

Et Tu, Brute?

Reputations, Foreign Direct Investment, and Contract Breach

Ekrem T. Başer*

June 27, 2020

Abstract

A central problem facing foreign investors is whether host-governments will violate their contracts once investments are made. According to conventional wisdom, states can overcome this commitment problem by cultivating reputations in the eyes of potential investors. Yet, the field has a static perspective on reputations: we do not know how states' current reputations shape their incentives to breach contract. I develop a dynamic model of FDI and breach behavior to show that better reputations for treating foreign investors generate stronger incentives to breach contract. To test this prediction, I leverage the as-if random occurrence of natural disasters to show that states with good reputations for treating foreign investors are more likely to breach contract following disasters. For the broader cooperation literature in IR, this suggests that the reputation mechanism is a double-edged sword. Reputational incentives boost cooperation by enabling states to convince their audience to trust them, but it also generates opportunities to take advantage of this trust.

*Political Science Department, University of Illinois Urbana-Champaign, baser2@illinois.edu. I would like to thank Çağlayan Başer, Rob Carroll, Stephen Chaudoin, Nuole Chen, Xinyuan Dai, Brian Gaines, Chris Grady, Alice Iannantuoni, Jim Kuklinski, Charla Waeiss, Matt Winters and seminar participants at University of Illinois for helpful comments and discussions. I also would like to thank Institute for Humane Studies for providing funding for this research through the Humane Studies Fellowship. All errors remain my own.

1 Introduction

The central problem hindering international cooperation is the difficulty states face in persuading their partners that they are committed to carry out their promises. From the perspective of foreign investors, such as transnational firms, this problem manifests itself as the risk of host governments expropriating their assets or otherwise breaching contract. As the international arena is anarchical, and thus lacks strong contract-enforcing institutions, both scholars of International Relations (IR) and foreign direct investment (FDI) have long highlighted reputational concerns as the primary vehicle through which states resolve their credible commitment problems and achieve cooperation (Axelrod and Keohane 1985; Cole and English 1991; Downs and Jones 2002; Thomas and Worrall 1994). To the extent that host governments value future investment flows, they can forgo short-term incentives to breach contract in favor of fostering good relations with investors.

The field of IR has a static view of the reputation mechanism, because our theories of reputations assume that states' preferences and abilities are fixed. States either find it worthwhile to cultivate their reputations and never falter, or decide that today's opportunity costs of reputation building supersede the future dividends and never exert the effort. Yet, casual observation suggests that states frequently tarnish their hard earned reputations. They lie, cheat, default on their debts, violate foreign investors' property rights, and otherwise renege on their promises. Moreover, states are often able to rebuild their tarnished reputations. They are able to improve their credit after defaulting, or are able to attract investment again after expropriating foreign assets. Therefore, the prevailing static perspective severely limits our understanding of reputation dynamics, and thus international cooperation. Should a state with an already stellar reputation be expected to further cultivate its reputation or not? Why, if ever, would states tarnish their hard earned reputations, and once they do, how are they able to rebuild them? And more generally, how do incentives to cultivate future reputations change as a function of current reputations?

In order to answer these questions, I concentrate on the laboratory of foreign direct investment and contract breach by host governments. I analyze a dynamic model of reputation building and spending, adapted to the foreign direct investment context, where foreign firms make investment decisions by observing the host state's prior treatment of foreign investors. To introduce a dynamic perspective on reputations, I assume that states' preferences and abilities regarding violating their

contracts with foreign investors are not fixed but change over time and across interactions. This captures the idea that a given state's returns from breaching contract respond to a multitude of factors such as leader turnover, changes in the domestic legal system, shifts in public opinion, and state of the economy. Since states' preferences are subject to change, this restricts how much foreign firms can infer states' future behavior based on their past observations, which in turn leads to both persistent doubt and benefit of doubt. Persistent doubt limits the usefulness of reputations, and benefit of doubt opens up the possibility of rebuilding tarnished reputations, all for the same reason: things might be different *this* time.

The main argument I make via the formal model is that, if a state is motivated by reputational incentives, then the better its reputation, the more it will shirk — in the FDI case, shirking would coincide with contract breach. The reason is twofold. One, states know that tarnishing their reputation today does not prevent future reputation building. The fact that states' preferences might change, together with their partners' desire to achieve cooperative outcomes (e.g. foreign firms' would like to invest and profit from their investment) mean that benefit of doubt will be extended in the future. Two, the knowledge that states' preferences might change prevents the audience's suspicions to be cleared completely. In other words, persistent doubt puts an upper limit on reputations, and in turn, puts an upper limit on the amount of benefit improving reputations will bring. Together, these two points imply that as a state's reputation gets better, there is both less room for improvement and reduced returns on further cultivating the reputation. Then, states with better reputations have lower incentives to exert further effort compared to states with poorer reputations. In the FDI context this means that if a state is motivated by reputational dividends, the better its reputations for investor-friendliness, the more likely it is to breach contract with foreign firms.

In order to test this argument, I leverage as-if random occurrence of natural disasters. Natural disasters function as exogenous adverse economic shocks, increasing states' temptation to expropriate foreign firms' assets or otherwise breach contract. I examine heterogeneous treatment effects of disasters based on states' prior reputation level, on a subset of states matched on the level of prior disaster risk calculated using gradient boosting machines on historical disaster data. The results provide strong support for the argument. states with better reputations for treating foreign investors are indeed much more likely to breach contract in the aftermath of natural disas-

ters compared to states with worse reputations. The result is robust across all specifications, and is both statistically and substantively significant. According to the results of matched designs, states with the best prior reputations have a 9 – 28% higher marginal probability of breaching contract as a result of facing a natural disaster, compared to states with poorest reputations.

This work makes several contributions to both the study of reputations and politics of FDI in international relations. It provides a dynamic reputation theory which, unlike previous work, can accommodate endogenous reputation building and spending behavior by states in equilibrium. This dynamic perspective enables characterizing how incentives to build reputations fluctuate as a function of current reputations. In particular, it shows, theoretically and empirically, that incentives to build reputations are stronger for states with poor reputations and weaker for those with better reputations. On the one hand, this result suggests that while the reputation mechanism does help in encouraging international cooperation, IR scholars might have been too confident in its effect on producing desirable behavior by individual states (Büthe and Milner 2008; Crescenzi 2018; Simmons 2000; Tomz 2007). Reputational incentives do constrain states, but this constraining power rapidly diminishes as reputations improve. On the other hand, the results also show that some of the critiques raised by reputation skeptics in IR — such as Press (2005)' argument that states with tarnished reputations should work extra hard to regain their lost advantage, and Mercer (2013) and Jervis (1997)' contention that states with good reputations should be able to afford behavior to the contrary — are in fact consistent with a rationalist reputation logic. Further, by heeding the call of Dafoe, Renshon, and Huth (2014) and providing evidence of an empirical implication of the reputation mechanism outside of whether reputations matter, this article strengthens the case that studying reputations remain important to understand patterns of behavior in world politics.

For the FDI literature, where scholars are comfortable lending credence to the reputation mechanism, the most important contribution is to highlight the limits of reputations to constrain state behavior. Importantly, it is not true that if a state is forward looking enough, or cares about future investment flows enough, it will always build and protect its reputation in the eyes of foreign investors a la Cole and English (1991) and Thomas and Worrall (1994). Even when a given state's preferences over contract breach remain fixed, potential investors believing that they might have changed is enough to generate reputation spending and building cycles in equilibrium. While it

is optimal for transnational firms to trust states with better reputations for investor-friendliness on average, this trust generates incentives for states to take advantage of it. In that sense, these results highlight the value of exploring channels alternative to reputational incentives that could constrain states from breaching contract, such as the foreign investors' links to the broader domestic economy (Henisz 2000; Johns and Wellhausen 2016) and political ties they establish with the governing elite or other actors that could provide "protection" from predatory state actions (Pyle 2011; Volkov 2002).

2 FDI, Contract Breach, and Reputation

Since the early 1990s, FDI flows have dramatically increased both in volume and importance. In 2015, FDI flows represented roughly 3% of GDP for OECD members and 4% of GDP for the least developed countries, up from 0.9% and 0.6% respectively.¹ Although this trend is showing signs of slowing down in the past few years, FDI has become a crucial component of global economic integration and development, providing a multitude of benefits to host countries such as more jobs, increased tax revenue, technological transfers, and productivity spillovers (Pandya 2016). Accordingly, governments have become increasingly eager to attract foreign investment, often providing fiscal and financial incentives for multinational companies to invest in their country (Li 2006; Li and Resnick 2003).

The primary obstacle host governments face in attracting FDI is persuading foreign investors that their property rights will be protected after the investment is made — a credible commitment problem. Governments have strong incentives to promise favorable treatment to foreign investors via formal or informal contracts, however, once investments are made, there is no higher authority to stop governments from breaking their promises should they wish to. Indeed, between 1990-2015, governments have been sued by foreign investors in international arbitration courts for violating their property rights more than 700 times (Wellhausen 2016). These violations, or breaches of contract, can include nationalization or otherwise forceful ownership change of foreign investors' assets, as well as "creeping expropriation." Creeping expropriation, which is now the most commonly observed type of contract breach, includes a broad set of behavior by host

1. According to World Bank's 2015 World Development Indicators Database

governments that can collectively be characterized as selective use of laws and regulations at the expense of foreign firms (Graham, Johnston, and Kingsley 2018; Jensen 2003; Jensen et al. 2019).

Given the substantial risks associated with contract breach, how do governments overcome their credible commitment problem to attract FDI? And despite the substantial benefits of ensuring continued FDI flows, what explains their contract breach behavior? A voluminous political economy literature has looked at the mechanisms that constrain and enable governments in violating contracts with foreign investors. These mechanisms can be categorized in two groups based on whether they emphasize (i) the short-run, or direct, consequences of breach behavior by governments or (ii) the long-run, or indirect, consequences of today's breach behavior on future interactions.

The mechanisms emphasizing direct consequences of breach behavior highlight factors that determine the ability of host-governments to break contracts and the net returns they will derive from a given act of violation. These include the host-government's domestic institutions, foreign firms' economic or political ties with other domestic actors, and issue linkage. Presence of strong domestic institutions protecting property rights, such as an independent judiciary or the existence of checks and balances on the executive branch, can constrain governments' ability to violate foreign investors' property rights (Jensen 2008; Levy and Spiller 1996; Li 2009; Li and Resnick 2003). Joint ventures with domestic firms (Henisz 2000), the depth of integration into domestic supply chains (Johns and Wellhausen 2016), as well as ties with the governing elite (Pyle 2011) or extralegal actors (Frye 2002; Volkov 2002) also make violating foreign investors' property rights costlier for governments. Similarly, linking undesirable government behavior toward foreign investors with other issue areas, such as lending from IMF or World Bank, is shown to reduce the temptation to break contracts (Haas 1980; Jensen et al. 2019).

The canonical explanation for how states could credibly commit to uphold their promises to foreign investors highlighting the direct, or long-run, returns to breach is reputation. The focus is on the market's ability to constrain governments via linking today's interactions with foreign investors to those of tomorrow (Cole and English 1991; Thomas and Worrall 1994). The core idea is that foreign firms considering investing in a given state are uncertain about what the costs and benefits of breaching contract are for that state. In other words, they have difficulty discerning how tempting, if at all, violating foreign investors' property rights might be for that state, which

breeds mistrust. Yet, the state knows that future investors can observe its past treatment of foreign investors, and derive inferences about the state’s future behavior from those observations. These inferences, which I call a state’s reputation, establish a link between interactions of today and tomorrow. The possibility of establishing a reputation in the eyes of potential investors create incentives for the state to forgo today’s temptation of violating contracts in order to bolster future FDI inflows. Then, if potential investors believe that future FDI flows are sufficiently important for the state, the state’s reputation immediately springs to life, enabling investments. If potential investors do not believe that future FDI flows are important enough for the state, then the credible commitment problem remains, disabling investments. Previous research has shown that a variety of factors including democratic institutions (Jensen 2003), being party to treaties regulating investment and trade (Allee and Peinhardt 2011; Büthe and Milner 2008), natural resources (Jensen and Johnston 2011), economic crises (Jensen et al. 2019), and the activities of co-national firms in the economy (Wellhausen 2015) can influence both the magnitude of these long-run reputation costs associated with market access and the extent to which they will be internalized by governments.

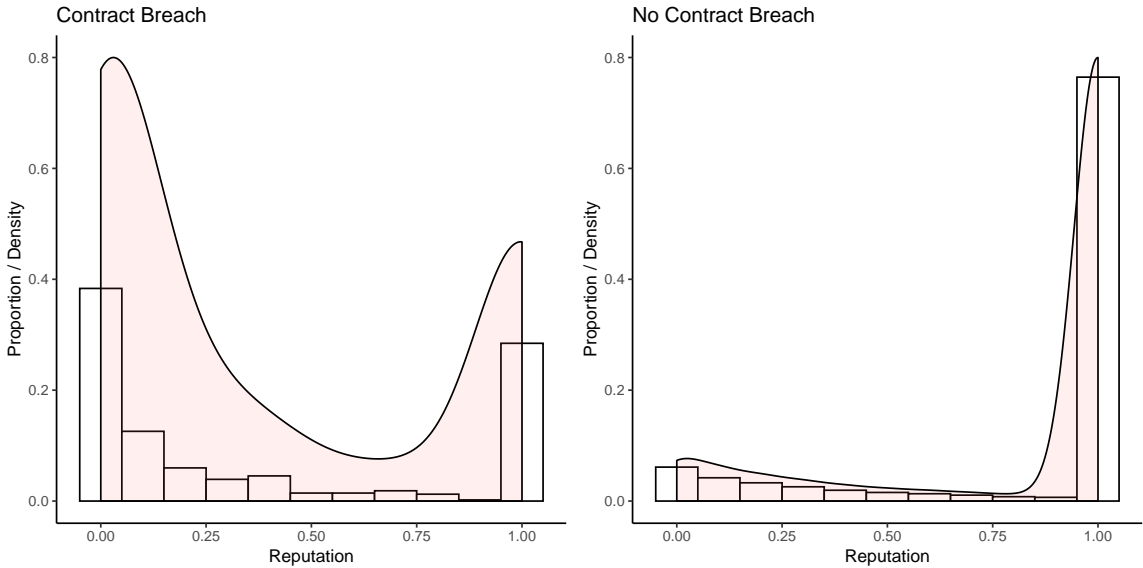


Figure 1: Reputation distributions by contract breach behavior (or lack thereof). Reputations range from 0 (bad) to 1 (good). The measure is set to 0 following a contract breach incident and increases until it reaches 1 — provided that the state does not engage in further contract breach. The kernel densities are scaled such that the highest point is 0.8 for ease of comparison across panels.

