Received: 12 July 2022 | Revised: 5 July 2023 | Accepted: 5 July 2023

DOI: 10.1111/ajps.12861

AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

ARTICLE

Banking bad? A global field experiment on risk, reward, and regulation

Michael G. Findley[1] | **Daniel L. Nielson**[1] | **J. C. Sharman**[2]

[1]Department of Government, University of Texas at Austin, Austin, Texas, USA

[2]Department of Politics and International Studies, University of Cambridge, Cambridge, UK

Correspondence
J. C. Sharman, 7 West Rd, University of Cambridge, Cambridge CB3 9DT, UK.
Email: jcs207@cam.ac.uk

[Correction added on 5 April 2024, after first online publication: The department of Michael G. Findley and Daniel L. Nielson in the affiliation has been changed.]

Abstract

Are banks sensitive to risk and reward in following global corporate transparency rules? Using a worldwide field experiment, this study evaluates competing predictions from expected utility, behavioralist, and institutionalist accounts. We incorporated a dozen companies around the world to make over 15,000 email solicitations asking for corporate accounts from 5000 of the world's internationally connected banks. Treatments randomize the risk profiles of different companies—by their countries' association with corruption, terrorism, and tax evasion—and vary rewards by stating differing amounts of business revenues. The outcomes are the rates at which banks offer accounts and comply with rules on customer identification. The results suggest that banks are moderately responsive to risk—though not reward—but the magnitude of the effects is small, providing mixed evidence for conventional models and suggestive support for institutionalist accounts.

Are banks sensitive to risk and reward in the context of international rules designed to screen out criminals' money from the financial system? We answer this question via a global field experiment. We randomly assigned over 15,000 email solicitations to the world's internationally connected banks inquiring after corporate accounts from one of twelve companies legally incorporated by the authors in different countries. The aim is to evaluate predictions about how banks approach varying potential risks and rewards, derived from expected utility and behavioralist approaches and then compared to an institutionalist approach premised on "organizational scripts." Banks are especially well suited for such a study given that their central purpose is calculating expected risk and reward, or business loss and profit, in allocating credit and financial services. In addition, the fundamental principle of the international banking regulatory regime is the "risk-based approach": the idea that regulators can achieve public policy goals such as preventing bank runs and money laundering by manipulating the incentives banks face in screening, accepting, or rejecting different types of business.[1]

In experimentally testing the effectiveness of financial soft law standards, we aim to contribute significant new knowledge on what has increasingly become policymakers' default tool of choice in the global governance of a variety of international scourges. Whether

[1] In this paper, we use the term risk as policymakers use it, which is more akin to a probability of being caught and facing severe penalties. That is, we use the term to refer to the probability of experiencing a negative outcome (e.g., being caught aiding illicit finance) relative to the probability of experiencing a positive outcome (e.g., benefiting from the flow of more money), otherwise thought of as an expected value. In contrast, expected utility theorists use the term *risk* to refer to how specific outcomes are valued (e.g., each dollar less than the next [risk acceptance], each dollar the same as the next [risk neutrality], or each dollar more than the next [risk aversion]). We do not investigate risk profiles in this formal-theoretic sense. See the "Expected Utility, Behavioralist and Organizational Scripts" section.

Verification Materials: The data and materials required to verify the computational reproducibility of the results, procedures, and analyses in this article are available on the American Journal of Political Science Dataverse within the Harvard Dataverse Network at: https://doi.org/10.7910/DVN/VBNCF2.

it is combating money laundering, corruption, nuclear proliferation, tax evasion, terrorism, environmental degradation, human-trafficking, or a range of other crimes, the United States, and other governments have more and more turned to exactly the kind of privately intermediated financial restrictions that are the focus of this paper (Morse, 2022; Zarate, 2013). Similarly, the development of economic sanctions has seen a move away from direct state-to-state measures to restrict trade and toward conscripting banks and other for-profit firms in freezing targeted entities out of the international financial system (Morgan et al., 2023). The costs of this new global regulatory apparatus are high but are largely hidden as they are imposed on not only firms but also consumers (Sharman, 2011). Yet, how and whether this new tool of statecraft and global governance works in practice is essentially unknown.

In conducting our study, we worked from regulators' manuals (e.g., Financial Action Task Force [FATF], 2012) to engineer alternative risk treatments by varying the jurisdiction of company incorporation. Because of these sharply varying risk profiles (e.g., high or low corruption or terrorism financing risk), we can measure respondents' risk sensitivity in the different rates at which they are willing to engage with would-be customers, and the degree to which they comply with international corporate transparency rules. To test sensitivity to varying rewards, the company solicitations randomly varied the level of business revenues at $500,000, $3 million, or $30 million.

The results show that banks are marginally sensitive to negative risk but relatively indifferent to positive reward. Surprisingly, obvious danger signs of corruption, terrorism, or tax evasion produced only substantively small variations from the low-risk placebo condition in banks' willingness to transact business with or verify the identity of high-risk corporate customers. Increases of one or two orders of magnitude in financial inducements made even less substantive difference. Though the results are not conclusive, they cast doubt on the effectiveness of the global regulations regime relying on banks applying a risk-based approach (RBA).

One possible interpretation of the surprisingly slight sensitivity to risk and reward may derive from the idea that individual human behavior cannot be unproblematically extrapolated to aggregate institutions. Even if the overall expected utility of an institution can be calculated by its principals, the principals must fashion a set of rules and procedures whereby agents are motivated to act in the interest of the institution. Slack, slippage, and agent discretion necessarily alter organizational outcomes away from the goals of the principals (Coase, 1937; Hawkins et al., 2006; Nielson & Tierney, 2003). Thus, an alternative explanation that may partially account for the small effect sizes relates to "organizational scripts": generic diagnostic templates representing the codification of shared expectations that shape routine institutional responses (Barnett & Finnemore, 2004: 3, 9; March & Olsen, 1989: 21–26; 1998: 948; Powell & DiMaggio, 1991: 15).

Beyond academic debates, the study is squarely focused on a topic of great real-world importance: When banks get their risk calculations wrong, the results can be highly damaging or even catastrophic, whether it is allowing criminals access to the financial system or precipitating general financial crises. The paper provides much more systematic evidence than previously available in answering an important question: Do banks, the gate-keepers in transnational finance, follow global rules in assessing customer risk before granting access to the international financial system? The results here cast doubt on the risk-based regulatory approach. Though the rate of outright noncompliance with identification standards is low in absolute terms, the ease of approaching many banks simultaneously online may mean that criminals seeking to access the international financial system anonymously can do so relatively easily. More broadly, over the last 30 years banks have been a key point of regulatory intervention for international policymakers seeking to stabilize and reform the global political economy, yet we lack a robust sense of the causal impact of these interventions.

Field experiments have recently been the target of strong ethical critiques (e.g., McDermott & Hatemi, 2020; Phillips, 2021), a challenge this paper takes seriously. It is important to stipulate that, unlike many other audit studies, this experiment employed no active deception. When contacting banks, researchers used their own names and represented actual, legally formed companies pursuing a business interest. In continuing the study, the authors have done business with banks and paid bank fees, so the study did not unnecessarily waste bank representatives' time. In addition to avoiding harms, a major positive ethical rationale for this study is the strong policy relevance of the knowledge generated, which in this case is much more than fanciful notions of scientific applicability, as earlier iterations of the research have led to demonstrable improvement in national and international policies. Perhaps even more importantly, this experiment is "studying up" in focusing on rich and powerful institutions (Nader, 1974). We emphasize here that, in making judgments about the ethics of field experiments, researchers should weigh beneficence alongside concerns for autonomy and justice (Belmont Report; United States, 1978). This is especially the case in domains where significant harms already occur.

