

**Banking Bad:  
A Global Field Experiment Testing  
Money-Laundering Risks in International Finance**

Michael Findley (University of Texas at Austin),  
Daniel Nielson (University of Texas at Austin) and  
Jason Sharman (University of Cambridge)

Prepared for presentation at the online  
Global Research in International Political Economy seminar,  
31 March 2021

NOT FOR FURTHER CIRCULATION

**ABSTRACT**

This article bridges the disciplinary divide between behavioralism and institutionalism via a global field experiment assessing the effects of organizational scripts on banks' and intermediary firms' reactions to requests for anonymous access to the financial system. We thus test the effectiveness of global banking regulations designed to screen out criminal money based on risk assessment. Directly contrary to both the international regulatory regime and scholars' surveyed expectations, both banks and other intermediary firms prove remarkably insensitive to risk in opening accounts. While our strongest risk treatments cause small variations, in the main gate-keepers to the global financial system do not discriminate between high- and low-risk customers. Contrary to baseline rational choice expectations, they are similarly insensitive to varying rewards. We attribute this insensitivity to the power of organizational scripts: shared, standardized patterns of institutionally constrained agent behavior for diagnosis and response. Because this study's experimental conditions and subject pool operate at the level of institutions (firms), rather than individuals, it mitigates the disciplinary divide between experimental methods testing behavior and theories of institutions. As a global field experiment, the study also addresses common external-validity limitations.

An intractable ontological and epistemological divide appears to cleave political science. On one side, behavioralists emphasize causal identification through random assignment of interventions targeting individual behavior, such as voting, activism, and opinion formation (Druckman et al. 2009, Gerber and Green 2012). On the other, institutionalists counter that many important phenomena in politics – from interest groups, parties, legislative committees, and bureaucracies to war, trade, and aid – occur at the level of aggregations of people governed collectively by rules and procedures (Olson 1965, Ostrom 1990, Cox and McCubbins 1993, Shepsle and Weingast 1987, Fearon and Laitin 2003, Martin and Simmons 1998, Nielson and Tierney 2003). Experiments on group-level outcomes are usually seen as impractical, unethical, or both, so institutionalists generally test arguments on observational data using regression analysis and thus are limited to statistical correlations.

The strong implication of the behavioralist-institutionalist fissure is that researchers can achieve *either* compelling causal identification *or* analysis of institutions, but not both. In answer, we argue here that, for many – though far from all – questions in politics, this is a false dichotomy. A significant and largely unexplored domain exists for which studies can be designed to identify the causes of institutional behavior by theorizing about the constraining effects of organizational rules and procedures and employing institutions as subjects in randomized experiments.

The present study demonstrates a way to bridge the behavioralism-institutionalism divide via a global field experimental study of banks and financial intermediaries as institutions. The experiment examines the interaction of institutions rather than individuals at both the treatment and outcome level, i.e. through institutional treatments from companies we legally incorporated for the study and a subject pool of banks and intermediary firms, also known as corporate service providers (CSPs). It is theoretically institutional, being centered on the concept of organizational scripts, which are diagnostic templates representing the codification of shared expectations that shape routine collective behavioral responses (March and Olsen 1989: 21-26; 1998: 948; Barnett and Finnemore 1999: 717; Seabrooke and Kentikelenis 2017). As such, our explanation does not depend on whether or not individuals can be plausibly assumed to approximate rational utility-maximizers, nor does it assume that institutions act as individual rational decision-makers writ large (Hafner-Burton et. al 2017, Powell 2017).

Aside from reconciling these divergent research programs, this paper provides systematic evidence answering a critical question in international political economy: do banks, the crucial gate-keepers in transnational finance, follow global rules in assessing customer risk before granting access to the international financial system? The risk-based approach is the fundamental regulatory principle in this area. Both regulators and scholars believe the gatekeepers do follow a risk-based approach, but a lack of credible, systematic evidence means that no one really knows. At a time of great stress on the global economic and financial systems, this ignorance is dangerous.

Using a global field experiment, this paper tests the effectiveness of international rules mandating that banks adopt a risk-based approach to customers so as to screen out criminals' money from the financial system. It does so by randomly assigning email solicitations to 12,000 banks and other financial intermediaries/CSPs from every country in the world to one of twelve companies set up by the authors. Because these companies and their solicitations

are deliberately created to exhibit sharply varying risk profiles, we can detect respondents' degree of risk sensitivity in the different rates at which they are willing to offer corporate accounts.

Banks are large, highly regulated organizations with complex internal governance hierarchies. Thus, in our pre-registration document and pre-analysis plan we predicted that banks would manage complexity through scripts. Yet these standardized and collective templates are incompatible with the sort of highly differentiated individual judgments the risk-based approach demands of those lower-level employees who decide whether and under what conditions corporate customers can open accounts.

Contrary to the risk-based approach, the central regulatory principle of international banking, and opposed to scholars' expectations as tested in a survey to forecast the results of this study, we find that banks are remarkably insensitive to risk. Working from regulators' manuals, we engineered radically different risk treatments by varying the jurisdiction of company incorporation and the language of the solicitation approach. Yet these varying customer-risk profiles made almost no difference to banks' willingness to offer a corporate account. Banks are also surprisingly insensitive to reward: varying the gross dollar value of business turn-over in the solicitations by two orders of magnitude – from \$500,000 to \$30 million – made no significant difference to outcomes.

The significance of this finding is that shell companies with bank accounts are perhaps the single most common mechanism for engaging in money laundering, transnational corruption, tax evasion, and other related crimes (World Bank 2011; FATF 2016, 2018). At a time when the world financial system is under enormous stress, and in which governments have entrusted banks with hundreds of billions of dollars of cheap or free credit, it is more important than ever to find out whether the rules that are meant to regulate banks actually work. Our study raises considerable skepticism on this score.

In substantiating the claims made above, we first discuss the behavioral-institutionalism approach with its focus on organizational scripts and its field experimental methods. We then provide a brief explanation of the risk at issue: the criminal misuse of bank accounts held through shell companies, the single most important mechanism for tax evasion, money laundering, cross-border corruption, sanction-busting and the like. We then briefly cover the global regulatory regime designed to counter this threat. Regulators have sought to conscript for-profit financial firms into policing the banking and financial system by imposing a duty to make complicated risk judgements about prospective and actual customers. Shifting from policy to scholarship, the next section explains how this study reconciles experimental methods and institutional theories, and how as a global field experiment it advances efforts to improve the external validity of experimental social science. Although much of the experimental design is reserved for appendices, it is important to discuss the subject pool, treatments, outcomes coding, and the expert survey. Finally, we analyze the reasons for and implications of the failure of the risk-based approach and the significance of organizational scripts. The study illuminates a critical domain of international political economy and simultaneously demonstrates the utility of the behavioral institutionalist approach.

## BEHAVIOR, INSTITUTIONS, AND BEHAVIORAL INSTITUTIONALISM

The present study seeks to break new theoretical and empirical ground in the wide gap between behavioralism and institutionalism. On the one hand, behavioral approaches leverage causal identification through randomized experiments on individuals as subjects. For behavioralists, often drawing on theory and methods from political psychology, interventions targeting motives and beliefs can be randomly assigned to individuals in survey, lab, and field experiments. Studying individuals as subjects enables causal inferences across an array of important political behaviors such as voting, activism, and opinion formation (Gerber and Green 2000; Gerber, Karlan and Bergan 2009; Druckman et al. 2009).

On the other hand, institutional approaches often grounded in political economy derive hypotheses from formal theory and test empirical implications on institutions as actors through observational statistics. Important aspects of politics, such as social movements, party organization, congressional committees, bureaucratic policymaking, civil war, global regimes, and many others, occur at the level of organizations (Chong 1991; Cox and McCubbins 1993; Shepsle and Weingast 1987; Thies 2001; Fearon and Laitin 2003; Martin and Simmons 1998).

And organizational rules can tightly constrain the behavior of agents within the institutions. Indeed, many important political outcomes can only be understood as the collective actions of many individuals behaving together under institutional constraints (Olson 1965; Ostrom 1990). So even if the effects of treatments are identified at the level of individuals, the causal inferences may be irrelevant to important political outcomes in which groups of individuals, subject to organizational rules and procedures, behave jointly. This is the aggregation problem (Powell 2017), and it undergirds institutionalist approaches.

Despite the divide between these two dominant approaches in the discipline, numerous recent studies have shown sensitivity to the twin concerns of causal identification on the one hand and deeper understanding of institutional behavior on the other. An influential set of studies performs field experiments not on individual citizens at large but on political leaders in their roles as policymakers. Researchers randomly assign interventions to sitting politicians probing their propensities for racial representation (Butler and Broockman 2011, Butler 2014, Broockman 2013), favoritism toward campaign donors (Kalla and Broockman 2016), immigrant prejudice (Gell-Redman et al. 2018), biases in behavioral choice (Sheffer et al. 2018), and attitudes toward foreign aid, taxes, and oil revenues (Findley et al. 2017, De la Cuesta et al. 2019).

Because the subjects in these studies are politicians with policymaking power, the findings should reflect on institutions more broadly. However, while these studies consider subgroup effects for party and other institutional parameters, these efforts are typically made empirically and without systematic theoretical attention to how politicians might be variably constrained by institutional rules and procedures (Sniderman and Levendusky 2007). They especially neglect how those individual politician decisions might aggregate into institution-wide outcomes.

A potential solution to the aggregation problem therefore designs studies around institutions, not individuals, as subjects in which the outcomes are measured for the collectives post-aggregation. Along these lines, an important genre of field experiments that partly inspired

the present study targets firms as collective actors. These studies enable direct tests of manipulations in employment attractiveness on institutional behavior as they might be driven by racial prejudice (Bertrand and Mullainathan 2014, Carlsson and Rooth 2007, Kaas and Manger 2011), anti-immigrant bias (Oreopoulos 2011), unemployment duration (Kroft, Lange, and Notowidigdo 2013), homophobia (Tilcsik 2011), and age discrimination (Lahey 2008).

Other studies have considered how firms that provide corporate services respond to possible criminality or information about international transparency standards (Findley, Nielson, and Sharman 2013, Findley, Nielson and Sharman 2014, Findley, Nielson and Sharman 2017). These field experiments on firms as actors address the aggregation problem by treating responses from company agents as representative of the firm at large. While this approach makes no assumptions about individual agent behavior independent of corporate rules and norms, it also leaves the origin of those organizational scripts undertheorized.

Other studies address the aggregation problem more directly by studying political elites not as individuals but as collective actors, such as staffs of congressional offices (Butler, Karpowitz and Pope 2012) and local governments as units (Chen, Pan, and Xu 2015, Worthy, John, and Vannoni 2017, Jensen, Findley, and Nielson 2020). While these experiments all probe subgroup effects to learn if the treatments display heterogenous effects across different types of institutions, much more could be done to theorize about how the institutions differ substantively in their organizational rules and procedures that constrain the agents that represent them. Do the organizational scripts vary systematically across institutions in discernible ways that might be predicted *a priori*?