Notice that conceptualizing the reputation mechanism this way leads to a static picture: the

state should either (i) decide to adhere to promises made to foreign investors and never give in to the temptation of breaching contract, or (ii) decide that reputation building is too costly and never try.² There are two issues I want to highlight. First, states often take reputation-damaging actions after building their reputations. The left panel of Figure 1 shows that, among all instances of contract breach cases brought to international arbitration since 1990, the reputation distribution of the perpetrator states exhibit a bimodal pattern. A significant portion of the perpetrators had good reputations for investor-friendliness based on prior contract breach behavior. Second, once states engage in such reputation-tarnishing behavior, we often observe that they are able to rebuild their reputations. They are able to find international lenders to borrow from or able to attract investment. The static understanding of reputations would have little to say, without appealing to exogenous shocks, about why a state would tarnish a reputation after extending the effort to build it, or how a state would rebuild a damaged reputation. More generally, the inability to explain endogenous tarnishing and rebuilding behavior means that we also do not have explanations about how states' incentives to build and maintain reputations change as their reputations change.

It is important to note that both the FDI and the broader IR literature share the static perspective on reputations, with a few notable exceptions (Crescenzi 2018; Sartori 2005; Tomz 2007). There is currently a debate in IR, where proponents and critics argue over whether reputations matter (for a summary of the debate see Crescenzi 2018; Tomz 2007, 14-36; Weisiger and Yarhi-Milo 2015). A number of the issues raised by critics to highlight the shortcomings of the reputation mechanism are about how current reputations are related to reputation building incentives, showing that the static perspective on reputations is at least partly to blame for fueling the controversy. Press (2005), for instance, argues that when states engage in reputation-damaging actions, such as backing down in a crisis, they sometimes work extra hard to counteract the reputation theory. In similar fashion, Mercer (2013) and Jervis (1997) argue that if a state can muster a good reputation, then it should be able to take advantage of the trust it generates, and afford some behavior to the contrary. These points, labeled by their authors as paradoxes, cannot be consistent with the static

2. The model presented by Cole and English (1991), and others dealing with expropriation within the class of defaultable debt models (e.g. Tomz and Wright 2010) do generate expropriation behavior in equilibrium. These models still have a static perspective of reputations for two reasons. First, expropriation happens depending on the realization of an exogenous stochastic variable (e.g. production level), not reputation dynamics themselves. These models can explain expropriation by showing that certain states of the world make reputation building more or less conducive for states. Second, expropriation events cause future investors to shun the state forever, meaning that these models include no reputation rebuilding.

perspective on reputations, since the static perspective is silent about how incentives to extend further effort for reputations change as a function of prior reputations. If states with good reputations cannot be relied upon to exhibit desirable behavior and if states with poor reputations cannot be expected to exhibit undesirable behavior, what does this mean for the reputation mechanism?

In the next section, I lay out my theoretical argument and introduce a dynamic model of reputation building. In particular, I show that reputations critics are both right and wrong in making the points highlighted above. They were right that those states with tarnished reputations should work harder, and those with good reputations do have stronger incentives to take advantage of the trust their reputations generate. They were wrong in claiming that these represent paradoxes. Indeed these paradoxes turn out to be consistent with the dynamic rational reputation logic presented here. Reputations help states achieve higher levels of cooperation, but they also generate opportunities for states to take advantage of the trust generated by good reputations.

3 Theoretical Argument

Given a commitment problem such as the predicament faced by states and foreign investors, reputational incentives arise because potential investors have incomplete information about states' returns from renegeing on its promises, or breaching contract. That is, in order for a state to find reputation building worthwhile, it needs to be able to manipulate potential investors' beliefs about its temptation to breach contract. In so doing, the state changes potential investors' predictions about what the state will do next. Incentives for reputation building arise, because by manipulating those beliefs in its favor, the state elicits greater trust from future investors, and thus more investment. Then, reputational incentives require the audience to be uncertain about whether the state stands to gain from breaching contract.³

If uncertainty is necessary for reputational incentives to arise, it follows that certainty destroys those incentives. According to the static perspective on reputations, expropriation behavior resolves foreign investors' uncertainty about the temptation faced by the state to breach contract.

3. Scholars have also used the word reputation to describe the logic of reciprocity that characterizes strategy profiles in repeated complete information games based on community punishment, as opposed to the incomplete information logic outlined above (e.g. Axelrod 1981). I follow much of the recent literature on reputations in IR and take reputations to be about an audience's beliefs about an agent's stage-game preferences, or "type," and the ensuing credibility of their commitments (Sartori 2005; Tomz 2007). See Fudenberg and Maskin (1986, 534-535) and Mailath and Samuelson (2015, 167-168) for comments about the relationship between the two.

Having observed contract breach, investors do not doubt whether the state is tempted to breach contract anymore — they are certain of it, and they never invest in that country again. This is also true of reputation theories in the broader IR literature. Herein lies the problem with the static perspective. On the one hand, if cooperation of the state is assured by the ultimate threat of market exclusion, states would never tarnish their reputations outside of exogenous shocks forcing their hands, and in any case there would be no rebuilding. On the other hand, if uncertainty about a state’s payoffs from breaching contract is the engine behind reputations, how can long-run reputational incentives be sustained notwithstanding the flow of informative signals, such as breach behavior?

The answer is that persistent reputational concerns require persistent uncertainty. If short-term gains from breaking contracts with foreign investors were fixed, given enough informative signals, a rational audience comprised of foreign multinationals would eventually learn them, preventing reputational incentives to emerge ever again.⁴ Yet, we already know that gains from breaking contracts are not fixed. Existing FDI literature provides ample evidence which suggests gains from breaking contracts vary over time, since important variables determining these gains themselves vary over time: e.g. domestic and global economic variables (Tomz and Wright 2010), public opinion (Graham, Johnston, and Kingsley 2018), governing factions and polarization (Pinto and Pinto 2008), formal and informal institutions (Li 2006), and globalization (Jensen et al. 2019). The temptation faced by states to violate contracts with foreign investors also varies across situations depending on variables pertaining to the foreign investor. Again, existing work provides a number of such variables, such as the depth of integration with a supply chain (Johns and Wellhausen 2016), and the presence of co-national firms active in the economy (Wellhausen 2015, 2016).

These findings suggest two things regarding the uncertainty surrounding a state’s direct returns from a given contract breach behavior. First, these costs are not fixed, they fluctuate over time. Second, observing a state’s breach behavior may be more or less informative for a potential investor depending on the similarities it shares with the violated party. This is not to argue that there is nothing investors can infer from past breach behavior about what will happen to them should they invest. On the contrary; reputation dynamics would not emerge if that was true.

4. See Ekmekci, Gossner, and Wilson (2012) and Mailath and Samuelson (2015), and in particular Cripps, Mailath, and Samuelson (2004), for detailed arguments and formal proofs about why long-run reputational incentives require a mechanism through which uncertainty is replenished.

The point is that the uncertainty underlying reputation dynamics should not adhere to a fixed parameter but a changing one. Investors are uncertain about the degree to which a state would be tempted to renege on its promises should they invest, but they also know that, compared to today, the state might be more or less tempted tomorrow. This means, whatever information potential investors acquire today by observing the state's behavior, this information will gradually lose its value for interactions that happen further into the future. This is the primary ingredient of the theory I present here, and is the main engine behind the theoretical results.

In addition to introducing additional realism to the reputation mechanism, focusing on shifting uncertainty directly addresses a core complaint raised by reputation critics. That reputation theories in IR assume different interactions across time and space are equally informative for the audience (Hopf 1994; Mercer 1996).⁵ Further, thinking of states as having changing as opposed to fixed short-run returns from contract breach also overlaps with the structural Realist assertion that states' intentions can never be fully known (Mearsheimer 2001; Waltz 1979). If anarchical international system is associated with persistent doubt about others' intentions, uncertainty should adhere to changing parameters.

What happens to reputation dynamics if what states stand to gain and lose from breach is subject to change? Shifting uncertainty makes investors think that the incentives facing a given state today might differ from those of yesterday, which impact reputation dynamics via two related channels. First, it leads to benefit of doubt. Suppose potential investors observe a given state to have breached contract with a foreign firm yesterday. The knowledge that the state's incentives might have changed today, even if it is unlikely, encourage investors to extend the benefit of doubt, if slowly. Adding the potential gains foreign firms would accrue through their investment, the fact that things might be different *this* time renders benefit of doubt irresistible and the market exclusion threat incredible. Once states recognize that market exclusion is not a credible punishment for contract breach, that benefit of doubt will be extended—if slowly—even when they misbehave, this reduces the reputational costs associated with contract breach behavior.

Second, because shifting uncertainty decreases the usefulness of yesterday's information for today, it causes reputations to decay. The idea that reputations are supposed to decay over time

5. On this point, see Crescenzi (2018) for an excellent non-strategic reputational learning model which takes seriously the variation in information content across different interactions.

parallels a number of prior work on reputations. In the model presented by (Sartori 2005), for instance, reputations for honesty are set to last for only two interactions, after which they dissipate. (Weisiger and Yarhi-Milo 2015), who provide empirical evidence that reputations for resolve deter future challenges, measure reputation by incorporating a decay function. Similarly Crescenzi (2018) imposes a decay function in his reputational learning model. In the current context, decaying reputations suggests that even if potential investors are certain that a given state faces no temptation to violate foreign firms' property rights today, there will still be a sliver of doubt tomorrow that perhaps those incentives may have changed to the detriment of investors. Here, the fact that things might be different *this* time puts an upper bound on how good reputations can become, and thus how much benefit could be derived from building reputations, determined by how quickly and how much those incentives can be expected to change. From the perspective of a given state, as the state's reputation in the eyes of potential investors improve, the returns from exerting further effort for its reputation decrease.

Together, these suggest that if a state is indeed facing a commitment problem pitting its short-term temptation of contract breach with long-run dividends of good behavior in the form of greater FDI flows, then that state should have weaker incentives to build its reputation for the future, the better its reputation gets. In other words, all else equal, the better that state's reputation gets, the higher the likelihood that this state will break its contract with a foreign firm, and the worse that state's reputation is, the lower the likelihood that this state will engage in contract breach.

The first thing to highlight here is that, this argument does not imply that those agents who have exhibited desirable behavior before are less likely to do so again on average. Note that the reputation mechanism adheres only to those states which have incentives to break contract in the short-run. If a state does not have positive expected returns from a given contract breach behavior in the short-run, that state will not breach contract, not due to reputational concerns but because the state has nothing to gain from it. Not all states have short-run incentives to violate foreign investors' property, and those who do have them, do not have those incentives all the time. In fact, this is precisely what is motivating the premise of shifting uncertainty. The studies cited in the previous section highlighting the short-run costs and benefits of contract breach provide ample mechanisms which might nullify such short-run temptations outside of the constraining power of

reputational incentives. These include domestic formal and informal institutions, foreign firms' economic or political ties with other domestic actors, and issue linkage (Jensen et al. 2019; Johns and Wellhausen 2016; Levy and Spiller 1996; Li 2009; Pyle 2011). If there is no temptation to breach contract, there will not be a breach of contract. No temptation also means there is no trade-off between short and long run incentives to give raise to reputational incentives. Potential investors might still have their doubts, but the decision problem for a state which does not face a commitment problem today is trivial from the perspective of reputations.

The second point to highlight is that the theoretical argument presented here *does not* imply that reputations are a collective bad. It is quite the opposite. The existence of reputational incentives is substantially improving the welfare of both the foreign investors and the states which want to attract them by making mutually beneficial outcomes attainable. However, since the constraining power of reputational incentives is greater for states with poor prior reputations and weaker for states with good prior reputations, this improvement in welfare is produced by the agents who are at the left tail of the reputation distribution. To put it differently, reputation mechanism is a double-edged sword. While it leads to greater FDI flows by incentivizing states with poor reputations to convince investors to trust them with their deeds, it also generates opportunities for states with good reputations to take advantage of this trust.

In the next section, I make these arguments with the help of a formal model.

3.1 Model

The goal is to examine how a state's incentives to break contracts change as a function of its current reputation for treating foreign firms in the eyes of potential investors. I focus on the interactions a given state (a host-government) G has with a new investor I in each period, for infinitely many periods $t = 0, 1, \dots, \infty$. The timeline of the stage game is as follows:

1. The investor, I , decides whether to invest in that country, $k \in \{0, 1\}$, where 1 indicates the decision to invest.
 - If I decides not to invest ($k = 0$), it receives its outside option payoff $b > 0$ and G receives its status-quo payoff, 0. The game proceeds to the next period.

2. If I decides to invest, then G decides whether to violate the contract with that investor, $v \in \{0, 1\}$, where 1 indicates the decision to violate the contract.
 - If G decides not to violate the contract ($v = 0$), then both G and I benefit from the investment, receiving $z > 0$ and $a > 0$ respectively.
 - If G decides to violate the contract ($v = 1$), it receives $z + \theta$, and I receives 0.
3. After G 's decision, the game proceeds to the next period, where G interacts with a new investor I .

I assume $a > b$, that is, given no breach of contract, returns from the investment are greater than the payoff from not making an investment. If this was not true, investors would never invest. Following an investment, if G decides to break contract, it receives $z + \theta$, where $\theta = \{\underline{r}, r\}$ and $\underline{r} < 0 < r$, that is, the value of θ determines whether G 's returns from contract breach is positive or negative. A positive return, $r > 0$, means that there are short-run advantages of breaking contract for G over upholding its promises. Each new investor I makes its investment decision having observed the history of play, in particular, acts of breach, and G decides whether to break its commitments toward the current investor knowing that future investors will adjust themselves accordingly. The stage-game is depicted in Figure 2.