WHY IT MATTERS: BANKS, SHELL COMPANIES, AND GLOBAL BANKING RULES

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 international financial system while remaining undetected. Banks are 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.), the "Know Your Customer" (KYC) rule. Riskier clients should be subject to greater scrutiny by banks and, if necessary, excluded. Some banks that recklessly pursued profit by taking on risky customers have been hit with multibillion-dollar fines, or in extreme cases have been deliberately destroyed by sanctions (Zarate, 2013).

A serious obstacle to this goal of banks' knowing their customers has been the use of shell companies. These are companies that have no substantive business activities but that can be set up online for between a few hundred and a few thousand dollars in a matter of days. Shell companies 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. Thus, global rules mandate that banks must assess the risk of their corporate customers by identifying the real people in control of companies that hold corporate bank accounts.

These global rules are promulgated by the FATF (2012, 2014), the world's anti-money-laundering standard-setter and enforcer, comprised of 37 of the world's most powerful countries. FATF standards have been endorsed by and incorporated within the rules of other international organizations, including the United Nations Security Council and the International Monetary Fund and have been mandated worldwide through FATF-associated regional organizations whose members include nearly every country. Individual states are responsible for legislating and implementing the FATF rules, and governments' compliance is publicly assessed by the FATF. States have legislated that banks in their jurisdictions must follow international rules centered on the RBA (FATF, 2012). Regulators seek to ensure compliance from banks by manipulating banks' risk-reward calculations (e.g., imposing fines). If too many banks in a given jurisdiction fail to comply, the national government may be blacklisted and sanctioned by the FATF (Morse, 2022).

Because governments and international organizations have generally tried to measure the effectiveness of the rules by reading laws and regulations on the books and counting inputs, rather than ascertaining the standards applied in practice, the effectiveness of these rules is unknown. Anecdotal evidence and recurrent scandals, however, give grounds for concern (ICIJ, 2020; Obermayer & Obermaier, 2016). Academic literature on the subject has noted the rapid global spread of anti–money-laundering standards (Morse, 2022), but has often been skeptical of the actual effectiveness of these standards (Sharman, 2011; Tsingou, 2018). Much of this literature laments the lack of hard evidence, however, which complicates related judgments concerning both causal inference and policy effectiveness. An exception is the work of Findley et al. (2014), which uses a field experiment to assess levels of compliance with FATF rules among businesses, such as corporate law firms and incorporation services, that form and on-sell shell companies. Yet to our knowledge, there is no audit or experimental study of banks' compliance with these rules, even though banks are unquestionably the most important private-sector actors in the global financial system. Our study fills this gap.

EXPECTED UTILITY, BEHAVIORALIST, AND ORGANIZATIONAL SCRIPTS

The FATF rules are centered on the RBA (FATF 2014). Banks that fail the KYC rule may face fines running into billions of dollars (Zarate, 2013). Yet, the chances of even egregious noncompliance by banks being detected, and then fines being imposed as a result, are quite small. Though the exact odds are impossible to calculate, a United Nations report estimates that only about 1% of criminal money in the international banking system is detected by authorities (UNODC, 2011: 7). A huge leak of bank due diligence material in 2020 revealed that international banks had moved $2.1 trillion of suspicious funds in 1999–2017, often in transactions involving shell companies whose ownership was unknown (ICIJ, 2020). In all but a few cases, banks received no penalties.

Conversely, it is standard practice that bankers who bring in more lucrative business will be rewarded with bonuses and promotion. Banks are, after all, run to maximize profit. The net effect of these contrary pressures—minimizing the likelihood of fines for noncompliance while maximizing the likelihood of profit through more business—is essentially unknown and thus motivates this study. The dominant regulatory principle of the RBA bears important similarities with the expected-utility account of behavior, although the term *risk* is used differently. Actors choose between prospective outcomes by comparing expected utility values. That is, actors compute the sum of the products of the probabilities of all realized outcomes (expected value) and apply a preference ordering to the alternatives (Becker, 1968).

The main trade-off consists of evaluating a potential customer for whether it will increase the probability of being caught and facing fines or, alternatively, not being caught and yielding substantial profit. Given the low likelihood of detection and penalties, expected utility approaches hold that bank directors and frontline agents will focus on maximizing reward, and only pay attention to any uncharacteristically high probability of being caught. Taken together, if bank directors and agents are both motivated to take on as many clients as possible to maximize profits and promotion possibilities, then banks as organizations are more likely to reflect this orientation. From an expected utility perspective, we would thus expect low sensitivity to risk but high sensitivity to reward.[2] All else equal, bank agents would be less likely to ignore or refuse solicitations from high value potential clients than lower value clients, and would be less scrupulous in requiring proof of identity from the former than the latter.

In contrast to an expected utility approach, behavioralism suggests that an array of cognitive biases—in particular, and with special relevance to this study, loss aversion and the overweighting of low probabilities—drive people's judgments away from expected-utility predictions (Davis & McDermott, 2021: 153, 165; Kahneman & Tversky, 1979: 279; Thaler, 2015; Thaler et al., 1997: 647; Tversky & Kahneman, 1992: 303).[3] In contrast to expected utility, these biases predict that bank directors and frontline agents will exhibit greater concern about the probability of being caught and facing fines than to the probability of not being caught and receiving rewards. Overweighting of probabilities suggests that bankers will behave as if the low odds of being fined for violating regulations by not complying with due diligence rules and taking on potentially unsavory customers were in fact significantly higher.

Loss aversion suggests that bank agents will exhibit higher sensitivity to potential loss than to gain and therefore ignore, refuse, or apply extra scrutiny to would-be unsavory customers who might lead to the bank being fined, because such penalties would damage or ruin the career prospects of the agents as well as the directors who bear ultimate responsibility and therefore occur in the psychological realm of losses.[4] Tversky and Kahneman suggest that loss aversion is particularly strong with low probability losses (1992: 297). From a behavioralist perspective, unless would-be customers stand to produce particularly substantial gains for the banks, we would expect high sensitivity to the prospect of being caught (meaning substantially more nonresponses, refusals, and compliance with KYC rules) and low sensitivity to the prospect of substantial gains, unless the expected gains are particularly great.

Overall, expected utility and behavioralist perspectives expect sensitivity to reward and risk, though they differ in the degree to which this sensitivity is manifest. Organizational bank sensitivity (or insensitivity) is a product of the aggregated behavior of bank staff, necessitating assumptions or arguments about banks' organizational dynamics. Standard expected utility models generally predict that institutions behave like individuals, maximizing expected benefits and minimizing expected costs. Behavioralist models could also be projected similarly, though they may have a more difficult time mapping cognitive biases to an organizational level.

Institutionalist scholars, however, are skeptical that a linear aggregation of the behavior of individuals characterizes institutional dynamics. Many important political outcomes may only be able to be understood as the collective actions of many individuals behaving together under institutional constraints. Institutions are more than merely the sum of the individuals within them (DiMaggio & Powell, 1991; March & Olsen, 1989; North, 1990; Ostrom, 1990).

