Causes of Non-Compliance with International Law:
A Field Experiment on Anonymous Incorporation

Michael G. Findley, Daniel L. Nielson, J. C. Sharman

18 February 2013

* The research design for this experiment was registered on March 2, 2011 with the Institute for Social and Policy Studies at Yale University. Of those interventions registered, we report on the FATF, Premium, Corruption, and Terrorism conditions in this paper. All other interventions outlined in the registered document are reported in other work. In our registration, we indicated that we would report results dichotomously as compliant or non-compliant, given a response. We still report response and non-response followed by a compliance level, but we expanded the set of possible types of compliance (non-response, non-compliance, partial compliance, compliance, and refusal). Presenting the information this way is more precise and is also consistent with the registry document because the fuller set of outcomes contains all information the dichotomized measures capture. While the full multinomial reporting is more precise, we report the dichotomized results (as originally registered) as well as selection models on the dichotomized results in the supporting information appendix. University and Institutional Review Board Clearances were received on 7 July 2010.

We would like to thank officials in the U.S. Department of Justice, Internal Revenue Service, Financial Action Task Force (FATF), World Bank, International Monetary Fund (IMF), Senate Permanent Subcommittee on Investigations, and United Nations Office on Drugs and Crime for their assistance in this project. We thank Drew Chapman, Zach Christensen, Madeleine Gleave, James Juchau, Brock Laney, Robert Morello, Catie Nielson, Wayne Sandholtz, Brittany Thorley, and Dane Thorley, for remarkable research assistance. We offer special thanks to Dustin Homer and Brock Laney for so ably leading the data collection. Funding provided by the Australian Research Council and the Department of Political Science; the College of Family, Home, and Social Sciences; and the David M. Kennedy Center for International Studies at Brigham Young University.
Abstract

Using two field experiments we probe the efficacy of international rules mandating that incorporation services establish their customers’ true identities. The standards were designed to prevent anonymous “shell” corporations central to money laundering, corruption, and many other crimes. Posing as consultants seeking confidential incorporation, we randomly assigned six experimental conditions in emails varying information about monetary reward, international and domestic law, and customer risk to 1,259 incorporation services in 181 countries and 1,405 firms in the United States. We find that firms in tax havens obey the rules significantly more often than in OECD countries, while services in poor nations sometimes prove more compliant than in rich countries. Offers to “pay a premium” induce non-compliance. Risks of corruption and terrorism decrease response rates but, alarmingly, also decrease compliance rates. Treatments raising relevant international and domestic law have no significant effect.
In 2002 the government of Kenya invited bids to replace its antiquated passport system. A French firm proposed €6 million, but the Kenyan government secretly awarded the contract to a British corporation, Anglo-Leasing Finance, which had tendered €30 million. Upon the acceptance of its inflated bid, Anglo-Leasing promptly subcontracted the work to its French competitor for €6 million and pocketed the remainder. A government official leaked word of the transaction to the press, which provoked outcry and recriminations in Britain and Kenya. Investigation revealed that the contracting firm, Anglo-Leasing, was merely a postal address in Liverpool; it was an anonymous “shell” corporation. Despite suspicions that the other €24 million were bound for corrupt officials, the investigation effectively stopped because it was impossible to determine the corporation’s owners.

The identities of the perpetrators were hidden notwithstanding international standards stipulating that all companies should be able to be traced to the real person in control. Dubbed “Anglo-Fleecing” by the press, this scandal provides merely one of many possible anecdotes underscoring the harm engendered by the lack of financial transparency and the non-compliance with international standards of disclosure (Wrong 2009; Kenya National Audit Office 2006). More broadly, debates in political economy about the extent to which the international system is rule-governed hinge on the causes of compliance and non-compliance. Dead-letter law is of little interest to either policy-makers or scholars; compliance is the key. Research on compliance with international rules has been hamstrung, however, by selection bias and endogeneity problems (Downs, Rocke and Barsoom 1996; von Stein 2005; Simmons 1998, 2010).

In principle, the best way to advance this research agenda would be to use field experiments (see Levitt and List 2007, 2009; Gerber and Green 2012). But to the extent that experiments have been used in international relations, they are usually laboratory experiments (for
excellent exceptions, see Hyde 2007, 2010). Until now researchers have assumed that it is impossible to use experiments to study the causes of compliance with international law. In upending this conventional wisdom, we present the first randomized field experiment probing compliance. To our knowledge, it is also the first fully global field experiment (for other path-breaking multi-national experiments, see Levine, Norenzayan, and Philbrick 2001 and Henrich et al. 2004).

Specifically, we assembled a subject pool of 2,664 incorporation services (for-profit firms that charge fees to set up companies for clients): 1,259 services in 181 countries for Experiment 1 and 1,405 firms in the United States for Experiment 2. We assigned the firms to treatment and placebo conditions that varied the rewards and risks associated with the potential transaction, and that manipulated information about domestic law enforcement and accepted international rules. After receiving IRB clearance, we used aliases, posed as consultants, and approached the firms via emails requesting confidential incorporation in order to avoid “excessive taxes” and limit legal liability. The study thus uses deception. Where social phenomena cause a great deal of harm, as with corruption, money laundering, tax evasion, sanctions busting, and the financing of terrorism, and where perpetrators are unlikely to report their behavior truthfully, the benefits of learning about their actions outweigh the costs of deception.

The two experiments evaluate response and compliance rates elicited by randomly assigned email treatments compared with a Placebo. The Placebo emails originate from aliases purportedly based in innocuous, low-corruption OECD countries. The treatments derive from international standards stipulated by the Financial Action Task Force (FATF) – the international institution charged with overseeing corporate transparency. Nearly every country has assented to
the FATF rules, which require full disclosure of company owners’ identities and mandate that corporate service providers employ a “risk-based approach” in scrutinizing potential customers.

In Experiment 1, the first treatment mentions the FATF standard requiring identifying documents. The second treatment employs language about which the FATF explicitly cautions incorporation services, offering to “pay a premium to retain confidentiality.” The third treatment probes the effects of the FATF’s injunction against corruption – the aliases originate from eight relatively indistinguishable countries that all rank high on scales of corruption. Finally, the fourth treatment examines the efficacy of FATF warnings regarding terrorism: the aliases claim citizenship in nations associated with terrorism but consult in Saudi Arabia for Islamic charities. In Experiment 2, performed only on firms in the United States, we drop the premium condition and substitute a fifth condition in which we inform subjects that U.S. law requires identity disclosure and that the Internal Revenue Service enforces the requirement.