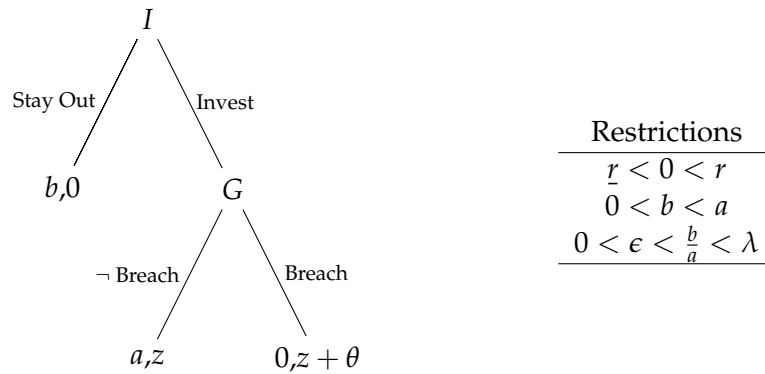


Figure 2: Stage Game. In all terminal nodes, the payoff listed on the left adheres to I .

Reputational incentives in this model arise due to private information G has about its returns from breach, captured by θ , whereas no uncertainty exists over the payoffs of investors I . I assume that G can be one of two types depending on whether contract breach is profitable, $\theta = \{\underline{r}, r\}$. If breach is not profitable $\theta = \underline{r} < 0$, and thus, G has no incentives to breach contract, I consider G

to be a *commitment type*. Alternatively, if G 's returns on breach is positive, $\theta = r > 0$, I consider it to be a *normal type*, in which case G will decide whether to breach contract or not depending on which behavior will maximize its expected utility, discounting future payoffs according to $\delta \in (0, 1)$. Investors I , do not know, but have beliefs over, whether G 's returns are positive or negative — that is, which type of G they might be facing. I believes G to be a commitment type (negative returns from breach) with probability μ , which corresponds to G 's reputation in the eyes of foreign investors. Throughout, we will focus on the behavior of the *normal type* of G , for whom breach is profitable in the short-run ($\theta = r > 0$).

The core premise underlying the arguments of this paper is that states' preferences are not fixed, but change over time. In the model, this is captured by allowing G 's returns from breach, its type, to change at the beginning of each period, following a simple Markov process. Importantly, investors know that G 's type might change but they do not directly observe when this happens. If G is a commitment type in period t , it will switch to a normal type with $1 - \lambda > 0$ probability, and with probability λ , it will remain a commitment type. If G is a normal type in period t , in $t + 1$ it will be a commitment type with probability ϵ , and will remain a normal type with probability $1 - \epsilon$. Together these determine the stationary probability for being a commitment type: $\frac{\epsilon}{1 - \lambda + \epsilon}$. The type transition matrix described above can be summarized as follows:

$$\begin{array}{l} \theta_{t+1} = Normal \quad \theta_{t+1} = Commitment \\ \theta_t = Normal \quad \left(\begin{array}{cc} 1 - \epsilon & \epsilon \\ 1 - \lambda & \lambda \end{array} \right) \\ \theta_t = Commitment \end{array}$$

Without loss of generality, I set $\mu_0 = \frac{\epsilon}{1 - \lambda + \epsilon}$ as the prior probability of G being the commitment type at the beginning of the game. The effect of the $1 - \epsilon$ term is twofold. One, it modifies how G views the future, making its effective discount factor $\delta(1 - \epsilon)$. Two, and more importantly, a positive $1 - \epsilon$ prevents the continuation value of the game for G to ever equal zero, and in so doing, introduces dynamism to the interaction that is necessary to generate building and spending behavior in equilibrium. Here I will assume the following about the transition probabilities:

$$0 < \epsilon < \frac{b}{a} < \lambda < 1 \tag{1}$$

This assumption makes sure that, if I is certain of G 's type today, the type replacements are rare enough that tomorrow's I will be able to act upon this knowledge. If this assumption is violated, type replacements are going to be too frequent, yesterday's information about G will not be relevant today, and thus reputational incentives will not arise. There are no other restrictions placed on the parameters, in particular, there are no restrictions on the discount factor δ .

The model is based on the adverse selection approach to reputations, hence it has the familiar structure of the canonical reputation models where a long-term player who has private information about its type sequentially interacts with short-term players *ad infinitum* (e.g. Alt, Calvert, and Humes 1988; Fudenberg and Levine 1989; Kreps and Wilson 1982). Different than those models, I relax the assumption that the long-term player's type is fixed, and let it vary according to a Markov process to generate persistent reputational concerns. In that regard, the model parallels the recent wave of research in economic theory on long-run reputation dynamics (Bohren 2013; Ekmekci, Gossner, and Wilson 2012; Faingold and Sannikov 2011; Liu 2011; Liu and Skrzypacz 2014; Phelan 2006; Wiseman 2008).

In recent years a few papers in economics have made the case that, similar to the dynamics presented here, a firm's incentives to invest in higher quality production decreases as it obtains a better reputation for quality production (Board and Meyer-ter-Vehn 2013; Bohren 2013). Phelan (2006) discusses similar dynamics, but focusing on the interaction between a government with a changing type which makes taxation decisions, and a continuum of citizens which individually decide how much to produce based on the government's prior actions. The model presented here is analogous to that of Phelan (2006), on which I rely extensively for both the presentation of the model and the formal proofs. Different than that paper, I model the commitment type of the long-run player, G here, based on its short-run returns to undesirable behavior, instead of a non-strategic automaton. Although as Weinstein and Yildiz (2013) show, these approaches yield the same results, modeling commitment types based on payoff uncertainty prevents confusions regarding the substantive interpretation of the type space. Additionally, in deriving the empirical implications below, I consider how the reputation dynamics presented here react to adverse exogenous shocks to the long-run player's stage payoffs. Most importantly, neither Phelan (2006) nor the other papers cited above discuss IR, FDI, or contract breach, and none of them include empirical tests of their arguments.

3.1.1 Strategies and Equilibria

I will focus on Markovian strategies and equilibria, with G 's reputation, μ , serving as the state variable. I restrict μ , I 's belief that G 's returns from breach is negative, in the following way: if I observes G to breach contract at any period, future investors will not think that G is more likely to be a commitment type.⁶ A mixed strategy for G is a mapping $\sigma_G(\mu, \theta): [0, 1] \times \{r, R\} \rightarrow [0, 1]$, which takes G 's reputation and returns from breach (its type) as inputs and gives the probability of contract breach. The commitment type of G has no incentive to breach contract, and given the restriction on beliefs mentioned above, its decision problem is trivial: if G has negative returns to breach, it never breaches contract, $\sigma_G(\mu, r) = 0$ for all μ . A mixed strategy for I is a function, $\sigma_I(\mu): [0, 1] \rightarrow [0, 1]$, from G 's reputation to the probability of investment. Then, a Markov strategy is defined as $(\sigma_I(\mu), \sigma_G(\mu, \theta))$.

Since each investor I enters the game for one period, investors care only about payoffs from the current period. An investor will decide to invest depending on how the probability of contract breach, σ_G , compares to a cutoff value $\sigma_G^* = 1 - \frac{b}{a}$; the point at which I is indifferent between investing and staying out. For any value of G 's reputation μ , if I plays a mixed strategy $0 < \sigma_I(\mu) < 1$, this implies $(1 - \mu)\sigma_G(\mu, N) = \sigma_G^*$. Note that if the reputation of G is high enough, that is, if I believes G to be the commitment type with sufficiently high probability, I will invest with certainty, since the commitment type of G never breaches contract. In particular, if $\mu > 1 - \sigma_G^*$, then regardless of the normal type of G 's strategy, thus for any value of σ_G , I will invest with certainty, $\sigma_I(\mu) = 1$. This means, beside σ_G^* , we can consider $\mu^* = 1 - \sigma_G^* = \frac{b}{a}$ as the cutoff reputation determining I 's actions.

Unlike investors, the state G cares about the future and can affect μ with its actions. First let's focus on how G can impact its reputation μ with its actions. Let μ_t denote the prior belief of I that G is the commitment type at time t , and μ'_t denote the updated posterior belief in that period. Define a function $M: [0, 1] \rightarrow [\epsilon, \lambda]$ where $M(\mu') = \lambda\mu' + \epsilon(1 - \mu')$. This function takes a posterior belief at the end of a given period t , and returns a prior belief for the next period based on the Markov transition probabilities. Since the commitment type of G never breaches contract,

6. That is, I prevent breaching contract to be a signal for G having no incentive to breach contract. Kreps and Wilson (1982) also appeal to a similar criterion for beliefs. See Cho and Kreps (1987) for a formal treatment of this restriction for sequential equilibria in signaling games.

if G breaches contract at time t , it reveals itself to be the normal type, setting G 's reputation to zero; $\mu'_t = 0$. This means the prior belief of I that G is a commitment type in the next period $t + 1$ is determined by the probability that a yesterday's normal type of G turns into a commitment type today; $\mu_{t+1} = M(\mu'_t = 0) = \epsilon$. If I invests and G does not breach contract at time t , the posterior belief μ'_t will be determined by σ_G , the probability of contract breach, and Bayes rule: $\mu'_t(\mu_t, \sigma_G) = \frac{\mu_t}{\mu_t + (1 - \mu_t)(1 - \sigma_G)}$. Since G 's type can change by the next period $t + 1$, the prior belief μ_{t+1} is determined as follows:

$$\mu_{t+1} = M(\mu'_t) = \lambda \left[\frac{\mu_t}{\mu_t + (1 - \mu_t)(1 - \sigma_G)} \right] + \epsilon \left[1 - \frac{\mu_t}{\mu_t + (1 - \mu_t)(1 - \sigma_G)} \right] \quad (2)$$

The first term on the right-hand side is the probability that G will stay a commitment type at time $t + 1$, if it was a commitment type at time t . The second term on the right-hand side of the equation is the probability that G will switch to a commitment type at time $t + 1$, if it was a normal type at time t . Finally, if I does not invest at time t , there is no updating to be done in that period, $\mu_t = \mu'_t$, and the prior belief at time $t + 1$ is determined by the type transition probabilities: $\mu_{t+1} = M(\mu_t) = \lambda\mu_t + \epsilon(1 - \mu_t)$.⁷

Since G can affect its future payoffs through changing its reputation, it will decide whether to breach contract or not at a given time depending on how this decision affects those future payoffs. Let $V(\mu)$ capture the continuation value for G , given a strategy $(\sigma_I(\mu), \sigma_G(\mu))$.⁸ Then we have:

$$V(\mu) = \sigma_G(\mu) \left[\sigma_I(\mu) [(z + r) + \delta(1 - \epsilon)V(\epsilon)] \right] + (1 - \sigma_G(\mu)) \left[\sigma_I(\mu) [z + \delta(1 - \epsilon)V(M(\mu'))] \right] \\ + (1 - \sigma_I(\mu)) [\delta(1 - \epsilon)V(M(\mu))] \quad (3)$$

The first term on the right-hand side of the equation captures G 's payoff from breaching contract and its continuation value after its reputation is spent. The second term captures the payoff from avoiding breaching contract, and the continuation value of the game with an improved rep-

7. The time index for μ is useful when trying to differentiate between the posterior at t from the prior at $t + 1$. Moving forward, if there is no time index specified, μ is the prior belief at the beginning of a period, and μ' is the posterior belief in the same period after the stage-game is resolved, but before the type transition probabilities are taken into account for the next period's prior.

8. Since the commitment type never breaches contract, I use $\sigma_G(\mu)$ instead of $\sigma_G(\mu, \theta)$ for notational convenience.

utation. The third term captures the continuation value of the game if I stays out. G will decide on whether to breach contract by comparing the continuation payoffs accrued from breaching contract $\sigma_I(\mu) [(z+r) + \delta(1-\epsilon)V(\epsilon)]$ and upholding its promises $\sigma_I(\mu) [z + \delta(1-\epsilon)V(M(\mu'))]$. This means if $0 < \sigma_G < 1$, then it must be that

$$\sigma_I(\mu) [(z+r) + \delta(1-\epsilon)V(\epsilon)] = \sigma_I(\mu) [z + \delta(1-\epsilon)V(M(\mu'))]$$

Given the description of the strategies of I and G , which depend only on the state variable μ , a Markov strategy $(\sigma_I(\mu), \sigma_G(\mu))$ is a Markov Perfect Equilibrium, such that σ_G and σ_I respect the optimization problems of G and I respectively, and are best-responses to each other for all μ .

The reason I focus on Markovian strategies are threefold. First, restricting strategies to be Markovian enables examining behavior that is substantively closest to the reputation dynamics I am trying to capture: How a state's reputation, formed via investors' learning and updating based on its past behavior, conditions the behavior of both the state and the investors. Thus I eliminate from consideration the strategies that involve complex "bootstrapping" where each player conditions its behavior on a certain configuration of play in the observed history by others.

Second, Markovian strategies demand a low cognitive load from the players, addressing prior concerns raised in the literature about whether the degree of sophistication dynamic models demand of its agents is appropriate (Downs and Jones 2002; Mercer 1996). Bhaskar, Mailath, and Morris (2013) show that in infinite horizon perfect information games, non-Markovian equilibria are not robust to the relaxation of infinite recall, and to adding small perturbations of the complete information game, displaying the fragility of non-Markovian folk theorems. This corroborates earlier results in economic theory which show that equilibria involving strategies based on infinite recall are not robust to the addition of an arbitrarily small noise to the monitoring structure (Mailath and Morris 2002, 2006). Further, the behavior prescribed by Markovian strategies, here conditioning behavior on the reputation of G , is also consistent with recent psychological research which highlights that agents often use simple rules of thumb, and that those can outperform more complicated strategies in achieving better outcomes particularly in highly complex decision making environments (Gigerenzer and Gaissmaier 2011).

Finally, this focus delivers significant analytical tractability and sharper predictions in dynamic

games by considerably reducing the set of equilibria, which is partly why Markovian equilibria are popular in the political economy literature (e.g. Acemoglu and Robinson 2001).