Bank directors may create organizational scripts to encode their expected risk/reward orientation to tightly constrain bank agents to behave in line with what bank directors value. Rather than each organizational choice 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 in aligning bank director views with bank agent behavior. But just as often it may reflect the informal and uncoordinated congealing of habits and routines, which may be deeply pathological for organizational function (Barnett & Finnemore, 2004).

Practically, 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, as demonstrated by the manuals that have been leaked to the public (Collin, 2021). Drawing on extensive participant-observation research among those in banks applying such rules, Tsingou finds that that money-laundering risk within banks is handled very differently to credit or market risk. In practice banks dispense with sophisticated models and simplify in line with a few prescriptive

[2] We have only articulated the context in which bank agents are risk neutral. It is possible that agents would be risk averse or risk acceptant, meaning that they would value each additional client or dollar of profit less than the last, if the averse, or more than the last, if acceptant.

[3] Behavioralist approaches have argued for a variety of effects, but we focus on two of the central insights, including loss aversion and overweighting of probabilities.

[4] Loss aversion bears some similarity, but is not the same, as risk aversion. Loss aversion is a behavioral phenomenon reflecting different sensitivity to loss than to gain, where an agent could be loss averse and at the same time both risk averse or risk acceptant. Risk aversion is a general bias against outcomes with high uncertainty even if the expected value of the uncertain and certain outcomes is the same, and at times if the expectation of the uncertain outcome is higher than the certain outcome.

rules, reducing "AML governance to legible categories through their use of lists and rules" (2018: 203–204).

In line with this observation, an institutionalist account of large, regulated institutions like internationally connected banks expects that organizational scripts will shape bank agent responses. Organizational scripts may be overly constraining and inadvertently promote standardized generic responses rather than facilitating complex judgments rooted in individual weighing of expected losses and profits from business partnerships and transactions. The scripts themselves may be overly constraining, or individual cognitive limitations may preclude execution of the scripts in line with their intended effects, but either way the organizational scripts approach expects little sensitivity to either varying expected losses or profits associated with client screening. While it is possible that scripts might be more nuanced and prescribe different scripted responses to different stimuli (stronger scrutiny to corruption risk, for example), in practice in the anti–money-laundering domain, simplified actions and generic responses predominate (Tsingou, 2018).

From a scripts perspective, we thus expect to observe similar rates of response, refusal, compliance, and noncompliance across treatment and control conditions. As such, institutional scripts generate different predictions—insensitivity to expected costs and benefits (risk and reward)—from the other two perspectives—sensitivity to expected costs and benefits (risk and reward).

RESEARCH DESIGN

Randomly allocating actors to control and treatment groups helps to address common obstacles to causal inference in boosting internal validity. Furthermore, our subjects do not opt into the study, and they do not know they are being observed, thus obviating social desirability bias and further strengthening internal validity. Because the study includes every reachable bank that is connected to the international wire transfer network—effectively the entire population of interest—the study also has uniquely strong representativeness to further enhance external validity. Research assistants represented agents for 12 shell companies incorporated by the authors and made email solicitations to thousands of banks around the world asking to set up bank accounts for these companies.

Subject pool and nonresponse checks

We first compiled a list of the world's international banks and their contact details. The list comes from the Society for Worldwide Interbank Financial Telecommunication (SWIFT), which allocates codes for the message system that underpins international bank wire transfers. Using this list, research assistants obtained the relevant email address for 5140 banks. These banks included all headquarter banks on the SWIFT list for which contact information could be obtained and a stratified random sample of the banks' branches in each country. We made three approaches to each bank and branch, for a total of 15,407 approaches, with a wash-out period between each approach to minimize the risk of detection. The approaches were in May–July 2019, March 2020, and July 2020, with an earlier nonresponse check in January–February 2019.

Email approaches

The standard approach template was a message from a representative agent (a research assistant) 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/author.[5] The researcher acting as company agent explained that the company would like to establish a corporate 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, if any. (An example displaying one of the ten base emails is shown in the Appendix C3, "Email Approaches," p. 4).

Approaches were made 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 vetted to verify that the monikers were not already well known. For greater authenticity, separate websites were established for each company. Communicating via email, rather than phone or video call, allows for identical treatments, more accurate coding, and a comprehensive record of correspondence. We vetted the emails with several parties, including practitioners in the incorporation and banking industries, to ensure that they address the key issues of concern to banks and were realistic. Our informants indicated that the email language evinced high levels of naturalism.

[5] The five topics are development assistance, education and training, impact assessment, feasibility studies, and information and communication technology.

Email is a standard medium for communication between banks and customers for both legitimate and rogue actors. Often, as in our research design, an agent emails on behalf of the principal. The prominence of email correspondence in this context is evidenced in the massive data leaks from banks and associated law firms, most famously the Panama Papers (Collin, 2021; ICIJ, 2020; Obermayer & Obermaier, 2016). These leaks have revealed email correspondence from those associated with corruption, tax evasion, money laundering, sanction-busting, and a multitude of other crimes. Nor is it only small-time crooks using email. Fugitive Malaysian Jho Low, who in cahoots with the Prime Minister of the day looted over $5 billion, approached a key bank via a Gmail account he formed for the purpose (Wright & Hope, 2018: 125–126).

Block randomization and treatments

We block randomized assignment to experimental conditions across two strata: bank host country and reply or not to the pretreatment nonresponse check. We preregistered further blocking by banks' status as headquarters or branches, their membership in the group of Global Systematically Important (GSI) Banks (see Appendix C5, p. 7), and their host country's category: OECD, offshore jurisdiction, or developing. Blocking by host country—by far the most relevant unit driving the regulatory environment—obviated the need for the stratum categorizing country type. The other two planned blocks were not employed due to the already small size of the blocks generated under country and prior response. These block randomization decisions represent deviations from the preanalysis plan. Nevertheless, we control for country type, headquarter versus branch, and GSI banks in estimation and show orthogonality/balance statistics for these covariates in the Appendix (C5, Table A.2, pp. 7–8). The presence or absence of these covariates does not substantively alter the experimental results reported here.[6] See Appendix E2d, Tables A.11–A.12, pp. 15–16.

The variation between control and treatments arises from the different jurisdictions of incorporation.[7] International and national standards mandate that banks assess the riskiness of potential clients in part based on their country of origin (FATF, 2012, 2014). 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 country perceived to have low corruption, all else equal. Countries have varying risk profiles for laundering money, financing terrorism, being a tax haven, and other concerns. In her participant-observation research with bank compliance officers, Tsingou finds that national origin is usually the dominant factor in assessing customer risk (2018: 202; see also Collin, 2021). The authors' own attendance at banking compliance industry events confirms this verdict. These sources further corroborate that bank agents work with a highly abstracted and simplified picture whereby countries are reduced to a few key risk factors via relatively crude metrics and rankings (Morse, 2022; Sharman, 2011). Crucially, according to the international standard for the RBA, high-risk solicitations should get fewer replies, more refusals, and stricter application of customer due diligence by banks than low-risk solicitations epitomized by our placebo companies from Australia and New Zealand.

The first jurisdiction treatment is designed to learn whether soliciting offers for bank accounts from companies formed in the United States, arguably the hegemon in global economic governance, affects the response and compliance rate relative to our placebo jurisdictions. The US government has been particularly aggressive in applying extra-territorial law enforcement measures against banks (Zarate, 2013). We formed one company each in Delaware and California. We also formed two companies in the United Kingdom, a leading financial center.