The aggregation dilemma is especially significant in international relations, which provides a formidable challenge to the behavioral revolution in IR capitalizing on the power of experiments, which as noted above are believed to require analysis at the level of individuals (first image) (Hafner-Burton et al. 2017: 18; Powell 2017: 265). Yet given that almost all IR theory operates at the level of institutions (second image) or systems of institutions (third image), not individuals, adopting experimental methods as suggested would seem to require dispensing with most of the accumulated scholarship in the field. It would also require building new theories capable of aggregating up from the level of individuals to international political outcomes, a task with which IR has had very little success in the past (Powell 2017).

Our study helps to resolve this apparent dilemma through the experimental study of institutions, rather than individuals. Specifically, the subject pool is made up of corporate entities (banks and intermediary firms), and the actors comprising the treatments are companies. The experiment is thus a study of relations between institutions in the international political economy, not relations between individuals. In this way, as students of politics we can have our cake (theories of institutions) and eat it too (experimental methods).

## EMPIRICAL AND POLICY BACKGROUND

Because most crime is motivated by profit – from tax evasion to corruption to fraud and human trafficking – criminals face the challenge of moving dirty money through the financial system while remaining undetected. Thanks to globalization, this dirty money increasingly crosses borders. Beyond relatively low thresholds, cash is bulky, conspicuous, and

impractical for rapid cross-border transfers, so criminals need to get their money into the banking system. Since the 1980s or 1990s, however, banks have been conscripted into a system of surveillance to prevent illicit wealth entering the financial system. Banks are now mandated to establish the true identity and the riskiness of their clients in connection with a range of crimes (terrorist financing, sanctions-busting, money laundering, tax evasion, etc.). Riskier clients should be subject to greater scrutiny by banks, and if necessary excluded.

For example, although it is not illegal for person to try to open a bank account by visiting a branch with a suitcase of cash, because this behavior is associated with the laundering of the proceeds of crime, it is a high-risk profile. Banks should subject such a prospective customer to extra scrutiny in terms of requiring proof of identity and evidence for the source of wealth, report this behavior to the authorities, and generally impose a higher bar in regulating this individual's access to the banking system. Banks that recklessly pursue profit by taking on risky customers are now routinely hit with multi-billion dollars fines, or in extreme cases deliberately destroyed by sanctions (section 311 of the Patriot Act, Zarate 2013).

A serious obstacle to this goal of banks' knowing their customers has been the use of shell companies, i.e., companies with no substantive business purpose that can be set up online for between a few hundred and a few thousand dollars in a matter of days, but that are nevertheless legal persons that can hold bank accounts and assets. Unless banks know the real person behind the shell company holding the account, the account is *de facto* anonymous, and inward and outward transfers are untraceable. Thus banks must assess the risk of their corporate customers at least as diligently as their individual account-holders.

The global rules in question are set by the Financial Action Task Force (FATF 2012, 2014), the world's anti-money laundering standard-setter and enforcer, an international club comprised of 35 of the world's most powerful countries. The FATF has spawned sister regional organizations promulgating the same rules with memberships incorporating nearly every country. FATF standards in this domain have been endorsed by and incorporated within the standards of additional international organizations, including the United Nations Security Council and International Monetary Fund. They have also been transposed into national legislation. International organizations, national regulators, and private providers produce specialized manuals (e.g. FATF 2019), which we draw from in designing our treatments.

It is important to clarify the relationship between international organizations, states, and firms with regards to these international rules. The standards we are testing have been collectively set down by the member states of the FATF. Individual states are responsible for implementing these rules at the national level, and governments' compliance is periodically and publicly assessed by the FATF. States have legislated that banks in their jurisdiction must follow international rules, centered on the risk-based approach (FATF 2012). In keeping with recent research on Regulatory Intermediary Theory, orchestration, and more general work on global governance, there is thus a chain of different actors involved in applying this international standard (Abbott et al. 2017; Abbott et al. 2015). As noted, if banks fail to comply, they may be sanctioned by national authorities. If too many banks and other firms in a given jurisdiction fail to comply, the national government may be blacklisted and sanctioned by the FATF (Morse 2020).

This regulatory chain linking an international organization to states to firms, in which firms provide “the last mile” of compliance (or non-compliance) with international standards, is common in much other international banking regulation (e.g., Basel III), and a wide range of other international agreements, from aviation safety to environmental standards (Abbott et al. 2017). The bottom line is that if banks and intermediary firms do not assess customer risk, no one else will.

Because governments and international organizations have generally tried to measure the effectiveness of the rules only by reading laws and regulations on the books, rather than by trying to ascertain the standards applied in practice, nobody really knows if they work. Anecdotal evidence, however, gives strong grounds for concern.

Recently, a string of the world’s biggest banks have been hit with multi-billion dollar fines, and in some cases criminal convictions, for deliberately or negligently aiding their customers’ tax evasion, sanctions busting, money laundering, and rigging of key markets setting inter-bank interest rates and foreign exchange rates. A survey of British banks by the regulator showed that over half failed to perform the mandated extra checks on high-risk customers, with some explaining away lapses because these customers were “from a respectable family” or were “very nice” (FSA 2011: 4, 32). The release of the Panama Papers confirmed that international banks were intimately involved in much of the misconduct revealed (Obermaier and Obermayer 2016). Further leaks in 2020 of banks’ private risk reporting to the US Treasury Financial Crimes Enforcement Network raise worrying questions about the efficacy of the rules designed to keep dirty money out of the financial system. But anecdotes may not aggregate into representative data, so we designed a systematic audit study and experiment to generate unbiased findings in expectation.

## THE METHODOLOGICAL CONTEXT: EXPERIMENTS

Proponents of experimental methods have made far-reaching critiques relating to the perils of advancing causal arguments on the basis of observational data (Druckman et al. 2006; Gerber and Green 2012). By conducting a global field experiment, we can test the effectiveness of international rules in a manner that obviates the problems of endogeneity and selection bias that have bedeviled observational research on global governance (Downs et al. 1996; von Stein 2005). In this way, we can get to grips with what in some ways is the fundamental question of international politics: whether a law-governed order can function in an anarchical international system.

Some scholars have used survey experiments to test public opinion related to international rules (e.g. Tomz 2007). Yet in the context of our study, there are serious problems in using survey experiments to reveal behavior where non-compliance is common, widely seen as inappropriate, or actually illegal (see Findley et al. 2017).

Field experiments are distinguished by the realism or naturalism of the treatments, outcomes, and settings, which together provide confidence in the ability to generalize from the results of the study to the wider world, i.e., stronger ecological validity (Gerber and Green 2012). Because every bank that is connected to the international wire transfer network (SWIFT) is included in our study, the entire population of interest, the study also has uniquely strong external validity. As there is no definitive list of intermediary firms, as they are often

unregulated, so it is important to acknowledge that our subjects are a convenience sample of firms with Internet profiles, yet with roughly 7,000 intermediaries from more than 170 countries, the sample has a far wider representation than most experimental studies.

This global coverage is in contrast to the tendency of political science field experiments to focus on single countries (e.g., Chong et al. 2016 in Mexico; Michelitch 2015 in Ghana; Gottlieb 2016 in Mali; Chen et al. 2016 in China; Avdeenko and Gilligan 2015 in Sudan; Fearon et al. 2015 in Liberia; Beath et al. 2016; Lyall 2019; Lyall, Zhou, and Imai 2020 in Afghanistan), especially the US (e.g., Gerber et al. 2015; Gerber et al. 2016; Nyhan and Reifler 2015; Green et al. 2015; Kalla and Broockman 2016). Often, in fact, these studies take place in one quite small region of a single country. But for questions of global governance and international relations, we need field experiments that include a large number of countries and that target the transnational phenomena of interest. However, in her review of the field, Hyde notes that, for all their popularity in other segments of political science, experiments are still rare in international relations, and field experiments rarest of all (2015: 413; see also Jensen et al. 2014).

Our field experiment provides other advantages. The random assignment of treatments drawn from regulators' risk-based mitigation manuals enables causal identification and thus well-founded internal validity. Further, participants do not self-select, or even know they are being observed, and thus the study is conducted with a high degree of naturalism in the treatments and outcomes. These features mitigate the danger that participants' responses will be tainted by social desirability bias that could reasonably be expected to arise by simply asking banks or corporate service providers in a survey whether they adhere to due-diligence rules (Findley et al. 2017). As such, together with its coverage of all the world's relevant banks, the naturalism of the exercise, and the fact that participants do not self-select, it has strong external validity compared to most other political science experiments. Thus, the study is well suited to answer big questions about the effect of global rules in an anarchical system.

## INSTITUTIONAL RESPONSES AND SCRIPTS

Given the claims above, it is vital to establish that our data (correspondence between banks and firms) do in fact represent institutional, rather than individual output. Banks' procedures for taking on new customers are set down in manuals, checklists, training guides, and standard operating procedures precisely to remove individual discretion and set a consistent institutional response. In this sense, though it may be an individual fielding our email ([johndoe@hsbc.uk](mailto:johndoe@hsbc.uk)), what we get is an institutional response. Banks spend much time and effort specifically making sure that their policies and procedures are not implemented in an idiosyncratically individual manner; if compliance officers are making decisions on which customers to accept and what identity documents are required on the fly, or according to individual whims, the system has failed. This explains why many email responses were often highly generic, clearly written according to common templates.

Interview material and guidelines produced by international organizations, regulators, and financial institutions specify that the primary unit of interest for corporate accounts is, not surprisingly, the relevant corporation itself. Certainly, establishing the identity of the individual beneficial owner(s) and other corporate officers is crucial in terms of the



international rules, but this fact does not change the inter-institutional character of the relationship. These individuals are of interest precisely because of their role within the institution (company).

Scholars suggesting the need for a behavioral revolution (among many others) have pointed out the shortcomings in modelling institutions as a single, rational individual writ large, or seeking to explain institutions as simply the sum of individual rational decisions (Powell 2017). Rather than seeking a solution at the level of the individual, according to the psychological maxims of behavioral economics, we change the level of analysis to institutions themselves. Under this view, institutions are more than just the sum the individuals that make them up, having emergent qualities (March and Olsen 1989). Furthermore, the logic of decision is much less the deliberate mean-ends calculations of the standard rational expected utility model. Rather than each decision being taken separately, decisions are clustered and categorized into generic classes, and they are linked with equally generic standard responses to form scripts. Partly this process of script-writing is deliberate, and sometimes it may aid efficiency. But just as often it reflects the informal, uncoordinated and anticipated congealing of habits and routines, which may be deeply pathological for organizational function (Barnett and Finnemore 2004). In this sense, institutional scripts are not functional, i.e. they do not exist because of the benefits they may create (Elster 1995).

