The Annals of the American Academy of Political and Social Science 628.1 (2010): 200-208.

Hamilton, B. H. and J. A. Nickerson (2003). "Correcting for endogeneity in strategic management research." Strategic Organization **1**(1): 51-78.

Harrison, G. W. and J. A. List (2004). "Field experiments." Journal of Economic Literature **42**(4): 1009-1055.

Heckman, James (1992). "Randomization and social policy evaluation". In Charles Manski and Irwin Garfinkel (Eds.), Evaluating Welfare and Training Programs. Harvard Univ. Press: Cambridge, MA: 201-230.

Helfat, C. E., et al. (2009). Dynamic capabilities: Understanding strategic change in organizations, Wiley. com.

Hutchinson, J. Wesley, Wagner A. Kamakura, and John G. Lynch Jr. "Unobserved heterogeneity as an alternative explanation for "reversal" effects in behavioral research." Journal of Consumer Research 27.3 (2000): 324-344.

Hymer, Stephen Herbert. (1976). The International Operations of National Firms: A study of foreign direct investment. The MIT Press: Cambridge.

Karlan, Dean, and Jonathan Zinman. "Observing unobservables: Identifying information asymmetries with a consumer credit field experiment." Econometrica 77.6 (2009): 1993-2008.

Kogut, B. and U. Zander (1996). "What firms do? Coordination, identity, and learning." Organization Science **7**(5): 502-518.

Levitt, Steven D and John A. List (2009). "Field experiments in Economics: The Past, The Present, and The Future." European Economic Review **53** (1):1-18.

List, John A (2008). "Informed consent in social science." Science 322 (5902): 672.

List, John A (2009). "The IRB is key in field experiments." Science 323 (5915): 713-714.

Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan (2011). "Mechanism experiments and policy evaluations." Journal of Economic Perspectives 25 (3): 17-38.

Lynch Jr, John G (1982). "On the external validity of experiments in consumer research." Journal of Consumer Research 9 (3): 225-239.

Lynch, John G. (1983). "The role of external validity in theoretical research." Journal of Consumer Research 10 (1): 109-111.

Lynch, John (1999). "Theory and External Validity." Journal of the Academy of Marketing Science, 27(3): 367-376.

Mezias, J. (2002). "Identifying liabilities of foreignness and strategies to minimize their effects: the case of labor lawsuits judgments in the United States." Strategic Management Journal 23 (2): 229-244.

Miller, S. and L. Eden. (2006). Local Density and Foreign Subsidiary Performance. Academy of Management Journal 49 (2): 341-355.

Miller, S. and A. Parkhe. (2002). "Is there a liability of foreignness in global banking? An empirical test of banks' X-efficiency." Strategic Management Journal 23 (1): 55-75.

Nachum, L. (2003). "Liability of foreignness in global competition? Financial service affiliates in the city of London." Strategic Management Journal 24 (12): 1187-1208.

Oakley, Ann, et al. "Process evaluation in randomised controlled trials of complex interventions." Bmj 332.7538 (2006): 413-416.

OReilly, C. A. and J. A. Chatman (1996). "Culture as social control: Corporations, cults, and commitment." Research in Organizational Behavior 18: 157-200.

Pew Foundation. (2007). Pew Global Attitudes Survey, available at<http://tinyurl.com/6tp2oka>

Rajan, R. G. and J. Wulf (2006). "The flattening firm: Evidence from panel data on the changing nature of corporate hierarchies." The Review of Economics and Statistics 88(4): 759-773.

Shadish, W., and Cook, T. (2009). "The Renaissance of Field Experimentation in Evaluating Interventions," Annual Review of Psychology 60: 607-629.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental design for generalized causal inference. Boston: Houghton-Mifflin.

Shih, M., et al. (1999). "Stereotype susceptibility: Identity salience and shifts in quantitative performance." Psychological Science 10(1): 80-83.

Siggelkow, Nicolaj. "Change in the presence of fit: The rise, the fall, and the renaissance of Liz Claiborne." Academy of Management Journal 44.4 (2001): 838-857.

Sørensen, J. B. (2002). "The Strength of Corporate Culture and the Reliability of Firm Performance." Administrative Science Quarterly 47(1): 70-91.

Spencer, Steven J., Mark P. Zanna, and Geoffrey T. Fong. "Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes." Journal of Personality and Social Psychology 89.6 (2005): 845.

Teece, D. J., et al. (1997). "Dynamic capabilities and strategic management." Strategic Management Journal **18**(7): 509-533.

Zaheer, S. (1995). "Overcoming the liability of foreignness." Academy of Management Journal **38**: 341-63.

Zaheer, S. (2002). "The Liability of Foreignness, Redux: A Commentary." Journal of International Management **8**(3): 351-358.

Zaheer, S. and E. Mosakowski. (1997). "The dynamics of the liability of foreignness: A global study of survival in financial services." Strategic Management Journal **18**: 439-64.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. Journal of Consumer Research, 37(2), 197-206.

