Political Risk Insurance: A New Firm-Level Dataset*

Vincent Arel-Bundock[†]

Clint Peinhardt[‡]

Amy Pond[§]

June 4, 2019

Abstract

When do governments impose costs on foreign firms? Many studies of foreign direct investment focus on incentives for government expropriation, but scholars are often forced to rely on indirect measures of expropriation to conduct empirical analyses. This paper introduces a dataset which includes information on over 5000 political risk insurance contracts issued by the U.S. Overseas Private Investment Corporation since 1961, and on all the claims filed by investors under these contracts. These detailed insurance data allow us to study the determinants of foreign investors' losses from a variety of sources, including expropriation, inconvertibility, and violent conflict. To illustrate the benefits of these data for hypothesis testing, we adopt a comprehensive empirical approach, and explore both shared and distinct causes across risk categories.

*The dataset described in this article can be downloaded from the authors' webpages: <http://arelbundock.com>
<http://people.tamu.edu/~apond/> <http://www.utdallas.edu/~clint.peinhardt/>

[†]vincent.arel-bundock@umontreal.ca

[‡]clint.peinhardt@utdallas.edu

[§]apond@tamu.edu

Private firms often suffer losses due to government action. Regulations can reduce profits as firms bring themselves into compliance; restrictions on currency convertibility can undermine the profitability of an investment for foreign shareholders; political conflict can spill onto the streets, preventing employees from going to work and delaying production. In the extreme, outright expropriation can undermine a firm’s ownership of its assets and the associated profits.

The possibility that government actions will reduce firm profits is one of the leading explanations for underinvestment in countries with weak property rights protection (Lucas, 1990; Li and Resnick, 2003; Jensen, 2003). Unfortunately, documenting firm losses in a systematic fashion has proven difficult for political-economists, because information about such events is not always made public. As a result, we know surprisingly little about the conditions that lead host governments to take actions that hurt foreign investors.

To fill that gap, we introduce a direct measure of firms’ overseas losses. Our new dataset includes information on over 5000 political risk insurance contracts issued by the U.S. Overseas Private Investment Corporation (OPIC), along with matching information on all the claims filed by investors under these contracts. In addition to insurance contracts and claims, the dataset includes information on over 1500 financing projects entered into by OPIC. These detailed firm-level data – which cover 162 host countries over the 1961-2017 period – allow us to study the determinants of foreign investors’ losses in a wide variety of political and economic settings, while holding constant the risk-mitigation strategies of those firms.

In the following, we briefly review prior works on the concept of political risk, and we outline the main measurement strategies used in this field. We discuss the market for political risk insurance and the benefits of firm-level insurance data for the study of political risk. We describe a new dataset documenting OPIC financing projects, insurance contracts, and insurance claims. Finally, we estimate several regression models drawing on extant theories about political risk. Our goal is not to produce precise estimates of causal effects. Rather, we treat this statistical exercise as a test of construct validity, which reinforces our belief that insurance data hold much potential for the study of foreign direct investment.

1 Political risk

Foreign investment flows have the potential to increase growth and efficiency in poor countries and to equalize factor returns around the world. Nevertheless, large disparities in factor returns remain, and political risk helps explain their persistence. Even if the returns to investment are high, foreign investors may not enter a market if they cannot be sure that their property rights will be protected (Lucas, 1990; Caselli and Feyrer, 2007).

The property rights of foreign investors are central to the literatures on international investment agreements, domestic institutions, and cross-border supply chains. International investment agreements often codify the rights of foreign investors and specify dispute resolution mechanisms that supersede domestic courts, in which foreign investors may not get a fair hearing (Elkins, Guzman and Simmons, 2006; Kerner, 2009). The costs of arbitration and future underinvestment add to the credibility of investment treaties (Allee and Peinhardt, 2011; Desai and Moel, 2008; Pelc, 2017). When agreements are absent, foreign investors may lean on political constraints, like veto players and democratic institutions, to guarantee their property rights (Jensen, 2003; Li and Resnick, 2003).¹ Multinational corporations may strategically place different components of production in ways that maximize leverage over host country governments (Henisz, 2000; Johns and Wellhausen, 2016).