## RESEARCH EXPECTATIONS

In contrast to this institutionalist account focusing on institutionally embedded norms and practices, the dominant regulatory principle of the risk-based approach bears important similarities with a rationalist account of behavior. Actors should calculate and compare probable utilities associated with alternative courses of action, choosing that which promises the highest expected utility. The alternative courses of action in this case are whether to respond an email solicitation, whether to offer an account, and then, crucially, what level of scrutiny to apply to a prospective corporate customer. The main trade-off is between the riskiness and rewards associated with the customer. As described above, our treatments manipulate risk, and our three levels of business turn-over (\$500,000, \$3 million, \$30 million) manipulate reward, as larger turn-over is associated with higher fees.

To what extent are these incentives institutional, or individual? Both banks and regulators seek to align the two. Individuals bankers may also be prosecuted for excessive risk-acceptance, though it is more likely to lead to being fired. One senior banker described the standard response to such lapses as to “reach out and throttle an underling” (Author interview). Major scandals can lead to whole compliance departments being sacked. Conversely, it is standard practice that bankers who bring in more business or more lucrative business will be rewarded with bonuses and promotion. The net effect of these contrary pressures is essentially unknown – this is the basic uncertainty concerning the effectiveness of regulations that this study seeks to resolve.

In contrast, in terms of institutional scripts, the larger and more regulated the institution, the more likely organizational scripts will shape the behavior of agents. It is commonsensical that while collaboration between just a few individuals can be coordinated on an ad hoc basis, the larger the number of those involved, the greater the diseconomies of scale applying to

informal approaches. Governance and coordination of large numbers requires some measure of hierarchy and specialization (Williamson 1975). Socialization in accord with scripts is much more likely to occur in large, durable organizations than in smaller, ad hoc and transient groupings (March and Olsen 1989).

Given the above, the expectation of a standard rationalist logic and risk-based approach is that banks and bankers will be highly attuned to risk and reward, and hence that the various experimental manipulations will generate large and significant changes in the outcome variables (response rate, refusal, compliance, and non-compliance). In contrast, to the extent that organizational scripts are important, complex judgments weighing risk and rewards will give way to standardized, internally-set responses that are unresponsive to variations in external contacts.

Intermediary firms that set up shell companies and liaise with banks to procure corporate accounts are much smaller, perhaps even just a single person, and less regulated, or in the United States and some other countries, unregulated (Findley et al. 2014). This means that they may have less need for scripts, while they also face different incentives compared to banks. They are much less scrutinized by regulators and are far less likely to attract penalties even when caught out aiding and abetting downright illegal activities (US Senate 2010; Global Witness 2014). A rationalist logic might thus suggest that intermediaries should be relatively insensitive to customer risk, given that they face a much lower likelihood of suffering the consequences of doing business with a corrupt official or money launderer. However in terms of reward, these firms should be at least as sensitive to banks in differential responses to more lucrative prospective customers.

An organizational scripts logic, however, would predict that the smaller, nimbler, and less-regulated intermediaries ought to be much more responsive to risk if for no other reason than to avoid the moral quandary of aiding and abetting high-risk customers who may undertake criminal activity. Intermediary firms may also fear punishment from regulators (even if they fly under the radar to a large extent) and their less-scripted organizational structures may enable them to respond to risk more fully.

Thus the rationalist, risk-based approach and the organizational scripts accounts have directly opposite predictions on how much banks' and intermediaries' response, refusal, and compliance will vary in response to different risk treatments.

## ESTABLISHING EXPECTED RESULTS VIA EXPERT SURVEY

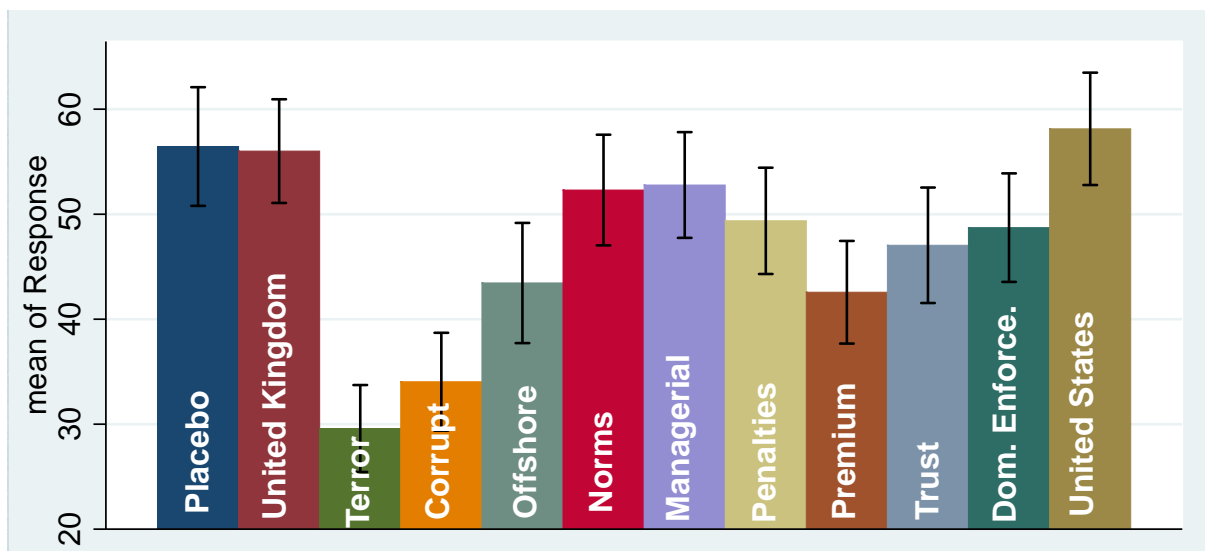
Is it accurate to suggest that the firm baseline presumption for banks is acute sensitivity to customer risk, and hence that findings of script-driven insensitivity to risk would represent a novel and surprising result? Rather than relying on assertion, or a possibly cherry-picked sampling of the literature, we conducted a survey of scholars to establish baseline expectations (see Milgram 1963). We surveyed 116 international political economy scholars.

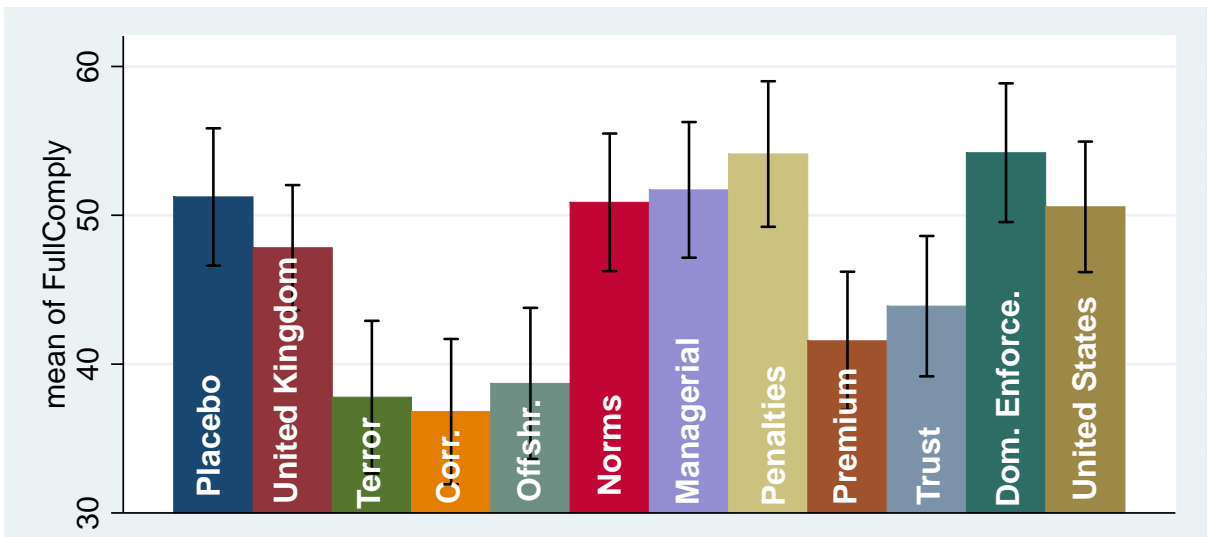
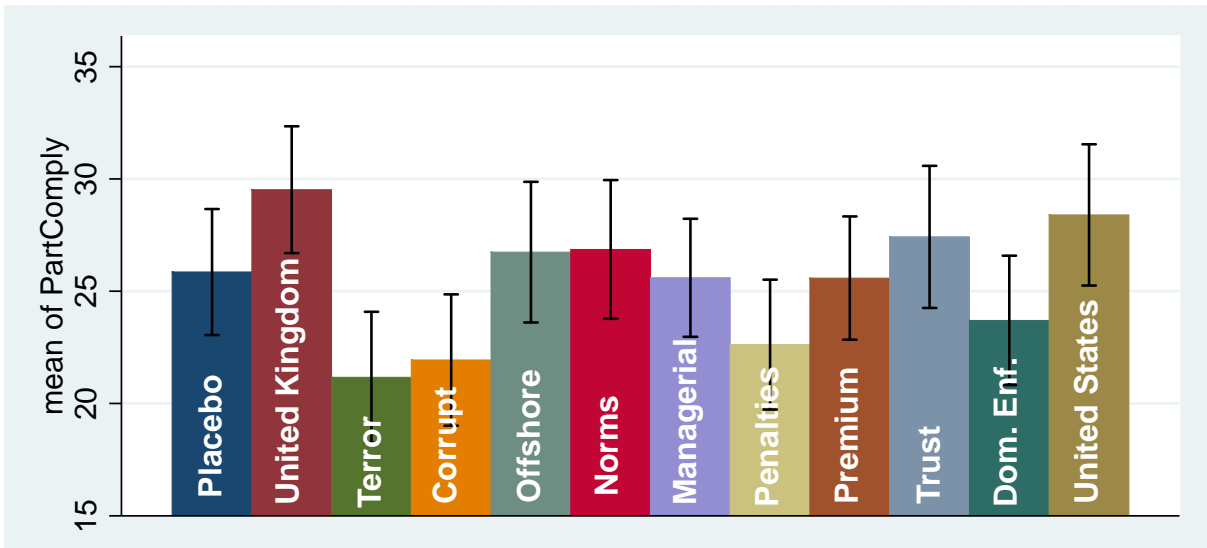
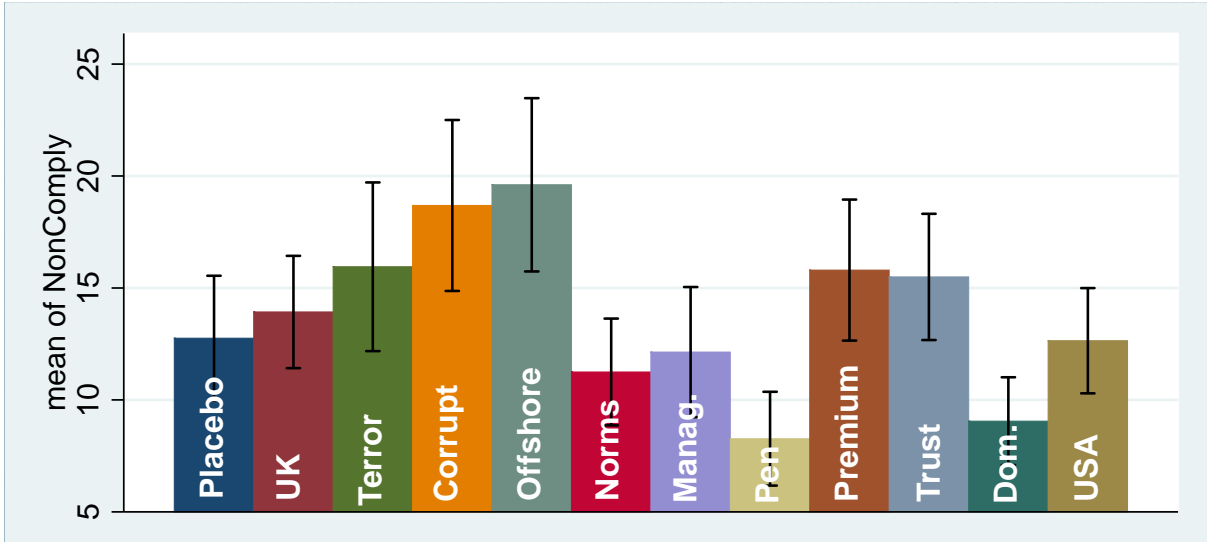
In the survey, we presented respondents with the study's experimental conditions and asked them to estimate the likely response rates, as well as levels of compliance. More specifically, we asked them to estimate rates across a five-point categorical outcome scale (non-compliance, partial compliance, compliance, refusal, non-response). After this survey and