In what follows, we situate the research in the relevant literature, provide background on shell companies, describe international standards, discuss the ethics of deception, and describe the research design. Aside from descriptive statistics, we use difference-in-means tests and multinominal logit estimates to analyze the findings (with multinomial probit and a selection model as robustness checks, which appear in the online appendix).

The results of the experiments are often counter-intuitive. Incorporation services based in tax havens comply with international standards at significantly greater rates than those in OECD countries. Also, providers in developing countries are sometimes significantly more compliant than in wealthy nations. Disturbingly, approaches from clients posing corruption and terrorism risks tended to reduce both response and compliance rates in both experiments. The offer of premium payment lowered compliance rates. Identifying the applicable international rules and
rule-maker (FATF) had no significant effects on response or compliance rates in either experiment. Similarly, raising the specter of the IRS had no significant effect on U.S. firms.

**Background and Literature**

Provocative stories like “Anglo-Fleecing” prompt inquiry into the causes of adherence to international corporate transparency standards to which nearly every country has agreed. Appropriately, the general subject of compliance has received prominent scholarly attention (see Chayes and Chayes 1993; Downs, et al. 1996; Simmons 2000, 2010; Raustiala and Slaughter 2002; Von Stein 2005). In their foundational article, Chayes and Chayes (1993, see also Henkin 1979) conclude that compliance with international standards is the norm. This “managerial school” holds that non-compliance generally arises as a result of ignorance, ambiguities in agreements and treaties, and administrative shortcomings, as opposed to deliberate attempts to defy such standards. If the managerial logic holds, learning about international rules should induce compliance, and deliberate attempts to violate standards should meet with greater enforcement. Alternatively, rationalists hold that compliance results from concerns about international reputation and is reinforced by geographic diffusion, but international law still has constraining effects (Simmons 2000).

Responding to the optimism regarding the efficacy of international law, Downs et al. (1996) brought to light the challenges posed by endogeneity and selection problems. Compliance with international standards might be high precisely because states agreed to those standards where compliance proves easiest (Raustiala and Slaughter 2002; Drezner 2007; Von Stein 2005). If this is so, selection effects – and not the inherent constraining power of international law – explain compliance.
Other fields, especially economics, have addressed the problems of selection and endogeneity by employing field experiments using random assignment to treatment and control conditions. Any difference in outcomes between groups can be causally attributed to the intervention, because in expectation randomization balances – and therefore neutralizes – the effects of all other observable as well as unobservable factors. This approach has achieved prominent success in behavioral and development economics, where the units of analysis are ordinary individuals who can be effectively treated as research subjects in experiments (Chattopadhyay and Duflo 2004; Banerjee et al. 2007; Levitt and List 2009; Humphreys and Weinstein 2009). The problem in international relations is clear: the objects of inquiry are typically sovereign governments, which are difficult to manipulate both practically and ethically.

Yet in many important areas of IR, including financial transparency, governments are not the main locus of compliance with international standards. Instead, important contributors to related debates agree that ordinary citizens and firms make the specific decisions that ultimately aggregate to a pattern of nationwide compliance or violation (e.g Drezner 2007: 13; Keohane et al. 1993: 16). Referencing the compliance literature specifically, Simmons (2010) emphasizes the need for scholars to study non-state actors, which can better capture the actual locus of compliance.

In financial transparency the relevant actors are corporate service providers (CSPs) – for-profit firms that specialize in setting up businesses for others. International standards require CSPs to establish the true identity of individuals seeking incorporation. CSPs meet this requirement by obtaining a notarized copy of the picture page of the individual’s passport and proof of address, such as an electricity bill. Yet prior scholarship has suggested that such standards are quite variably enforced (Author 2010; World Bank 2011). Non-compliance with financial trans-
pacity standards enables the formation of shell corporations that cannot be traced to identifiable individuals, which in turn facilitates corruption, money laundering, tax evasion, and even terrorism (OECD 2001; FATF 2006; Schott 2006; World Bank 2011).

An inter-governmental institution, the Financial Action Task Force (FATF), both sets and monitors enforcement of regulations to counter money laundering and terrorist financing, and 180 countries have assented to the standards. The FATF has published 40 Recommendations directing countries to avoid harboring unsavory financial activity within their borders (FATF 2012). Specifically, the key provision states: “Countries should ensure that there is adequate, accurate and timely information on the beneficial ownership and control of legal persons [i.e. companies] that can be obtained or accessed in a timely fashion by competent authorities” (FATF 2012).

The FATF has publicly blacklisted non-compliant jurisdictions and compelled nearly all countries to align their domestic laws with the international standards (Drezner 2007: 142-145; Author 2011). The FATF also issues prescriptive instructions that firms should follow in order to minimize the possibility of enabling financial crime, called the “risk-based-approach.” This enjoins businesses to be on the lookout for certain customer profiles – including offers to pay a premium, origin in corruption-prone countries, work for “charities,” and association with havens for terrorism (FATF 2008). We base three of our experimental interventions on these recommendations.

These FATF standards were enacted far in advance of member states’ domestic provisions (UN 1998; OECD 2001; FATF 2006; World Bank 2011). So acceding to FATF standards was likely not a case of selection bias. In the detailed assessments of each FATF members’ laws
only one, Italy, had legislation that met FATF standards on identifying company owners before the FATF passed this standard (World Bank 2011).¹

We argue that the best way to test effectiveness of the FATF standards involves minor deception. Both experimental guidelines and federal regulations allow exceptions to informed consent under certain conditions: (1) the costs are minimal, (2) the subjects are not exposed to emotional or physical pain, (3) the research cannot be performed in another way, and (4) the benefits are significant (Belmont Report 1979; CFR 46.116(d)). The present research qualifies under each condition.

We estimate that it took subjects roughly five to ten minutes to respond to our queries, so costs were low. Subjects were clearly responding within the context of their normal day-to-day routines and therefore did not face any harm from the study (see Singleton et al. 1985, 452). All identifiable information about incorporation service providers has been eliminated to protect subjects’ privacy. Given the great damage caused by money laundering, corruption, tax evasion, sanctions busting, and terrorist financing, the potential benefits of unbiased findings on corporate transparency are significant. And the conventional methods of social science, including surveys and interviews, will produce biased results because non-compliant actors will almost certainly hide their true actions from researchers. We could thus conceive of no other way to achieve unbiased results without deception. This research builds on important precedents using a similar method in economics and political science to learn about discrimination (Bertrand and Mullainathan 2004; Butler and Brookman 2011).