3.2 Analysis

I will construct a Markov perfect equilibrium where the state G is increasingly willing to breach contract as its reputation increases and breaches contract with certainty if its reputation is high enough, whereas each investor I is increasingly willing to invest as G 's reputation increases and I invests with certainty if G 's reputation is high enough. Throughout the focus is on the behavior of the normal type of G .⁹ First let's restate the assumptions over the parameters:

Assumption 1. The following restrictions are assumed to be true:

1. $\underline{r} < 0 < r$
2. $0 < b < a$
3. $0 < \epsilon < \frac{b}{a} < \lambda < 1$
4. $Pr(\theta = \underline{r} | v = 0) \geq Pr(\theta = \underline{r} | v = 1)$

The first item means that breaching contract is advantageous for the normal type of G in the short run $r > 0$, but not for the commitment type of G , $\underline{r} < 0$. The second item means that for I investing is better than not investing, if G does not breach contract. The third restriction ensures that type replacements are not too frequent. Otherwise reputations decay or improve too fast, even when G takes no action, and destroy reputational incentives. The fourth and final item restricts I 's beliefs such that observing a breach does not make I think that G is more likely to be a commitment type.

I will start with the equilibrium strategies of the state G . Claim 1 describes the strategy G will follow in equilibrium.

Claim 1. In equilibrium, if $\mu \geq \mu^* = \frac{b}{a}$, then G will breach contract with certainty. If $\mu_t \leq \mu^*$, then

9. If the type of G is not specified, I am always referring to the normal type of G .

G will play a mixed strategy to induce indifference between investing and staying out for I :

$$\sigma_G(\mu, N) = \begin{cases} 1 & \text{if } \mu \geq \mu^* \\ \frac{1-\frac{b}{a}}{1-\mu} & \text{if } \mu < \mu^* \end{cases}$$

The claim will have been proven once I construct the equilibrium. If the state is following the strategy described in Claim 1, this σ_G could be used to examine the evolution of G 's reputation. Setting $\sigma_G = \frac{1-\frac{b}{a}}{1-\mu}$ simplifies the equation for the equilibrium evolution of the prior beliefs at time $t + 1$, given investment and no contract breach at time t , as follows:

$$\mu_{t+1} = \left[\frac{\lambda - \epsilon}{\mu^*} \right] \mu_t + \epsilon \quad (4)$$

Before proceeding to the investor's equilibrium strategies, I need to show that, starting from any reputation, G can surpass the cutoff reputation level by upholding its promises.

Lemma 1. *Given Claim 1, μ can exceed μ^* in finite periods where investment is realized and the state does not breach contract.*

Proof. To prove Lemma 1, we need to consider two cases: (i) $\frac{\lambda - \epsilon}{\mu^*} \geq 1$, and (ii) $\frac{\lambda - \epsilon}{\mu^*} < 1$. If the first case is true, this means equation (4) is linear in μ^t , has a positive intercept, ϵ , and a slope greater than or equal to 1, which means G 's reputation can exceed μ^* in finite steps. In the latter case, the slope is smaller than 1, therefore we need to show that the fixed point of (4) is greater than μ^* . The assumption that $\lambda > \frac{b}{a} = \mu^*$ implies that the fixed point of equation (4) is $\frac{\mu^* \epsilon}{\mu^* - \lambda + \epsilon} > \frac{\mu^* \epsilon}{\lambda - \lambda + \epsilon} = \mu^*$. That is, the fixed point of (4) is greater than μ^* , indicating that L 's reputation can surpass μ^* in finite steps. \square

Next is the equilibrium strategy of investors. Similar to G 's strategy described in Claim 1, investors I will also set their probability of investment to make the state G indifferent between breaching contract and not. Let $\mu(0), \mu(1), \dots, \mu(N)$ be the sequence of prior probabilities I assigns to G being the commitment type, where $n = 0, 1, \dots, N$ indexes the number of consecutive periods where G upholds its promises given investment. Let $\mu(0) = \epsilon$, which is the same as G 's reputation after I observes contract breach in the previous period. These $\mu \in \{\mu(0), \mu(1), \dots\}$ chart the evolution of G 's reputation upon investors observing that the state upholds its promises for

n consecutive periods. I 's strategy will be specified based on those n consecutive steps tracking the improvement in G 's reputation. Let $\mu(N)$ be the first probability to equal or exceed the cutoff reputation level $\mu^* = \frac{b}{a}$. Let $V(n)$ be the continuation value of the game for G , when the prior is $\mu(n)$, and let \hat{V} be the continuation value when $\mu = \mu(0) = \epsilon$. Since I will randomize to make G indifferent between breaching contract and upholding its promises, for $n = 0, 1, 2, \dots, N - 1$, we have the following set of Bellman equations:

$$V(n) = (1 - \delta(1 - \epsilon))(z)\sigma_I(\mu(k)) + \delta(1 - \epsilon)V(n + 1) \quad (5)$$

$$= (1 - \delta(1 - \epsilon))(z + r)\sigma_I(\mu(k)) + \delta(1 - \epsilon)\hat{V} \quad (6)$$

The first expression is the continuation value for G upon upholding its promises, when its reputation is $\mu(n)$. The second expression is the continuation value when G breaches contract, when its reputation is $\mu(n)$. The payoffs are normalized by multiplying with the constant $(1 - \delta(1 - \epsilon))$. Finally, since Claim 1 specifies that the state will breach contract with certainty once its reputation hits the cutoff μ^* , we also have:

$$V(N) = (1 - \delta(1 - \epsilon))(z + r) + \delta(1 - \epsilon)\hat{V} \quad (7)$$

This represents the continuation value of the game for G once it builds up its reputation at a level at or above the cutoff μ^* and then breaches contract. Equations 5-7 describe a system of Bellman equations with $2N + 1$ equations and $2N + 1$ unknowns, and thus the system has a unique solution. Solving for σ_I we find that:

$$\sigma_I(\mu(n)) = \frac{\sum_{i=0}^n \left(\frac{r}{z\delta(1-\epsilon)}\right)^i}{\sum_{i=0}^N \left(\frac{r}{z\delta(1-\epsilon)}\right)^i} \quad (8)$$

which is between 0 and 1. For periods where G 's reputation surpasses μ^* , $n \geq N$, we set $\sigma_I(\mu(n)) = 1$. For periods $n < N$, both players G and I are indifferent between their actions, and best respond to each other. I strictly prefers to engage in any period t if $\mu_t > \mu^*$ (and indifferent if equality), so $\sigma_I(\mu(n)) = 1$ for $n \geq N$ is also a best response. In order to prove that these strategies constitute a Markov perfect equilibrium, we need to show that G prefers breaching contract at

period $n \geq N$, when its reputation is at or above μ^* .

Lemma 2. *For $n \geq N$, G strictly prefers breaching contract:*

$$V(n) > (1 - \delta(1 - \epsilon))(z) + \delta(1 - \epsilon)V(n + 1)$$

Proof. Proof is by contradiction. Suppose at period n , we have the prior $\mu(n) \geq \mu^*$, and G weakly prefers upholding its promises:

$$V(n) \leq (1 - \delta(1 - \epsilon))(z) + \delta(1 - \epsilon)V(n + 1)$$

Given that $\sigma_I = 1$ for all such priors and that G weakly prefers breaching contract, for all $n \geq M$ it should be that $V(n) = V(n + 1)$. Facing any such prior, the continuation of the game is identical. G weakly preferring to uphold its promises for $n \geq N$ means that:

$$\begin{aligned} (1 - \delta(1 - \epsilon))(z + r) + \delta(1 - \epsilon)\hat{V} &\leq (1 - \delta(1 - \epsilon))(z) + \delta(1 - \epsilon)V(N + 1) \\ V(N + 1) &\geq \frac{1 - \delta(1 - \epsilon)}{\delta(1 - \epsilon)}r + \hat{V} \end{aligned}$$

However, by setting $n = N - 1$ in (6) and solving for $V(N)$ using the equality of (5) and (6) we get:

$$V(N) = \frac{1 - \delta(1 - \epsilon)}{\delta(1 - \epsilon)}r\sigma_I(\mu(N - 1) + \hat{V})$$

Since $0 < \sigma_I(N - 1) < 1$, this implies $V(N + 1) > V(N)$, a contradiction. \square

How much the state G is tempted to breach contract depends on the probability that the investors are willing to invest. Once the reputation of the state reaches the cutoff reputation level, investors are willing to invest with certainty. This means sacrificing the opportunity cost of contract breach will not result in a higher likelihood of investment as it did when the state's reputation was lower. Further, good reputations are decaying, as can be seen in equation (4). A state with a perfect reputation yesterday sees its reputation reduced to $\lambda < 1$. Add this the fact that $\hat{V} > 0$, that is, contract breach does not preclude rebuilding reputations, the temptation to breach contract for periods $n \geq N$ becomes irresistible.

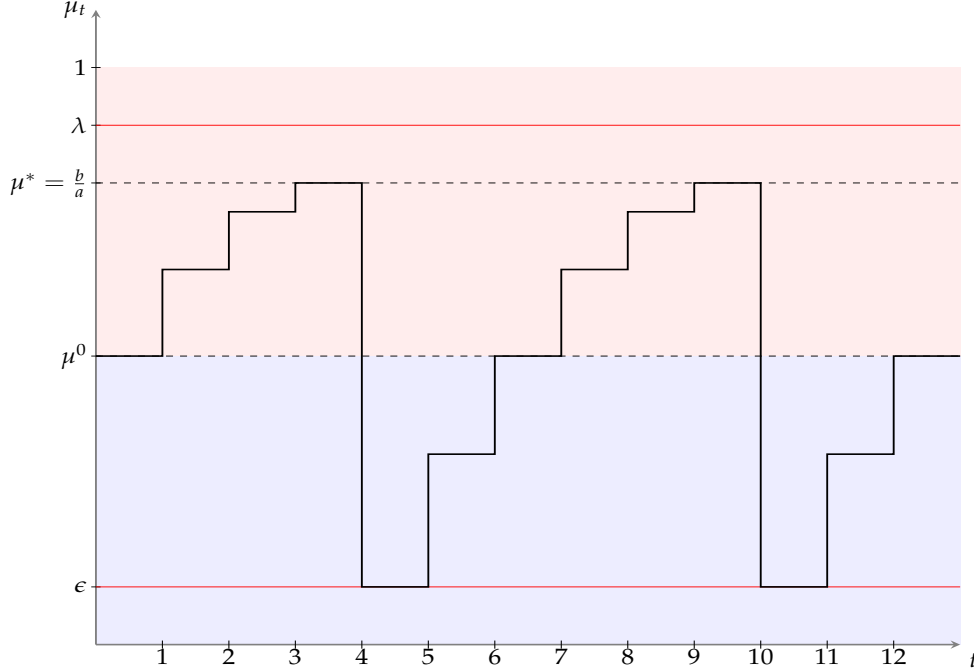


Figure 3: Evolution of G 's reputation in a scenario consistent with the equilibrium play.

Together these conclude the proof that the strategies described above constitute a Markov Perfect Equilibrium, which is stated formally in Proposition 1.

Proposition 0. *Suppose Assumption 1 holds, then the following strategies constitute a Markov Perfect Equilibrium:*

- If $\mu < \mu^* = \frac{b}{a}$, then G plays a mixed strategy, where it breaches contract with probability $\sigma_G(\mu, r) = \frac{1-\mu^*}{1-\mu}$. If $\mu \geq \mu^*$, then G breaches contract with certainty; $\sigma_G(\mu, r) = 1$.
- Let $\mu(0), \mu(1), \dots, \mu(N)$ be a sequence of prior probabilities, where $n \in \{0, 1, \dots, N\}$ is the number of consecutive periods before the current one in which G did not breach contract given I 's investment. Then if $\mu < \mu^*$, I plays a mixed strategy such that it invests with probability $\sigma_I(\mu(n)) = \frac{\sum_{i=0}^n \left(\frac{r}{z\delta(1-\epsilon)}\right)^i}{\sum_{i=0}^N \left(\frac{r}{z\delta(1-\epsilon)}\right)^i}$. If $\mu \geq \mu^*$, then I invests with certainty; $\sigma_I(\mu) = 1$.

Figure 3 displays the evolution of G 's reputation in a scenario consistent with the equilibrium dynamics laid out in Proposition 1, where the state G builds its reputation up only to spend it once again. Next, Corollary 1 states how increasing and decreasing G 's reputation affects (i) the probability that G breaches contract today, and (ii) the probability that I invests in G .

Corollary 0.1. *In the Markov Perfect Equilibrium described in Proposition 1,*

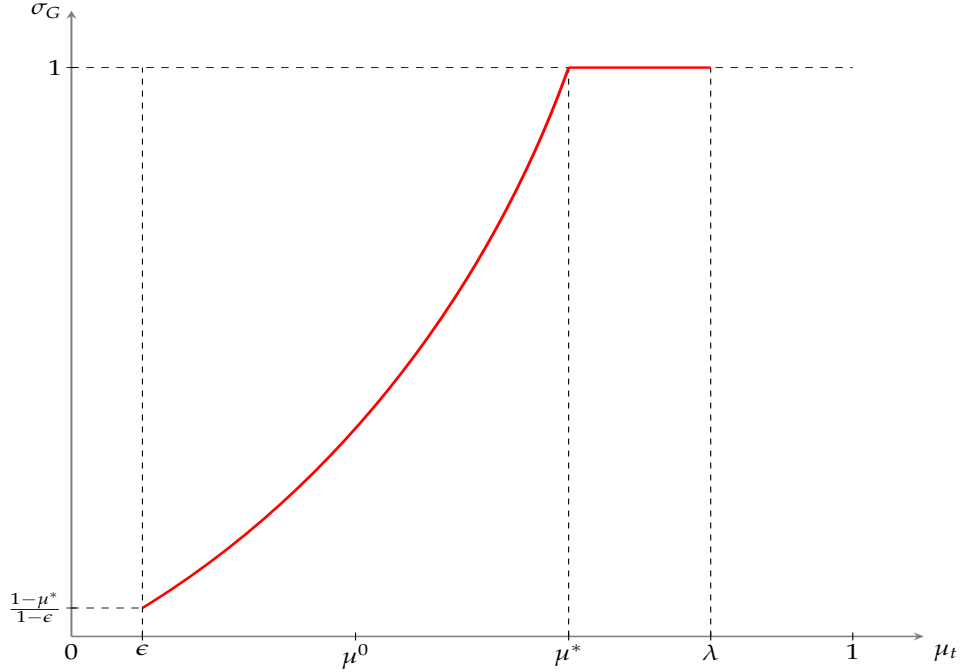


Figure 4: Probability of G breaching contract as a function of its reputation.