¹Recent work has questioned the importance of political institutions (Arel-Bundock, 2017).

These works have improved our understanding of the political-economy of foreign direct investment. Still, measurement remains a major stumbling block on the path to scientific progress in this field. In the next section, we consider the measurement strategies adopted in prior works, and discuss how firm-level insurance data can complement and improve on these strategies.

1.1 Measurement strategies

To understand how our new data can help researchers, we briefly survey the most common measurement strategies in the field.

Aggregate flows. Many studies of political risk use aggregate FDI data to *indirectly* assess their theories. For example, several authors have argued that democratic institutions foster strong property rights protection which, in turn, lowers political risk and attracts investment (Jensen, 2003; Li and Resnick, 2003). These authors look for evidence in support of their arguments by regressing aggregate FDI flows on a measure of democracy (or some associated concept).²

These analyses are extremely useful, as they relate two variables of great interest. However, focusing on macro-level quantities such as FDI flows also black boxes the causal mechanisms at work in those theories, namely that (a) institutions reduce risk, and (b) that this reduction in risk attracts new investments. Our new firm-level data can thus be a useful complement to this approach, since it allows us to *directly* measure the losses that accrue to foreign firms in different countries.

Catalog of expropriation events. Another common strategy is to build catalogs of expropriation events, drawing on information from news reports, trade publications, etc. Following the path-breaking work of Kobrin (1984), several authors have used such datasets to study the political factors that induce governments to seize foreign assets (Minor, 1994; Li, 2009; Guriev, Kolotilin and Sonin, 2011; Hajzler, 2012; Biglaiser, Lee and Staats, 2017). These studies yield valuable insights, but the data they rely on is limited in some respects.

First, the temporal and geographic scope of the available data is often limited. Second, since data collection on individual events worldwide is very costly, researchers may not always be able to identify the full population of relevant events. This could potentially introduce biases in the statistical analysis. Third, existing data on firm losses often track expropriations, but ignore less direct forms of firm losses. Recent studies recognize that other types of loss have become increasingly important as governments abstain from expropriation; they instead emphasize currency convertibility for profit repatriation, violent conflict, and costly regulation that reduces firm profits (Graham, Johnston and Kingsley, 2016, 2018; Pelc, 2017).

Our dataset improves on the “catalog of events” strategy by covering a large temporal and geographic scope, using a well-defined universe of firms, and a broader array of event types.

International investment disputes. Disputes initiated under investor-state arbitration are also useful as a direct measure of firm losses (Allee and Peinhardt, 2011; Freeman, 2013; Wellhausen, 2015b; Johns and Wellhausen, 2016; Kim, 2017). Unfortunately, we know that investment disputes are often kept private and that many firms do not have access to arbitration, since the cost of initiating such disputes is high, and because investment arbitration is often the last step in an escalating conflict between firms and host governments. As a result, it is not clear if investment disputes are

²Li, Owen and Mitchell (2018) recently performed a meta analysis to identify common trends and biases in the regression results of democracy on FDI from over 200 estimates drawn from 39 published empirical studies.

representative of the population of interest. Our dataset improves by precisely identifying the full population of firms with an OPIC insurance contract, as well as the full population of firms that file claims.

Expert perceptions and insurance premiums. Scholars have also used country ratings, published by political risk insurance providers and expert agencies, to infer the occurrence of firm losses: where premiums are highest, firms must suffer large losses (Jensen, 2008; Jensen and Young, 2008; Jensen, 2012; Wilson and Wright, 2017). These ratings are direct measures of the *ex ante* perception of the risk to which firms are exposed. However, they fail to measure actual firm losses, and their use as a proxy risks backward engineering insurers’ methods for calculating risk, rather than locating the sources of firm loss. The OPIC data that we introduce here allow us to directly measure firm losses.