¹ A partial exception to this question of timing relates to tax havens. In 2000 the FATF blacklisted several tax havens and other standards violators, especially focusing on the identification of shell company owners. Despite complaints of double-standards, every targeted country subsequently changed their laws to meet the FATF’s demand (FATF 2007).
Research Design

Subject Pool

We carried out the experiments on a large pool (N = 2,664) of incorporation services worldwide. Experiment 1 targeted 1,259 firms based in 181 countries. Experiment 2 treated an additional 1,405 firms in the United States. All data collection and correspondence for both experiments took place between January and August 2011. We built a convenience sample of CSPs from information available on the Internet using systematic Google searches. Some incorporation service providers exist mainly as Internet entities; others are specialized law firms offering incorporation as one of several services. Each service offers to incorporate new businesses within a specified set of countries for a fee usually ranging between $500 and $3,000. We acknowledged that the firms listed online may prove more likely to comply with international standards than firms that are “off the radar.” Therefore, the data presented here may actually overstate the degree of compliance with global transparency standards.

Experiment 1 Block Randomization

We administered a blocking procedure on the subject pool to improve covariate balance across experimental conditions. Subjects are grouped according to values of observable covariates and the randomization then takes place within each blocking stratum. As Gerber and Green (2004) emphasize, blocking ensures that the covariates are not collinear with assignment to experimental conditions. It generates balanced proportions of subjects in each condition for each block and thus rules out certain “rogue” randomizations by design, leading “to substantially more precise estimates than simple randomization” (Gerber and Green 2012, 114).

For Experiment 1, we blocked by company type (incorporation service vs. law firm) and country group. After creating separate categories for OECD countries and tax havens, we used
the World Bank’s *Ease of Doing Business Index* (2011) to differentiate among developing countries. The five country categories were therefore (1) OECD members; (2) tax havens; and developing countries grouped according to (3) high, (4) medium, and (5) low “friendliness to business.” Countries in each of the five categories are listed in Online Appendix A. We divided these five strata again by whether or not subjects were incorporation services or law firms, leaving 10 strata within which we ultimately made the random assignments to experimental conditions.

Within each blocking stratum, we randomly assigned a treatment or the Placebo condition. We also randomly assigned an alias (and associated country of origin, 20 in total), the text of the email (among 33 different possibilities), and the subject line of the email (10 options). See Online Appendix A for examples of the experimental conditions.

**Experiment 1 Treatments**

We sent emails from aliases posing as consultants to each of the 1,259 service providers in the international subject pool for Experiment 1. All emails request confidential incorporation. The main outcome of interest is the degree to which subjects comply with international law by demanding certified identity documents.

Experiment 1 subjects were randomly assigned to one of five conditions:

---

2 In statistical analysis we found that the categories of developing countries were statistically similar across all business friendliness levels for nearly all analyzed outcomes, with one minor exception (medium-business-friendly developing countries were significantly more part-compliant than the low-friendliness group). Given this, we pooled the developing-country categories for convenience in reporting the data and ease of interpretation.
1. **Placebo: Low-Corruption OECD Country**
   The email originates from an alias based in “Norstralia,” or one of eight countries (Australia, Austria, Denmark, Finland, Netherlands, New Zealand, Norway, and Sweden), ranked at the top of the list of nations that are perceived to effectively control corruption.

2. **FATF/International Law**
   The email references the Financial Action Task Force and its requirements for information disclosure. We note here that, for firms with no prior knowledge of these rules, this treatment acts to inform them and thus presents a direct test of the effects of knowledge of the law on compliance. Our interviews and a comprehensive survey we conducted of more than 300 CSPs suggest that knowledge of FATF standards is widely lacking (more than 70 percent had not been briefed on the standards), so this direct effect should be common. For firms with prior knowledge of the standards, the treatment acts to “prime” them by calling to mind the standards. Both mechanisms are consistent with the managerial logic, but only the first presents a direct test of managerialism.

3. **Premium**
   The email offers to pay a premium for confidential incorporation. The condition thus probes the effectiveness of the FATF’s injunction to companies that they screen customers who offer “to pay extraordinary fees for services which would not ordinarily warrant such a premium” (FATF 2006, 22). If firms are following international law, compliance should increase under this condition. Alternatively, if firms on balance are behaving opportunistically, this condition should decrease compliance rates.

4. **Corruption**
   The email originates from a consultant working in “government procurement” and hailing from “Guineastan” – eight relatively indistinguishable countries ranked by Transparency International as high in perceived corruption: Equatorial Guinea, Guinea, Guinea Bissau, Papua New Guinea, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan. The corruption treatment seeks to learn the efficacy of the FATF risk guidelines warning of nations “identified by credible sources as having significant levels of corruption, or other criminal activity” (2006, 21). This condition should increase rates of compliance.
5. **Terrorism**

The email originates from a citizen of Lebanon, Pakistan, Palestine, or Yemen living in Saudi Arabia and consulting for a Muslim charity. All four origin nations were identified as key sites of suicide terror by Pape (2005). The terrorism treatment thus examines the efficacy of two FATF risk factors: terrorist countries and charities. The FATF warns against “[c]ountries identified by credible sources as providing funding or support for terrorist activities that have designated terrorist organisations operating within them” (2006, 21). Likewise, the FATF enjoins companies to screen “[c]harities and other ‘not for profit’ organisations which are not subject to monitoring or supervision (especially those operating on a ‘cross-border’ basis)” (2006, 22). The terrorism condition should also increase compliance rates.3

Each email was sent by a putative consultant who expressed a desire to form an offshore corporation to enhance confidentiality while limiting legal liability and tax payments (examples of each are included in Online Appendix A). While legitimate consultancy arrangements are widespread, consultancy fees are a common alibi for funds derived from criminal activities (Author 2010: 2-4; World Bank 2011).

We acknowledge here that FATF risk-based standards regarding extraordinary payments, corruption, and terrorism are grounded in real-world concerns that extend beyond the FATF’s reach. The treatments therefore probe the effects of conditions about which the FATF has warned, not simply the FATF rules themselves. This introduces some potential confounds to the conditions implying that any treatment effects will include the effects of factors beyond the force

---

3 We note that these last three treatments draw on a similar causal logic invoking potential opportunism in corporate service providers, so causal mechanisms are not mutually exclusive. However, each intervention is based on a different provision of the FATF’s risk-based approach, so they are conceptually distinct. Moreover, treatment effects are different across the three interventions, suggesting that the inclusion of all three conditions is warranted.
of international law. We would argue that the treatment effects of terrorism, corruption, or offering a premium are interesting in their own right, independent of international rules, so they thus warrant study. However, since the FATF has made all three conditions central to its standards, any treatment effects should nevertheless reflect on international law, even if subjects are not responding exclusively to the global rules per se.