- *The probability that the normal type of G breaches contract weakly increases in μ . It strictly increases for $\mu < \mu^*$.*
- *The probability that I invests in G weakly increases in μ . It strictly increases for $\mu < \mu^*$.*

Corollary 1 follows directly from Proposition 1, and by checking the first order condition of σ_G with respect to μ , and the construction of σ_I , and lays out the answer to the question posed in the beginning. If a state is facing a commitment problem, where short-term incentives for breaching contract and long-run incentives to uphold promises to secure greater FDI flows coexist, then that state's willingness to build its reputation diminishes, the better its reputation gets. In other words, all else equal and given the commitment problem, the better a state's reputation is, the higher the likelihood that this state will break its contract with a foreign firm. The reverse is also true: the worse a state's reputation is, the lower the likelihood that this state will break its contract with a foreign firm. The equilibrium probability of G's breach behavior as a function of its reputation is depicted in Figure 4

Interestingly, it is also true that the better a state's reputation for treating FDI is, the more likely that foreign investors will invest in it. In other words, the fact that states have increasing

incentives to take advantage of their reputations is not inconsistent with the fact that investors are willing to trust states with better reputations. While the existence of reputational incentives help parties achieve mutually desirable outcomes, this is primarily due to reputations' constraining effect on states with poor reputations, and the side effect is that the reputation mechanism creates opportunities for states to take advantage of their reputations. Investors are not unaware of this problem: they are willing to endure the cost of an odd betrayal for the benefit of constraining those with poor reputations.

Next, I will lay out the empirical strategy to test the first item in Corollary 1, that those states with better reputations are more likely to breach contract. Before that, however, I will discuss an important threat to inference, a problem that the researcher shares with the investors in the game: inability to distinguish which states have the short-run incentives to breach contract, and which do not.

4 Uncertainty over Returns to Breach and Empirical Inference

The primary challenge to address to find evidence for the mechanism where a reputation for investor-friendliness increases incentives to breach contract is the uncertainty over returns to breach behavior. Recall that reputational incentives arise in the game due to investors' uncertainty about states' preferences: Investors cannot be sure whether host-governments have short-run incentives to breach contract or not, as that is host governments' private information. The researcher is also uncertain about this, much like the investors. Put another way, we do not know, at any given time, the distribution of "commitment types," who do not stand to gain from breaching contract in the short-run, and "normal types" who do. Since reputation dynamics, and thus the model's predictions, pertain only to "normal types," who do face the short-run versus long-run incentives dilemma, and not to "commitment types," who do not, how these two groups are distributed across reputation levels is consequential.

In order to see the problem, consider the following simulated data based on the game depicted in Figure 2. The steady state distribution of types is set to be 0.25 for commitment types and 0.75 for normal types. The type transitions are based on the following transition probability matrix:

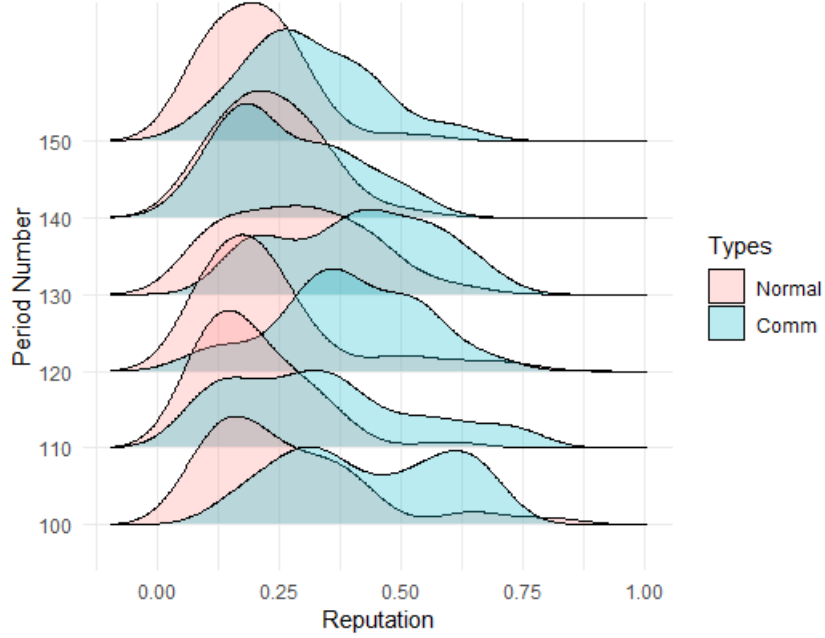


Figure 5: Distribution of reputations by type in periods 100-150

$$\begin{array}{l}
 \theta_{t+1} = \text{Normal} \quad \theta_{t+1} = \text{Commitment} \\
 \theta_t = \text{Normal} \\
 \theta_t = \text{Commitment}
 \end{array}
 \left(\begin{array}{cc}
 1 - \epsilon = 0.9 & \epsilon = 0.1 \\
 1 - \lambda = 0.3 & \lambda = 0.7
 \end{array} \right)$$

These probabilities are selected to reflect two points. One, commitment types are rarer than normal types. Two, commitment types transition into normal types more frequently than normal types transitioning into commitment types. Other parameters are set to the following values, which are selected to fit the inequalities in Assumption 1: (i) $b = 2$, which is investors' outside option; (ii) $a = 3$, which is I 's payoff given entry and no contract breach; (iii) together these imply that $\mu^* = \frac{b}{a} \approx 0.67 > \mu^0 = 0.25$; (iv) $\delta = 0.9$, which is the discount factor; and finally (v) G 's payoff from breaching contract is set at $z + r = 3$ and the payoff from upholding its promises at $z = 2$, which together satisfy the requirement $\delta(1 - \epsilon) = 0.81 > 0.33 = \frac{r}{z}$. In the simulation, 100 states G each interact with a different investor I over 500 rounds.

Figure 5 displays the reputation distribution by type at six different points in time between rounds 100 and 150. Unsurprisingly, the central tendency of the commitment types' reputation distribution tends toward better reputations, while the opposite is true for the normal types. An-

alytically, the normal type of G in the game determine its probability of breach by taking into account the likelihood that it will be mistaken as a commitment type; effectively taking into account the reputation distribution among commitment types. This means, investors I , who do not know which type they are facing, will always face the same mixed action σ_G^* . In other words, if we regressed a variable capturing G 's breach behavior on G 's reputation without having a way of taking types into account, the coefficient of reputation would be zero in expectation according to the theoretical model. Since type transitions are probabilistic, it is easy to see that in the absence of a way to account for these distributions, regressing an outcome variable, such as the one capturing states' breaches of contract, on their reputations, would provide a biased answer to our query. In fact, the result could be the opposite of what we would expect.

Using the simulated data, I create a proxy variable for reputations by counting the number of observed prior interactions (that is, given investment) where the state upheld its promises since the state's last contract breach behavior. Since the data is simulated, values of the underlying parameters are known, and thus it is possible to compare this proxy for reputation with the actual reputation of G (prior belief at time t about G being the commitment type). The pairwise correlation between the two variables is 0.94. I then estimate two simple OLS models, both using as the outcome the same binary indicator variable which captures whether divesting behavior occurred in that period. In the first model, I regress this outcome on the proxy reputation variable. In the second model, I do the same but only within the subset of "normal type" states. The results are in Table 1.

Table 1: OLS Regressions: Reputation and Contract Breach

| | Pr(Breach Contract) | Pr(Breach Contract Normal Type) |
|------------------|---------------------|----------------------------------|
| Reputation Proxy | -0.001 (0.001) | 0.048 (0.002) |
| Constant | 0.337 (0.004) | 0.385 (0.004) |
| Observations | 27,939 | 20,304 |
| R ² | 0.00003 | 0.040 |

The first model displays a negative substantive relationship between reputations and con-

tract breach behavior. The coefficient of the reputation variable, although associated with a low p -value, indicates that states with a better reputation are less likely to tarnish their reputations. The second model estimates the same model within the “normal type” subset of the simulated data and recovers the actual relationship between reputations and breach behavior. Clearly, conducting the analysis on the normal type subset of states is only possible in simulations. Then, it is necessary to develop an empirical strategy that would enable tracing the existence of this mechanism among states who do face a commitment problem vis-a-vis foreign investors, due to the existence of short-run incentives to breach contract.

Next, I will discuss the empirical strategy to circumvent this problem: via leveraging the as-if random occurrence of natural disasters as exogenous economic shocks manipulating the short-run incentives to breach contract.

5 Empirical Strategy

Table 1 shows that a naive comparison between the contract-breach behavior of lower-reputation and higher-reputation states can lead to empirical results that are null or even opposite the expectations, even when the data are generated directly from the model. In order to seek evidence for the positive relationship between reputations and contract breach behavior, I leverage the occurrence of severe natural disasters, which represent adverse economic shocks and are exogenous to reputation dynamics. The idea is that the occurrence of a natural disaster will increase the short-run temptation to breach contract for a given state.¹⁰

The effect of a natural disaster can travel through two channels: (i) through increasing returns on breach (r), and (ii) through the state’s reputation itself (μ).¹¹ I will discuss empirical implications pertaining to each of these channels below.

10. If after disasters there were substantial foreign aid inflows counteracting the damages, then natural disasters would not function as adverse shocks increasing states’ short-run temptations to breach contract. However, this does not seem to be the case. Noy, Becerra, and Cavallo (2012) find that while aid flows do increase following disasters, the typical surge in aid covers less than 3% of total estimated damages.

11. Note that there are no restrictions on the discount factor δ to generate the equilibrium behavior discussed above, apart from it being positive. In particular, σ_G does not depend on δ . Therefore, I am not considering “patience” as one of the channels through which the disaster effect can travel.

5.1 Returns Channel

A natural disaster can increase a given host state's returns (type) from violating foreign investors' property rights (r). To put it differently, it increases the state's opportunity costs of upholding its promises vis-a-vis foreign investors. Suppose a disaster strikes at time t . Then, I take the returns channel to imply that, if a state was a commitment type ($r < 0$) entering time t , the disaster's effect on returns to breach will turn that state into a normal type ($r > 0$) with some probability. If a state was already a normal type, that is, if that state already had short-run temptation to breach contract ($r > 0$), this shock will not cause the expected behavior to change, since the state's equilibrium probability of breaching contract, $\sigma_G(\mu) = \frac{1-\mu^*}{1-\mu}$, does not depend on r .

We do not know how exactly commitment and normal types are distributed in the world by reputation level. However, the model implies two things about the relationship between normal and commitment types' probability distributions over reputation levels, both of which can be gleaned at from Figure 5. One, at any time, there is positive probability associated with any reputation level for both normal and commitment types. Because types can change, a normal type with a poor reputation can turn into a commitment type tomorrow, or vice versa; a commitment type with a good reputation can turn into a normal type. Two, on average, commitment types should have better reputations. This is because commitment types never breach contract in equilibrium, but normal types do. This implies that the reputation distribution of the commitment types first order stochastically dominate that of normal types:

$$F_C(\mu) \leq F_N(\mu), \quad \forall \mu$$

Then, it should be that a relatively higher proportion of states who receive the disaster shock will turn into normal types the further we move toward the right tail of the overall reputation distribution. This means, change in types will be more pronounced in aggregate as we move toward the right tail of reputations. Moreover, we know that normal types breach contract with probability $\sigma_G(\mu) = \frac{1-\mu^*}{1-\mu}$. This probability is increasing in reputations ($\sigma'_G(\mu) > 0$). This means, if a state switches from being a commitment type to a normal type due to the disaster shock, change in its probability of breach will be greater the higher its reputation. Together, these imply the following expectation for the returns channel:

Table 2: Change in the probability of contract breach following a disaster

| Reputation | Returns | Beliefs | In Aggregate |
|------------|---------|---------|--------------|
| High | ++ | - | + |
| Low | + | -- | - |

Expectation 1. Following disasters, states with better reputations are more likely to breach contract. This effect is greater for higher reputations, and muted for lower reputations.

5.2 Reputation Channel

A natural disaster can also affect state behavior by directly impacting its reputation. Recall that I define the reputation of G as the probability I attaches to G 's returns from breach being negative. I claimed above that disasters are a shock to G 's returns from breach, turning a commitment type of G into a normal type with some probability. Also note that, unlike the underlying type changes in the model which are not observed by potential investors, future investors do observe the occurrence of natural disasters. Together, these imply that, having observed the disasters, investors will assign a lower probability for G being the commitment type following disasters. In other words, disasters are an adverse shock to reputations as well.

Having received this adverse shock on its reputation, the commitment type of G will not change its behavior. The normal type of G , however, will be less likely to breach contract, as reducing reputations increases the probability that G will breach contract ($\sigma'_G(\mu) > 0$). Additionally, given that the probability distribution of reputations among commitment types first order stochastically dominates that of normal types, the behavioral change will be more pronounced in aggregate as we move toward the left tail of reputations. This implies the following expectation:

Expectation 2. Following disasters, states with poor reputations are less likely to breach contract. This effect is greater for lower reputations, and muted for higher reputations.

5.3 Hypothesis

Table 2 lays out the expectations regarding a state's propensity of contract breach, given that it faced a natural disaster, based on the returns and reputations channel. These expectations can be summarized with the following hypothesis:

Hypothesis 1. States with better reputations for investor-friendliness are more likely to breach contract after disasters compared to states with worse reputations.

In order to evaluate this hypothesis, I follow two different approaches. First, I estimate a set of linear models with and without state and year fixed effects to examine whether states' breach behavior following disasters vary by their prior reputation for investor-friendliness. Second, I match states based on the level of ex-ante disaster risk they face. Since some states are geographically more prone to disasters than others, it is likely that firms make their investment decisions by taking such risks into account. However, once the ex-ante disaster risk is accounted for, the actual occurrence of natural disasters should satisfy the assumption of strongly ignorable treatment assignment Rosenbaum and Rubin (1983). Following the matching procedure, I examine heterogeneous treatment effects of natural disasters on contract breach behavior based on states' prior reputation levels.

6 Data and Measures

The dataset I employ consists of a yearly panel of all countries over the time period 1990-2016. Information about contract breach, the outcome of interest, come from cases filed against governments by foreign firms in international arbitration courts (Wellhausen 2016).¹² As such, this data do not contain information about breach incidents which have not been brought to international arbitration, or was not publicly disclosed. Notwithstanding this limitation, this dataset includes the most comprehensive set of records for a wide range of breach behavior, many of which are of the "creeping expropriation" type, and is widely used in the political economy literature to capture breach behavior (e.g. Allee and Peinhardt 2011; Wellhausen 2015). The dependent variable is a binary indicator, which takes the value of 1 if a foreign firm filed a case against the host-government for breaching contract in an arbitration venue and zero otherwise.