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 (Morse, 2022; Obermayer & Obermaier, 2016). We formed companies in the British Virgin Islands, the most popular single offshore shell company jurisdiction with roughly 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 countries perceived to have major corruption problems. We formed companies in Papua New Guinea (ranked 142nd of 180 countries on the 2020 Corruption Perceptions Index) and Bangladesh (ranked 146th). 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, more than 3 years later

[6] While including the covariate for GSI banks does not substantively alter results, as seen in Tables A.11 and A.12, subsetting only on the GSI banks themselves suggests that the large global banks may be more sensitive to some risks, but the results are not strong as the extremely small number of observations preclude convergence of the maximum likelihood estimator. We include the script/results with the replication data so that the basic (nonconverged) results can be observed.

[7] Additional experimental conditions varied the language in the emails, providing additional information "primes" regarding relevant international law. Descriptions of these treatments and empirical findings are reported elsewhere.

and counting, of forming a company in Pakistan.[8] We expected both Pakistani entities to signal a high terrorism-financing risk. Pakistan has earlier played host to the Taliban and leadership of al-Qaeda, as well as a variety of other terrorist groups. Moreover, Pakistan ranked fifth on the 2017 Global Terrorism Index. The placebo condition was an inquiry originating from either the Australian or New Zealand company.

We note here that there are many other features of the countries in our treatment categories that differ; Australia and New Zealand are hardly equivalent on many dimensions, and the same is true for the British Virgin Islands and Seychelles, and Papua New Guinea and Bangladesh. Our contention is that for the purposes of signaling financial risk, bank agents will likely see the paired countries as similar. It is important to acknowledge, however, that we cannot entirely rule out the possibility that these jurisdictions may trigger cultural biases among some respondents aside from our intended risk treatments. We investigate possible differences in treatment effects between countries in the treatment categories in Appendix Table A.13 (E2c, p.14) and Figure A.2 (Appendix, E2e, p.18) and find general similarity.

As well as treatments creating low- and high-risk corporate customer profiles, further treatments created variation between lower- and higher-value customers. This reflects the pervasive assumption that banks trade off risk against potential reward, and that customers offering larger returns to banks may receive more favorable treatment compared to those offering smaller returns. Specifically, banks may be more likely to reply positively to wealthy companies and less likely to scrupulously apply relevant rules relative to lower-value customers.

We varied the reward by randomizing the recent turnover of the business disclosed in the initial email approach, given that higher revenues can be expected to generate higher fees for banks. The research assistant specified 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 author owner's total revenue in research grants in the preceding years and thus reflects truthful information).

Outcomes and coding

Coding was completed on the basis of the email correspondence. The first outcome was 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 nonresponse 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 could indicate a willingness to open an account, but require photo ID verification documents. This was coded as compliant. This coding was a simplified version of the documents required by FATF rules (FATF, 2012, 2014, 2019). Most important, and therefore key to our coding rules, is identification 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.

Third, we coded required ID for the person ultimately in charge of the company, called the beneficial owner, as narrow compliance. We separately coded required ID for any company representative as broad compliance. The descriptive statistics and analysis shown below are for narrow compliance (see Appendix E2e, Table A.14, p. 19, for descriptive statistics for broad compliance). The experimental findings are qualitatively similar for broad compliance (shown in Appendix E2e, Figure A.3, p. 20). We use narrow compliance here to align with FATF standards requiring identification of the beneficial owner. Alternatively, verifying only the credentials of a company agent or nominee as in broad compliance, even with photo identification, removes the ability of law enforcement to connect suspicious transactions to the person ultimately in charge of the company, and thus violates FATF rules.

Fourth, banks coded as noncompliant failed to specify required identification after repeated follow-up. Banks that are willing to open an account without this supporting documentation to establish the customer's true identity were coded noncompliant because they are breaking international rules by offering what amounts to an anonymous or untraceable bank account. Without customer identification, it is difficult or impossible to find the 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 noncompliant. Emails were independently coded by two research assistants for accuracy and consistency, with discrepancies adjudicated by a third, senior coder. The frequencies and percentages across the outcome conditions for the different categories of bank host countries, bank types, and experimental conditions are shown in Table 1. We note that some substantive differences emerge across descriptive categories, especially for offshore jurisdictions compared to OECD host countries, with smaller but still meaningful differences across branch

[8] The fact that the company was still in the process of being formed was disclosed to banks and intermediaries in correspondence.

TABLE 1 Outcomes across categories and experimental conditions.

Category	Subcategory		Outcome				
			Noncompliant	Compliant	Refusal	No reply	Total
Descriptives for observational indicators							
Bank host country type	OECD	*N*	208	141	1444	5613	7406
			2.81%	1.90%	19.50%	75.79%	100.00%
	Offshore	*N*	42	104	232	665	1043
			4.03%	9.97%	22.24%	63.76%	100.00%
	Developing	*N*	245	389	782	5542	6958
			3.52%	5.59%	11.24%	79.65%	100.00%
Branch vs. headquarters	Branch	*N*	247	244	1307	7462	9260
			2.67%	2.63%	14.11%	80.58%	100.00%
	Headquarters	*N*	248	390	1141	4334	6113
			4.06%	6.38%	18.67%	70.90%	100.00%
Global systematically important (GSI) bank	Not GSI bank	*N*	451	600	2178	11,105	14,334
			3.15%	4.19%	15.19%	77.47%	100.00%
	GSI bank	*N*	44	34	280	715	1073
			4.10%	3.17%	26.10%	66.64%	100.00%
Descriptives for experimental conditions							
Jurisdiction conditions	Jurisdiction placebo	*N*	173	241	875	3859	5148
			3.36%	4.68%	17.00%	74.96%	100.00%
	Corruption	*N*	33	40	171	995	1239
			2.66%	3.23%	13.80%	80.31%	100.00%
	Terrorism	*N*	38	42	218	1012	1310
			2.90%	3.21%	16.64%	77.25%	100.00%
	Offshore	*N*	52	34	189	977	1252
			4.15%	2.72%	15.10%	78.04%	100.00%
	US origin	*N*	36	51	198	1048	1333
			2.70%	3.83%	14.85%	78.62%	100.00%
	UK origin	*N*	163	226	807	3929	5125
			3.18%	4.41%	15.75%	76.66%	100.00%
Money conditions	Money $500k	*N*	172	217	775	3983	5147
			3.34%	4.22%	15.06%	77.38%	100.00%
	Money $3 M	*N*	160	217	852	3891	5120
			3.12%	4.24%	16.64%	76.00%	100.00%
	Money $30 M	*N*	163	200	831	3946	5140
			3.17%	3.89%	16.17%	76.77%	100.00%
Total		*N*	495	634	2458	11,820	15,407
			3.21%	4.12%	15.95%	76.72%	100.00%

Note: Outcome frequencies and percentages across categories and experimental conditions. The table displays cell sizes and proportions for each outcome for the different values of treatments and covariates. Country origins used for jurisdiction experimental conditions were Australia and New Zealand for the placebo, Bangladesh and Papua New Guinea for corruption, Pakistan (one company and one trust) for terrorism, and British Virgin Islands and Seychelles for offshore.

versus headquarter banks and Global Systematically Important Banks compared to the rest. While, on average, only 3.2% of the sample including nonresponses proved noncompliant by failing to demand ID documents, this number was 13.8% of responses received, suggesting opportunity for anonymous banking. Difference in proportions tests analyzing the effects of treatments on outcomes are shown in Appendix E2b, Table A.9, pp. 13–14.