To execute the experimental conditions, we created fictitious identities based on the most common male names in each of the countries. The names were carefully vetted to insure that no extraordinary connotation would be applied to any alias, such as with a famous actor, athlete, or politician. Twenty aliases with associated email accounts were created; each corresponded to one of the countries used in the Placebo, Corruption, and Terrorism conditions.

Different texts for 33 unique emails were created and randomly assigned to the subjects. All 33 emails were written according to the same criteria, but they were infused with different language, style, grammar, and syntax to ensure uniqueness. For the emails originating from aliases in non-English-speaking countries, two small spelling errors were introduced to improve authenticity. The many diverse email texts both minimized the potential for detection and mitigated the potential outlier effects of any one email text.

**Coding Protocol**

Responses to the control and treatment emails were coded as Refusal, Compliant with international standards, Partially Compliant, or Non-Compliant (see Online Appendix A for examples of replies). We categorized the remainder as No Response.\(^4\) International standards mandate

\(^4\) Since subjects in the No Response category are not enabling anonymous shell corporations, an argument might be made that No Response is akin to Refusal or Compliance in preventing shady incorporation. The data,
that service providers require a certified copy of at least one official photo identity document (notarized copy of a passport picture page or national identification card), along with proof of address (such as an original utility bill or a notarized copy), before forming a company for the customer. Service providers should then keep this documentation on file so that the company can be traced back to its true owner by law enforcement should the need arise.

When providers did not respond to the alias email within five business days, the researcher posing as the consultant prompted the subject again with a standardized, brief second email. Where service providers’ first response to the approach email did not specify the identity documentation required (if any), researchers drew from a standardized set of response scenarios to draft an appropriate follow-up email.

If firms declined service, we coded them as “Refusal.” Services were categorized as “Compliant” if they required notarized photo identification. We coded subjects as “Partially Compliant” if they required a copy of photo identification but failed to demand notarization or certification of the document. Finally, services that did not request documentation of any kind were classified as “Non-Compliant.” Requirements for identity documentation are outlined by the FATF and clarified by the Basel Committee (2001).

Once the specified information on identity documentation was obtained, researchers informed providers that “needs have been met” and that they no longer sought to incorporate. To

however, suggest that soft refusers amount to only 9 to 13 percent of the subject pool, leaving 42 percent (worldwide) to 73 percent (in the U.S.) failing to respond in the less-risky Placebo condition. Suspicious that many non-responses were due to incapacity or indifference, not soft refusal, we followed up with multiple rounds of correspondence from different aliases. The follow-ups culminated with brief non-response checks that essentially asked whether the firms were still in business and assisting customers. In the end, among the non-responses, 52 percent (worldwide) to 84 percent (in the U.S.) still failed to respond to this very low risk inquiry. These findings suggest that conflating No Response with the Refusal and/or Compliance categories would prove problematic methodologically and may lead to bias; the majority of non-responding firms do not reply to any inquiry, even the most innocuous we could design.
preserve the security of the exercise, all correspondence took place through specially created Internet email accounts. Proxy servers that randomly assign IP addresses throughout the globe (with a concentration in Europe and East Asia) were used to prevent service providers’ determining that emails in fact came from within the United States.

A skeptic might worry that company providers could employ a “bait and switch” strategem that involves initially promising anonymous incorporation but then asking for identity documents further along the process. To demonstrate the validity of our approach here we rely on a closely-related audit study (Author 2010). Closely mirroring our approach here, the audit study was based on email solicitations for shell companies from a purported consultant to CSPs, and again was focused on determining what identity documents (if any) were necessary to establish a company. However, the audit study went through the whole process of incorporation – bar the final transfer of funds – with 42 separate providers. In every single case, the initial email specification of whether or not identification documents were required was an accurate reflection of the requirements at each subsequent stage of the process. Furthermore, in three cases the author actually purchased shell companies from providers. Once again, the identification requirements remained consistent from the initial email contact until the conclusion of final stage when the money had changed hands. These results thus provide strong evidence for the contention that the email correspondence we received from providers does in fact constitute a valid indicator of compliance with corporate transparency standards.

**Experiment 1 Data and Results**

**Observational Data on Country Categories**

We begin with observational and descriptive statistics for Experiment 1. These statistics reveal two surprising results: first, service providers in tax havens are far more diligent in ob-
serving international incorporation rules than those in OECD countries, and second that firms in OECD members are not significantly more compliant than those in developing countries (and when there are significant differences, they generally favor greater compliance by firms in poor nations).

Figure 1 displays the proportions and differences for each of the five outcomes across the three country categories: OECD members, tax havens, and developing nations. In the upper panel, Figure 1 displays proportions for each outcome across the country groupings. In the lower panel, the figure displays differences from the OECD and 95-percent confidence intervals for those differences generated in difference of means tests (indicating statistical significance where they fail to overlap zero).

Because we used this categorization in our block randomization procedure, the treatments – confirmed by randomization checks – were balanced across the three types of countries, which enables comparison without undue concern that these descriptive statistics are biased by treatment effects. But given that country type cannot be manipulated, we emphasize that the data shown in Figure 1 and Table 1 are observational and not experimental.
Figure 1: Outcome Proportions by Country Group
Table 1: Contingency Table of Outcomes across Country Groups

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>No Response</th>
<th>Non-Compliant</th>
<th>Part-Compliant</th>
<th>Compliant</th>
<th>Refusal</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD Members</td>
<td>297</td>
<td>151</td>
<td>44</td>
<td>33</td>
<td>37</td>
<td>32</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>50.8%</td>
<td>14.8%</td>
<td>11.1%</td>
<td>12.5%</td>
<td>10.8%</td>
</tr>
<tr>
<td>Tax Havens</td>
<td>299</td>
<td>106*</td>
<td>14*</td>
<td>52*</td>
<td>107*</td>
<td>20</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>35.5%</td>
<td>4.7%</td>
<td>17.4%</td>
<td>35.8%</td>
<td>6.7%</td>
</tr>
<tr>
<td>Developing Nations</td>
<td>662</td>
<td>364</td>
<td>53*</td>
<td>109*</td>
<td>62</td>
<td>74</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>55.0%</td>
<td>8.0%</td>
<td>16.5%</td>
<td>9.4%</td>
<td>11.2%</td>
</tr>
</tbody>
</table>

Significant in difference-in-means tests compared to OECD Countries: * .05 level.