Figure 1: Published Field Experiments in Strategy/Management vs. Economics
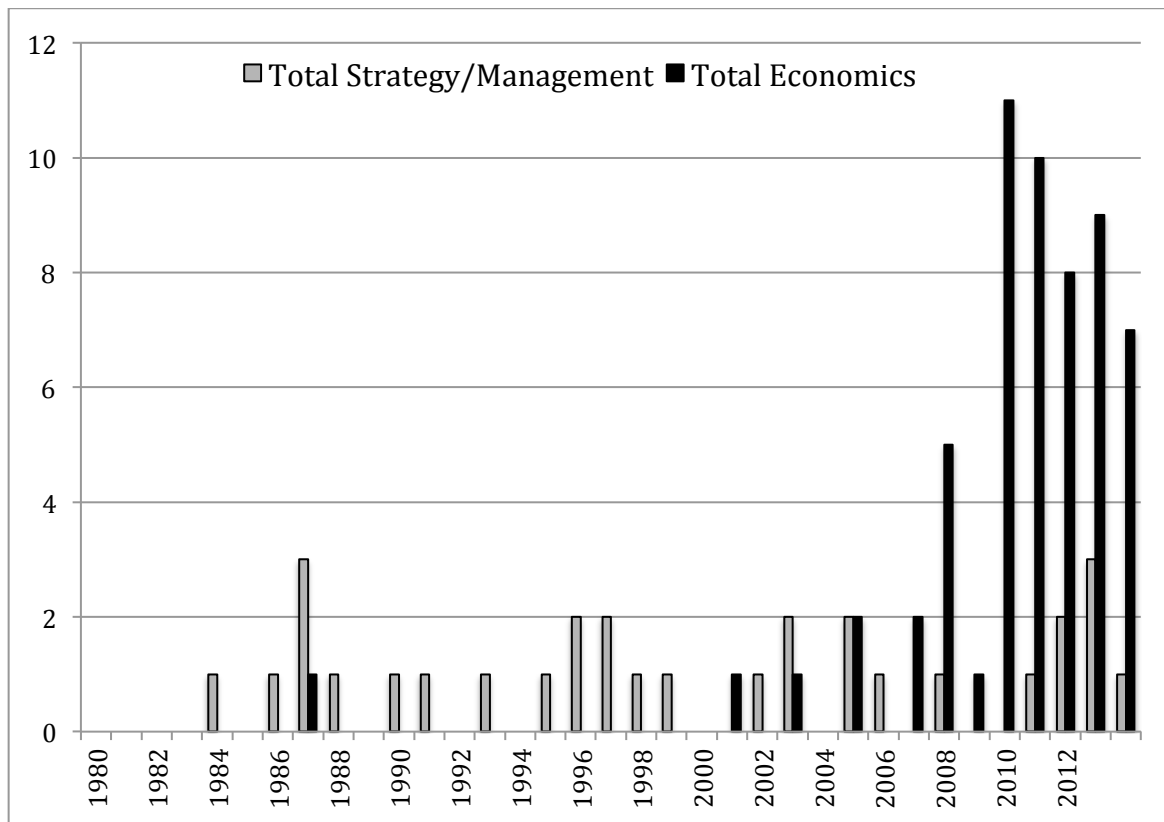
Notes: Figure counts articles published in top strategy/management and top economics journals between 1980 and 2013. The following journals were considered as strategy/management: *Academy of Management Journal*, *Academy of Management Review*, *Administrative Science Quarterly*, *European Management Review*, *Journal of Economics & Management Strategy*, *Journal of Management*, *Management Science*, *Organization Science*, *Strategic Management Journal*, and *Strategic Organization*. In economics: *American Economic Journal: Applied Economics*, *American Economic Journal: Economic Policy*, *American Economic Journal: Macroeconomics*, *American Economic Journal: Microeconomics*, *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Journal of the European Economic Association*, *Quarterly Journal of Economics*, and *Review of Economic Studies*.

Table 1: Treatment Differences in Experiment 1

Treatment Country	US	Japan	China
No of observations	1,047	1,034	1,048

Panel A: Response Rates

N Responded	184	152	165
Response Rate	0.176	0.147	0.158
Test of Differences:			
USA vs. Foreign	-0.024*	[-0.051, 0.004]; p=0.091	
USA vs. Japan	-0.029*	[-0.060,0.003]; p=0.0749	
Japan vs. China	-0.011	[-0.042,0.020]; p=0.5017	

Panel B: Proportion of Cities Offering Incentives

Offered Incentives	103	101	87
Proportion	0.098	0.098	0.083
Test of Differences:			
USA vs. Foreign	0.008	[-0.013, 0.030]; p=0.461	
USA vs. Japan	0.001	[-0.025, 0.026]; p=0.957	
Japan vs. China	0.015	[-0.010, 0.039]; p=0.243	

Panel C: Means of Natural Logs of Grant Dollars Offered

Mean ln(dollars)	0.157	0.156	0.165
Test of Differences:			
USA vs. Foreign	-0.004	[-0.090, 0.082]; p=0.930	
USA vs. Japan	0.001	[-0.097, 0.098]; p=0.987	
Japan vs. China	0.009	[-0.110, 0.091]; p=0.853	

Panel D: Means of Natural Logs of Grant Dollars Offered

Mean ln(dollars)	1.953	2.270	2.088
Test of Differences:			
USA vs. Foreign	0.219	[-0.768, 1.206]; p= 0.663	
USA vs. Japan	0.317	[-0.843, 1.477]; p= 0.590	
Japan vs. China	0.182	[-1.022, 1.386]; p= 0.765	

Notes: Table shows the outcome variables for the three treatments in Panel A-D. Each panel also shows the differences between the various treatments, 95% confidence intervals and p-values. Non-response is coded as 0 in Panel C and excluded in Panel D.
* p < 0.1, ** p < 0.05, *** p < 0.01

Table 2: Effect of Identity Treatments on Opening of Email

	(1) All	(2) All	(3) Low Engagement	(4) High Engagement	(5) Not HQ	(6) HQ
Firm Identity x Treatment Periods	0.022*	0.029**	0.043***	0.008	0.019	0.044**
	(0.013)	(0.012)	(0.015)	(0.018)	(0.015)	(0.019)
BU Identity x Treatment Periods	0.011	0.026**	0.031**	0.018	0.026*	0.027
	(0.013)	(0.012)	(0.015)	(0.018)	(0.015)	(0.019)
BU Competition x Treatment Periods	-0.000	0.019	0.015	0.025	0.017	0.023
	(0.014)	(0.012)	(0.015)	(0.019)	(0.015)	(0.020)
Treatment Periods (=1)	0.022*	0.011	0.008	0.013	0.019*	-0.003
	(0.012)	(0.008)	(0.010)	(0.013)	(0.010)	(0.014)
Time Trend	-0.010***	-0.009***	-0.010***	-0.008***	-0.009***	-0.010***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Constant	0.717***	0.716***	0.709***	0.726***	0.732***	0.690***
	(0.006)	(0.004)	(0.006)	(0.006)	(0.005)	(0.007)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. of mailings	1-5	1-8	1-8	1-8	1-8	1-8
R squared	0.000	0.006	0.008	0.005	0.007	0.006
No. of observations	20,925	33,480	19,888	13,592	20,480	13,000
No. of clusters	4,185	4,185	2,486	1,699	2,560	1,625

Notes: Coefficients of OLS regressions with individual fixed effects. Robust standard errors in parentheses clustered at the individual level. Dependent variable: 1 if opened email and 0 otherwise. Column (1) excludes post-intervention periods. Columns (3) and (4) split the sample at the median level of engagement index. Columns (5) and (6) show results for employees who work either in the Headquarters (HQ) or not.
Significance Level: *p < 0.1, **p < 0.05, ***p < 0.01