1.2 Kobrin’s challenge: Political events vs. firm losses

The empirical strategies described above have allowed political economists to make important contributions to the study of political risk. Our firm-level insurance data will allow researchers to build and improve on those contributions.

Over 30 years ago, Kobrin (1979) criticized the concept of “political risk” for its polysemic nature: it can refer to government actions that reduce the profitability of an investment (*political events*), or to the unwanted consequences of political activities (*firm losses*). Although closely related, conflating these two concepts is a mistake, and using a single measure to represent both raises serious conceptual and empirical issues. Kobrin’s solution to this problem was to distinguish the business consequences of host state actions from the political events that drive those actions, and to study political risk using data on expropriations (Kobrin, 1984). In the same spirit, we introduce firm-level insurance data below to empower researchers to disentangle the adverse events that firms experience from the background political conditions that cause those events.

2 Political Risk Insurance: Firm-level Data

As the previous section makes clear, empiricists have been extremely creative in finding measures to study the problem of political risk. In this paper, we argue that a focus on firm-level insurance claims data can complement and improve upon prior approaches. Now, we briefly discuss the market for political risk insurance, before describing the organization which is the source of our new data.

2.1 The Overseas Private Investment Corporation

Foreign investors can purchase political risk insurance (PRI) from many providers. Dozens of private carriers offer insurance and reinsurance,³ and many countries offer public PRI through their export credit agencies or separate entities. Indeed, almost all OECD members and many emerging markets offer some form of PRI to firms that meet citizenship requirements. Moreover, the World Bank’s Multilateral Investment Guarantee Agency has since the mid-1980s offered PRI to firms that cannot obtain it from their home government.

The American government was the first public entity to issue PRI, originally as part of the Marshall Plan (Whitman, 1965; Lipson, 1985). PRI services were initially offered by the Export-Import Bank of Washington, and then by the U.S. Agency for International Development (USAID). In 1971, OPIC was created to take over those programs. It provides financing, guarantees, and

³For a discussion of private insurance providers see Jensen 2012, 14.

political risk insurance to American firms who want to invest in emerging markets. OPIC operates on a self-sustaining basis, by charging for its services. Since its inception, the organization has supported over \$200 billion of investment in more than 150 countries, and it has regularly deposited profits in the U.S. treasury.⁴

Since OPIC, USAID, and the Export-Import Bank are public entities, they are subject to requests under the Freedom of Information Act. As a result, we were able to obtain firm-level data on over 1500 financing projects and 5000 insurance contracts issued between 1961 and 2017. Then, we used OPIC’s annual reports to identify every claim for which a monetary settlement occurred since 1966,⁵ and we matched each claim to the original insurance contract manually, using background information from various sources.⁶ Our dataset also includes information on the sector of each project, as well as dollar amounts for the level of insurance guarantee or financing. To our knowledge, this paper is the first to study the full sample of OPIC insurance contracts and claims at the investor-country-year level.⁷ These data offer an unparalleled look into the determinants of politically-related losses experienced by firms.

2.2 Insurance claims and firm losses

OPIC currently covers five categories of risk: (i) inconvertibility of funds; (ii) expropriation; (iii) losses caused by material changes in project agreements unilaterally imposed by the host state; (iv) political violence; and (v) terrorism.⁸ We code dummy variables to identify the category in which insurance claims are filed. In the analysis that follows, we combine categories (ii) and (iii), as well as categories (iv) and (v), in order to produce three claims categories: inconvertibility, expropriation, and political violence. We do this because claims are relatively rare events, and because these groupings are conceptually coherent. By extending the scope of measurement beyond expropriation, this insurance data opens up many fruitful avenues for future research.