That said, the contrasts are stark. Firms in tax havens were significantly more likely to respond to inquiries (64 percent) compared to services in OECD (49 percent, \( p = .000 \)) and developing countries (45 percent, \( p = .000 \)). Thus, as Figure 1 and Table 1 show, outcomes for No Response were significantly different in tax havens compared to the other two country groups, and firms in OECD and developing countries were not significantly different from one another.

Relating directly to Compliance with international law, tax-haven services were significantly more likely to demand notarized identification than services in either OECD or developing countries. The Compliance rate for tax-haven firms of 36 percent was nearly three times greater than OECD firms (12 percent, \( p = .000 \)) and nearly four times greater than developing-country services (9 percent, \( p = .000 \)). Also, OECD firms were no more Compliant than developing-country services by conventional standards of statistical significance. This is counter-intuitive because non-compliance with international standards is often said to be a product of inability to comply, because of a lack of resources, rather than a product of unwillingness (Chayes
and Chayes 1993; Author interviews FATF 2007, IMF 2010, World Bank 2011). Considered a different way, more than 55 percent of tax-haven services responding to our inquiries required certified identification documents compared to only 25 percent in OECD countries and 21 percent in developing countries.

Likewise, at a mere 5 percent, tax-haven firms were significantly less likely than OECD companies (15 percent, \( p = .000 \)) and developing-country services (8 percent, \( p = .061 \)) to offer incorporation without identifying documents and thus be found in Non-Compliance. Surprisingly, services in wealthy countries displayed willingness to violate international law significantly more often than firms in poor countries (15 vs. 8 percent, \( p = .001 \)).

Both the Compliance and the Non-Compliance results undermine the conventional wisdom that firms in tax havens are pariahs that ignore corporate transparency standards. Tax-haven firms were also significantly more likely to be found in Partial Compliance by requiring non-notarized identity documents than firms in OECD countries (17 vs. 11 percent, \( p = .028 \)). Firms in developing countries were also significantly more likely to be found Part-Compliant than in OECD nations. Because such a large proportion of firms in tax havens complied with international law, fewer may have been left over for Refusal of service, an outcome in which a significantly smaller proportion of tax-haven services were found compared to OECD (\( p = .077 \)) and developing-country (\( p = .030 \)) companies. This may also reflect the greater competitiveness of tax-haven services in their increased willingness to do business with most potential clients. But again, these findings are based on observational data, so the question remains open as to whether specific interventions derived from the international standards can cause firms to alter their compliance levels. This brings us to the results from the two experiments.
**Experiment 1 Results**

Two randomization checks – using both individual logistic and multinomial logistic regressions – suggest balance of covariates among experimental conditions. We find that neither Company Type (incorporation service vs. law firm) nor Country Group (OECD, tax haven, developing) was significantly related to the probability that a given firm would be assigned to a specific condition.

To analyze the data, we elected to treat each outcome category – No Response, Non-Compliant, Part-Compliant, Compliant, and Refusal – as independent outcomes and employ contingency tables, difference of means and proportions tests, and multinomial logistic analysis to assess the results. We argue that the fewer assumptions in this approach makes it particularly appealing, so we feature it here and use alternatives as robustness checks reported in Online Appendices C and D.

We thus begin with simple descriptive statistics by experimental condition on the outcomes across the five different categories: No Response, Non-Compliance, Part-Compliance, Compliance, and Refusal. In expectation the balance induced by randomization enables very simple statistical analysis, so we lead with these results and emphasize that difference-in-means tests should reveal the principle treatment effects. Figure 2 corresponds to Table 2. In the upper panel, Figure 2 displays proportions for each outcome across experimental conditions, and in the lower panel the figure displays treatment group differences from the Placebo (and 95-percent confidence intervals).
Figure 2: Outcome Proportions across Experiment 1 Conditions
Table 2: Contingency Table of Outcomes across Experiment 1 Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>No Response</th>
<th>Non-Compliant</th>
<th>Part-Compliant</th>
<th>Compliant</th>
<th>Refusal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>268</td>
<td>115</td>
<td>20</td>
<td>51</td>
<td>51</td>
<td>31</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>42.9%</td>
<td>7.5%</td>
<td>19.0%</td>
<td>19.0%</td>
<td>11.6%</td>
</tr>
<tr>
<td>FATF</td>
<td>232</td>
<td>106</td>
<td>24</td>
<td>36</td>
<td>42</td>
<td>24</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>45.7%</td>
<td>10.3%</td>
<td>15.5%</td>
<td>18.1%</td>
<td>10.3%</td>
</tr>
<tr>
<td>Premium</td>
<td>235</td>
<td>117</td>
<td>20</td>
<td>42</td>
<td>29*</td>
<td>27</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>49.8%</td>
<td>8.5%</td>
<td>17.9%</td>
<td>12.3%</td>
<td>11.5%</td>
</tr>
<tr>
<td>Corruption</td>
<td>276</td>
<td>152*</td>
<td>31</td>
<td>35*</td>
<td>42</td>
<td>16*</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>55.1%</td>
<td>11.2%</td>
<td>12.7%</td>
<td>15.2%</td>
<td>5.8%</td>
</tr>
<tr>
<td>Terrorism</td>
<td>248</td>
<td>132*</td>
<td>16</td>
<td>30*</td>
<td>42</td>
<td>28</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>53.2%</td>
<td>6.5%</td>
<td>12.1%</td>
<td>16.9%</td>
<td>11.3%</td>
</tr>
<tr>
<td>Total</td>
<td>1,259</td>
<td>622</td>
<td>111</td>
<td>194</td>
<td>206</td>
<td>126</td>
</tr>
<tr>
<td></td>
<td></td>
<td>49.4%</td>
<td>8.8%</td>
<td>15.4%</td>
<td>16.4%</td>
<td>10.0%</td>
</tr>
</tbody>
</table>

Significant in difference-in-means tests compared to Placebo condition: * .05 level.

In Table 2, we report cell size and conditional proportions for each treatment and outcome category. Using asterisks, we also report statistical significance at the .05 level in simple difference-in-means tests for the treatments compared to Placebo when the specific outcome in question is scored “1” and all other outcomes across the row are scored “0.”

The results in Table 2 suggest that informing incorporation services about international law has no statistically significant effect on their propensity to Comply either by demanding identity documents or Refusing service. However, offering to pay a Premium or origin in a country associated with Corruption or Terrorism did produce statistically significant differences, but those differences undermine the argument that international law proves effectual. Instead,
the results on balance support the counter-argument that a significant share of materialist actors will pursue their own self-interest despite explicit cues that their actions are inappropriate.