ONLINE APPENDIX for

Field Experiments in Strategy Research

Content:

A1. Published and Forthcoming Field Experiments in *SMJ*

Author(s)	Title	Issue	How the paper addresses a debate in the literature:
Shardul Phadnis, Chris Caplice, Yossi Sheffi, Mahender Singh	Effect of scenario planning on field experts' judgment of long-range investment decisions	forthcoming	The authors indicate that their work contrasts with previous results on the impact of scenario planning on judgment of investment decisions. Nicely summarized in their abstract, they indicate: "Our results show, contrary to the past findings, that the use of multiple scenarios does not cause an aggregate increase or decrease in experts' confidence in their judgment." Their empirical results point to alternative mechanisms for how expert judgments evolve. We categorize their contribution as providing new empirical evidence on an important topic that runs counter to the results of prior work.
Lisa Jones Christensen, Enno Siemson, and Sridhar Balasubramanian	Consumer behavior change at the base of the pyramid: Bridging the gap between for-profit and social responsibility strategies	forthcoming	The authors situate this work as addressing competing theories of philanthropy versus a for-profit strategy in emerging markets. The authors note in the introduction that their work is one of the few to directly test these competing theories. They note: "Yet, no research of which we are aware has taken a longitudinal view to investigate how each strategy relates to adoption and retention of consumers in the context of non-durables." They find evidence for benefits from a "deeply discounted" for-profit strategy. We categorize their contribution as adjudicating an important scholarly debate with a new research design.
Giada Di Stefano, Andrew A. King, Gianmario Verona	Kitchen confidential? Norms for the use of transferred knowledge in gourmet cuisine	forthcoming	The authors study how norms affect knowledge transfer, an important research domain. Their main results reinforce existing work on the importance of norms, but they also provide new theory and additional evidence of how norms are transferred. We categorize this paper as testing existing theory, along with generating and testing new theory.
J. Robert Mitchell, Dean A. Shepherd, Mark P. Sharfman	Erratic strategic decisions: when and why managers are inconsistent in strategic decision making	32 (7), 2011	This paper presents a new model of erratic decision-making. The authors note that few studies examine this form of decision-making and thus their theory and empirical contributions are novel. Their main findings reveal the conditions under which erratic strategic decisions are more or less common. We categorize this paper as making largely new theoretical and empirical contributions to strategy research.

A2. Practical Considerations for Conducting Field Experiments

In this section, we point out a number of practical considerations for conducting field experiments (for a more comprehensive guide about running field experiments – mainly illustrated in the area of development, see Glennester and Takavarasha (2013)).

       - *Initiating the experiment*: During the course of designing our field experiments we learned some key lessons: First, for field experiments within firms to be successful, the researchers need the "buy-in" of important decision-makers within the firm and dedicated staff that are responsible for working on the project. How does the researcher achieve "buy in"? First, the field experiment must provide important and actionable implications for the firm. If the research question is not a first-order business concern for the company, the project is unlikely to be a priority for internal staff. To answer an important question for the firm, the researchers need to first listen carefully to corporate leaders to understand what important opportunities and threats the company is facing. Second, one needs a strong champion that can move the project forward at the senior level, especially with regard to internal legal processes concerning consent forms and protections of employee identity. Ideally this internal champion will have direct reports that can handle internal data requests and assist with implementation of the experiment.

       - *Ethical/legal considerations and IRB*: Field experimenters face a special duty to ensure the ethical treatment of their subjects, particularly since they often withhold the knowledge that subjects are involved in a social scientific study (see List (2008) for a discussion about informed consent). Both the *Belmont Report*, which lays out the ethical guidelines for the treatment of human subjects, and the U.S. Department of Health and Human Services' *Common Rule*, which formally regulates the domain, allow for the suspension of fully informed consent when four conditions are met: the benefits of the research are significant, the risks are minimal, no physical

or emotional pain is inflicted, and the research cannot be carried out in another way. If those criteria are fulfilled it is even possible that some mild deception is acceptable (for the role of Institutional Review Boards (IRB) on field experiments, see List (2009)). As can been seen in Field Experiment 1, our design involved some mild deception.[33] The IRB of the universities that employed the authors all reviewed the research protocol before its execution. All three universities exempted this research. Field Experiment 2 was also granted an exemption by IRB at the author's university.

With regards to experiments inside firms, it is often the case that the company wants to keep the results of a study private (especially from their competitors) while the researchers want to disseminate their findings. One practical 'solution' to this problem is to contractually agree on a period during which the results of the study will not enter the public domain (but during which the researchers can still discuss the results in seminars and submit papers for review). Given the lag between first drafts and journal submissions and then to eventual publication, these contracts may often be mutually acceptable.

- *Sample Size and Power Calculation*: An important first practical step is to determine the size of the study. Conducting a power calculation is often a necessary step to deciding on the design of the study. Power calculations provide the researcher with information about whether a design has enough statistical power given a certain sample size per treatment group. When the sample size is given, as in our experiments, power calculations provide information about the number of treatments a researcher can run and still have statistical power to detect an effect. Power calculations ask the question: Given a certain number of observations per treatment, how

---

[33] Field experiment 1 uses about the same or less severe deception than prior important studies that identified employment, housing, and political discrimination by adopting aliases (e.g., Bertrand and Mullainathan 2004, Carpusor and Loges 2006, Butler and Broockman 2011).

big does the effect have to be to be statistically significant? Power calculations are "part science and part art" because they rely on assumptions about the variance in the outcome variables and estimations of the likely treatment effect. Those assumptions are more reliable as the baseline data improves and the researchers learn more about the outcome variable of interest. As such, historical data about the dependent variable and estimations of treatment effects are very helpful.

- *Randomization*: The next key decision involves the method of randomization. The simplest way to randomly assign individuals into the different treatment groups is to literally randomize the selection. However, a more sophisticated approach is to stratify according to certain characteristics, which was the approach followed in both experiments. Stratifying, or block randomization, is advisable for at least two reasons: First, it ensures that individuals with certain characteristics are represented in all treatments in equal proportions. Second, it demonstrates that the researcher has considered heterogeneous treatment effects on the stratified dimension. This feature substantially reduces the danger of data mining (see next point).