In this paper, we treat PRI claims as a proxy for firm losses. However, it is important to acknowledge that this interpretation is somewhat imprecise. Indeed, many of the OPIC-covered companies who suffer losses due to political events eventually receive compensation under their insurance contracts.⁹ Nevertheless, we stress that the adverse events that our data document are

⁴For a historical overview of OPIC, see Peinhardt and Allee (2016).

⁵One difficulty arises because the annual claims report lists settlements under the fiscal year when they were issued. However, because there can be processing delays, the fiscal year of settlement does not typically correspond to the year during which the actual event took place. We found the year of event on a case-by-case basis using the claims determinations documents published in Kantor, Nolan and Sauviant (2011). We were only able to find detailed accounts of the events in question for about half of the claims. For the rest, we lag the fiscal year by one and use this lag as the event year. One is the modal and median number of years between events and the fiscal year in the subset of claims for which we found detailed accounts.

⁶A primary resource was Kantor, Nolan and Sauviant (2011), but in many cases companies changed names or went through mergers and acquisitions between the issuance of a contract and the filing of a claim. We thus had to conduct several web searches and we visited many company websites to track down those changes.

⁷In some versions of an unpublished working paper, Jensen et al. (2012) consider a subset of 23 expropriation claims. We extend this sample considerably. The main practical difficulty that we encountered in building our dataset is that a very large share of the claims settled by OPIC were filed against contracts issued when OPIC activities were under the purview of USAID or the Export-Import Bank. We filed separate Freedom of Information Act requests with those organizations, and obtained several lists of contracts in PDF format. These lists revealed the existence of over 2000 previously unobservable contracts, for a total of more than 5000. Extending the dataset in this way allowed us to match a far greater number of claims to the investor and year when contracts were issued. Instead of only 23 expropriation events, our dataset thus includes 60 expropriations, 98 inconvertibility claims, and 48 claims related to political violence.

⁸Firms who are worried about a particular kind of risk can tailor their insurance coverage accordingly. OPIC categories have evolved over time; terrorism insurance, for example, was added in 2003.

⁹It is important to note that typical catalogs of expropriations suffer from a similar problem, since they do not

real, and that the losses are often felt for a period of months or years before insurance payouts. According to Peinhardt and Allee (2016, 216), the average time to resolution for OPIC claims is over 700 days, a period during which firms experience considerable risk and uncertainty. Although no measurement strategy is perfect, we believe that insurance data are one of the best (and only) ways to document the adverse events experienced by a well-defined universe of “at-risk” firms, operating in a broad range of countries, over a long period of time.

The occurrence of such adverse events is closely related to the concept of “political risk,” but it is not identical. Political risk is the likelihood that a government will implement policies which are costly for foreign investors; this concept is prospective in the sense that it captures whether a firm anticipates losses due to government actions in the future. In contrast, our measure retrospectively captures firm-identified incidents where actual losses occurred.¹⁰

2.3 Are OPIC data representative?

One of the biggest challenges in the empirical study of political risk is selection bias. For instance, empirical work using data on ICSID claims often ignores two major sources of bias. First, access to ICSID is typically limited to deep-pocketed firms from countries whose governments have an investment agreement with the host.¹¹ Second, ICSID reporting is notoriously opaque, since governments may opt to keep cases confidential. Similarly, catalogs of expropriation events can suffer from important blind spots because of their limited geographic and temporal scope, and because they tend to be “compiled from secondary sources and [to] cover nationalizations that have received the most public attention” (Biglaiser, Lee and Staats, 2017, 318).