Association with Corruption or Terrorism did cause significantly greater proportions of services to ignore email inquiries, suggesting some soft compliance. The Corruption and Terrorism treatments increased the rates of No Response from 42.9 percent in the Placebo condition to 55.1 percent for Corruption \( (p = .005) \) and 53.2 percent for Terrorism \( (p = .024) \). This suggests that firms exercise some discretion in responding to inquiries.

A prima facie expectation for the treatments would be a lower Response rate matched with a higher Compliance rate. Some providers should react to riskier clients by failing to reply, while others might be more punctilious in requiring identity documents, in accord with the FATF’s prescribed risk-based approach. Thus, the statistically lower response rate for Corruption and Terrorism, and lower response rates generally across the conditions, indicate that the treatments may induce some soft compliance. But why do the Corruption and Terrorism conditions simultaneously make providers less likely to request identifying documents?

It seems that the providers’ initial choice of whether or not to reply selects responding firms that are more risk-tolerant than the average subject. According to this logic, those most likely to be compliant with international standards and most attuned to the dangers of providing anonymous shell companies choose not to respond in the first place. As we suspected, it appears that there may also be a set of incorporation services that, regardless of risk, employ a standard operating procedure that requires little identity disclosure from potential clients. Thus, when the risk-averse withdraw through No Response, relatively more risk-acceptant companies may be left in the subject pool, thus altering the outcome proportions for some of the treatments.
From the point of view of individuals seeking to evade international law, this response pattern may make their task easier. Indeed, clients offering to pay a Premium caused significantly lower Compliance rates than in the Placebo condition – dropping one third from 19.0 percent to 12.3 percent ($p = .041$). Likewise, Refusal rates in the Corruption condition decreased by half, from 11.6 to 5.8 percent ($p = .017$). Finally, both the Corruption and Terrorism conditions lowered Part-Compliance rates by roughly one third from the 19.0 percent seen in the Placebo condition to 12.7 ($p = .043$) and 12.1 ($p = .031$), respectively.

**Multinomial Models**

We use multinomial logit and probit models to analyze the conditional probabilities of subjects choosing a given outcome compared to a base outcome (Long and Freese 1996). Multinomial models enable us to capture all possible categories of outcomes without a loss of information from collapsing the data. Online Appendix Table B1 reports the coefficients and robust standard errors for the multinomial models. Those receiving the Premium prompt compared to the Placebo were significantly less likely to be found Compliant than in the No Response category ($p = .022$). The Corruption condition caused significant decreases in the proportion of Part-Compliant ($p = .005$) and Compliant ($p = .041$) subjects compared to the Placebo. And the Terrorism treatment likewise caused a lower rate of Part-Compliance ($p = .002$) and Compliance ($p = .083$) compared to the Placebo. The FATF treatment does not appear to cause significant differences from the Placebo condition.$^5$ The results are broadly similar qualitatively when we include controls for Company Type, OECD, and Tax Haven. We display the coefficients, standard

---

$^5$ We note here that, as might be expected, these results shift when we rotate the base outcome. Tables displaying the results of these rotations can be found in Tables B2-B5 in Online Appendix B, but they generally support the conclusions reported.
errors, and significance levels for these specifications using control variables in Table B9 in Online Appendix B.

To further facilitate interpretation we calculate a set of predicted probabilities for each of the different conditions, displayed in Table 3. We rely on the models that used the control dummy variables for Company Type (incorporation service = 0, business law firm = 1), OECD Members, and Tax Havens (with Developing Countries as the comparison group). OECD membership was sometimes significantly related to the outcomes, but these results suggest that rich countries generally proved significantly less compliant than the poor countries in the reference group. Tax Haven status was significant in ways consistent with the results reported in Figure 1 and Table 1 above. Company Type was also often significant as reported in Table B8 in Online Appendix B, with incorporation services demonstrating significantly more Part- and full Compliance than law firms.

Table 3 displays changes to the predicted probabilities for the Placebo condition vs. each of the treatments, again with No Response set as the base outcome for comparison with the other outcomes (and Compliance set as the base outcome for comparison with No Response). This analysis employs covariates for company type and country group but omits the covariate results for ease of display (see Online Appendix Table B9 for the full results). For two of the conditions, Corruption and Terrorism, the treatment caused significant increases in No Response, a potential indicator that subjects may have been complying in a “soft” way through ignoring the email.

**Table 3: Predicted Probabilities of Outcomes for Experiment 1**

<table>
<thead>
<tr>
<th>Conditions</th>
<th>No Response</th>
<th>Non-Compliant</th>
<th>Part-Compliant</th>
<th>Compliant</th>
<th>Refusal</th>
</tr>
</thead>
<tbody>
<tr>
<td>FATF Placebo</td>
<td>48.3%</td>
<td>6.2%</td>
<td>28.8%</td>
<td>11.6%</td>
<td>5.1%</td>
</tr>
<tr>
<td>Treatment</td>
<td>51.8%</td>
<td>9.5%</td>
<td>21.4%</td>
<td>10.5%</td>
<td>6.9%</td>
</tr>
<tr>
<td>Change</td>
<td>3.4%</td>
<td>3.3%</td>
<td>-4.0%</td>
<td>-1.1%</td>
<td>1.8%</td>
</tr>
</tbody>
</table>

| Premium   |       |      |       |       |      |
| Placebo   | 44.3% | 7.0% | 27.5% | 13.3% | 8.0% |
| Treatment | 50.4% | 8.3% | 22.2% | 8.8%  | 10.4%|
| Change    | 6.2%  | 1.3% | -5.3% | -4.5% | 2.4% |

| Corruption|       |      |       |       |      |
| Placebo   | 42.0% | 8.0% | 28.0% | 13.6% | 8.3% |
| Treatment | 51.8% | 13.1%| 17.1% | 10.5% | 7.5% |
| Change    | 9.8%  | 5.2% | -10.9%| -3.2% | -0.8%|

| Terrorism |       |      |       |       |      |
| Placebo   | 44.8% | 5.7% | 29.9% | 14.7% | 4.9% |
| Treatment | 57.8% | 5.9% | 16.1% | 12.2% | 8.0% |
| Change    | 13.1% | 0.1% | -13.8%| -2.5% | 3.1% |

* p < 0.05

Receiving the Premium treatment decreases Compliance (by 4.5 percent, p = 0.054) with international standards compared to the Placebo, and we might expect this significant decrease given that a monetary offer is being made to maintain anonymity. Likewise, the Corruption condition decreases the predicted probability for Part-Compliance from 28 percent in the Placebo condition to 17 percent – a statistically significant 11 percent decrease (p = .005). Corruption also decreases the predicted probability for full Compliance by 3.2 percent, from 13.6 to 10.5 percent (p = .067). Likewise, the Terrorism condition causes a drop by nearly half in the predicted probability of Part-Compliance compared to the Placebo condition (from 29.9 to 16.1, p = .002). The Terrorism condition also leads to a 2.5 percent decrease in Compliance, from 14.7 to 12.2 percent (p = .089). Robustness analysis using a selection model results generally corroborate the multinomial results reported above, with the exception that the Terrorism condition is
associated with significantly lower Response rates but not with a significant reduction in Compliance. See Online Appendix Table C1.