- *Subgroup Analysis / Pre-registration*: The analysis of field experiments is relatively straightforward as the randomization process does not require any sophisticated correction for differences in observable and unobservable characteristics between control and treatment groups. In expectation, all confounds are balanced across experimental conditions. Analysis grows more complicated when researchers want to analyze heterogeneous treatment effect, i.e. whether the treatment has a stronger effect for certain subgroups. One needs to be careful with heterogeneous treatment analysis for at least two reasons. First, the effects cannot be interpreted as causal, because the subgroups are not randomly assigned. For example, if one wants to see whether a certain treatment works differently for employees working in HQ than in other locations, one needs to be cautious in interpreting this effect as a "HQ" effect, because employees at HQ might

differ in multiple observable (which we could control for) and non-observable characteristics (which we cannot control for).

Second, the other problem with analyzing subgroups in order to detect heterogeneous treatment effects is that doing so creates the potential problem of "data mining." If the experiment does not create a main effect, the temptation is to split the sample in multiple dimensions to see whether one finds at least one significant "treatment effect." To minimize the second problem (without solving the first), the subgroup analysis should either be based on a strong theoretical prediction and/or the treatment should be stratified by subgroup to signal that the researcher wanted to conduct the subgroup analysis even before seeing the results of the main effect. Another way to be transparent about the planned research before seeing the data is to pre-register the design and the planned data analysis. This is becoming common in medical research and could also become standard in the social sciences.

A3. Field Experiment 1 Details

This section provides details and additional analysis for field experiment 1 "Liability of Foreignness – country of origin and U.S. inward investment promotion." This section contains the following subsections:
    A3.1. Example Email
    A3.2. Web Form
    A3.3. Web Tracking
    A3.4. Additional Table

A3.1. Example Email

"I am an associate with GLOBEUS Consulting (see our website here [insert hyperlink]). GLOBEUS is a new consulting firm that specializes in matching cities with prospective firms. I work in the GLOBEUS group focusing on investors based in [the United States / Japan / China] and am contacting you to see if your city would be a good match for a client I am representing.

Our client is considering an expansion of a manufacturing plant producing electrical grounding products. The company is looking to make a decision and announce the investment in [two months before next election / one month after next election]. Based on specs from another facility, we project that the plant would create 19 full-time hourly jobs at around $12 an hour plus benefits and 6 salaried jobs at around $40,000 per year.

The company is looking to buy or lease a 15,000 to 20,000 square-foot building. The total investment would be $2,000,000 ($1,750,000 on building and equipment and $250,000 on other various moving expenses). Previous plants have taken 6 months from the time of the announcement to being fully operational.

To examine the feasibility of your city for this proposed project we are asking for you to fill out this web form (available here [insert hyperlink]) on the type of incentives you could potentially offer this investor and what types of incentives you have offered in the past.

As you might expect, this offer is not binding and we realize any formal offer would require due diligence and direct interaction with our client. Our goal at this stage is to present a detailed analysis to our client on the feasibility of relocating to your village.

We regret that we are not authorized to provide any more details about our client at this point, but if you have any questions please feel free to contact us via email. We look forward to your response.

[Associate Name]
[us / japan / china]_client_team@globeusconsulting.com
Selection & Incentives Associate Globeus Consulting—[U.S. / Japan / China] Client Team
Team www.globeusconsulting.com

Field Experiments in Strategy Research

A3.2. Web Form

Globeus Consulting
Selection & Incentives Department

Introduction

This data you enter into this webform will be used by our client to narrow down their location decision. Your answers are not binding, but any concrete details you can provide will help us evaluate the feasibility of your ${e://Field/type} as a site for the plant relocation.

In this form we will ask about:
a)   Grants and loans for relocation provided on a per job basis.
b)   Tax abatements (on property and earnings taxes).
c)   Any other local incentives provided.

Grants and Loans

Please indicate the availability of grants and loans.

Local grant dollars for relocation (dollars per job)

Local loans for relocation (dollars per job)

Please enter additional comments or information about grants and loans below.

Real Property Tax

Does your ${e://Field/type} have local real property taxes?

◯ Yes

◯ No

Please indicate below the local real property tax abatement or refund your ${e://Field/type} is able to offer.

Not Applicable

0   10   20   30   40   50   60   70   80   90   100

9

7/30/13   Qualtrics Survey Software

	0	10	20	30	40	50	60	70	80	90	100	
Local real property tax abatement or refund (%)												☐

Please indicate below the number of years you are able to offer this abatement or refund.

											Not Applicable	
	0	5	10	15	20	25	30	35	40	45	50	
Years												☐

Please enter additional comments about real property taxes below.

```
[                                                              ]
```

Personal Property Tax

Does your ${e://Field/type} have personal property taxes?

○ Yes
○ No

Please indicate below the local personal property tax abatement or refund your ${e://Field/type} is able to offer.

											Not Applicable	
	0	10	20	30	40	50	60	70	80	90	100	
Local personal property tax abatement or refund (%)												☐

Please indicate below the number of years you are able to offer this abatement or refund.

Not
Applicable

Field Experiments in Strategy Research

Qualtrics Survey Software

	0	5	10	15	20	25	30	35	40	45	50	Applicable
Years												☐

Please enter additional comments about personal property taxes below.

Local Earnings Tax

Does your ${e://Field/type} have local earnings taxes?

○ Yes

○ No

Please indicate below the local local earnings tax abatement or refund your ${e://Field/type} is able to offer.

	0	10	20	30	40	50	60	70	80	90	Not Applicable 100	
Local real property tax abatement or refund (%)												☐

Please indicate below the number of years you are able to offer this abatement or refund.

	0	5	10	15	20	25	30	35	40	45	Not Applicable 50	
Years												☐

Please enter additional comments about local earnings taxes below.

A3.3. Web Tracking

Out of the 3,117 cities, we received hits from a total of 1,031 unique IP addresses and a total of 7,822 page views. Visitors averaged 6.9 actions per visit and spent an average time on our website of 4 minutes and 50 seconds. While we can track unique IP addresses, it is impossible for us to (other than inferences based on location), to identify individuals. But by providing unique links for each treatment group, we can compare how treatment groups used our websites and the contact that individuals accessed on our website.

First, we explored website visits by treatment group. While sizeable numbers of visitors located our website through a web search (and thus we could not track the treatment group) a total of 601 visitors viewed our website by clicking on unique web links associated with the different treatment groups. These visits were almost perfectly balanced across groups where 31.44% of visitors were from the U.S. treatment group, 33.11% from the Japan treatment group, and 35.44% from the China treatment group.[34] Thus, we found no evidence of discrimination against foreign investors based on evidence from our website. We also find no specific discrimination against Chinese investors.