What about the danger of a "bait and switch" tactic, whereby banks might indicate one set of requirements in their initial response and then later change from

compliant to noncompliant (or vice versa) further along in the process? To guard against this, we selected a random sample of 30 banks offering compliant and noncompliant responses to be engaged in extended correspondence after the initial round. Through multiple replies and responses via email but also on phone calls, this process aimed to go to the penultimate step of account opening, and it sought to press those banks specifying a need for customer identification to relax this requirement. None of these 30 banks flipped their response, though two did indicate some slight softening of their compliant responses (changing an initial request for proof of identity of all shareholders to identifying only the majority shareholder, both of which count as compliant in the coding scheme). Finally, we set up a company bank account in Belize, paying all fees and transferring money into the account; the (stringent) KYC requirements remained consistent throughout the process, as initially advertised.

Though banks are responsible for ongoing monitoring of accounts, huge leaks of banks' confidential risk management correspondence in September 2020 confirmed that customer risk assessment is mainly concentrated at the point of first engagement (ICIJ, 2020). The initial decision as to whether customers are granted bank accounts becomes even more important given that these same leaks show that ongoing monitoring reports sent to the authorities on existing customers are often not even read.

DATA ANALYSIS

The estimation of results faces some unique challenges, specifically that we have four outcomes, three of which (compliance, noncompliance, and refusal) are contingent on another (response). Treating this as a selection process comes with challenges, and we consider that possibility below. However, current best analysis practices indicate that it is preferable to "flatten" the sequential outcomes and treat them as four categorically distinct possibilities (Slough, 2021). As preregistered and deposited prior to data collection in our preanalysis plan, data analysis employs multinomial probit estimation to consider treatment effects alongside covariates on all four outcomes simultaneously.

As preregistered, in the Appendix we also report difference-in-proportion tests comparing each treatment condition to the control. The results are substantively similar (see E2b, Table A.9, pp. 13–14). Also as preregistered, we report selection models in the Appendix. The logic behind the selection model is that three of the other outcomes—compliance, noncompliance, 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 instruments to identify separate stages of a selection model that satisfy the nonexcludability requirement, so we draw on the estimation technique developed by Sartori (2003), which allows the same variable to identify both stages. Again the results are similar to those reported here (see Appendix E2c, Table A.10, pp. 13–14). Below we discuss the most important elements of our findings: the high proportion of nonresponses, but most especially the limited sensitivity of banks to both varying risk and varying reward among their corporate customers.

Nonresponse: Soft-screeners or never-responders?

Before experiment launch we used five different email scripts, randomly assigned to the banks, to perform preexperiment response checks, each asking innocuous questions about international transfers and potential upgrades to a merchant account to process credit-card payments. Both are widely offered services but not necessarily discernible from bank websites. The response checks revealed that more than 75% of banks did not respond to our innocuous inquiry, setting a critical baseline level of subject responsiveness to any email inquiry.

Excluding auto replies, the response rate for banks to our experimental email solicitations was 23.3% (i.e., including compliant, noncompliant, and refusal responses): 3587 responses from 15,407 solicitations. The response checks had predicted roughly this level of replies. A 23.3% response rate is strong relative to much experimental research in business, but it still warrants investigation, both on its own account, and in terms of what it means for the conclusions that can be drawn from the replies that we did receive.

A key question is whether the banks and firms not replying to the solicitations were "soft-screeners," that is, 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 noncompliant 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 that system. We took some additional steps, which suggest that the large majority of

the nonresponses are "never-responders," and hence not a particularly relevant population for the aims of this study. Indeed, in our regression models, by far the strongest predictor of lack of reply in the experiment is nonresponse in the earlier check. This ameliorates, but does not entirely solve, the challenges to interpreting results.

Also prior to the launch of the experiment, we randomly sampled 95 of the banks that did not respond to these earlier checks and phoned them to ask 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 a.m. to 5 p.m. in the institution's local time. Many calls continued to ring with no answer. Several banks had the call dropped before any contact was made. A large majority of banks (more than 75%) could not be contacted on the phone number they provided on their website. The phone call approach primarily revealed that many international banks do not answer their phones.

As a final step, in the analysis of the data, we made reasonable assumptions about the group of likely soft compliers and soft refusers. We provide more details in the Appendix (E2d, Tables A.11 and A.12, pp. 13–14). To identify possible soft compliers or soft refusers, we identified all banks that responded in at least one of the experiment rounds or nonresponse checks and then identified the rounds in which those same banks did *not* respond and recoded those outcomes. More specifically, we recoded nonresponse to either compliance (for one analysis) or to refusal (for a second analysis), leaving the nonresponse category for banks that never responded in any round. This reduces nonresponse by about 20 percentage points and thereby increases compliance and refusal. In conducting the analysis with the recoded data, the results are substantively and statistically similar to the main results reported here (see E2d, Tables A.11 and A.12, pp. 13–14).

Treatment effects

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 some jurisdiction treatments. In our preregistered hypotheses, we predicted increases in non-response for all jurisdiction conditions save the United Kingdom (which was expected to be similar to placebo). We further expected decreased compliance rates for the corruption and terrorism conditions but increased compliance for the US and offshore treatments. We also anticipated increased refusal and decreased noncompliance rates for the US, terrorism, and offshore conditions along with increased refusal but no change from baseline for noncompliance in the corruption condition. Several of these predictions found support in the results, but others did not, producing mixed findings for our hypotheses. See Figure 1.

All the jurisdiction treatments caused significant increases in nonresponse rates. The corruption condition caused a statistically significant 5.1 percentage point increase in nonresponse, which represents a 6.4% increase from the base nonresponse rate of 76.7%. The terrorism, offshore, and US origin conditions likewise caused roughly 3–4 percentage point increases in nonresponse (4%–5% increases from the base rate). The UK treatment caused an unanticipated 1.8% increase in nonresponse (2.3% from the base). All of these increases in nonresponse are small substantively, but they nevertheless provide statistically significant evidence of soft-screening by some banks. This is evidence in favor of the behavioral model especially, though it is weakened by the relatively small substantive effects. The small substantive effects provide some qualified support for the organizational scripts arguments, suggesting that banks script contingent sensitivity to risk only to a limited degree.

The corruption condition caused statistically significant decreases in the rates of compliance and refusal of 1.6 and 3.0 percentage points, respectively. While these effects again are small substantively, when considered as 7%–13% of all responses, the results appear less trivial. The terrorism and offshore conditions likewise caused significant decreases in the compliance rates of 1.8 and 2.1 percentage points respectively (8%–9% of all responses). These results suggest that the banks that would have been compliant (or refused service) in the control condition are most likely failing to respond to the higher risk customers in treatment conditions. We read this as evidence of soft screening. This represents some good news for the risk-based approach to Know Your Customer compliance, but arguably only on a relatively small scale. These effects are substantively smaller than might be expected given the FATF guidelines. The substantively small results also stand in contrast to predictions made by international political economy scholars in a survey conducted prior to fielding the study (see Appendix E1, Figure A1, p. 11, and Tables A5 and A6, pp. 11–12).