Do these results hold when we focus specifically on a large subject pool of firms in the globally dominant country? We sought to answer that question with Experiment 2.

Experiment 2 Data and Results

Treatments in United States

Experiment 2 employed the Placebo condition and the FATF, Corruption, and Terrorism treatments on a subject pool of 1,405 companies in the United States, including 233 incorporation services and 1,172 business law firms. The isolation of a single country allowed a comparison of firms’ behavior in response to international law versus domestic regulation, and thus we replaced the Premium treatment with a treatment based on national enforcement.

6. IRS/Domestic Enforcement

The email notes that United States law requires identity disclosure and that the Internal Revenue Service enforces this requirement.\(^6\)

Block Randomization

In the U.S. subject pool, blocking strata were formed according to state-by-state business friendliness and again by the type of CSP: incorporation service or law firm. The state groupings were created first by taking the states with the greatest number of subject firms – California and Nevada – and next the states with the reputations for greatest ease of incorporating anonymously – Delaware and Wyoming – and making each of the four states into individual strata. The rest of

\(^6\) We thank Jessica Preece for suggesting this treatment.
the states were blocked according to the Beacon Hill Institute’s State Competitiveness Report (2010). Specifically, we used the measure for “Business Incubation,” which captures a number of factors relating to the ease of setting up new businesses in a given state. We then further subdivided the strata by company type. This created 14 strata in which the random assignment to the five experimental conditions took place. Email inquiries were sent according to the same protocol and coded by the same criteria described above for Experiment 1. As in Experiment 1, randomization checks for Experiment 2 suggest that the covariates were balanced across conditions.

**Results**

Figure 3 and Table 5 display the results for Experiment 2. The patterns display some differences from Experiment 1. A much higher proportion of U.S. firms failed to reply to our inquiries compared to international companies: an average of nearly 77 percent fell into the No Response category across conditions. In particular, the Terrorism condition saw nearly 84 percent of treated firms fail to reply – a difference of 11 percent from the Placebo’s No Response rate of 73 percent ($p = .002$). The IRS treatment showed a nearly 6 percent increase in No Response compared to the Placebo, but the difference falls just short of statistical significance at the .1 level ($p = .109$) in the tests. None of the other treatments were different from the Placebo for Response rates to a statistically significant degree. That the Corruption treatment received nearly the same proportion of replies as the Placebo stands in marked contrast to the results in Experiment 1, where international firms were significantly less likely to respond to the Guinean aliases.
Figure 3: Outcome Proportions across Experiment 2 Conditions
Table 5: Contingency Table of Outcomes across Experiment 2 Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>No Response</th>
<th>Non-Compliant</th>
<th>Part-Compliant</th>
<th>Compliant</th>
<th>Refusal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>270</td>
<td>197</td>
<td>29</td>
<td>7</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>73.0%</td>
<td>10.7%</td>
<td>2.6%</td>
<td>0.0%</td>
<td>13.7%</td>
</tr>
<tr>
<td>FATF</td>
<td>293</td>
<td>212</td>
<td>33</td>
<td>6</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>72.4%</td>
<td>11.3%</td>
<td>2.0%</td>
<td>0.7%</td>
<td>13.7%</td>
</tr>
<tr>
<td>IRS</td>
<td>283</td>
<td>223</td>
<td>24</td>
<td>8</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>78.8%</td>
<td>8.5%</td>
<td>2.8%</td>
<td>0.4%</td>
<td>9.5%</td>
</tr>
<tr>
<td>Corruption</td>
<td>271</td>
<td>204</td>
<td>35</td>
<td>6</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>75.3%</td>
<td>12.9%</td>
<td>2.2%</td>
<td>0.4%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Terrorism</td>
<td>288</td>
<td>241*</td>
<td>21</td>
<td>4</td>
<td>0</td>
<td>22*</td>
</tr>
<tr>
<td>Proportion</td>
<td></td>
<td>83.7%</td>
<td>7.3%</td>
<td>1.4%</td>
<td>0.0%</td>
<td>7.6%</td>
</tr>
<tr>
<td>Total</td>
<td>1,405</td>
<td>1,077</td>
<td>142</td>
<td>31</td>
<td>4</td>
<td>151</td>
</tr>
</tbody>
</table>

Significant in difference of proportions and difference of means tests compared to placebo condition:
* .05 level.

Perhaps the greatest substantive differences between Experiments 1 and 2 involve the rates of Part-Compliance and Compliance. While the proportions for Non-Compliance and Refusal were similar across the two experiments, in Experiment 2 only 35 of the 1,405 firms in the United States asked for any document establishing identity. A meager 4 firms required that the identification documents be notarized. This means that the rate of Part-Compliance was merely 2.2 percent, and the rate of full Compliance was an astonishing 0.3 percent.

Likely due in part to the relatively small cell sizes, few of the differences between the treatments and Placebo were statistically significant. However, while rates of Part- and full
Compliance for Terrorism were similar to the other conditions, the 7.6 percent rate of Refusal for Terrorism was 6 percent below the Placebo rate ($p = .020$ in difference-in-means tests). This result is particularly alarming, especially given the fact that Refusal was virtually the only active response U.S. firms employed that was consistent with international standards (with passive refusal through No Response being the alternative). The Refusal rate for the Corruption condition was 4.5 percent lower than the Placebo, but the difference fell just shy of statistical significance at the .1 level ($p = .102$).

**Multinomial Models**

Again, however, it is likely that firms’ choice of a given response strategy is conditional on the other options, so we once more analyze the results using a multinomial logit model. However, due to the very small cell sizes for Compliance, the models employing all five outcomes would not converge. Two alternatives presented themselves: either we could collapse the few Compliant subjects with the Refusals, or we could drop the four subjects from the analysis. The results are qualitatively similar with either approach; for presentation we opted to preserve the subjects in question through collapsing the categories, and we display the multinomial results in Online Appendix Table D1.