Second, we created two links on our website under a tab called "Clients." Visitors had the option of clicking on "International Clients" and "US Clients," which included a general description of the types of clients we want to represent (but with no names of companies). Our "Clients" tab received 1,424 clicks. A total of 726 visitors continued and clicked on the US client link (50.9%) while 415 (29%) clicked on the International Clients link. What is striking about this result is that in our experiment 2/3 of the cities were treated with an email from our

---

[34] In the U.S. condition, 189 of 1,051 cities visited the website in ways we could track. For Japan and China, 199 of 1,047 and 213 of 1,051 cities visited respectively. The proportions of the overall treatment condition were thus 18.0% for the U.S., 19.0 for Japan, and 20.3 for China.

Japan or China group and only 1/3 from our US group. We take this finding to be at least some indication that cities are more interested in U.S. firms relative to foreign firms.

However, we find it curious that the experimental results did not reinforce this impression from the observational data on the links subjects clicked. Indeed, when it came down to actual incentive offers, there were no differences across experimental conditions that suggested foreign firms are the targets of discrimination.

A3.4. Additional Tables

A3.4.1 Effect of Treatment Variables on Incentive Offers, Controlling for Blocking Covariates

Country	Incentive Offer	Response Rate	Logged Dollars
Japan	0.0214	-0.178*	0.871
	(0.123)	(0.106)	(0.989)
China	-0.0434	-0.0436	1.321
	(0.123)	(0.103)	(0.916)
After Next Election	-0.0531	-0.0447	-0.844
	(0.101)	(0.0856)	(0.781)
Government Type	0.251*	0.0865	1.753
	(0.145)	(0.120)	(1.262)
Population Block	0.193*	0.177*	-1.342
	(0.107)	(0.0918)	(0.855)
1st Quarter	-0.868**	-0.805***	-0.806
	(0.383)	(0.298)	(3.400)
2nd Quarter	-0.205	-0.0668	-3.255*
	(0.220)	(0.184)	(1.795)
3rd Quarter	-0.345	-0.180	-4.026
	(0.578)	(0.479)	(3.051)
Constant	-2.027***	-1.129***	0.833
	(0.247)	(0.192)	(1.809)
Pseudo R Squared	.0773	.0519	.365
N	1,125	1,182	131

Notes: Coefficients of probit (Incentive Offer and Response Rate) and OLS (Logged Dollars) regressions. Robust standard errors in parentheses clustered at the individual level. Dependent variable: 1 offered incentives and 0 otherwise. Coefficients for fixed effects for individual states were included in the regression but omitted here for simplicity in presentation. The variable "elected" was also omitted due to collinearity.
Significance Level: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table A3.4.2: Response rate, incentive offered, and logged dollars with Japan and China interactions

IVs	Response rate		Incentive offered		Logged dollars	
Japan	-0.341*	-0.379*	-0.324	-0.388	0.106	0.078
	(0.19)	(0.195)	(0.228)	(0.242)	(0.153)	(0.152)
Japan × Population	0.196	0.245	0.421**	0.507***	-0.04	-0.019
	(0.147)	(0.15)	(0.169)	(0.176)	(0.118)	(0.117)
Japan × 1st Half	0.069	0.05	-0.041	-0.078	0.027	0.01
	(0.158)	(0.163)	(0.18)	(0.189)	(0.127)	(0.126)
Japan × Northeast	0.094	0.153	0.022	0.082	-0.134	-0.095
	(0.243)	(0.251)	(0.302)	(0.32)	(0.18)	(0.18)
Japan × Midwest	0.106	0.122	0.142	0.17	-0.079	-0.068
	(0.206)	(0.211)	(0.24)	(0.253)	(0.171)	(0.17)
Japan × South	0.047	0.076	0.073	0.12	-0.168	-0.127
	(0.208)	(0.216)	(0.246)	(0.265)	(0.173)	(0.172)
China	-0.105	-0.139	-0.134	-0.219	0.167	0.122
	(0.184)	(0.188)	(0.226)	(0.238)	(0.154)	(0.153)
China × Population	0.009	0.055	0.111	0.197	-0.113	-0.088
	(0.144)	(0.148)	(0.169)	(0.176)	(0.118)	(0.117)
China × 1st Half	0.197	0.181	0.149	0.0999	0.035	0.042
	(0.155)	(0.16)	(0.181)	(0.189)	(0.127)	(0.127)
China × Northeast	-0.036	0.01	0.079	0.137	-0.097	-0.063
	(0.239)	(0.246)	(0.294)	(0.312)	(0.181)	(0.179)
China × Midwest	-0.01	0.003	-0.17	-0.114	-0.147	-0.116
	(0.201)	(0.205)	(0.242)	(0.252)	(0.171)	(0.17)
China × South	-0.12	-0.103	-0.056	0.031	-0.135	-0.103
	(0.204)	(0.21)	(0.246)	(0.262)	(0.172)	(0.171)
Population	0.119	0.097	-0.002	-0.04	0.089	0.073
	(0.1)	(0.102)	(0.116)	(0.12)	(0.083)	(0.083)
1st Half	-0.125	-0.363***	-0.108	-0.322**	0.014	-0.13
	(0.109)	(0.135)	(0.126)	(0.16)	(0.09)	(0.104)
Region 1 (Northeast)	-0.421**	—	-0.315	—	-0.113	—
	(0.163)		(0.206)		(0.126)	
Region 2 (Midwest)	0.142	—	0.358**	—	0.02	—
	(0.14)		(0.167)		(0.12)	
Region 3 (South)	0.099	—	0.22	—	0.115	—
	(0.142)		(0.172)		(0.122)	
State dummies	NO	YES	NO	YES	NO	YES
Constant	-0.92***	-0.283	-1.357***	-1.044*	0.111	3.107***

	(0.128)	(0.467)	(0.155)	(0.586)	(0.108)	(0.863)
Japan interactions	6.25	6.58	6.4	8.49	0.19	0.1
F-test	[0.396]	[0.361]	[0.38]	[0.204]	[0.98]	[0.997]
China interactions	3.35	2.97	3.38	3.88	0.27	0.17
F-test	[0.764]	[0.813]	[0.76]	[0.693]	[0.953]	[0.985]
Pseudo R-squared	0.026	0.061	0.033	0.087	0.000	0.019
N	2690	2684	2697	2614	2697	2697

Notes: Coefficients of Probit models (for response rate and incentive offered) and OLS regression (for logged dollars). Robust standard errors in parentheses clustered at the individual level. The "Quarter" variables are combined to avoid biased standard errors; they appear here as 1st and 2nd Half (where 2nd Half is the omitted variable). Region 4 (West) is the omitted category for the region dummies. P-values for joint significance are shown in brackets under the F-statistics.