Banks also proved largely insensitive to the reward treatments, save for the unexpected result that the intermediate level of turnover—$3 million—caused a statistically significant decrease of 2.1 percentage points in nonresponse and a 1.8% increase in refusals compared to the baseline of $500,000. However, the $30 million in turnover produced responses statistically indistinguishable from baseline (and the

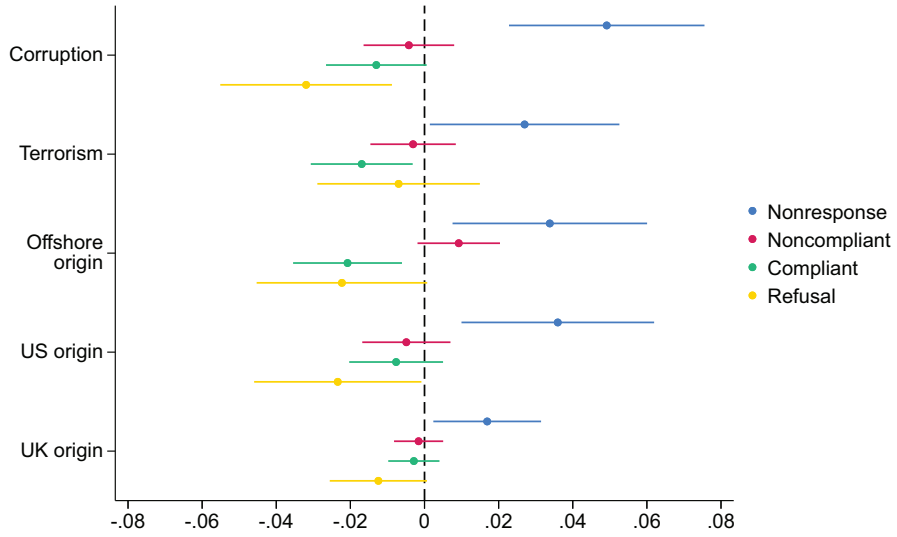

FIGURE 1 Effects on bank responses from jurisdiction treatments. *Notes*: The figure displays a coefficient plot of point estimates (dots) and 95% confidence intervals (lines). Confidence intervals that overlap the vertical line at zero are not significant statistically at the 95% level.

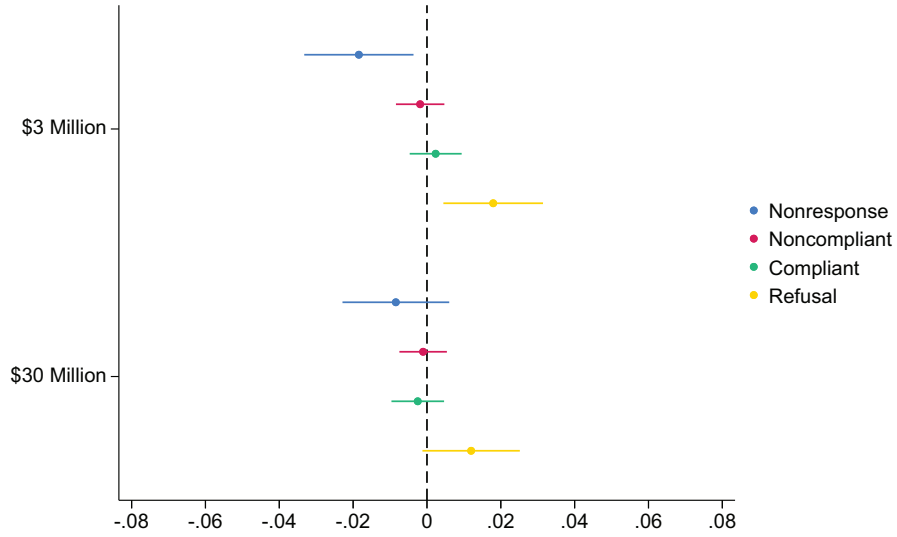

FIGURE 2 Effects on bank responses from money treatments. *Notes*: The figure displays coefficient plots with dots representing point estimates and lines representing 95% confidence intervals. Confidence intervals overlapping the vertical line at zero are not significant statistically.

$3 million condition) at the .05 level of significance.[9] See Figure 2. Still, the substantive effect sizes for the $3 million treatment were again small, and the decreased rates of nonresponse and increased refusal are—especially when considering the null results for the $30 million condition—nonsensical at worst and, at best, difficult to interpret.

The jurisdiction findings indicate marginal sensitivity to increased risk, though the effects of the money treatments suggest little sensitivity to reward. Given that the money conditions produced small effects that are difficult to interpret, the results may, on balance, be said to favor behavioral over expected-utility accounts. This is reinforced by the vanishingly small risk of criminal prosecution and the fact that the potential rewards from customers with $30 million in revenues are nontrivial.

These same substantively limited findings are furthermore important for policymakers, in that they cut against the central principle of the global regulatory regime designed to combat transnational financial crime: As a rule, banks appear to only marginally assess customer risk in granting or withholding

[9] We note that the increase in refusal rate in the $30 million condition is significant at the .1 level. Again, this result was unanticipated and difficult to explain from an expected-utility perspective. Regardless, it is small in substantive terms.

shell companies' access to the international banking system. Perhaps even more surprisingly, banks are relatively indifferent to increased customer monetary value in fielding solicitations.

It might be argued that the relative insensitivity to risk and reward is more in line with the expectations of organizational scripts: the idea of generic, standardized institutional responses to fulfilling recurrent and repetitive organizational tasks, such as the procedure for opening new corporate accounts. Scripts tend to be resistant to the imposition of finely grained differentiation and considerations of individual judgment. Indeed, much of the essence of scripts is to encourage consistency and conformity (March & Olsen, 1989). Regulators have sought to transfer to banks complex and delicate questions of differentiating between customers to establish different levels of customer risk and differential treatment. The evidence suggests that banks have responded by, at best, producing operating procedures that respond to risk mostly at the margins.

Aligned with these results, Tsingou finds that sophisticated risk models to counter money laundering are ignored in practice as bankers default to simple, generic rules of thumb in performing customer due diligence (2018). The authors' attendance at banking compliance industry events supports this impression of conformity and standardization, with one attendee summarizing the ethos as "if you're average, you're awesome." We emphasize, however, that the evidence in favor of organizational scripts here is merely suggestive and is found mostly in the breach. A more dispositive test of organizational scripts would require fuller examination of the response texts themselves to examine evidence of textual patterns of similarity across replies.

Increasingly critical public comments from regulators themselves lament that banks fail to distinguish between high- and low-risk customers, and instead adopt a generic "tick-box" (check-box) approach. Thus, the FATF President complained to the G20 finance ministers about "a widespread failure in effective supervision and compliance of anti-money laundering measures… Supervisors and compliance officers in banks and companies are meant to understand and mitigate the financial crime risks they face. But at the moment, most take a basic tick-box approach: they make sure forms are filled in correctly but don't focus on the real risks" (Pleyer, 2021). The FATF Executive Secretary has spoken in similarly damning terms: "When we look at the measures, the preventative measures that we expect banks to take… There's a 100 percent failure rate… all too often it just becomes a tick-box process. People do it because they think they have to, their supervisors say they have to… They're not applying a proper risk-based approach" (Lewis, 2019). Perhaps surprisingly, major banks themselves agree, openly asserting that they are regulated on the basis of ritualized compliance with the letter of the regulations, which are divorced from and indifferent to detecting and responding to the real risks (Wolfsberg Group, 2019). Our experimental results support this gloomy picture.