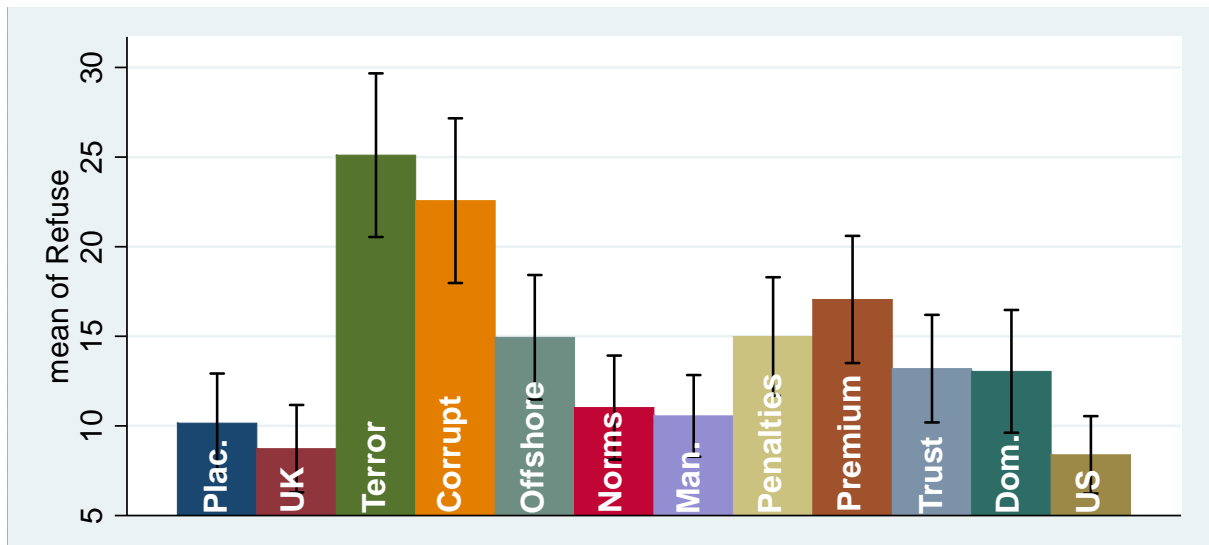
before depositing our pre-analysis plan, we decided to collapse the partial compliance and non-compliance categories (by summing the responses). However, the main distinction between compliance and partial compliance involved the notarization or certification of the identification documents. We thus softened the definition of compliance to encompass all ID demands of the beneficial owner. For purposes of transparency, we include the original part-compliance and non-compliance responses from the scholars in the survey. See Figure 1 for IPE scholars' expectations of bank behavior across the outcomes. Note that refusals, non-, part-, and full compliance are rescaled as a percentage of anticipated responses. Point predictions are shown as average predicted proportions; 95-percent confidence intervals are shown by error bars.

As displayed, many of the sizable expected effects are significant statistically compared to the placebo condition. The key takeaway from the scholar survey results accords with the prediction of the rationalist risk-based response model: the strong expectation that banks will be highly sensitive to risk.

**Figure 1: Results from Expert Survey Displaying IPE Scholars' Average Predictions of Study Treatment Effects**







## RESEARCH DESIGN

To probe the extent and variability of scripts across treatment conditions and responding institutions' organizational types, the experiment is based on legally incorporating a variety of companies and having agents working on behalf of these shell companies make email solicitations to thousands of banks and intermediary firms around the world asking to set up a bank account for these companies.

The study does not involve active deception: we have legally incorporated the shell companies seeking bank accounts, and we have a sincere interest in opening bank accounts and making wire transfers. The research design pre-registered the eventual acquisition of multiple bank accounts and the international transfer of funds between them. Many of the banks and intermediary firms will thus add to their profits by working with us and our companies as clients. This sharply diminishes ethical concerns about the potential wasting of banks' and intermediaries' time and effort. Additionally, lying to banks in writing can constitute bank fraud, a criminal offence, so untruths in communication are formally illegal in many jurisdictions. Thus, the experiment was designed to minimize deception, and no false information was involved in the correspondence. It is important to note, however, that researchers did not disclose to banks and intermediaries that the inquiries were part of an academic study. Finally, we delete the names of all individual banks, firms and employees, and otherwise thoroughly de-identify the data, to protect participants from any adverse consequences from the study.

The outcomes of interest are, first, whether banks and other intermediaries reply to our solicitations, second, whether they are willing to offer accounts and, third, whether they follow international rules in verifying the owner of the company. This is in line with our earlier discussion of private firms as the proximate agents of compliance with rules set by international organizations, with states playing an intermediary role linking the two.

The experiment creates different treatments in two ways. The first is to set up shell companies of varying risk profiles, as signaled by the jurisdiction of incorporation. According to international rules, companies and individuals acquire the risk profile of their home

countries. For example, a company from a country with a high perceived corruption prevalence should itself be judged a high corruption risk. The second is to insert variations into the language of our approach email to the banks. In expectation, the random assignment of thousands of banks to different company risk profiles and approach emails will enable us to isolate the causes of banks' compliance and non-compliance with international know-your-customer (KYC) rules.

## SUBJECT POOL AND NON-RESPONSE CHECKS

The first step in our study was to compile a list of the world's banks and their contact details. The most complete list is that of the Society for Worldwide Inter-bank Financial Telecommunication (SWIFT), the organization that allocates codes for the message system that underpins international bank wire transfers. SWIFT has information on banks (including subsidiaries and branches), in almost every country of the world. Using this list, which comprised our sampling frame, research assistants obtained the relevant email address for roughly 5,000 banks. These 5,000 banks included all headquarter banks on the SWIFT list and a random sample of bank branches in each country. We made three approaches to each bank and branch, for a total of roughly 15,000 approaches, with a wash-out period of three months between each approach to minimize the risk of detection.

The pool of intermediary firms was collected through comprehensive Internet searches. The final sample includes roughly 7,000 firms from more than 200 jurisdictions that set up and sell shell companies and also commonly assist in establishing bank accounts. We likewise made approaches to each of these providers in three randomly segmented tranches of one third each, again with a three-month wash-out period, for a total of approximately 7,000 approaches to the intermediaries.

An important aspect of the findings discussed below is the number of banks and intermediary firms simply not making any response at all to the email solicitations. Ahead of the experiment itself, we sent "non-response" checks to all banks in the sample in order to gauge which banks are responsive at baseline and which are not. It is difficult to sort out exactly what non-response means (a matter we consider in a dedicated section below); the non-response checks allow us to gather some initial information on the banks and also include that information as blocking criterion to better balance the sample and a covariate to improve precision in estimation. We used five different email scripts, randomly assigned to the banks and intermediaries, for the non-response checks. We asked the firms if they served international customers and if the corporate account could later be upgraded to a merchant account to process credit-card information if needed. The non-response checks revealed that roughly 75 percent of banks and roughly 60 percent of the intermediary firms did not respond to our innocuous inquiry, setting critical baseline levels of subject responsiveness to any email inquiry.

## EMAIL APPROACHES

The standard approach template was a message from a representative agent (a researcher) corresponding on behalf of the legally incorporated company, stating the company name and

jurisdiction of incorporation. In the approach email, the representative specified that the company is a consultancy concerned with two of five randomly assigned business areas drawn from the actual consulting topics of the beneficial owner, who is one of the principal investigators of the study (the five topics are development assistance, education and training, impact assessment, feasibility studies, and information and communication technology). The RA disclosed the amount of recent business randomly assigned at “more than” three levels: \$500,000, \$3 million, and \$30 million, the last of which represents the beneficial owner’s total revenue in research grants in the preceding years and thus reflects truthful information. The RA explained that the company would like to establish an account with the bank in the local currency and that the account needed to be able to receive and send international wire transfers. The email then asked how much money establishing the account would cost, how long the process would take, and, crucially for judging rule compliance, what verification documents would be required to set up the account. The email approaches to the intermediary firms requesting bank accounts adopted the same language with slight modifications to match the intermediary context.

It is important to reiterate that all of the information in these base emails was truthful: researchers used their real names; they represented actual, legally incorporated companies based in the stated jurisdictions; and the information presented related truthfully the substance and financing amounts of prior projects by the beneficial owner/author.

Approaches were made via email using specially created email accounts with domain names reflecting the name of the company formed. Company names were generated using random four-letter acronyms and thoroughly vetted to verify that the monikers were not already well known. Communicating via email, rather than phone or video chat, allows for identical treatments, more accurate coding, and a comprehensive record of correspondence.

Email is a standard medium for all but the highest net worth individuals and companies and has been a documented approach across more than 20,000 accounts in one bank alone (US Senate 2014: 83,87,88). We vetted the emails with several parties, including practitioners in the incorporation and banking industries, to ensure that they actually address the key issues of concern to banks and were realistic.

## TREATMENTS

The variation between control and treatments arises from the different jurisdictions of incorporation and different language in the approach email. We employ these conditions in both the direct approaches to banks and those via intermediary firms in a fully crossed factorial design involving five information conditions – (1) invoking the international legal standards, (2) noting penalties for non-compliance, (3) referencing norms of compliance, (4) recognizing domestic central bank laws, (5) and demanding banking secrecy – for a combined total of 32 unique conditions. The absence of any of the language conditions is the baseline comparison and was assigned to half of all subject approaches.