The main independent variable, reputation, is constructed using past contract breach behavior and takes values between 0 and 1. In any given year, I take the existence of contract breach

12. These venues include International Centre for Settlement of Investment Disputes (ICSID), the International Chamber of Commerce (ICC), the Stockholm Chamber of Commerce (SCC), the Permanent Court of Arbitration (PCA), the London Court of International Arbitration (LCIA), regional arbitration centers, and ad-hoc tribunals. See Wellhausen (2016) for further information.

behavior as bad news for that country's reputation, and the lack of contract breach behavior as good news. If a state breached contract, I set its reputation to zero for the next year. Then, for each year the state did not breach contract, I increase its reputation by $\frac{1}{h}$ until it reaches 1, where h is the horizon for the reputation updating. Theory does not provide clear guidance about how long it should take a reputation of 0 to reach 1 given consecutive good news, so I construct reputation measures with $h = \{3, 5, 10\}$ year horizons. All models reported on the body of the paper employ the reputation measure based on a 10-year horizon, but results are robust to conceptualizing reputations with 3- and 5-year horizons as well.¹³

Data on natural disasters come from the publicly available Emergency Events Database (EM-DAT). The information that make up the dataset is provided by a variety of sources such as UN agencies, governments, insurance companies and research institutions, as well as non-governmental organizations such as the Red Cross and Red Crescent (Caruso 2017). Further, the dataset also includes information about the magnitude of each disaster in terms of the number of people killed, and the financial damage incurred. I exclude industrial disasters as well as disasters with plausible links to state infrastructure and short-run human activity, such as epidemics. Following the previous literature examining the effect of disasters on political and economic outcomes of interest, I concentrate on storms, floods, volcanic activities, earthquakes, and landslides (Cavallo et al. 2013; Escaleras and Register 2011; Ramsay 2011).

As discussed in Cavallo and Noy (2011) and Cavallo et al. (2013), there is an increasing trend in the occurrence of natural disasters attributed to the improved recording of mild events over time. Following Cavallo et al. (2013), I concentrate on large disaster events, defining "large" based on whether a given event falls above the cutoff value based on the 99th, 90th, 75th percentiles of the world distribution in terms of casualties as a share of the total population.¹⁴ Further, the focus on large disasters is also theoretically warranted, since these events are supposed to represent economic shocks significant enough to potentially affect state behavior. Here I present results

13. The choice of presenting reputations with 10-year horizons is based on case studies regarding how investors continue to be wary of investing in countries which have expropriated ten years earlier, and in some cases even more. See for example the case of Philippines and Manila Airport case with the German firm Fraport AG, where the government of Philippines continued to have difficulties attracting foreign investment in 2013 after breaching its contract with Fraport AG in 2002 (Johnston, Janz, and Mahdavi 2015).

14. The increasing trend in disaster occurrence disappears for large disasters conceptualized this way. See Cavallo et al. (2013) for a more involved discussion. I am focusing on casualties instead of monetary damages to determine severe disasters, because casualty data are more widely available and deemed more reliable (Cavallo et al. 2013). The results are robust to defining disasters based on financial damages.

based on severe disasters above the 90th percentile cutoff, but results are robust to conceptualizing disasters via the other cutoffs. Finally, for the first set of models presented here, I consider a state which experienced a disaster at time t as “treated” for the years t to $t + 2$. I also use alternative conceptualizations of the treatment period with +1, and +4 years. I show the results based on all three conceptualizations of the treatment period for the matched designs. The results are not sensitive to these different conceptualizations of the treatment. Figure 6 shows that those states which faced severe disasters have similar GDP per capita and population figures to those which did not.

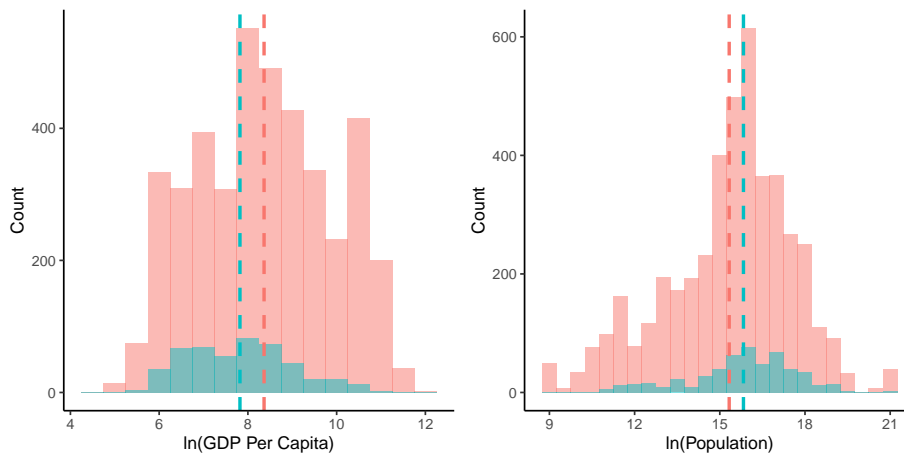


Figure 6: GDP per capita and population distributions by disaster occurrence (90th percentile). Green histograms are treated observations.

For the designs that involve matching states based on ex-ante disaster risk, I create a measure of disaster risk at time t for a given state, using historical disaster data pertaining to years $t - 1$ to $t - 10$ from EM-DAT. This means, the disaster data used to calculate ex-ante risk is from 1980-2016. I create the measure by using gradient boosting machines (GBM), with country and year indicators, GDP, populations, and population density information.¹⁵ Unsurprisingly, the predictive performance of the risk measure is more limited for more severe disasters, country indicators carry the heaviest weight, and the measure is highly skewed. In order to conduct exact matching on disaster risk, I collapse the risk measure into 5-categories based on its quintiles ranging from 0 (very low risk) to 4 (very high risk).

15. See Kuhn (2008) for an accessible introduction of tree-based models, including GBMs, to a political science audience. I used the R package caret for using GBMs to calculate ex-ante disaster risk (Kuhn 2008)

Information about GDP, GDP per capita, population, and population density are used in the creation of disaster risk measures, since the severity of the disaster is expressed in casualties as a percentage of the country's population. This information comes from the World Bank Development Indicators Database and Penn World Tables (Feenstra, Inklaar, and Timmer 2015), combined and augmented by Graham and Tucker (2019). I use the natural log transformations of GDP, GDP per capita, and population variables. Additionally, I also use net FDI inflows as a percentage of GDP both as a control variable in linear models, and in matching. This variable captures net FDI in a given economy from foreign sources minus the net FDI by the reporting economy to the rest of the world. The idea is that those states relying on a greater volume of FDI inflows might have more opportunities to breach contract, and that they would also put a higher value on accessing future FDI flows. This data come from the World Bank Development Indicators. Table 3 displays summary statistics for the variables used in the empirical analysis.

Table 3: Summary statistics

| Statistic | N | Mean | SD | Min | Max |
|-------------------------------------|-------|---------|-----------|---------|----------|
| Contract breach | 4,992 | 0.097 | 0.296 | 0 | 1 |
| Reputation(10-year horizon) | 4,992 | 0.773 | 0.376 | 0 | 1 |
| Disaster (≥ 90 th percentile) | 4,801 | 0.107 | 0.310 | 0 | 1 |
| Net FDI inflows (% of GDP) | 4,353 | 4.799 | 14.641 | -82.892 | 451.716 |
| Population density | 4,768 | 248.519 | 1,295.547 | 1.406 | 18,865.5 |
| ln(Population) | 4,992 | 15.384 | 2.209 | 9.105 | 21.039 |
| ln(GDP) | 4,801 | 23.781 | 2.441 | 16.881 | 30.440 |
| ln(GDP per capita) | 4,620 | 8.307 | 1.504 | 4.749 | 11.886 |
| Disaster Risk | 4,992 | 1.925 | 1.486 | 0 | 4 |

Figure 7 includes a set of histograms which show the distribution of reputations by breach behavior, and by whether or not breach happened following a disaster. Raw data seem to accord with my expectations. Comparing the top left and bottom left panes, we see that the right side of the distribution, belonging to those states with higher reputations, represents a greater proportion of the distribution after disasters (top left) compared to non-disaster years(bottom left).

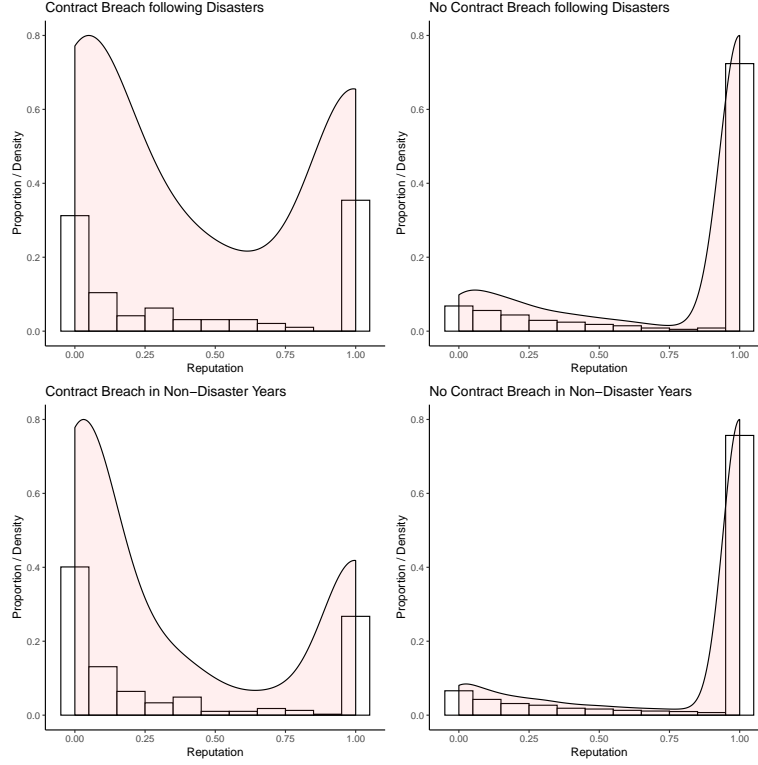


Figure 7: Reputation distributions by contract breach behavior (or lack thereof) and by whether this behavior happened during disaster years. The kernel densities are scaled such that the highest point is 0.8 across the panels for ease of comparison.

7 Empirical Analysis

Are states with better reputations more likely to breach contract? In the first part of the empirical analysis, I present OLS regressions of the following form:

$$Breach_{it} = \pi Reputation_{it-1} + \alpha_i + \eta_t + \mathbf{X}'_{it-1}\beta + \epsilon_{it} \quad (9)$$

$$Breach_{it} = \pi_0 Reputation_{it-1} + \pi_1 Reputation_{it-1} \cdot Disaster_{it} + \pi_2 Disaster_{it} + \alpha_i + \eta_t + \mathbf{X}'_{it}\beta + \epsilon_{it} \quad (10)$$

$Breach_{it}$ is a binary measure capturing whether country i in year t engaged in contract breach. $Reputation_{it-1}$ is the reputation of country i pertaining to the previous year $t - 1$. α_i are the set of country fixed effects, absorbing all unobserved time-invariant country-specific factors, and η_t are

the year fixed effects, absorbing global trends affecting breach behavior and reputations. These are included in all specifications presented in Table 4. \mathbf{X}_{it} is the matrix of control variables, which include one-year lagged measures of logged GDP, logged GDP per capita, logged population, population density and net FDI inflows (as % of GDP). In all regressions, I report heteroskedasticity-robust standard errors clustered at the country level.¹⁶ The first two models reported in Table 4 present estimates based on Equation 9, with and without the control variables \mathbf{X}_{it} . Third and fourth models in Table 4 present estimates based on Equation 10, with and without the control variables.

Table 4: Reputation and Contract Breach

| | (1) | (2) | (3) | (4) |
|---|-------------------|-------------------|-------------------|-------------------|
| Reputation _{t-1} | -0.079 (0.024) | -0.060 (0.025) | -0.098 (0.028) | -0.077 (0.028) |
| Reputation _{t-1} • Disaster _t | | | 0.060 (0.038) | 0.053 (0.039) |
| Disaster _t | | | -0.061 (0.035) | -0.050 (0.036) |
| Controls | ✗ | ✓ | ✗ | ✓ |
| Observations | 4,797 | 4,158 | 4,797 | 4,158 |
| Clusters | 195 | 187 | 195 | 187 |
| Adjusted R ² | 0.216 | 0.217 | 0.217 | 0.218 |

Reported are OLS coefficients and CR2 standard errors clustered by country in parentheses. All variables include country and year fixed effects. Control variables are $\ln(\text{GDP})_{t-1}$, $\text{GDP per capita}_{t-1}$, $\ln(\text{Population})_{t-1}$, Population density_{t-1}, and Net FDI inflows_{t-1}.

The results in columns 1-2 of Table 4 display a negative and statistically significant correlation between reputations and breach behavior, where higher reputations are associated with lower probability of contract breach — substantively similar to the results presented in Table 1, column 1, based on simulated data. These results do not take the disaster shock into effect, suffer from the inference problem discussed in Section 4, and thus are unable to recover the proposed positive re-

16. Standard errors clustered by country are calculated based on the bias reduced linearization method (CR2) following McCaffrey and Bell (2002) and Imbens and Kolesar (2016).

relationship between reputations and breach. That is, these correlations ignore both the uncertainty around whether states' returns from breach is positive or negative, and the resulting strategic behavior by states which face the temptation to breach contract.

Results in Table 4, columns 3-4 take the disaster shock into account. The primary quantity of interest for the hypothesis is the interaction term, which captures the change in the relationship between reputations and breach behavior as a result of the disaster shock. For ease of interpretation, the left panel of Figure 8 displays three sets of effects based on models 3 and 4: (i) the relationship between reputations and contract breach behavior in non-disaster years, (ii) the same relationship in disaster years, and (iii) the difference between the two. The bands represent 90% (thick line) and 95% (thin line) confidence intervals. The panel on the right (based on Table 4, model 3) displays the marginal effect of disasters on the probability of contract breach by reputation level.