Banks' relative insensitivity to reward is important in helping to 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 take their chances given the low probability of sanctions. Such an explanation might accommodate accepting likely criminals at the same rate as low-risk corporate customers, but it cannot explain banks' tendency to accept high-rollers at the same (or even lower) rate as firms with much smaller turnover. Banks are also refusing potential customers at relatively high rates (especially compared to compliant and noncompliant responses), but they only marginally discriminate in their refusals based on either risk or reward.

CONSIDERATIONS INVOLVING RESEARCH ETHICS

Are these scientifically interesting and policy-relevant findings justified ethically? Strong criticisms have been raised about the social harms of field experiments and their apparent violations of the bedrock principles of justice and autonomy (McDermott & Hatemi, 2020). We are sensitive to these criticisms, but we also note that a deontological preoccupation with violations of rules may overlook pragmatic matters related to the principle of beneficence—that research should be undertaken for the social good—and that minimal harms to research subjects may be outweighed by major benefits of knowledge acquisition. The principle of beneficence motivated this study since systematic, credible information on banks' compliance with global transparency rules to counter serious crime is almost entirely missing and, thus, hamstrings governments' ability to see problems and fashion policy remedies.

We undertook multiple measures in this study to minimize potential harms and otherwise ensure ethical research conduct. First, we employed no active deception in our communication with banks. Rather than use aliases that may have enabled greater experimental control and minimized expense, we legally incorporated 12 companies in 9 jurisdictions. Researchers using their real names acted as representatives of the companies in their communication with banks. No correspondence from researchers contained untruthful statements.

Second, components of the study involve the establishment of bank accounts, the depositing of money, and the payment of bank fees, so some banks receive financial benefits from business with our companies. The study is not wasting bank representatives' time since, as part of their normal routines, they communicate with potential customers under the expectation that the correspondence may—but not necessarily will—lead to earnings. That was the case here. What is more, we estimate that bank representatives on average took less than 5 minutes to respond to our inquiries, and they generally did so with boilerplate language used in frequent correspondence.

Third, all communication is kept strictly confidential and all data are thoroughly anonymized before deposit in a replication archive. Such precautions help to minimize any potential harms to subjects from the study. Nevertheless, reducing harms should not cause us to lose sight of researchers' duty to undertake studies aimed at the public good. Social scientists as citizen investigators should feel an obligation to perform such research, especially by "studying up" in scrutinizing powerful institutions and organizations. Such studies can do much to enhance knowledge and inform policymaking.

CONCLUSION

The results of our global field experiment provide only limited support for the predictions of conventional expected-utility perspectives on risk and reward, but also new alternative mechanisms taken from behavioral economics and political psychology. Despite epitomizing the capitalist orientation to profit and loss in popular discourse, this experiment shows that banks are only marginally sensitive to reward and risk in dealing with corporate customers seeking bank accounts. One potential option for drawing firmer implications for what these results mean for a more institutionalist view premised on scripts may be to disaggregate outcomes to the level of phrases and words in email responses. Rather than coding emails into one of a few categories, and thereby potentially losing much of the information available, it may be possible to better identify scripts and test for their possible effects by analyzing the text as data to look for systematic variation in phrasing and content.

Our findings are important in policy terms in demonstrating that the major principle of banking regulation, the RBA, has limited effectiveness, supporting earlier skepticism on this score (Findley et al., 2014; Sharman, 2011; Tsingou, 2018). The faith shown by international organizations and governments that they can delegate complex risk judgments to banks in screening out criminal money is only weakly supported in the experimental results. There is some reassurance to be taken from the fact that the overall level of noncompliance is low, in that only a small minority of banks (3.2% of all contacted banks, which was 13.8% of received replies) were willing to provide corporate bank accounts without identifying the beneficial owners (in line with Morse, 2022). On the other hand, given the ease of web searches and email correspondence, criminals determined to find one of these minority of noncompliant banks would, on average, be able to do so by sending 30 emails or so—and thereby enter the financial system anonymously.

More broadly, these findings suggest caution in seeking to explain institutional-level behavior on the basis of the individuals within them, or via an analogy that institutions are simply individuals writ large. Despite their traditional focus on institutions and collective actors rather than individuals, and their general aversion to single-site studies, IR scholars have rarely embraced the advantages of field experiments. However, this study shows how institutions can be experimental subjects and that such experiments can achieve truly global coverage. Institutionalist theory can be tested and enriched experimentally.

ACKNOWLEDGMENTS

We thank many scholars for their valuable feedback on the research design and drafts of the manuscript. They include Lisa Argyle, Michael Barber, Bret Barrowman, Celeste Beesley, Daniel Berliner, Andrea Binder, Jake Bowers, Adam Brown, Terry Chapman, Stephen Chaudoin, Bill Clark, Alex Cooley, Cesi Cruz, Ryan Davis, Eric Dickson, Adam Dynes, Joshua Gubler, Jay Goodliffe, Don Green, Darren Hawkins, Kirk Hawkins, S.P. Harish, John Heathershaw, Macartan Humphreys, Wade Jacoby, Nate Jensen, Niels Johannesen, Bryan Jones, Chris Karpowitz, Josh Kertzer, Charles Littrell, Elisa de Anda Madrazo, Eddie Malesky, Dan Maliniak, Tom Mayne, Pat McDonald, Christoph Mikulachek, Helen Milner, Quinn Mecham, Quin Monson, Julia Morse, Chad Nelson, Karen Nershi, Irfan Nooruddin, Kelly Patterson, Lauren Prather, Jeremy Pope, Ryan Powers, Jessica Preece, Tenna Prelec, Joel Selway, Ricardo Soares de Oliveira, Ken Stiles, Mike Tierney, Dustin Tingley, Maurits Van Der Veen, Jim Vreeland, Rachel Wellhausen, Tim Werner, Sven Wilson, Scott Wolford, Cornelia Woll, and participants in research seminars and conferences at Brigham Young University, the College of William & Mary, Columbia University, Copenhagen, Experiments and Governance and Politics, Global Research in International Political Economy, the Harvard Government Department Workshop in International Relations, the International Political Economy Society, the London School of Economics, New York University - Abu Dhabi, Strategy and Business Environment, Northwestern University, the