We vary the language embedded in the approach email to inform (or, at the very least, prime) subject banks and intermediary firms about international legal standards or a demand for secrecy. These treatments are designed to test the causal impact of providing information to banks and firms about international law of which they may be unaware or, at a minimum, to

test the causal effects of priming the banks and firms about the international standards on which they may have already been briefed. Both of the implications – direct information or prime – are consistent with the argument that information about international standards has a causal effect. Each of the information conditions is assigned independently in a fully crossed factorial design (32 different combinations) with each condition roughly capturing the following ideas.

- Standards: Invokes the requirement that institutions demand documents establishing the identity of the beneficial owner.
- Norms: Expresses interest in following global standards in order to establish and maintain a strong reputation.
- Penalties: Acknowledges that there may be penalties for failure to follow global rules.
- Domestic Enforcement: Notes that the domestic central bank requires identity disclosure and that noncompliance may result in penalties.
- Secrecy: Makes clear that the owner will not disclose his identity because secrecy is of the utmost importance.

We also include several treatments based on the jurisdiction of incorporation. International and national standards mandate that banks assess the riskiness of a potential client in part based on their country of origin. Thus, a company from a country ranked poorly on Transparency International’s Corruption Perceptions Index should be assigned a higher risk than a firm from a low-corruption country, all else equal. Countries have varying risk profiles for money laundering, terrorist financing, being a tax haven, and other concerns. Crucially, according to the international standard for the risk-based approach, high-risk solicitations should get fewer replies, more refusals, and more diligent application of customer due diligence by banks and intermediaries than low-risk solicitations such as our placebo email.

The first treatment is designed to learn whether soliciting offers for bank accounts from companies formed in the United States, the dominant and arguably hegemonic country in the governance of the global economy, affects the response and compliance rate relative to our placebo jurisdictions, Australia and New Zealand. The United States government has been particularly aggressive in applying extra-territorial law-enforcement measures and tax regulation to pierce the corporate veil and scrutinize foreign corporate bank accounts. We formed one company each in Delaware and California. Delaware is home to one in three publicly listed US corporations, but it is also often accused of being America’s secrecy haven and leading a race to the bottom in incorporation standards, hence the “Delaware effect.” We formed a second company in California, another of the most popular incorporation jurisdictions in the United States, which has in the past been reputed to have higher standards than Delaware (GAO 2006).

The second set of treatment companies were incorporated in the UK. In addition to London’s being a leading financial center, the British government has sought to take the lead in enforcing corporate transparency initiatives in the G7 and G20. In 2015, the UK Parliament tightened the country’s Know Your Customer regulations.<sup>1</sup>

---

<sup>1</sup> We expect that the US and UK treatments will be considered differently by banks and therefore we employ them as separate treatments. With that said, we acknowledge that banks could see these jurisdictions as



Our offshore treatment aims to learn whether approaches to banks from companies incorporated in stigmatized tax-haven jurisdictions are more or less likely to elicit a compliant response from banks. Offshore centers have been targeted by various multilateral regulatory initiatives and have suffered extensive reputational damage with adverse media coverage. We formed companies in the British Virgin Islands, the most popular single offshore shell company jurisdiction with around 500,000 active companies on its registry, as well as the Seychelles, another stereotypical offshore center hosting more than 100,000 shell companies.

High corruption risk is signaled by companies incorporated in two countries perceived to have major corruption problems: Papua New Guinea (ranked 137<sup>th</sup> of 180 countries on the 2019 Corruption Perceptions Index) and Bangladesh (ranked 146<sup>th</sup>). This signal is in accord with the FATF guideline that the risk posed by companies should be in part assessed by the risks in their country of incorporation.

On the same principle that, according to international standards, the country risk is transferred to its corporate citizens, we formed a trust and are still in the process, two years later and counting, of forming a company in Pakistan.<sup>2</sup> We expected both entities to signal a high terrorism-financing risk. The two conditions also give us leverage on the potential differences in risk between trusts, which are believed to be especially high-risk vehicles, and limited liability companies. Pakistan has earlier played host to the Taliban and leadership of al-Qaeda, as well as a variety of other terrorist groups. Moreover, Pakistan ranks fifth on the 2017 Global Terrorism Index.

The placebo condition was an inquiry originating from either the Australian or New Zealand company and containing only the language from the base email with no additional text from the information treatments. Half of all inquiries were assigned to the placebo information condition and one fourth to the placebo jurisdiction condition (another quarter were assigned to the UK companies in anticipation of pooling UK with Australia and New Zealand as an additional placebo jurisdiction). This design choice directly addresses the multiple-comparisons problem by increasing the statistical power of comparisons to the control group whose parameters can thus be estimated with higher precision. The response and compliance rates to the basic placebo control emails form a baseline against which to measure what difference, if any, various risk profiles and information prompts make to banks' willingness to open an account and enforce know-your-customer rules.

Variations in the approach language test the mechanisms and observable implications associated with various theoretical schools of thought (though we hasten to acknowledge that we certainly cannot test the theories themselves directly in their entirety).

---

fundamentally similar in some respects and therefore may end up concluding that we should collapse them for purposes of analysis.

<sup>2</sup> The fact that the company was still in the process of being formed was disclosed to banks and intermediaries in correspondence.

## OUTCOMES AND CODING

Coding was completed on the basis of the email correspondence. The first outcome could be no response at all, which could indicate either disorganization, a commercial judgment that the inquiry is not worth answering, or a form of “soft screening” risk management to turn away undesirable potential clients. Given the high proportion of non-response revealed in the results, we devote a separate section to this in the discussion of the findings.

The second outcome was a refusal to do business, with or without a reason. This was coded as an independent, nominal category.

Third, banks and providers could indicate a willingness to open an account, but require verification documents for the company’s beneficial owner. This was coded as compliant. This coding simplified the documents required by FATF rules (FATF 2012, 2014, 2019). The FATF’s required documents are proof that the company exists (e.g. certificate of incorporation, which shows jurisdiction of incorporation), a copy of the company’s by-laws (articles or memoranda of incorporation), and its registered address. Most important, and therefore key to our coding rules, is proof of the beneficial owner’s identity, evidenced by a verified copy of a government photo identity document, usually a passport, or an in-person visit to the bank to establish identity. If the bank or intermediary required a passport copy but did not specify that it must be for the beneficial owner (or might instead apply to a nominee director or agent, both of which can enable anonymous banking), researchers followed up to reduce the ambiguity. Banks and firms coded as non-compliant failed to specify required ID of the beneficial owner after repeated follow-up.

Banks and providers that are willing to open an account without this supporting documentation material to establish the owner’s true identity were thus coded as non-compliant, as they are breaking international rules by offering what amounts to an anonymous or untraceable bank account. Without ID of the beneficial owner, it is difficult or impossible to find the real person who controls the account via the company and thus that person can commit financial crimes with impunity.

Hence there are four outcomes: no response, refusal, compliant, and non-compliant. Emails were independently coded by two research assistants for accuracy and consistency, with discrepancies adjudicated by a third, senior coder.

## DATA ANALYSIS

As pre-registered and deposited prior to data collection in our pre-analysis plan, data analysis employs multinomial probit models comparing each treatment condition to the control, holding constant all other experimental condition and the blocking covariates. The multinomial probit analysis enables consideration of treatment effects alongside covariates on all four outcomes simultaneously.

As pre-registered, we check robustness with selection models. The logic behind the selection model is that three of the outcomes – compliance, non-compliance, and refusal – can only be observed if a subject responds. As such, response acts as a hurdle or gate-keeping stage in estimation. In our case, we do not have different variables to identify separate stages of a selection model, so we draw on the estimation technique developed by Sartori (2003), which

allows the same variable to identify both stages. The substantive results from the selection models are consistent with the displayed results from the multinomial probit analysis.

Below we discuss the three most important elements of our findings: the high proportion of non-responses, the uneven effectiveness of the Know Your Customer rule, but most especially the insensitivity to risk.

#### NON-RESPONSE: SOFT-SCREENERS OR NEVER-RESPONDERS?

Excluding auto replies, the response rate for banks to our email solicitations was 22.7 percent (i.e. including compliant, non-compliant and refusal responses), 3,532 responses from 15,563 solicitations. The earlier response checks had predicted roughly this level of replies. Nevertheless, the high rates of non-response demand an explanation, both on their own account, and in terms of what it means for the conclusions that can be drawn from the replies that we did receive.

The non-response category poses significant inferential challenges because non-response could be capturing a wide array of information. The most important question is whether the banks and firms not replying to the solicitations were “soft-screeners,” i.e. deliberately managing risk by taking decisions to ignore the messages, or “never-responders,” who simply never respond to any international email approaches, perhaps because they are unable or unwilling to cater to foreign customers. Soft-screeners are relevant to our study because they could potentially provide access to the international banking system in a compliant or non-compliant manner. Even if they did not reply to our solicitations, other different approaches from prospective foreign clients might induce them to do so.

On the other hand, because “never responders” never provide access to the international banking system, by definition they are less relevant to a test of the conditions under which it is possible to access to that system. While a full understanding of non-response is elusive, we did take some additional steps which suggest that the large majority of the non-responses are “never-responders.”

As noted, ahead of the experiment itself, we sent non-response checks to all banks in the sample in order to gauge which banks are responsive at baseline. The non-response checks allow us to gather some initial information on the banks and use that information as blocking criteria to better balance the sample. We used five different email scripts, randomly assigned to the banks, for the non-response checks. We asked the banks if they served international customers and if the corporate account could later be upgraded to a merchant account to process credit-card information if needed. The non-response checks revealed that roughly 75 percent of banks and 55 percent of intermediaries did not respond to our innocuous inquiry. In the multinomial and selection analysis, we include a covariate for response in the initial non-response check, which controls for baseline response propensity and improves precision of estimation.