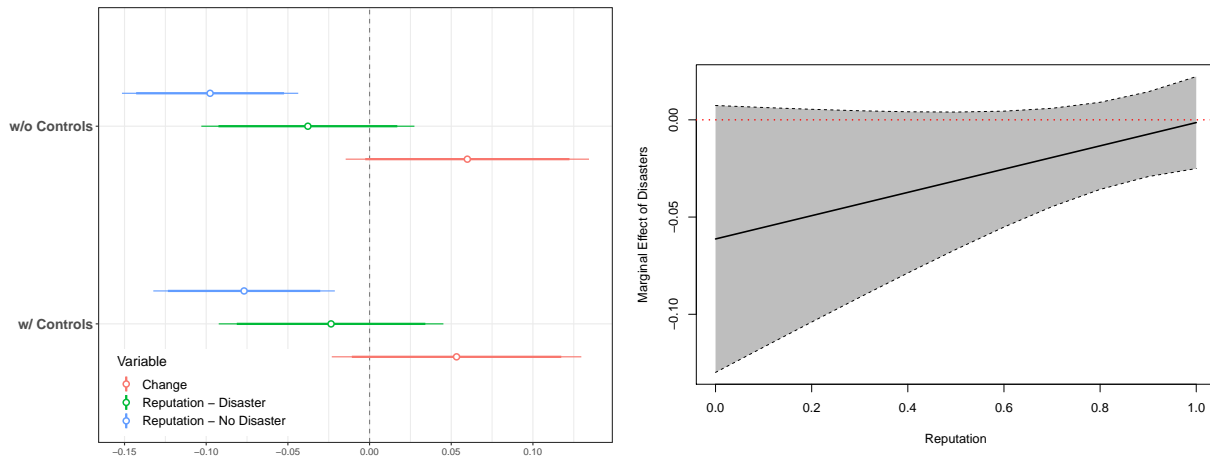


Figure 8: Left panel displays the relationship between reputation and contract breach by disaster occurrence based on models 3-4 of Table 4. The bands represent 90% (thick line) and 95% (thin line) confidence intervals. The right panel is based on model 3 of Table 4 and displays the marginal effect of disasters on the probability of contract breach by reputation level.

The first thing to observe in the left panel of Figure 8 is that the overall correlation between reputations and contract breach is negative, regardless of whether we look at disaster years or non-disaster years. In other words, in either case, higher reputations are associated with lower probability of breach. This negative correlation weakens in disaster years (green lines) compared to non-disaster (blue lines) years, and becomes statistically indistinguishable from zero for conventional levels. This change and the uncertainty around it are captured by the red lines in the figure. Although the change is imprecisely estimated, the results indicate that the relationship

between reputations and contract breach shift in the hypothesized direction. The right panel of Figure 8 shows that disasters have a greater marginal effect on the probability of breach for higher reputation levels. In particular, facing a severe disaster is 6% more likely to lead to contract breach among countries with very good reputations ($Reputation_{t-1} = 1$) compared to countries with poor reputations ($Reputation_{t-1} = 0$).

7.1 Matched Design

The preceding analysis provided supportive evidence for the hypothesis that states which faced severe disasters are more likely to violate foreign investors' property rights the better their reputation. However, some areas in the world are more prone to receiving disasters than others. While the exact timing of disasters are difficult to predict, both foreign firms and states could be adjusting their behavior according to the overall ex-ante risk of facing a disaster. Further, although Figure 6 shows that states which faced disasters have similar population and GDP per capita distributions to those states which did not face disasters, ignoring ex-ante risk might still lead to a biased comparison.

In order to alleviate the concern that the treatment assignment is not random, I employ a matched design to clarify the comparison of breach behavior between states which did and did not face a disaster. The goal is to compare post-disaster behavior of states which did face a disaster to behavior of states which did not face a disaster in the same time period, had otherwise similar ex-ante risk of facing a disaster, had similar reputations, and FDI inflows. Given the interest in post-disaster behavior, the operationalization of the treatment period is consequential for the matching procedure. For this reason, I report the results for three different operationalizations of the treatment period: +1, +2, +4 years after disaster strikes (that is, 2-, 3-, and 5-year treatment periods).

I performed exact matching on the ex-ante risk of facing a disaster. The idea is that, once the ex-ante disaster risk is accounted for, the actual realization of natural disasters will be as-if random, and thus satisfy the strong ignorability assumption (Rosenbaum and Rubin 1983). As discussed in Section 6, the disaster risk measure takes five values from 0 (very low risk) to 4 (very high risk). Treated and control observations were also exactly matched on the year, in order to control for

the time trends affecting breach behavior and reputations. To improve balance, states were also matched on reputations and net FDI inflows, both lagged one year, based on how similar they are on the Mahalanobis distance metric and by imposing a 0.5 standard deviation caliper. The matching procedure excluded within-unit matches to prevent matching the same state in different time periods with itself. I perform optimal full matching of the treatment and control groups using the `optmatch` package in R (Hansen and Klopfer 2006; Hansen 2007). This procedure treats matching as an optimization problem balancing the requirement that there be dispersion in the treatment variable, and the need for uniformity in matched variables across treatment and control groups. This procedure optimally generates non-overlapping matched sets where a given treated unit can be matched with multiple control units or vice versa. I restricted the procedure to prevent assigning more than one treated unit to a single control unit.

Table 5 displays the standardized mean differences across treatment and control groups before and after matching, for the three different operationalizations of the treatment period. In all three cases, the matching procedure considerably increased balance. The χ^2 tests indicate that, before matching, there was considerable evidence against the null that the treatment and control groups come from the same distribution, but after matching, we are unable to reject this null. Moreover, matching yields balance across the treatment and control groups for each individual variable, for all three matching procedures.

The OLS regressions I estimate on the matched samples have the structure in Equation 11:

$$Breach_{it} = \pi_0 Reputation_{it-1} + \pi_1 Reputation_{it-1} \cdot Disaster_{it} + \pi_2 Disaster_{it} + \lambda_k + \epsilon_{it} \quad (11)$$

$Breach_{it}$ and $Reputation_{it-1}$ are the same as in the previous specifications. λ_k is the set of block fixed effects where k is the index for the matched sets. In each regression, I report heteroskedasticity-robust standard errors calculated using leverage-adjusted residuals (HC2). Again, the main parameter of interest for the hypothesis is the coefficient of the $Reputation_{it-1} \cdot Disaster_{it}$ interaction term, π_1 .

Table 6 reports the results from analyses after conducting matching. Each model reported in Table 6 pertain to a different treatment time interval: I take the treatment period following disas-

Table 5: Balance Statistics for Matching

| | 2-year treatment | | 3-year treatment | | 5-year treatment | |
|--------------------------------|------------------------|-------------------|------------------------|-------------------|------------------------|-------------------|
| | Before | After | Before | After | Before | After |
| Disaster Risk _t | -0.17*** (-3.48) | 0.0 (0.0) | -0.19*** (-3.56) | 0.0 (0.0) | -0.19*** (-3.40) | 0.0 (0.0) |
| Net FDI Inflows _{t-1} | -0.11** (-2.15) | -0.01 (-0.90) | -0.12** (-2.37) | -0.01 (0.50) | -0.14** (-2.42) | -0.01 (-0.60) |
| Reputation _{t-1} | -0.05 (-1.01) | 0.01 (0.95) | -0.10* (-1.96) | 0.01 (1.52) | -0.13** (-2.36) | 0.01 (1.34) |
| Year _t | 0.01 (0.27) | 0.0 (0.0) | 0.06 (1.21) | 0.0 (0.0) | 0.08 (1.39) | 0.0 (0.0) |
| Overall Test (χ^2) | 16.20 ₄ *** | 1.76 ₂ | 20.70 ₄ *** | 2.57 ₂ | 21.37 ₄ *** | 2.16 ₂ |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standardized mean differences (treatment minus control) are reported with associated z-scores in parentheses. Final row displays the results of χ^2 tests evaluating whether the two groups come from the same distribution, given the set of covariates. Subscript is the degrees of freedom.

Table 6: Reputation and Contract Breach following Disasters on Matched Samples

| | 2-year treatment | 3-year treatment | 5-year treatment |
|--------------------------------------|-------------------|-------------------|-------------------|
| Reputation _{t-1} | -0.845 (0.308) | -0.541 (0.440) | 0.424 (0.612) |
| Disaster | -0.066 (0.053) | -0.121 (0.069) | -0.236 (0.072) |
| Reputation _{t-1} × Disaster | 0.089 (0.057) | 0.164 (0.075) | 0.286 (0.081) |
| Effective Sample Size | 612 | 538 | 468 |
| Adjusted R ² | 0.164 | 0.162 | 0.172 |

Reported are OLS coefficients and HC2 standard errors in parentheses. All models include block fixed effects for matched sets.

ters to be +1 year for the first model, +2 in the second, and +4 in the third. The results provide strong evidence for the hypothesis. The coefficient of the interaction term, which represents the change in the relationship between breach behavior and reputations following disasters, is positive across all three models, and is statistically significant at $\alpha = .05$ level in the second and third columns. This means, consistent with the hypothesis, states with better reputations are more likely to breach contract following disasters compared to states with worse reputations.

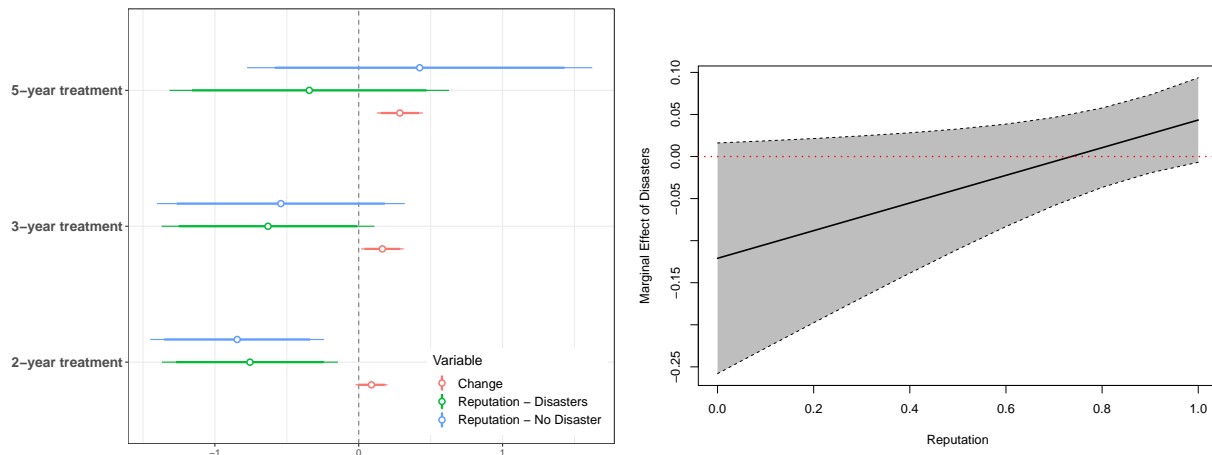


Figure 9: Left panel displays the relationship between reputation and contract breach by disaster occurrence based on Table 6. The bands represent 90% (thick line) and 95% (thin line) confidence intervals. The right panel is based on the 3-year treatment column of Table 6 and displays the marginal effect of disasters on the probability of contract breach by reputation level.

The left panel in Figure 9, which is based on Table 6, displays three effects: (i) the relationship between reputations and breach behavior during disaster years (green lines), (ii) the same relationship in non-disaster years, and (iii) the change between the two. The results indicate that reputations are negatively correlated with breach behavior in disaster years, although not statistically distinguishable from zero in 3- and 5-year treatment models. This relationship is more ambiguous in non-disaster years: better reputations are associated with a lower probability of breach in 2- and 3-year treatment models, but are associated with a higher probability of breach in the 5-year treatment model. Further in 3- and 5-year treatment models, this relationship is not statistically significant at conventional levels. Notwithstanding this ambiguous picture, the change we observe between the reputation-breach relationship following disasters, depicted by the red lines, is systematically positive across the models and statistically significant ($p < 0.05$).

The right panel in Figure 9 shows that the marginal effect of facing a disaster on the probability of breach is increasing in prior reputation level. In particular, it shows that states with poor prior reputations for investor-friendliness are less likely to breach contract after disasters, and states with good reputations for investor friendliness are more likely to breach contract after disasters. Substantively, facing a severe disaster is 16% more likely to lead to contract breach among countries with very good reputations ($Reputation_{t-1} = 1$) compared to countries with poor reputations ($Reputation_{t-1} = 0$).

8 Conclusion

How can states achieve cooperation in an anarchical environment? Perhaps the most common answer to this enduring IR problem have been that reputational incentives constrain states' temptations to lie, cheat, and renege on their promises, and thus bolster cooperation. However, we do not know how states' incentives to build and maintain their reputations are affected by their current reputations. Is a state with an excellent reputation or a very poor reputation have the greatest incentive to exert further effort for its reputation? To put it differently, we do not know how the extent to which reputational incentives constrain undesirable behavior depend on existing reputations. This is due to the field having a static perspective on reputations which takes states' preferences as fixed.

In this paper, I present a model of reputations where states' preferences and abilities are subject to change over time and across interactions, if slowly. In so doing, I show that reputations are a double-edged sword. On the one hand, reputational incentives are beneficial to all parties, because they make desirable cooperative outcomes possible. On the other hand, states with better reputations have diminished incentives to extend further effort. In other words, reputations primarily constrain those with poor reputations. The reason is simple: if states' preferences change, and everybody is aware of this fact, then things might be different *this* time. Two implications follow from situating reputational incentives within a changing preferences framework. One, there is a limit to how much benefit states can derive from improving their reputations, because the knowledge that preferences can change precludes extending complete trust. Two, rebuilding reputations after tarnishing them is possible, because the fact that preferences can change causes the audience to extend the benefit of doubt. Together, these reduce the incentives to extend further effort as reputations improve, because the return on that effort is expected to be lesser in magnitude the better one's reputation is.

I also test this key argument by concentrating on the laboratory of foreign direct investment and contract breach. A chief problem of foreign investors is trusting the host-governments that they will not renege on their promises once the investment is made. Hence, states who stand to gain from breaching contracts in the short-run face a commitment problem. They would be better off if they could commit to protecting the sanctity of their contracts with foreign firms, for this would facilitate valuable inward FDI flows, but investors' suspicions about states' short-run temptations make committing difficult. Then, states' concerns about their reputation in the eyes of future potential investors should help resolve this commitment problem. Given the theoretical argument advanced here, states who do stand to gain from breaching contracts in the short-run should be more likely to give in to their short-run temptation the better their reputations.