In contrast, the dataset that we introduce covers a well-defined universe of firms, and it records the full set of insurance claims settled under those contracts. The geographic scope of the dataset is vast and the time coverage extensive. Moreover, PRI settlements are generally less controversial and smaller than awards from other investor-state disputes. When OPIC claims are settled, firms receive around 90% of total claims on average, and over half of awards total less than \$400,000 (in 2005 USD, authors’ calculations). Firms thus appear to use PRI to manage smaller downside risks than the extraordinary adverse events that trigger ISDS. It thus seems reasonable to treat the observations in our data as representing a more “normal” set of assets and risks.¹²

At the same time, OPIC insurance data raise selection problems of their own. By definition, our dataset can only include U.S. firms that have purchased OPIC insurance or financing. Risk insurance from the U.S. government may cause domestic governments to calculate costs carefully before taking action that could result in a claim,¹³ since insurance settlements often involve the transfer of assets from the covered firm to the U.S. government. More generally, PRI covers a relatively small fraction of total FDI,¹⁴ and it is more likely to be purchased by firms who are more at risk than is typical. Thus, empirical findings about the determinants of OPIC claims may not be generalizable to the full population of foreign investors.

distinguish between events for which investors were compensated and events for which no compensation was received (Hajzler, 2012).

¹⁰The predicted values of logistic regression models estimated using claims as a dependent variable could potentially be interpreted as a measure of political risk.

¹¹Those agreements must also delegate arbitration to ICSID.

¹²Another distinction is that PRI requires firms to have the foresight to insure themselves prior to experiencing loss, whereas ISDS can be accessed after a loss occurs, regardless of the firm’s risk management strategy.

¹³The deterrence imposed by the need to compensate or fear of retaliation is often called ‘regulatory chill’, and bias induced by it would also be present for ISDS claims (Pelc, 2017; Betz and Pond, 2019).

¹⁴The importance of PRI fluctuates considerably from year to year, but the ratio of PRI coverage to FDI has been estimated in the 10-20% range World Bank (2011).

Unfortunately, since our firm-level dataset is the first of its kind, and since we have limited information about the firms themselves, we do not have good external benchmarks against which to measure the representativeness of the data we observe.¹⁵ As such, we encourage those who use this dataset to be cautious in their interpretation. One reasonable approach would be for researchers to make modest claims to generalizability, by limiting the scope of inference to the subset of firms who are likely candidates to buy PRI. A more ambitious approach would be to claim that conclusions extend beyond this limited subset. However, this would require making the untestable assumption that firm characteristics — beyond sector and project size — are conditionally independent from the probability of investing or suffering a loss.

In the online appendix, we use directed acyclic graphs to build intuition about the potential selection problems that researchers face when using OPIC data. In the next section, we use regression analysis to assess the construct validity of our measure (Carmines and Zeller, 1979, Chapt. 2). We show that several observable patterns in our data accord with core theoretical expectations from prior works on political risk. This exercise increases our confidence in the validity of the measure.

3 Descriptive statistics

Before moving to regression analyses, we offer a descriptive portrait of the dataset. Figure 1 shows the evolution of OPIC financing projects, insurance contracts, and claims over time. In the left column, we see yearly total amounts; in the right columns we see the number of observations. There are three types of events. In the first row, we see the insurance claims that OPIC settled. In the second row, we see the projects that OPIC financed. In the last row, we see the projects that OPIC insured.

Several points are noteworthy. First, OPIC’s financial assistance program has increased in size over time, while the insurance program peaked in the early part of its existence under USAID. Second, insurance settlements are often concentrated in time, with many happening during the 1980s.

Figure 2 shows the geographic distribution of claims, finance projects, and insurance contracts. This graph shows that the largest share of OPIC-backed projects are in Latin America. Another interesting pattern to note is the large number of claims emerging from Sub-Saharan Africa relative to the number of insurance contracts issued to cover projects in that region.

Figure 3 shows the distribution of projects and claims across sectors of activity. These sector categories follow the ones recorded by OPIC. In terms of the number of insurance contracts issued, manufacturing is by far the most important category.