Also prior to the launch of the experiment, we randomly sampled 95 of the banks that did not respond to these non-response checks and called them by phone to seek to understand why they failed to reply. Many banks lacked a phone number and we were unable to contact them. For the rest, we used a Google Voice number to contact each bank, and each were called

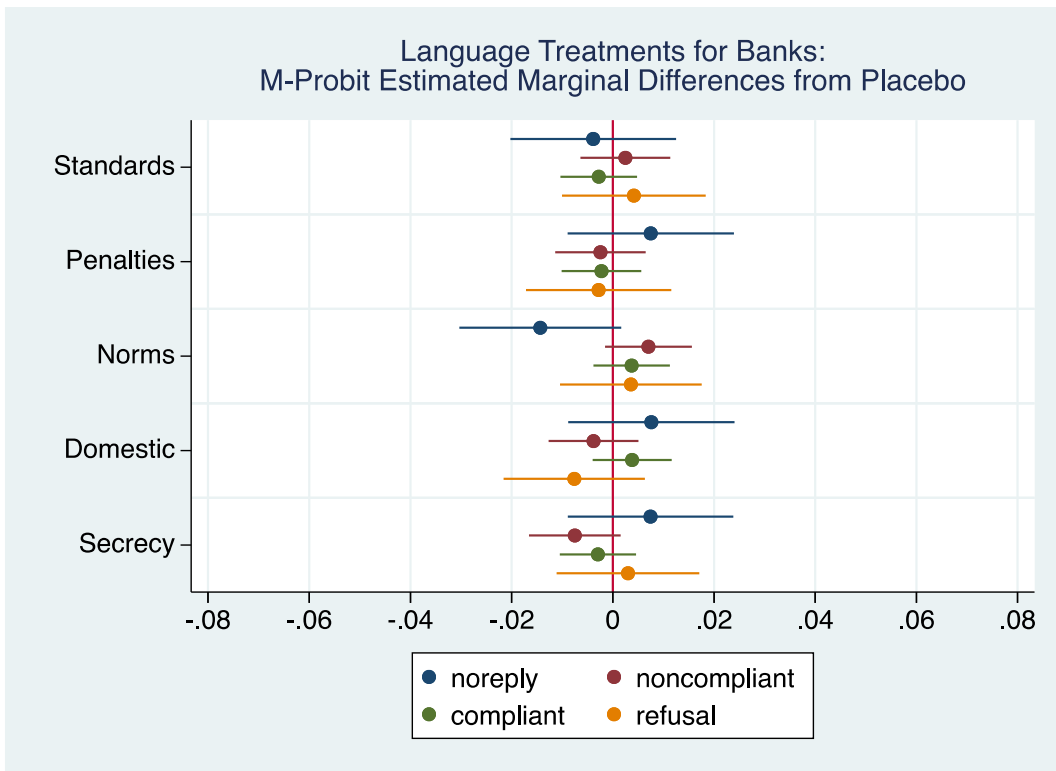
between Monday and Friday, from 9 am to 5 pm in the institution's local time. When contacted, many calls continued to ring with no answer. Several banks had the call dropped before any contact was made. A large majority of banks (over 75 percent) could not be contacted on the phone number they provided on their website.

Of the banks that were contacted successfully, several had logistical issues, such as a language barrier when English was not the primary language used among the staff. We received a few responses in which the bank chains claimed inability to open accounts internationally or required in-person verification. They were insistent that these were new international restrictions placed on the institution. These banks were located in Papua New Guinea, Switzerland, and Austria. The phone-call approach primarily revealed that many international banks did not answer their phones.

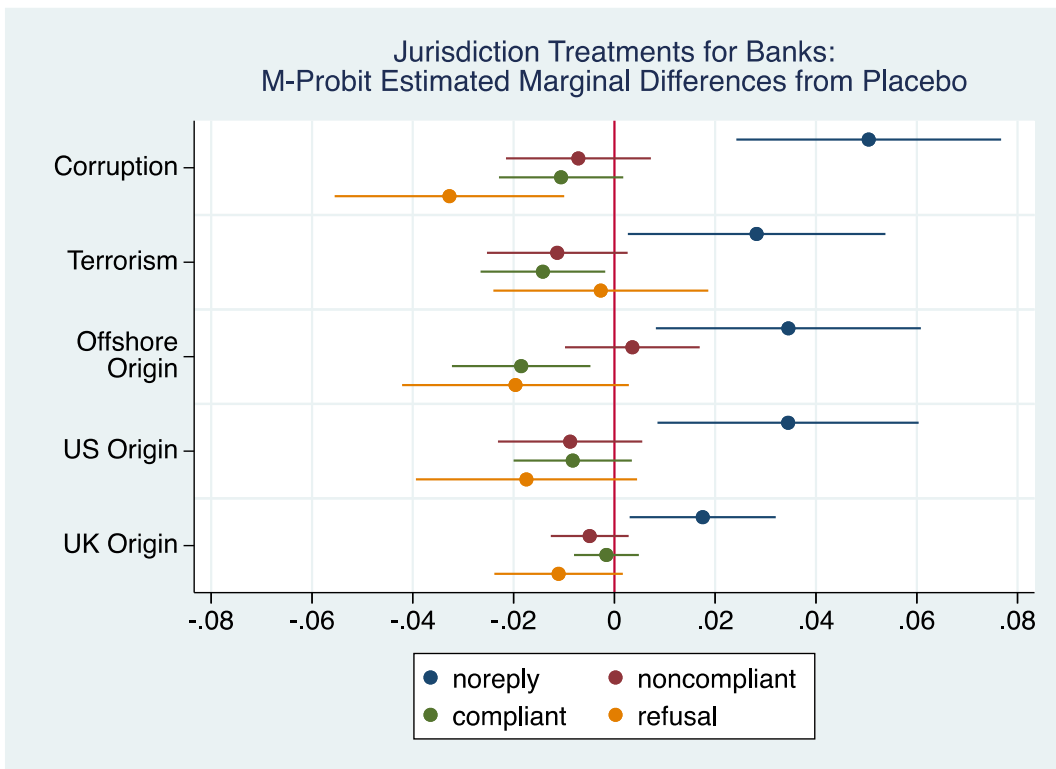
### SCRIPTS NOT RISK: THE FAILURE OF THE RISK-BASED APPROACH

Beyond the baseline result of having a viable number of replies, the second major endorsement of the research design is that in some instances banks did react to at least some treatments, though effects for information treatments (see Figure 2) are essentially null. Jurisdiction treatments (see Figure 3) are sometimes significant statistically – especially for non-response – but always small substantively. In this sense, enough banks are reading and responding to the different emails to show some treatment effects. But those effects are minor in substantive terms.

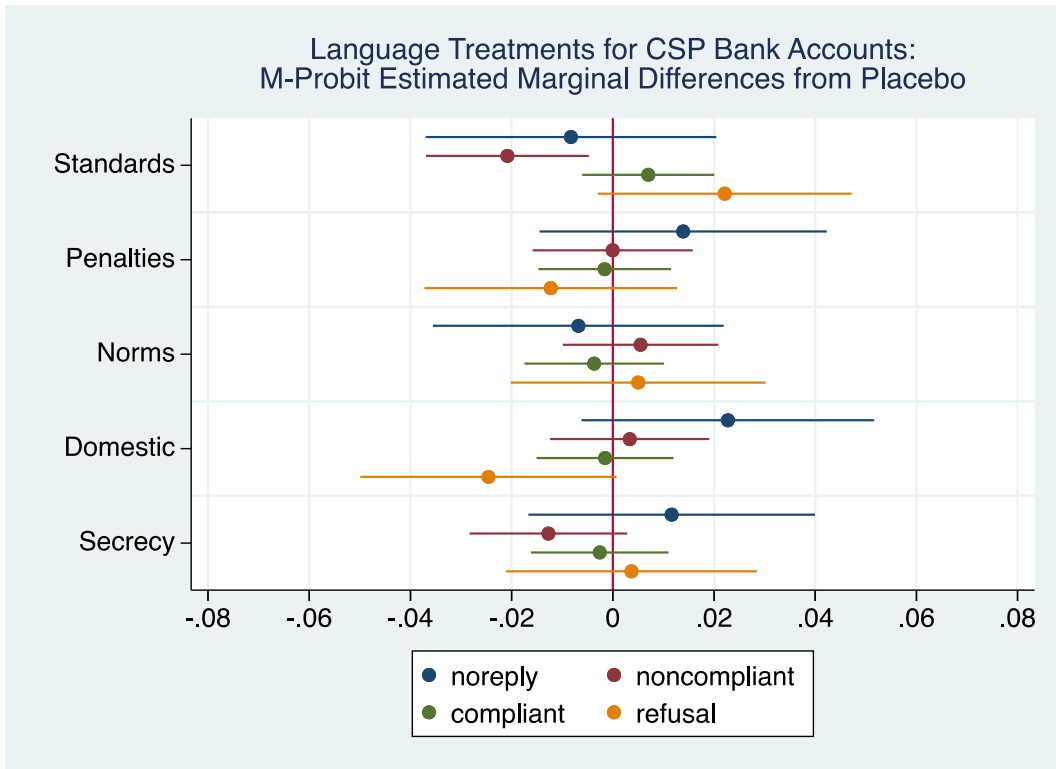
**Figure 2**



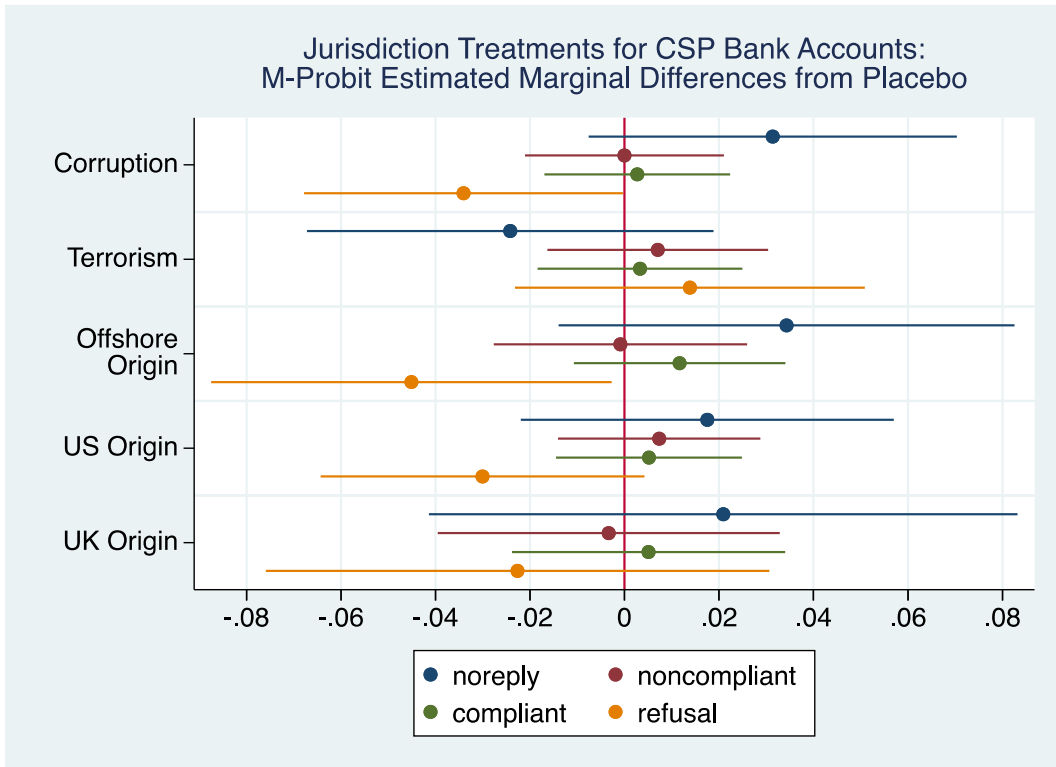
**Figure 3**



**Figure 4**



**Figure 5**



Nevertheless, with only minor exceptions, the major conclusion is that banks and corporate service providers are surprisingly indifferent to customer risk. With only rare and generally minor exceptions, even strong and obvious differences in customer risk profile generally did not make a significant difference to banks' refusal, compliance, and non-compliance rates. All of the jurisdiction treatments caused increases in the banks' non-response rate compared to placebo, but those differences are substantively quite small: five percentage points or less. We know that this finding is indeed surprising to scholars, because our expert survey, as reported in Figure 1, clearly predicted many more and much larger treatment effects than were actually observed. We observe even fewer significant differences for intermediary corporate service providers (Figures 4 and 5), though the smaller sample size for these subjects increases the size of confidence intervals and necessarily makes estimation less precise.