The core problem for testing the argument is that, like the investors, the researchers are also uncertain about which states stand to gain from breaching contracts and when. Moreover, states who face short-run temptations to engage in breach behavior build their reputations by manipulating their audience into thinking that they do not, in fact, face these temptations. In effect, what they are doing is to "hide" among states who do not face such temptations. Together, these mean that naive comparisons of behavior between states with good reputations and poor reputations

cannot recover the stipulated relationship in this paper.

I show in Section 5 that, adverse exogenous shocks increasing states' temptation to breach contract in the short-run provide a solution. States with good reputations for investor-friendliness should be more likely to engage in contract breach as a result of the shock, compared to states with poor reputations. I leverage the as-if random occurrence of severe natural disasters, which function as adverse exogenous shocks, and employ matching designs to provide evidence for the core argument. I show that, consistent with the expectations, states with good reputations ($Reputation_{t-1} = 1$) are 9 – 28% more likely to engage in contract breach following a severe natural disaster compared to states with poor reputations ($Reputation_{t-1} = 0$).¹⁷

If reputational incentives primarily constrain those with poor reputations, as is corroborated by the evidence, there are implications for all work which invoke reputational incentives to explain the existence of cooperative outcomes. This is true whenever commitment problems coexist with long-run relationships and stable actors, and thus these results are of relevance not just to IR scholars, but to political scientists in general. In IR, scholars have long considered facilitating the better functioning of the reputation mechanism as a primary function of international institutions, beside issue linkage. If reputational incentives fail to constrain those with good reputations, then a fruitful area of future work is to examine how institutions can be designed to optimally induce incentives to reduce such inefficiencies that I show are inherent in reputation mechanisms. Further, if those states with good reputations have the least incentive to further invest in them, it should also be the case that they have greater incentives to hide their undesirable behavior from their audiences. This parallels, for instance, findings in Kono (2006) where democracies, which otherwise have liberal trade policies, tend to erect non-tariff barriers to trade, because they are more difficult to detect. Similarly, Bazillier, Hatte, and Vauday (2017) show that multinational firms with well-established reputations for environmental responsibility are more likely to have locations in “dirty” countries with laxer environmental regulations. Future work can explore the extent to which states and other actors hide their undesirable behavior from their audiences, and the technologies they develop for this purpose.

17. According to the point estimates of the interaction terms reported in Table 6.

References

- Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91 (4): 938–963.
- Allee, Todd, and Clint Peinhardt. 2011. "Contingent Credibility: The Impact of Investment Treaty Violations on Foreign Direct Investment." *International Organization* 65 (3): 401–432.
- Alt, James E., Randall L. Calvert, and Brian D. Humes. 1988. "Reputation and Hegemonic Stability: A Game-Theoretic Analysis." *American Political Science Review* 82 (2): 445–466.
- Axelrod, Robert. 1981. "The Emergence of Cooperation among Egoists." *American Political Science Review* 75 (2): 306–318.
- Axelrod, Robert, and Robert Keohane. 1985. "Achieving Cooperation under Anarchy: Strategies and Institutions." *World Politics* 38 (1): 226–254.
- Bazillier, Rémi, Sophie Hatte, and Julien Vauday. 2017. "Are environmentally responsible firms less vulnerable when investing abroad? The role of reputation." *Journal of Comparative Economics* 45 (3): 520–543.
- Bhaskar, V., George J. Mailath, and Stephen Morris. 2013. "A Foundation for Markov Equilibria in Sequential Games with Finite Social Memory." *Review of Economic Studies* 80 (3): 925–948.
- Board, Simon, and Moritz Meyer-ter-Vehn. 2013. "Reputation for Quality." *Econometrica* 81 (6): 2381–2462.
- Bohren, J. Aislinn. 2013. "Stochastic Games in Continuous Time : Persistent Actions in Long-Run Relationships."
- Büthe, Tim, and Helen V. Milner. 2008. "The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements?" *American Journal of Political Science* 52 (4): 741–762.
- Caruso, Germán Daniel. 2017. "The Legacy of Natural Disasters: The Intergenerational Impact of 100 Years of Disasters in Latin America." *Journal of Development Economics* 127:209–233.
- Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *Review of Economics and Statistics* 95 (5): 1549–1561.
- Cavallo, Eduardo, and Ilan Noy. 2011. "Natural Disasters and the Economy — A Survey." *International Review of Environmental and Resource Economics* 5 (1): 63–102.
- Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 102 (2): 179–221.
- Cole, Harold L., and William B. English. 1991. "Expropriation and Direct Investment." *Journal of International Economics* 30 (3-4): 201–227.
- Crescenzi, Mark J. C. 2018. *Of Friends and Foes: Reputation and Learning in International Politics*. Oxford: Oxford University Press.
- Cripps, Martin, George J. Mailath, and Larry Samuelson. 2004. "Imperfect Monitoring and Impermanent Reputations." *Econometrica* 72 (2): 407–432.

- Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. "Reputation and Status as Motives for War." *Annual Review of Political Science* 17 (1): 371–393.
- Downs, George W., and Michael A. Jones. 2002. "Reputation, Compliance, and International Law." *Journal of Legal Studies* 31 (1): 95–114.
- Ekmekci, Mehmet, Olivier Gossner, and Andrea Wilson. 2012. "Impermanent types and permanent reputations." *Journal of Economic Theory* 147 (1): 162–178.
- Escaleras, Monica, and Charles A. Register. 2011. "Natural Disasters and Foreign Direct Investment." *Land Economics* 87 (2): 346–363.
- Faingold, Eduardo, and Yuliy Sannikov. 2011. "Reputation in Continuous-Time Games." *Econometrica* 79 (3): 773–876.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105 (10): 3150–3182.
- Frye, Timothy. 2002. "Private Protection in Russia and Poland." *American Political Science Review* 46 (3): 572–584.
- Fudenberg, Drew, and David K. Levine. 1989. "Reputation and Equilibrium Selection in Games with a Patient Player." *Econometrica* 57 (4): 759.
- Fudenberg, Drew, and Eric Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica* 54 (3): 533–554.
- Gigerenzer, Gerd, and Wolfgang Gaissmaier. 2011. "Heuristic Decision Making." *Annual Review of Psychology* 62 (1): 451–482.
- Graham, Benjamin A. T., Noel P. Johnston, and Allison F. Kingsley. 2018. "Even Constrained Governments Take." *Journal of Conflict Resolution* 62 (8): 1784–1813.
- Graham, Benjamin A. T., and Jacob R. Tucker. 2019. "The International Political Economy Data Resource." *Review of International Political Economy* 14 (1): 149–167.
- Haas, Ernst B. 1980. "Why Collaborate? Issue-Linkage and International Regimes." *World Politics* 32 (3): 357–405.
- Hansen, Ben B. 2007. "Optmatch: Flexible, Optimal Matching for Observational Studies." *R News* 7 (2): 18–24.
- Hansen, Ben B., and Stephanie Olsen Klopfer. 2006. "Optimal Full Matching and Related Designs via Network Flows." *Journal of Computational and Graphical Statistics* 15 (3): 609–627.
- Henisz, Witold J. 2000. "The Institutional Environment for Multinational Investment." *Journal of Law, Economics, and Organization* 16 (2): 334–364.
- Hopf, Ted. 1994. *Peripheral visions: Deterrence Theory and American Foreign Policy in the Third World, 1965-1990*. Ann Arbor: University of Michigan Press.
- Imbens, Guido W., and Michal Kolesar. 2016. "Robust Standard Errors in Small Samples: Some Practical Advice." *Review of Economics and Statistics* 98 (4): 701–712.
- Jensen, Nathan M, and Noel P Johnston. 2011. "Political Risk, Reputation, and the Resource Curse." *Comparative Political Studies* 44 (6): 662–688.

- Jensen, Nathan M. 2003. "Democratic Governance and Multinational Corporations: Political Regimes and Inflows of Foreign Direct Investment." *International Organization* 57 (3): 587–616.
- . 2008. *Nation-States and the Multinational Corporation: A Political Economy of Foreign Direct Investment*. Princeton: Princeton University Press.
- Jensen, Nathan M., Noel P. Johnston, Chia-yi Lee, and Hadi Sahin. 2019. "Crisis and Contract Breach: The Domestic and International Determinants of Expropriation." *The Review of International Organizations*, 1–30.
- Jervis, Robert. 1997. *System Effects: Complexity in Political and Social Life*. Princeton: Princeton University Press.
- Johns, Leslie, and Rachel L. Wellhausen. 2016. "Under One Roof: Supply Chains and the Protection of Foreign Investment." *American Political Science Review* 110 (1): 31–51.
- Johnston, Noel P, Nicole Janz, and Paasha Mahdavi. 2015. "How Taking from Foreigners Affects Domestic Human Rights." *Blavatnik School of Government Working Paper Series* 2015/004.
- Kono, Daniel Y. 2006. "Optimal Obfuscation: Democracy and Trade Policy Transparency." *American Political Science Review* 100 (3).
- Kreps, David M, and Robert Wilson. 1982. "Reputation and imperfect information." *Journal of Economic Theory* 27 (2): 253–279.
- Kuhn, Max. 2008. "Building Predictive Models in R Using the caret Package." *Journal of Statistical Software* 28 (5): 1–26.
- Levy, Brian, and Pablo T. Spiller. 1996. *Regulations, Institutions, and Commitment: Comparative Studies of Telecommunications*. Edited by Brian Levy and Pablo T. Spiller. New York: Cambridge University Press.
- Li, Quan. 2006. "Democracy, Autocracy, and Tax Incentives to Foreign Direct Investors: A Cross-National Analysis." *Journal of Politics* 68 (1): 62–74.
- . 2009. "Democracy, Autocracy, and Expropriation of Foreign Direct Investment." *Comparative Political Studies* 42 (8): 1098–1127.
- Li, Quan, and Adam Resnick. 2003. "Reversal of Fortunes: Democratic Institutions and Foreign Direct Investment Inflows to Developing Countries." *International Organization* 57 (1): 175–211.
- Liu, Qingmin. 2011. "Information acquisition and reputation dynamics." *Review of Economic Studies* 78 (4): 1400–1425.
- Liu, Qingmin, and Andrzej Skrzypacz. 2014. "Limited records and reputation bubbles." *Journal of Economic Theory* 151 (1): 2–29.
- Mailath, George J., and Stephen Morris. 2002. "Repeated Games with Almost-Public Monitoring." *Journal of Economic Theory* 102 (1): 189–228.
- . 2006. "Coordination Failure in Repeated Games with Almost-Public Monitoring." *Theoretical Economics* 1 (3): 311–340.
- Mailath, George J., and Larry Samuelson. 2015. "Reputations in Repeated Games." In *Handbook of Game Theory with Economic Applications*, 4:165–238. 1.

- McCaffrey, Daniel F., and Robert M. Bell. 2002. "Bias Reduction in Standard Errors for Linear and Generalized Linear Models with Multi-Stage Samples." *Survey Methodology* 28 (2): 169–181.
- Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York: W. W. Norton.
- Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca: Cornell University Press.
- . 2013. "Emotion and Strategy in the Korean War." *International Organization* 67 (2): 221–252.
- Noy, Ilan, Oscar Becerra, and Eduardo A Cavallo. 2012. "Foreign Aid in the Aftermath of Large Natural Disasters." *Inter-American Development Bank Working Paper* 333.
- Pandya, Sonal S. 2016. "Political Economy of Foreign Direct Investment: Globalized Production in the Twenty-First Century." *Annual Review of Political Science* 19 (1): 455–475.
- Phelan, Christopher. 2006. "Public trust and government betrayal." *Journal of Economic Theory* 130 (1): 27–43.
- Pinto, Pablo M., and Santiago M. Pinto. 2008. "The Politics of Investment Partisanship and the Sectoral Allocation of Foreign Direct Investment." *Economics & Politics* 20 (2): 216–254.
- Press, Daryl. 2005. *Calculating Credibility: How Leaders Assess Military Threats*. Ithaca: Cornell University Press.
- Pyle, William. 2011. "Organized Business, Political Competition, and Property Rights: Evidence from the Russian Federation." *Journal of Law, Economics, and Organization* 27 (1): 2–31.
- Ramsay, Kristopher W. 2011. "Revisiting the Resource Curse: Natural Disasters, the Price of Oil, and Democracy." *International Organization* 65 (3): 507–529.
- Rosenbaum, Paul R, and Donald B Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.
- Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton: Princeton University Press.
- Simmons, Beth A. 2000. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *American Political Science Review* 94 (04): 819–835.
- Thomas, Jonathan, and Tim Worrall. 1994. "Foreign Direct Investment and the Risk of Expropriation." *Review of Economic Studies* 61 (1): 81–108.
- Tomz, Michael. 2007. *Reputation and International Cooperation: Sovereign Debt across Three Centuries*. Princeton: Princeton University Press.
- Tomz, Michael, and Mark Wright. 2010. "Sovereign Theft: Theory and Evidence about Sovereign Default and Expropriation." In *The Natural Resources Trap: Private Investment without Public Commitment*, edited by W. Hogan and F. Sturzenegger, 69–110. Cambridge: MIT Press.
- Volkov, Vadim. 2002. *Violent Entrepreneurs: The Use of Force in the Making of Russian Capitalism*. Ithaca: Cornell University Press.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. New York: McGraw-Hill.
- Weinstein, Jonathan, and Muhamet Yildiz. 2013. "Robust Predictions in Infinite-Horizon Games—an Unrefinable Folk Theorem." *Review of Economic Studies* 80 (1): 365–394.

- Weisiger, Alex, and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69 (02): 473–495.
- Wellhausen, Rachel L. 2015. "Investor–State Disputes: When Can Governments Break Contracts?" *Journal of Conflict Resolution* 59 (2): 239–261.
- . 2016. "Recent Trends in Investor-State Dispute Settlement." *Journal of International Dispute Settlement* 7 (1): 117–135.
- Wiseman, Thomas. 2008. "Reputation and impermanent types." *Games and Economic Behavior* 62 (1): 190–210.