Significance Level: *p < 0.1, **p < 0.05, ***p < 0.01

A4. Field Experiment 2 Details

This section provides details and additional analysis for field experiment 2 "The influence of organizational culture – Corporate Identity and Prosocial Behavior". This section contains the following subsections:

A4.1. Instructions: Subject Line and Text for Different Treatments

T0 (Control Group)
Subject Line: Engage with Colleagues to Solve FIRM's Questions and Challenges
Text:

> Please take a few minutes now to review and reply to these selected items recently posted in Ask FIRM
>
> or on the platform as a Question. You just may have the answer to help a colleague.

T1 (FIRM corporate identity email):
Subject Line: Think about a time where a FIRM colleague helped you solve a challenging problem
Text:

> ## What Makes FIRM Special?
>
> **[FIRM] employees work together to answer business and technical questions**
>
> Please take a few minutes now to review and reply to these selected items recently posted in Ask
>
> FIRM or on the platform as a Question. You just may have the answer to help a colleague. **All FIRM**
>
> **employees are on 1 team. Let's work together to make it the best one in the world.**

T2 (Business Unit identity email):
Subject Line: Think about a time where a [BUSINESS UNIT] colleague helped you solve a challenging
problem
Text:

> ## What Makes [BUSINESS UNIT] Special?
>
> **[BUSINESS UNIT] employees work together to answer business and technical questions**
>
> Please take a few minutes now to review and reply to these selected items recently posted in Ask
>
> FIRM or on the platform as a Question. You just may have the answer to help a colleague. **All**
>
> **[BUSINESS UNIT] employees are on 1 team. Let's work together to make it the best one in the**
>
> **world.**

T3 (Business Unit competition email):
Subject Line: Which business unit is the most engaged in solving challenging problems?
Text:

> ## What Makes [BUSINESS UNIT] Special?
>
> **[BUSINESS UNIT] employees work together to answer business and technical questions**
>
> Please take a few minutes now to review and reply to these selected items recently posted in Ask
>
> FIRM or on the platform as a Question. You just may have the answer to help a colleague. **Show that**
>
> **[BUSINESS UNIT] is the most engaged business unit in the firm! For two months we will keep**
>
> **track of who is the most 'helpful' business unit on 'Ask FIRM' and 'The Platform'.**

A4.2. Summary Statistics and Randomization Check

Table A4.2.1: Summary Statistics and Randomization Check

Variable	Mean (S.D.) of T0	Difference T0 - T1	Difference T0 - T2	Difference T0 - T3
Engagement Index (0-4)	1.394 (1.011)	-0.066 [0.136]	-0.004 [0.931]	-0.044 [0.316]
Open Email (=1)	0.713 (0.452)	-.0013 [0.499]	0.013 [0.499]	-0.004 [0.858]
Click Email (=1)	0.085 (0.279)	-0.003 [0.818]	0.008 [0.539]	0.011 [0.361]
Question Asked (=1)	0.069 (0.253)	-0.009 [0.425]	0.005 [0.658]	-0.012 [0.285]
Headquarter (=1)	0.383 (0.486)	-0.020 [0.346]	0.033 [0.116]	0.006 [0.767]

Notes: Table provides mean and standard deviations for the control group in column (1). The remaining columns show the difference between T0 and the three treatments. In brackets are *p*-values of an OLS regression with the "Variable" as the dependent variable and three dummies for the treatments as the independent variables. All information is for time period 1.

A4.3. Additional Figures and Tables

This section shows the raw data in Figure A4.3.1 and Figure A4.3.2. It also shows regression

results as in Table 2 in the text but with 'clicking a link' and 'answering a question' as the

dependent variable (Table A4.3.1 and Table A4.3.2). Finally, we present a Table with results on

post-treatment behavior.

When interpreting the results, it is important to keep the prevalence of the three outcome

variables in mind. When looking at the pre-treatment periods, 70% opened the email, 8% clicked

on a link, and 0.41% responded to a question. That means that for clicking links and especially

responding to a question, the baseline incidence is relatively low and our study might not have

enough power to detect differences.

Figure A4.3.1 shows three results worth discussing: First, although no *statistically*

significant difference exists between opening rates *prior* to the intervention (in time 1, 2, and 3)

between the different treatment groups (T0: .698, T1: .680, T2: .719, T3: .693), slight differences

exist between the treatment groups (see also Table A4.2.1 for a randomization check). Those

differences can be taken into account to estimate the effect of the intervention more precisely.

We can see how important such pre-treatment differences are when we look at the treatment

periods. When we just look at the treatment periods, the raw differences show that "BU Identity"

(T2) has the largest effects compared to the control group. The opening rate is 72.8% in T2

compared to 69.1% in T0 (p=0.062 in a regression in which the standard errors are clustered at

the individual level). However, as we saw in the periods before the intervention, employees in T2

are the most active even before the treatment periods, which we account for in our analysis.

This point leads to the second observation: when we look at the behavioral change from

pre-treatment to treatment period that is caused by the intervention, T2 (Firm Identity) appears to

have the biggest effect. Individual fixed-effects regressions can estimate this effect even more precisely. Third, after the intervention, cooperative behavior of employees, as measured by opening the email with questions from colleagues, drops for almost all treatments. This effect could be interpreted as the continuation of an apparent downward trend or as the effect of the treatment. For example, when looking at "BU Competition," the drop in opening of emails is either the almost-perfect continuation of the downward trend or is caused by the competition treatment. In the analysis in the main text, we show results both with and without taking the post-treatment periods into account.

Table A4.3.3 shows regressions on opening the emails for mailings after our intervention, i.e. all individuals receive the same, standard email. Column (1) confirms the impression from Figure A4.3.1: across the three post-intervention mailings, individuals who were in 'Firm Identity' and 'BU Competition' the opening rates are lower by 3.2 ($p<0.1$) and 4.3 percent ($p<0.05$), respectively. Column (2) indicates that the time trends in the post-treatment period are not different in any statistically significant way. However, for 'BU competition' the time trend is most negative.