University of Texas at Austin, and Rutgers University. Many people performed valuable research assistance on the project, including Majid Ali, Wajid Ali, Brittni Anderson, Sarah Austin, Rachel Boles, Jena Burgess, Felicia Caten, Holly Christensen, Jackson Freeman, George Gonzales, Mary Harris, Amber Hutchinson, Tayla Ingles, Samiya Javed, Michella Kopti, Robby Lindsay, Isaac Miles, James Morrow, Lydia Moss, Junaid Muksadur, Rachel Pourre, Meg Price, Tom Rehnquist, J.P. Reppetto, Abdus Salamup, Ronak Shah, Samuel St. Clair, Amerins Tolman, Tayzlie Tripple, Zac Whitehead, Sasha Woffinden, Abigail Woodfield, Byrson Xu, Scott Zimmerman, Adityamohan Tantravahi, Alex Walheim, Allie W Gretzinger, Amy Jeans, Ansh A Samdaria, Arijit Paladhi, Arvind Ashok, Ben Layton, Benjamin Christensen, Bethan Owen, Bianca Peña, Billy Matthias, Brandon J Gajeton, Brittni Anderson, Brooke Hunsaker, Carlo Moran, Chris Joyner, Chris Zimmer, Daniel Baldwin, Domingo A Salerno, Drew Burd, Durant Sellers, Elijah J Hance, Eliza Riley Oak, Elizabeth Whatcott, Emily Rose Verdoia, Eric Reuben, Erik Hansen, Erin P Eggleston, Estela Zuzunaga, Frederik Winguth, Gavin C Loyd, George Gonzalez, Hannah N Greer, Harumi Shiraishi, Ishaana Talesara, Jackie Clark, Jacob R Cosgrove, Jade G Tucker, Jess Allred, John K Vaughan, John P Reppeto, Jolynn Hall, Jonah Isaac, Jonathan Velasquez, Kanika Varma, Karina Bertelsmann, Kathlina M Brady, Kelsie L Arlin, Kimberly A Schuster, Kwinn Sutton, Kyler Morris, Laura Boyer, Lauren Barden, Leah Lewis, Lilly Andersen, Lydia J Moss, Maci Hiatt, Margaret Price, Mariana Caldas, Mia H Weissbluth, Mike Denly, Naomi Dorsey, Nathan Allred, Nathan Mcquarrie, Nikola Skerl, Nina Nomura, Reagan Curtis, Reilly Andreasen, Samantha Shoff, Sara J Lowe, Sean Christensen, Sofia Gonzalez, Stephen Hunsaker, Tayzlie Haack, Thomas Rehnquist, Tyler J Morrow, Wilo Zuzunaga, Zachary Whitehead, Zoe Morgan, Joyce Weng, Luca Labardini, Sage Griffee, Britta Thompson, and Vanna Chen. We especially thank David Bates, Joshua Bennion, Danny Cowser, Elden Griggs, Isaac St. Clair, and Michaela Bevan for managing important parts of the research effort. Critically, Rebecca Russavage and Eoin Power served as exceptional project managers.

REFERENCES

Barnett, Michael, and Martha Finnemore. 2004. *Rules for the World: International Organizations and World Politics*. Ithaca: Cornell University Press.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76(2): 169–217.

Coase, Ronald Harry. 1937. "The Nature of the Firm." *Economica* 4(16): 386–405.

Collin, Matthew. 2021. "What Lies Beneath: Evidence from Leaked Account Data on How Elites use Offshore Banking." Brookings Global Working Paper #156. https://www.brookings.edu/wp-content/uploads/2021/05/What-lies-beneath_Collin.pdf.

Davis, James W., and Rose McDermott. 2021. "The Past, Present, and Future of Behavioral IR." *International Organization* 75(1): 147–77.

DiMaggio, Paul, and Walter W. Powell, eds. 1991. *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.

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). 2019. Guidance for a Risk-Based Approach for Trust and Corporate Service Providers (TCSPs) (Paris).

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.

Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney (Eds). 2006. Delegation Under Anarchy: States, International Organizations, and Principal-Agent Theory. Cambridge: Cambridge University Press.

ICIJ (International Consortium of Investigative Journalists). 2020. "The FinCEN Files." https://www.icij.org/investigations/fincen-files/

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2): 263–91.

Lewis, David. 2019. FATF Executive Secretary Podcast Interview KYC360 18 November. https://kyc360.riskscreen.com/podcast/david-lewis-executive-secretary-fatf/

March, James G., and Johan P. Olsen. 1989. *Rediscovering Institutions: The Organizational Basis of Politics*. New York: Simon & Schuster.

McDermott, Rose, and Peter K. Hatemi. 2020. "Ethics in Field Experimentation: A Call to Establish New Standards to Protect to the Public from Unwanted Manipulation and Real Harms." *Proceedings of the National Academy of Sciences* 117(48): 30014–21.

Morgan, T. Clifton, Constantinos Syropoulos, and Yoto V. Yotov. 2023. "Economic Sanctions: Evolution, Consequences, and Challenges." *Journal of Economic Perspectives* 37(1): 3–30.

Morse, Julia C. 2022. *The Bankers' Blacklist: Unofficial Market Enforcement and the Global Fight Against Illicit Financing*. Ithaca: Cornell University Press.

Nader, Laura. 1974. "Up the Anthropologist—Perspectives Gained from Studying Up." In *Reinventing Anthropology*, edited by Dell Hymes, 284–311. New York: Vintage Books.

Nielson, Daniel L., and Michael Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57(2): 241–76.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Obermayer, Bastian, and Frederik Obermaier. 2016. *The Panama Papers: Breaking the Story of How the Rich and Powerful Hide Their Money*. New York: Oneworld.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Phillips, Trisha. 2021. "Ethics of Field Experiments." *Annual Review of Political Science* 24(1): 277–300.

Pleyer, Marcus. 2021. FATF President Speaks at Second G20 Finance Ministers and Central Bank Governors Meeting, 7 April. https://www.fatf-gafi.org/publications/fatfgeneral/documents/g20-fmcgb-april-2021.html

Powell, Walter W., and Paul DiMaggio. 1991. "Introduction." In *The New Institutionalism in Organisational Analysis*, edited by Walter W. Powell, and Paul DiMaggio, 1–40. Chicago: University of Chicago Press.

Sartori, Anne. 2003. "An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions." *Political Analysis* 11(2): 111–38.
Slough, Tara. 2021. "On Theory and Identification: When and Why We Need Theory for Causal Identification." Unpublished Manuscript, New York University.
Sharman, J. C. 2011. *The Money Laundry: Regulating Criminal Finance in the Global Economy*. Ithaca: Cornell University Press.
Thaler, Richard, Amos Tversky, Daniel Kahneman, and Alan Schwartz. 1997. "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test." *Quarterly Journal of Economics* 112(2): 647–61.
Thaler, Richard. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: Penguin.
Tsingou, Eleni. 2018. "New Governors on the Block: The Rise of Anti-Money Laundering Professionals." *Crime, Law and Social Change* 69: 191–205.
Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5(4): 297–323.
UNODC (United Nations Office on Drugs and Crime). 2011. Estimating Illicit Financial Flows Resulting from Drug Trafficking and Other Transnational Organized Crimes. Vienna.
United States. 1978. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Bethesda, MD: The Commission.
Wolfsberg Group. 2019. "Statement on Effectiveness: Making AML/CTF Programmes More Effective." https://db.wolfsberg-group.org/assets/6beba0cc-4bc1-4474-ac9d-5c388ce0ac14/Effectiveness%201%20pager%20Wolfsberg%20Group%202019%20FINAL_Publication.pdf.
Wright, Tom, and Bradley Hope. 2018. *Billion Dollar Whale: The Man Who Fooled Wall Street, Hollywood, and the World*. New York: Hachette.
Zarate, Juan C. 2013. *Treasury's War: The Unleashing of a New Era of Financial Warfare*. New York: Public Affairs.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Findley, Michael G., Daniel L. Nielson, and J. C. Sharman. 2025. "Banking bad? A global field experiment on risk, reward, and regulation." *American Journal of Political Science* 69: 545–59. https://doi.org/10.1111/ajps.12861