Though we did not survey regulators, the experimental results cut directly against the central principle of the global regulatory regime designed to combat transnational financial crime: as a general rule, banks and intermediaries/CSPs do not assess customer risk in granting or withholding shell companies access to the international banking systems.

This relative insensitivity to risk is in line with the expectations of institutional scripts: the idea of generic, standardized responses to fulfilling recurrent and repetitive organizational tasks, such as the procedure for opening new corporate accounts, that are resistant to the imposition of finely-grained differentiation and considerations of individual judgment. Regulators have sought to transfer complex and delicate questions of differentiating between customers to establish different levels of customer risk and differential treatment to banks. The evidence in favor of scripts suggests that regulators have mostly failed in this objective.

The highest risk jurisdiction treatment, Terrorism, caused statistically significant higher Non-Response and a decrease in refusals compared to the placebo jurisdictions of Australia and New Zealand, but the increase in non-response is only three percentage points (from a base of roughly 78 percent). We interpret this as a small proportion of banks that shifted from compliance with standards in placebo jurisdictions toward no response to inquiries from Pakistan, possibly as a form of soft compliance with international standards. A roughly similar pattern is seen in banks' increased non-response and decreased compliance in the Offshore condition originating in the Seychelles and the British Virgin Islands.

The Corruption condition – originating from Papua New Guinea and Bangladesh – caused a larger but still relatively small five-percentage-point increase in the non-response rate, which might be partially accounted for by the decrease in the refusal rate. Again, this might possibly be interpreted as minor soft compliance on the part of some banks. The U.S. and U.K. jurisdiction treatments caused small but statistically significant increases in non-response but had no other significant effects on the other outcomes.

The results suggest that a small proportion of banks respond to risk by ignoring would-be high-risk customers. This is in contrast to the reaction that might be expected, whereby high-risk customers would get more scrutiny relative to low-risk customers. Thus it appears that neither banks nor intermediary CSPs apply a risk-based approach to customer due diligence; KYC procedures are roughly constant despite widely varying levels of customer risk. This finding is especially noteworthy remembering that our treatments are generally directly

derived from the FATF checklist of risk indicators that are meant to set the standards for banks; to this extent, they are textbook, obvious risks and hence strong treatments.

Aside from our language and jurisdiction treatments, we also randomized the amounts of money our shell companies specified for their turn-over. The three figures provided: \$500,000, \$3 million or \$30 million over five years (the authors have in fact been involved with projects of these values, in accord with our no-deception rule for dealing with banks). The intuition here is fairly straight-forward: banks may treat rich customers differently than poorer ones. The greater attractiveness of a \$30 million turn-over firm might mean that banks are keener to accept such clients (lower Refusal and lower Non-Response). They might raise Compliance (as it is more worthwhile for banks to go to the trouble to conduct due diligence checks), or raise Non-Compliance (banks are more inclined to accept more risks in flouting the rules for greater rewards).

In fact, however, we found that even across these three orders of magnitude the potential value of the customer made little significant difference to the results. Banks were slightly (but statistically significantly) more likely to refuse customers with \$3 and \$30 million turnovers than \$500,000. And banks were slightly (just under 2 percentage points) more likely to respond to customers with \$3 million turnover than \$500,000 (by 1.5 and 2 percentage points, respectively). These findings once again suggests that banks are comparatively unresponsive to reward in deciding whether or not to accept customers, and whether or not to apply Know Your Customer standards. Curiously, the response to reward is opposite the prediction of rational expectations that banks should be more responsive to and accepting of wealthier customers.

Banks' relative insensitivity to reward is important in helping disconfirm another potential explanation: perhaps banks have no inherent interest in screening out criminals (after all they are profit-makers, not law enforcers), and are willing to trust their luck in surreptitiously defying regulators? Such an explanation might just accommodate accepting likely criminals at the same rate as low-risk corporate customers, but it is much harder to square with banks' tendency to refuse high-rollers at slightly greater rates as firms with a much small turn-over.

While it appears that banks proved relatively insensitive to risk or reward, the results suggest that intermediary corporate service provider (CSP) intermediaries prove even less sensitive, as seen in Figures 4 and 5. Among the information conditions compared to placebo, only the Standards condition caused a statistically significant – but substantively small two percentage-point – decrease in the in a single outcome measure by significantly decreasing the Non-Compliance rate.

The jurisdiction conditions also produced relative insensitivity among intermediaries. The Corruption and Offshore treatments caused statistically significant decreases of roughly four percentage points in the Non-Compliance rate but did not appreciably alter the other outcomes. Thus, contrary to researcher expectations, it appears that intermediaries, like banks, are relatively insensitive to customer risk.



## CONCLUDING THOUGHTS ON COMPLIANCE

A final matter concerns the baseline question of what counts as compliance. An important challenge for both our coding decisions and the contemporary policy debate inside the Financial Action Task Force concerns the “who” and “how” of establishing the identity of shell company owners. As discussed, the critical question for ensuring the transparency of shell companies and the corporate accounts held in their name is identifying the beneficial (i.e. real) owner. The most obvious implication from this is the need to verify that the customers are who they say they are with reference to government documents, typically official passports. Yet for the transparency rules to work, banks and other intermediaries must not only verify a person’s identity, but also verify that that person is in fact the beneficial owner, rather merely than a stand-in or proxy for the real owner. Thus, as well verifying an individual, for the system to work there is also a requirement to verify that individual’s relationship to the company, i.e. to verify that they are indeed the beneficial owner.

The significance of this double requirement, not just the “who” but also the “how” of beneficial ownership, matters critically for our research design, but also is currently a central question for policy-makers. Our solicitation emails are sent out on behalf of companies by research assistants who represent, but do not own, the companies. A reply from a bank or firm that “We need to verify your identity with a passport copy” raises the question of *whose* identity needs to be verified: the representative, which counts as a non-compliant response, or the owner, which is a compliant response?

For policy-makers, getting both parts of beneficial ownership verification right is no mere quibble, but rather goes to the heart of the effectiveness of the regime. In setting up shell companies, it is standard that key corporate roles like company director or secretary will be filled by a nominee, a professional stand-in. The nominee individual may be acting in this role for hundreds of companies without having any idea of what those companies actually do or even who owns them. Indeed, this nominee or corporate proxy function is one of the main services offered by intermediary firms that set up and service shell companies and connect them with corporate bank accounts. As such, even the most thorough and searching identity checks will be for naught if they are applied to the wrong person, i.e. the nominee rather than the actual owner.

In our study, about of third of the responses that specified the need for identity documents did not specify that these documents had to relate to the actual owner. Either explicitly or implicitly, they allowed that verification of the representative would be sufficient. Aside from being a worrying policy vulnerability, this may also feed back into our scripts story. Obtaining a passport copy for every applicant for an account is a relatively easy requirement to standardize, being a box to be checked. But ensuring that this person is the actual beneficial owner is more difficult, given that there are many different ways this ownership can be exercised.

## BIBLIOGRAPHY

- Abbott, Kenneth W., David Levi-Faur and Duncan Snidal (2017) "Introducing Regulatory Intermediaries." *Annals of the American Academy of the Political and Social Sciences* 670 (March): 6-13.
- Abbott, Kenneth W., Philip Genschel, Duncan Snidal and Bernhard Zangl (Editors) (2015) *International Organizations as Orchestrators*. Cambridge: Cambridge University Press.
- Abbott, Kenneth W., and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54 (3): 421-456.
- Avdeenko, Alexandra, and Michael J. Gilligan (2015). "International Interventions to Build Social Capital: Evidence from a Field Experiment in Sudan." *American Political Science Review* 109 (3): 427-449.
- Basel Committee on Banking Supervision. 2012. *Core Principles for Effective Banking Supervision*. Basel, Switzerland: Bank for International Settlements.
- Beath, Andrew, Fotini Christia, Georgy Egorov, and Ruben Enikolopov (2016). "Electoral Rules and Political Selection: Theory and Evidence from a Field Experiment in Afghanistan." *Review of Economic Studies* 83 (3): 932-968.
- Becker, Gary S. (1968). "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169-217.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4): 991-1013.
- Butler, Daniel M. 2014. *Representing the Advantaged: How Politicians Reinforce Inequality*. New York: Cambridge University Press.
- Butler, Daniel M., and David E. Broockman. 2011. "Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators." *American Journal of Political Science* 55 (3): 463-477.
- Butler, Daniel M., Christopher F. Karpowitz, and Jeremy C. Pope. 2012. "A Field Experiment on Legislators' Home Styles: Service Versus Policy." *The Journal of Politics* 74, (2): 474-486.
- Carlsson, Magnus, and Dan-Olof Rooth. 2007. "Evidence of Ethnic Discrimination in the Swedish Labor Market Using Experimental Data." *Labour Economics* 14 (4): 716-729.
- Chayes, Abram and Antonia Handler Chayes (1995). *The New Sovereignty: Compliance with International Regulatory Agreements*. (Cambridge Mass.: Harvard University Press).
- Chen, Jidong, Jennifer Pan, and Yiqing Xu (2016). "Sources of Authoritarian Responsiveness: A Field Experiment in China." *American Journal of Political Science* 60 (2): 383-400.