The results once more suggest that information about international standards, invoked with the FATF treatment, does not cause greater compliance. But would the same pattern hold for domestic law? The intuition behind Experiment 2 suggested that domestic law enforced by the well-known and widely feared IRS would induce greater compliance to stipulated identification standards. While the rate of No Response did increase in the IRS condition (though not statistically significantly), the rates of Non-Compliance, Part-Compliance, and Compliance/Refusal were not statistically different from the Placebo condition. We find this result surprising, and it
suggests that not only does information about international law fail to move incorporation services, explicit stipulation of U.S. federal law – at least as invoked by the IRS treatment – proves no more compelling to subject firms.

However, in these multinomial results we do encounter some evidence that one of the treatments improves compliance with international law. The Terrorism treatment decreases the Non-Compliance rate significantly (\(p = .043\)). This result provides some encouragement that U.S. firms may be exercising a degree of vigilance over their particularly risky prospective clients. However, the concomitant negative result for the Terrorism treatment on Refusal/Compliance (\(p = .019\)) along with the larger coefficient for that estimate may offset this positive result. The Corruption condition also led to lower Refusal/Compliance rates in the multinomial logit estimates. The Corruption treatment effect for the Compliant/Refusal category now appears insignificant, but the \(p\)-value is 0.101, just attenuating slightly.

Adding covariates to the analysis produced the predicted probabilities and rate changes from treatment to Placebo displayed in Table 6.\(^7\) The Corruption condition causes a 6 percent decrease in the rate of Refusal/Compliance compared to the Placebo (\(p = .073\)). But the biggest changes are seen in the Terrorism condition, where No Response increases by nearly 11 percent (\(p = .004\)), Non-Compliance decreases by 3.4 percent (\(p = .004\)), and Refusal/Compliance decreases by 7.2 percent (\(p = .013\)). The decrease in Non-Compliance rates for the Terrorism treatment, as above, provides some evidence for the efficacy of the FATF’s risk-based approach, but the decrease in Refusal/Compliance offsets this effect by more than double. Results for robustness checks using a selection model are displayed in Online Appendix D and again generally

\(^7\) See Table D3 in Online Appendix D for full results.
corroborate the findings above. The sole exception is for the Corruption treatment, where the co-efficient for the Compliance outcome is not statistically significant.

**Table 7: Predicted Probabilities of Outcomes for Experiment 2**

<table>
<thead>
<tr>
<th>Treatments</th>
<th>No Response</th>
<th>Non-Compliant</th>
<th>Part-Compliant</th>
<th>Compliant + Refusal</th>
</tr>
</thead>
<tbody>
<tr>
<td>FATF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>81.8%</td>
<td>3.0%</td>
<td>0.0%</td>
<td>15.2%</td>
</tr>
<tr>
<td>Treatment</td>
<td>80.5%</td>
<td>2.7%</td>
<td>0.0%</td>
<td>16.8%</td>
</tr>
<tr>
<td>Change</td>
<td>-1.3%</td>
<td>-0.3%</td>
<td>0.0%</td>
<td>1.6%</td>
</tr>
<tr>
<td>IRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>80.5%</td>
<td>4.9%</td>
<td>0.6%</td>
<td>14.1%</td>
</tr>
<tr>
<td>Treatment</td>
<td>85.8%</td>
<td>3.4%</td>
<td>0.6%</td>
<td>10.2%</td>
</tr>
<tr>
<td>Change</td>
<td>5.3%</td>
<td>-1.5%</td>
<td>0.0%</td>
<td>-3.9%</td>
</tr>
<tr>
<td>Corruption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>80.6%</td>
<td>4.8%</td>
<td>0.5%</td>
<td>14.1%</td>
</tr>
<tr>
<td>Treatment</td>
<td>85.6%</td>
<td>5.6%</td>
<td>0.6%</td>
<td>8.2%</td>
</tr>
<tr>
<td>Change</td>
<td>5.1%</td>
<td>0.8%</td>
<td>0.1%</td>
<td>-6.0%</td>
</tr>
<tr>
<td>Terrorism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Placebo</td>
<td>80.1%</td>
<td>5.9%</td>
<td>0.0%</td>
<td>14.0%</td>
</tr>
<tr>
<td>Treatment</td>
<td>90.8%</td>
<td>2.4%</td>
<td>0.0%</td>
<td>6.8%</td>
</tr>
<tr>
<td>Change</td>
<td>10.7%*</td>
<td>-3.4%*</td>
<td>0.0%</td>
<td>-7.2%*</td>
</tr>
</tbody>
</table>

* p < 0.05

**Conclusion**

At the heart of current debates about global governance and the nature of the international system is the question of whether international law causes better behavior. This paper has addressed the previous stalemate that has hobbled research on the topic: scholars have been well aware of the inherent problems of using observational data in terms of endogeneity and selection bias, but have until now been unable to employ experimental methods. This issue is consequen-
tial in policy terms because untraceable shell companies are the most common mechanism for several types of major financial crimes.

The dominant policy consensus on corporate transparency and international financial regulation more generally is that OECD states comply, while developing countries are often unable to comply, and tax havens are often unwilling to comply. Our results show this view to be wrong on each count. Corporate service providers in tax havens are far more compliant than those in OECD states. Overall, the significant differences between levels of compliance in rich and poor countries generally favor developing nations. In combination these results suggest that the relatively low levels of compliance in OECD states are a product of a deliberate unwillingness to comply, rather than a lack of capacity.

Analysis of the experimental data also gives grounds for concern about compliance. Contrary to views that non-compliance is a product of lack of knowledge or legal precision, prompting incorporation services about their responsibility to collect identity documents as per FATF and IRS standards made them no more likely to do so. Further, a significant number of services were willing to deliberately violate international standards when offered extra money. Moreover, the Terrorism and Corruption treatments provoked divergent responses compared with the Placebo. While one sub-group seemed to respond to the extra risk by refusing any contact, another group of incorporation services were in effect prepared to sell anything to anyone, no matter how dangerous and disreputable. Given this insensitivity to risk, the fact that the current regime relies on private, for-profit firms enforcing standards according to a risk-based approach does not inspire confidence.

More generally, our findings substantiate the reservations of those writing about compliance with international law who worry that endogeneity and selection effects have led to overly
sanguine conclusions about the effectiveness of international law. The results of our experiments suggest material self-interest remains an all-too-powerful temptation to violate international standards.
References


FATF. 2010. FATF Recommendation 5: Customer due diligence and record-keeping. Accessed 17 October 2010 at http://www.fatf-gafi.org/document/58/0,3343,en_32250379_32236920_43642938_1_1_1_1,00.html