4 Regression analysis

In this section, we test the construct validity of our data by estimating a series of regression models. Following Carmines and Zeller (1979), we derive theoretical expectations for how our new measure of firm losses relates to political and economic conditions. We then assess these expectations empirically. If the results are consistent with our expectations, it increases our confidence in the validity of the measure (Carmines and Zeller, 1979, 27). We build these empirical models using key intuitions from the literature on political risk. To be clear, our goal is not to produce credible

¹⁵Matching these firms to databases of publicly-listed firms, like COMPUSTAT or in the US Orbis, would exclude privately-listed firms. In Europe for example Orbis has broader private-firm coverage. Recent work has demonstrated that use of only publicly-listed firms biases empirical evidence (Dinlersoz et al., 2018).

Figure 1: OPIC projects and claims over time.

Figure 2: OPIC projects and claims across regions.

Figure 3: OPIC projects and claims across sectors.

causal identification. Rather, we want to check if the events recorded by OPIC are consistent with baseline intuition.

4.1 Theoretical expectations

The vast literature on foreign direct investment and political risk is fertile, and it offers many theories that help us develop expectations and hypotheses about the behavior of firms and host governments.

Political constraints make a government’s commitment to investors more credible and can reduce the probability that investors experience losses. Early work anticipated a ‘democratic advantage’ in attracting investment (Jensen, 2003), due to strong rule of law (Li and Resnick, 2003). Recently, scholars have concluded that this advantage likely stems from strong legal institutions (Biglaiser and Staats, 2012; Biglaiser, Lee and Staats, 2017) and constraints on executive power in democracies, which enable policymakers to make credible commitments to investors (Jensen, 2008; Li, 2009). Credible commitments reduce the likelihood that a government will expropriate assets or impose limits on currency convertibility, which can be a form of ‘creeping expropriation’ (Graham, Johnston and Kingsley, 2018). Political constraints also prevent power grabs and provide clear expectations for succession (Olson, 1993), making armed-political conflict less likely. We thus expect political constraints to reduce insurance claims due to expropriation, inconvertibility, and conflict.

Coalition turnover should increase firm losses. When a new ruling coalition seizes power, it often brings with it new priorities and can push leaders to renege on prior commitments (Leeds, Mattes and Vogel, 2009; Albertus and Menaldo, 2012). A new coalition may also be targeted by violent challengers, as it has not yet consolidated rule. Although these expectations may be tempered by a new government’s desire to build a reputation for property rights enforcement, we follow the extant literature in anticipating a positive association between coalition turnover and firm losses due to expropriation, inconvertibility, and conflict.

Capital intensive firms should be more likely to suffer losses, as capital intensive investments are less mobile. Firm owners are unable to move immobile investments abroad or withhold them from government extraction (Vernon, 1971; Frieden, 1994; Kerner and Lawrence, 2012; Kerner, 2014). Capital intensity then should be related to losses caused by expropriation and by inconvertibility.

The *Resource dependence* of a country should likewise increase firm losses from expropriation and inconvertibility. Even once the effect of capital intensity is accounted for, resource dependence may increase losses, as resource-related production is prone to populist appeals (Albornoz, Galiani and Heymann, 2012). Resource-dependent governments may also be less sensitive to reputation costs, as they have resource wealth to fall back on (Jensen and Johnston, 2011), and they may fail to make strong property rights commitments (Ross, 1999, 2012).

Economic crises have ambiguous effects on the likelihood that firms will experience losses. On the one hand, policymakers who face an economic crisis could place restrictions on convertibility to prevent capital flight. Crises can also reduce government receipts, and foreign investors can be attractive targets for revenue generation (Wellhausen, 2015a). On the other hand, Abiad and Mody (2005) find that crises are often connected to liberalization, and Jensen et al. (2012) argue that economic crises can make leaders more responsive to reputation concerns, which would make them less likely to expropriate.

Capital account openness should decrease firm losses. Inconvertibility claims document an investor’s inability to move capital abroad. Because measures of capital account openness report the absence of restrictions on convertibility (Quinn, 1997; Chinn and Ito, 2008), inconvertibility claims are less likely when capital accounts are open. Capital account openness could also reduce firm losses indirectly. Because capital account openness facilitates capital flight, the costs of expropria-

tion are higher when markets are open (Strange, 1996; Garrett, 1998; Pond, 2018). Capital account openness should reduce claims due to expropriation and inconvertibility.