- Chong, Alberto, Ana De La O, Dean Karlan and Leonard Wantchekon (2015). “Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification.” *Journal of Politics* 77 (1): 55-71.
- Chong, Dennis. 2014. *Collective Action and the Civil Rights Movement*. Chicago: University of Chicago Press.
- Cox, Gary and Mathew McCubbins. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley, CA: University of California Press.
- De la Cuesta, Brandon, Helen V. Milner, Daniel L. Nielson, and Stephen F. Knack. 2019. “Oil and Aid Revenue Produce Equal Demands for Accountability as Taxes in Ghana and Uganda.” *Proceedings of the National Academy of Sciences* 116 (36): 17717-17722.
- Downs, George W., David M. Roche and Peter N. Barsboom (1996). “Is the Good News about Compliance Good News about Co-operation?” *International Organization* 50 (3): 379-406.
- Druckman, James N., Donald P. Green, James H. Kuklinski and Arthur Lupia (2006). “The Growth and Development of Experimental Methods in Political Science.” *American Political Science Review* 100 (4): 627-635.
- Fearon, James D., Macartan Humphreys, and Jeremy M. Weinstein (2015). “How Does Development Assistance Affect Collective Action Capacity? Results from a Field Experiment in Post-Conflict Liberia.” *American Political Science Review* 109 (3): 450-469.
- Fearon, James D., and David D. Laitin. 2003. “Ethnicity, Insurgency, and Civil War.” *American Political Science Review* 97 (1): 75-90.
- Financial Action Task Force (FATF) (2012). *International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation* (Paris).
- Financial Action Task Force (FATF) (2014). *Transparency and Beneficial Ownership* (Paris).
- Financial Action Task Force (FATF) (2016). *FATF Report to the G20: Beneficial Ownership* (Paris).
- Financial Action Task Force (FATF) (2018) *Concealment of Beneficial Ownership* (Paris).
- Financial Action Task Force (FATF) (2019) *Guidance for a Risk-Based Approach for Trust and Corporate Service Providers (TCSPs)* (Paris).
- Findley, Michael G., Adam S. Harris, Helen V. Milner, and Daniel L. Nielson. 2017. “Who Controls Foreign Aid? Elite Versus Public Perceptions of Donor Influence in Aid-Dependent Uganda.” *International Organization* 71 (4): 633-663.
- Findley, Michael G., Daniel L. Nielson, and J.C. Sharman (2013). “Using Field Experiments in International Relations: A Randomized Study of Anonymous Incorporation.” *International Organization* 67 (October 2013), 657-693.
- Findley, Michael G., Daniel L. Nielson, and J.C. Sharman (2014). *Global Shell Games: Experiments in Transnational Relations, Crime and Terrorism* (Cambridge: Cambridge University Press).

- Findley, Michael G., Brock Laney, Daniel Nielson, and J.C. Sharman (2017). "External Validity in Parallel Global Field and Survey Experiments on Anonymous Incorporation." *Journal of Politics* 79 (4): 856-872.
- Financial Services Authority (FSA). 2011. *Banks' Management of High Money Laundering Risk Situations*. London.
- Gell-Redman, Micah, Neil Visalvanich, Charles Crabtree, and Christopher J. Fariss. 2018. "It's All About Race: How State Legislators Respond to Immigrant Constituents." *Political Research Quarterly* 71 (3): 517-531.
- Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94 (3): 653-663.
- Gerber, Alan S. and Donald P. Green (2012). *Field Experiments: Design, Analysis and Interpretation* (New York: W.W. Norton).
- Gerber, Alan S., Gregory A. Huber, Marc Meredith, Daniel R. Biggers, and David J. Hendry (2015). "Can Incarcerated Felons be (Re)integrated into the Political System? Results from a Field Experiment." *American Journal of Political Science* 59 (4): 912-926.
- Gerber, Alan S., Gregory A. Huber, Daniel R. Biggers, and David J. Hendry (2016). "A Field Experiment shows that Subtle Linguistic cues might not Affect Voter Behavior." *Proceedings of the National Academy of Sciences* 113 (26): 7112-7117.
- Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal: Applied Economics* 1 (2): 35-52.
- Global Witness. (2014). *The Great Rip-Off: Anonymous Company Owners and the Threat to American Interests*. Washington D.C.
- Gottlieb, Jessica (2016). "Greater Expectations: A Field Experiment to Improve Accountability in Mali." *American Journal of Political Science* 60 (1): 143-157.
- Government Accountability Office (2006). "Company Formations: Minimal Ownership Information is Collected and Available" (Washington).
- Green, Donald P., Jonathan S. Krasno, Costas Panagopoulos, Benjamin Farrer, and Michael Schwam-Baird (2015). "Encouraging Small Donor Contributions: A Field Experiment Testing the Effects of Nonpartisan Messages." *Journal of Experimental Political Science* 2 (2): 183-191.
- Guzman, Andrew T. (2010). *How International Law Works: A Rational Choice Theory* (Oxford: Oxford University Press).
- Hafner-Burton, Emilie and Kiyoo Tsutsui (2005). "Human Rights Practices in a Globalizing World: The Paradox of Empty Promises." *American Journal of Sociology* 110 (5): 1373-1411.

- Hafner-Burton, Emilie M., Stephan Haggard, David A. Lake, and David G. Victor. 2017. "The Behavioral Revolution and International Relations." *International Organization* 71(S1): S1-S31.
- Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science* 162 (3859): 1243-1248.
- Hyde, Susan D. (2015) "Experiments in International Relations: Lab, Survey, and Field." *Annual Review of Political Science* 18 (1): 403-424.
- Jensen, Nathan M., Bumba Mukherjee and William T. Bernhard. 2014. "Experiments in International Political Economy." *International Interactions* 40 (2): 287-304.
- Kaas, Leo, and Christian Manger. 2012. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment." *German Economic Review* 13 (1): 1-20.
- Kalla, Joshua L. and David E. Broockman (2016). "Campaign Contributions Facilitate Access to Congressional Officials: A Randomized Field Experiment." *American Journal of Political Science* 60 (3): 545-558.
- Kramer, Adam D.I., Jamie M. Guillory, and Jeffrey T. Hancock (2014). "Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks." *Proceedings of the National Academy of Sciences* 111 (24): 8788-8890.
- Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo. "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment." *The Quarterly Journal of Economics* 128 (3): 1123-1167.
- Lahey, Joanna N. "Age, Women, and Hiring: An Experimental Study." *Journal of Human Resources* 43 (1): 30-56.
- Lyall, Jason (2019). "Civilian Casualties, Humanitarian Aid, and Insurgent Violence in Civil Wars." *International Organization*, 73 (4): 901-926.
- Lyall, Jason, Yang-Yang Zhou, and Kosuke Imai (2020). "Can Economic Assistance Shape Combatant Support in Wartime?" *American Political Science Review* 114 (1): 126-143.
- March, James G. and Johan P. Olsen (1998). "The Institutional Dynamics of International Political Orders." *International Organization* 52 (4): 943-969.
- Martin, Lisa L., and Beth A. Simmons. 1998. "Theories and Empirical Studies of International Institutions." *International Organization* 52 (4): 729-757.
- Mearsheimer, John J. (1994/95). "The False Promise of International Institutions." *International Security* 19 (3): 5-49.
- Michelitch, Kristin (2015). "Does Electoral Competition Exacerbate Interethnic or Interpartisan Economic Discrimination? Evidence from a Field Experiment in Market Price Bargaining." *American Political Science Review* 109 (1): 43-61.
- Milgram, Stanley. 1963. "Behavioral Study of Obedience." *Journal of Abnormal and Social Psychology* 67(4): 371-378.
- Morse, Julia. 2020, "Blacklists, Market Enforcement, and the Global Regime to Combat Terrorist Financing." *International Organization*, forthcoming

- Nyhan, Brendan, and Jason Reifler (2015). "The Effect of Fact-Checking on Elites: A Field Experiment on US State Legislators." *American Journal of Political Science* 59 (3): 628-640.
- Obermayer, Bastian and Frederik Obermaier (2016). *The Panama Papers: Breaking the Story of How the Rich and Powerful Hide their Money* (New York: Oneworld).
- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.
- Powell, Robert. 2017. "Research Bets and Behavioral IR." *International Organization* 71(S1): S265-S277.
- Prina, Silvia (2015). "Banking the Poor via Savings Accounts: Evidence from a Field Experiment." *Journal of Development Economics* 115 (1): 16-31.
- Reus-Smit, Christian (Ed.) (2004). *The Politics of International Law* (Cambridge: Cambridge University Press).
- Shepsle, Kenneth A., and Barry R. Weingast. 1987. "The Institutional Foundations of Committee Power." *American Political Science Review* 81 (1): 85-104.
- Sheffer, Lior, Peter John Loewen, Stuart Soroka, Stefaan Walgrave, and Tamir Sheafer. 2018. "Nonrepresentative Representatives: An Experimental Study of the Decision Making of Elected Politicians." *American Political Science Review* 112 (2): 302-321.
- Sherman, Lawrence W. and Heather Strang (2004). "Experimental Ethnography: The Marriage of Qualitative and Quantitative Research." *Annals of the American Academy of Political and Social Science* Vol. 595: 204-222.
- Sikkink, Kathryn (2011). *The Justice Cascade: How Human Rights Prosecutions are Changing World Politics*. New York: W.W. Norton.
- Simmons, Beth A. and Daniel J. Hopkins (2005). "The Constraining Power of International Treaties: Theory and Methods." *American Political Science Review* 99 (4): 623-631.
- Simmons, Beth A. (2010). "Treaty Compliance and Violation." *Annual Review of Political Science* 13 (1): 273-296.
- Sniderman, Paul M. and Matthew S. Levendusky. 2007 "An Institutional Theory of Political Choice." In Russell J. Dalton and Hans-Dieter Klingeman (Eds) *Oxford Handbook of Political Behavior*. Oxford: Oxford University Press, pp.1-22.
- Thies, Michael F. 2001. "Keeping Tabs on Partners: The Logic of Delegation in Coalition Governments." *American Journal of Political Science* 45 (3): 580-598.
- Tilcsik, András. 2011. "Pride and Prejudice: Employment Discrimination Against Openly Gay Men in the United States." *American Journal of Sociology* 117 (2): 586-626.
- Tomz, Michael. (2007). "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61 (4): 821-40.

US Senate Permanent Subcommittee on Investigations. 2014. "Offshore Tax Evasion: The Effort to Collect Unpaid Taxes on Billions in Hidden Offshore Accounts."

von Stein, Jana (2005). "Do Treaties Screen or Constrain? Selection Bias and Treaty Compliance." *American Political Science Review* 99 (4): 611-622.

von Stein, Jana (2010). International Law: Understanding Compliance and Enforcement. In Robert Denemark (Ed.) *International Studies Association Compendium Project*. pp.1-18.

von Stein, Jana (2016). "Making Promises, Keeping Promises: Democracy, Ratification and Compliance in International Human Rights Laws." *British Journal of Political Science* 46 (3): 655-679.

Wendt, Alexander (1999). *Social Theory of International Politics* (Cambridge: Cambridge University Press).

Williamson, Oliver E. (1975) *Market and Hierarchies: Analysis and Antitrust Implications*. New York: Macmillan.

World Bank (2011). *The Puppet Masters: How the Corrupt Use Legal Structures to Hide their Stolen Assets and What to do About It* (Washington D.C.: World Bank).

Young, Oran R. 1979. *Compliance and Public Authority: A Theory with International Applications*. Baltimore: Published for Resources for the Future by the Johns Hopkins University Press.

Yousafzai, Shumaila Y., John Pallister, and Gordon R. Foxall (2005). "Strategies for Building and Communicating Trust in Electronic Banking: A Field Experiment." *Psychology & Marketing* 22 (2): 181-201.

Zarate, Juan C. (2013) *Treasury's War: The Unleashing of a New Era of Financial Warfare* (New York: Public Affairs).