Figure A4.3.1: Opening of Email for Treatment and Control Groups
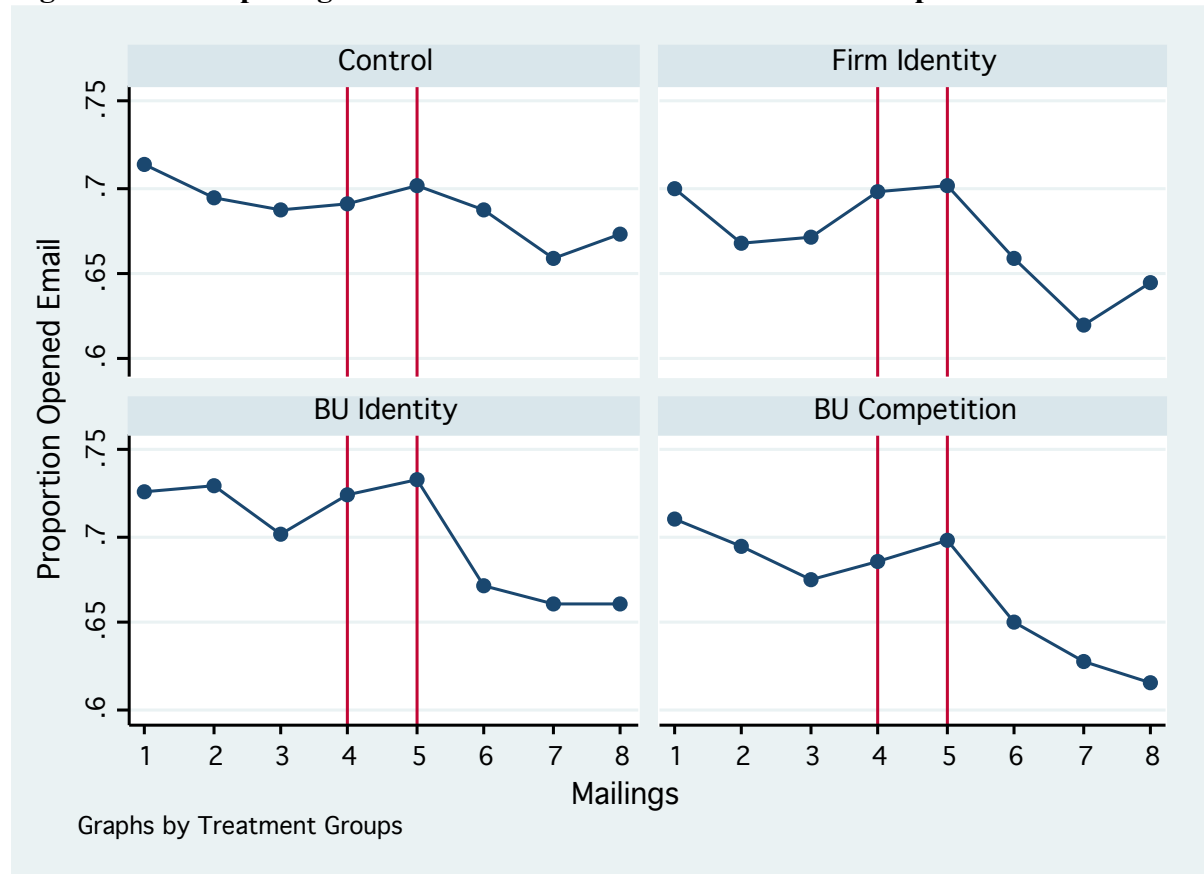
Notes: Figure shows graphically the effect of the treatments on opening of the email as an outcome variable. The four panels show for each treatment the proportion of employees opening the email with the featured questions. It shows three mailings before the experiment, the two mailings (mailings 4 and 5 indicated by the vertical red line) in which we changed the text of the mailings according to the treatment design, and the three mailings after the experimental intervention that returned to the standard text. Mailings are sent bi-monthly, i.e. we observe 1.5 month before and 1.5 month after the interventions.

Field Experiments in Strategy Research

Figure A4.3.2: Opening of Email for Above or Below Median Engaged

Graphs by Treatment Groups
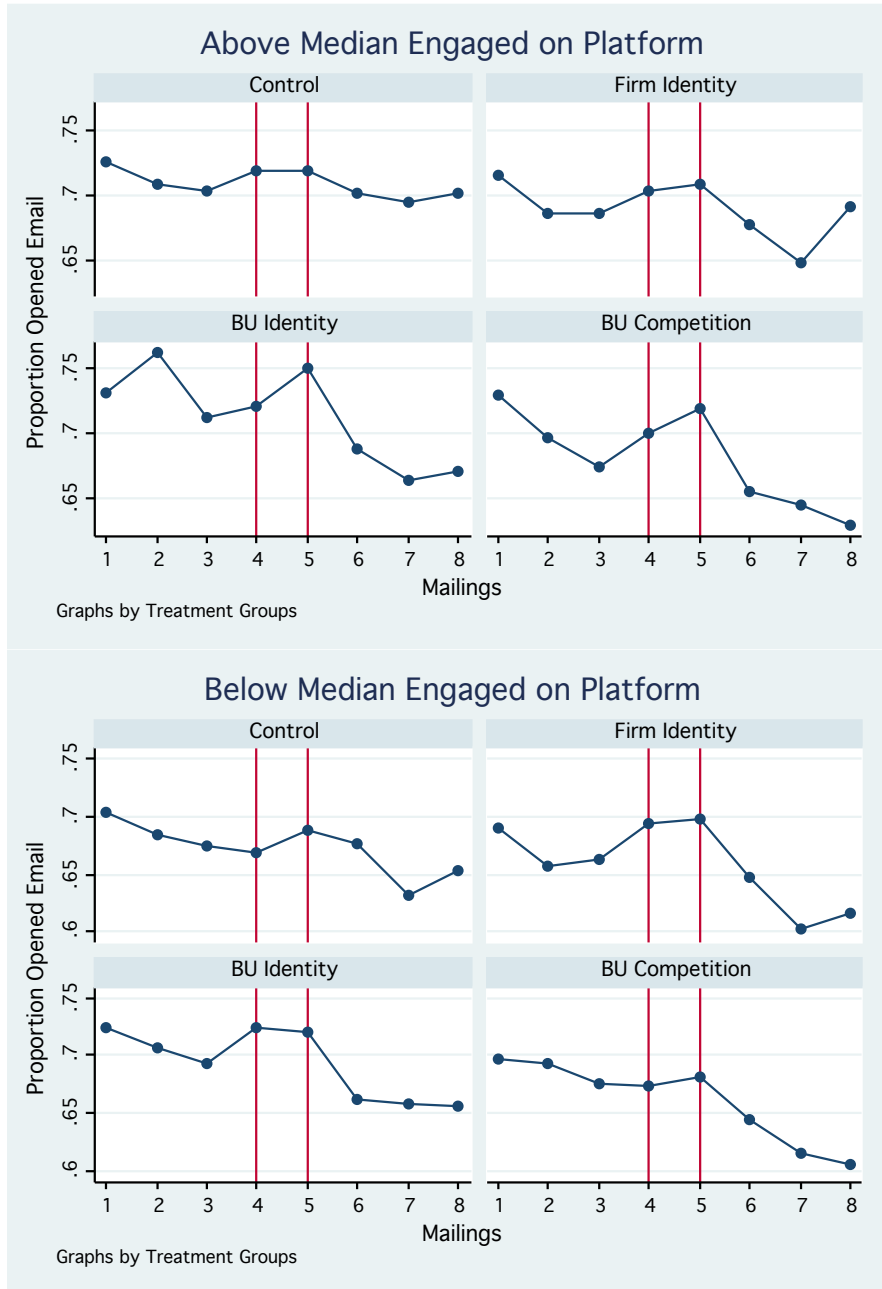
Notes: Figure shows proportion of employees who opened the email with questions from their co-workers by treatment group. Mailings 4 and 5 are the mailings with the treatment manipulations. Mailings are sent bi-monthly, i.e. we observe 1.5 month before and 1.5 month after the interventions.