The above hypotheses highlight one of the major benefits of our firm data: They allow us to test theories that relate to both national and firm-level factors, and we can disaggregate effects depending on claim type.

4.2 Model specification

To check if the above expectations are consistent with the OPIC data, we estimate several variations on this regression model:

$$\begin{aligned}\Phi(\text{Claim}_{icy}) = & \beta_1 \text{Political constraints}_{cy} + \beta_2 \text{Coalition turnover}_{cy} + \\ & \beta_3 \text{Capital-intensive}_i + \beta_4 \text{KA openness}_{cy} + \\ & \beta_6 \text{Economic crisis}_{cy} + \beta_7 \text{Conflict intensity}_{cy} + \\ & \beta_8 \text{Distance to US}_c + \beta_9 \ln \text{GDP}_{cy} + \\ & \beta_{10} \text{Resource rent}_{cy} + \beta_{11} \text{Contract duration} + \alpha + \varepsilon_i,\end{aligned}$$

where i is an investor index, c a country index, and y a year index. Claim_{icy} is the number of insurance claims that firm i files under OPIC insurance for adverse events in country c during year y . Political constraints is an index which measures the "feasibility of policy change by the host-country government" (Henisz, 2002).¹⁶ Coalition turnover is a binary variable which is zero when there is no change in the leader's support coalition, and 1 otherwise (Mattes, Leeds and Matsumura, 2016).¹⁷ KA openness is a measure of capital account openness from Karcher and Steinberg (2013). Economic crisis is a binary variable which equals 1 during years of crisis (Reinhart and Rogoff, 2009). Conflict intensity is the highest level of armed conflict intensity over the last 5 years (Themnér and Wallenstein, 2011). Distance is the population-weighted distance between the host country and the United States (Mayer and Zignago, 2006). GDP is from the World Bank (2017). Resource rent is total natural resources rents as a % of GDP (World Bank, 2017). Finally, Contract duration is the number of years elapsed since the insurance contract was issued, α is an intercept, and ε a disturbance term.

Since firm losses are relatively rare events, we use multiple imputation with the **Amelia** software ($m = 10$) to preserve all observations. The baseline results that we present below were obtained using negative binomial regression because some firms file multiple claims, which means that the dependent variable can be treated as count data. In the online appendix, we present several alternative specifications using logit and cox proportional hazard models. We also present results from models that substitute the Political constraints index for measures of Polyarchy or Liberal democracy (V-Dem), Judicial independence (Linzer and Staton, 2015), Political checks (Beck et al., 2001), and Constraints on the executive (Polity IV).

¹⁶We use Henisz' revised polcon3 index.

¹⁷We take the 3-year moving average of this dummy variable to account for the fact that changes in support coalitions can have effects in the periods immediately before (because of anticipations) and after the actual transition. This is also a useful way to transform the independent variable, because it lets us account for the fact that our data on the timing of OPIC claims are somewhat imprecise (see footnote 5).

4.3 Regression results

Table 1 presents the results of this exercise. Column 1 shows results for all three types of claims combined. The remaining columns disaggregate the results by claim type. Column 2 reports results for expropriation claims, column 3 for inconvertibility, and column 4 for violence.

Table 1: Negative Binomial regression models with multiple imputation.