Table A4.3.1: Effect of Identity Treatments on Clicking a Link

	(1) All	(2) All	(3) Low Engagement	(4) High Engagement	(5) Not HQ	(6) HQ
Firm Identity x Treatment Periods	0.006	0.008	0.012	0.002	0.007	0.010
	(0.009)	(0.008)	(0.009)	(0.014)	(0.010)	(0.014)
BU Identity x Treatment Periods	-0.001	-0.000	0.004	-0.005	0.006	-0.008
	(0.010)	(0.008)	(0.010)	(0.015)	(0.010)	(0.014)
BU Competition x Treatment Periods	-0.005	-0.003	0.016*	-0.030**	-0.005	0.001
	(0.010)	(0.009)	(0.010)	(0.015)	(0.010)	(0.015)
Treatment Periods (=1)	-0.014**	-0.007	-0.011	-0.001	-0.004	-0.011
	(0.007)	(0.006)	(0.007)	(0.010)	(0.007)	(0.010)
Time Trend		-0.004***	-0.003***	-0.006***	-0.004***	-0.004***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.085***	0.095***	0.070***	0.131***	0.085***	0.111***
	(0.001)	(0.003)	(0.003)	(0.005)	(0.003)	(0.005)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. of mailings	1-5	1-8	1-8	1-8	1-8	1-8
R Squared	0.001	0.002	0.002	0.003	0.002	0.002
N	20,925	33,480	19,888	13,592	20,480	13,000

Notes: Coefficients of OLS regressions with individual fixed effects. Robust standard errors in parentheses clustered at the individual level. Dependent variable: 1 if clicked on a link in the email and 0 otherwise. Column (1) excludes post-intervention periods. Columns (3) and (4) split the sample at the median level of engagement index. Columns (5) and (6) shows results for employees who work either in the Headquarters (HQ) or not. *Significance Level*: *p < 0.1, **p < 0.05, ***p < 0.01

Table A4.3.2: Effect of Identity Treatments on Responding to a Question

	(1) All	(2) All	(3) Low Engagement	(4) High Engagement	(5) Not HQ	(6) HQ
Firm Identity x Treatment Periods	-0.002	-0.002	0.000	-0.005	-0.001	-0.003
	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.003)
BU Identity x Treatment Periods	0.001	0.003	0.002	0.004	0.001	0.005
	(0.003)	(0.002)	(0.002)	(0.005)	(0.002)	(0.005)
BU Competition x Treatment Periods	-0.002	-0.001	-0.000	-0.002	0.001	-0.003
	(0.002)	(0.002)	(0.001)	(0.004)	(0.002)	(0.003)
Treatment Periods (=1)	-0.001	-0.000	-0.000	-0.001	-0.001	0.000
	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.003)
Time Trend		-0.000*	-0.000	-0.001	-0.000***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.004***	0.005***	0.001***	0.010***	0.005***	0.004***
	(0.000)	(0.001)	(0.000)	(0.002)	(0.001)	(0.001)
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. of mailings	1-5	1-8	1-8	1-8	1-8	1-8
R Squared	0.000	0.000	0.000	0.001	0.001	0.001
N	20,925	33,480	19,888	13,592	20,480	13,000

Notes: Coefficients of OLS regressions with individual fixed effects. Robust standard errors in parentheses clustered at the individual level. Dependent variable: 1 if responded to a question and 0 otherwise. Column (1) excludes post-intervention periods. Columns (3) and (4) split the sample at the median level of engagement index. Columns (5) and (6) show results for employees who work either in the Headquarters (HQ) or not.
Significance Level: *p < 0.1, **p < 0.05, ***p < 0.01

Table A4.3.3: Post-Treatment Behavior

	(1)	(2)
Firm Identity	-0.032*	-.0035*
	(0.017)	(0.070)
BU Identity	-0.008	-0.022
	(0.017)	(0.067)
BU Competition	-0.043**	0.027
	(0.018)	(0.069)
Time Trend	-0.009***	-0.007
	(0.003)	(0.007)
Firm Identity x Time Trend		0.000
		(0.010)
BU Identity x Time Trend		0.002
		(0.009)
BU Competition x Time Trend		-0.010
		(0.010)
Constant	0.737***	0.723***
	(0.026)	(0.048)
Individual fixed effects	Yes	Yes
No. of mailings	6-8	6-8
R squared	0.002	0.002
No. of observations	12,555	12,555
No. of clusters	4,185	4,185

Notes: Coefficients of OLS regressions. Robust standard errors in parentheses clustered at the individual level. Dependent variable: 1 if opened email and 0 otherwise. The regressions only look at mailings post-intervention period.
Significance Level: *p < 0.1, **p < 0.05, ***p < 0.01