	All	Expropriation	Inconvertibility	Violence
Political constraints	-3.27*** (0.82)	-2.15** (1.03)	-3.58*** (1.25)	-4.45*** (0.81)
Coalition turnover	2.25** (0.91)	2.84** (1.36)	1.36 (1.15)	3.29*** (0.83)
Capital-intensive	0.33 (0.29)	0.20 (0.26)	0.31 (0.35)	0.41 (0.55)
Resource rent	0.18 (0.13)	0.49** (0.21)	0.05 (0.19)	0.18 (0.20)
GDP	-0.34*** (0.08)	-0.02 (0.13)	-0.47*** (0.10)	-0.44*** (0.17)
Conflict	1.21*** (0.24)	1.21*** (0.46)	0.86** (0.40)	2.08*** (0.43)
KA openness	-0.62*** (0.11)	-0.49** (0.23)	-1.13*** (0.24)	-0.10 (0.13)
Economic crisis	0.67 (0.44)	0.11 (0.60)	1.35** (0.55)	-0.20 (0.64)
Distance	-0.07* (0.04)	-0.01 (0.05)	-0.04 (0.06)	-0.17** (0.08)
Contract duration	-0.12*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.12*** (0.04)
Constant	2.92 (1.92)	-7.54** (3.43)	4.66* (2.39)	4.78 (3.75)
Observations	82536	82536	82536	82536

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

As expected, some risk sources are shared. Claims are less common in countries with constrained governments: The Political constraints coefficients are negative and statistically significant across all the models. Changes to leadership increase the incidence of claims: Coalition turnover is positively associated with the likelihood that firms file claims of all types. GDP reduces the likelihood of claims.¹⁸ Claims of all types are more common in countries experiencing violent Conflict. The coefficient associated with Capital-intensity is of the expected sign but statistically insignificant.

Other risk sources are confined to specific claim types. Resource rents are positively associated with expropriation claims. Economic crisis is associated with inconvertibility claims but not with expropriation claims. These results seem inconsistent with Abiad and Mody (2005) and Jensen et al. (2012), who emphasize the liberalizing and constraining potential of economic crisis, and with

¹⁸Coalition turnover lacks significance in the regression on inconvertibility, and GDP lacks significance in the expropriation regression.

Wellhausen (2015*b*), who discusses the revenue benefits of expropriation. The results are broadly consistent with the use of capital controls to prevent capital flight during crisis.

As expected, we see fewer insurance claims related to inconvertibility or expropriation in countries with high levels of capital account openness. The negative association with inconvertibility is unsurprising, given that many of the legal instruments that count as capital controls are designed explicitly to restrict convertibility. The negative correlation with expropriation claims could reflect the idea that capital account openness increases the costs that host countries pay to expropriate, by facilitating capital flight and firm exit.

In short, the results described above are consistent with many of the core intuitions developed in prior works on political risk. They also suggest that even if some factors make all types of claims more likely, other factors are associated with only some claim types. Researchers should take seriously the distinction between risk types when deciding on the appropriate empirical strategy to test their theories.

5 Conclusion

In this paper, we introduce rich new firm-level data that measure multiple types of firm losses. Unlike most data on political risk, we document a complete population of potential risk events with a collection of over 5000 political risk insurance contracts and 1500 finance projects from 1961 to 2017. We also include information on every claim filed against those insurance contracts, and we augment the data by manually identifying capital-intensive firms.

The dataset, which we make available freely online, allows researchers to circumvent typical concerns about selection bias, because the data include information on the full universe of “at-risk” assets, including all the cases where claims were filed, but also all cases where no claims were filed. Because it includes data on over 6500 projects, our dataset also offers rich possibilities for future research on the determinants of government support for outgoing investment by a major government agency.

Many studies draw on aggregate FDI flows or political risk insurance ratings to indirectly measure political risk. The OPIC dataset offers an unprecedented opportunity to directly test old and new explanations of political risk and firm losses over a substantial time period. The data allowed us to distinguish between risk factors at the national level, such as resource dependence and political institutions, and risk factors at the firm level, such as vulnerability due to capital intensity.¹⁹ The data could also be used to study the incidence of violence and inconvertibility, as newly identified categories of firm loss (Graham, Johnston and Kingsley, 2016, 2018); to assess the determinants of government support for investment outflows (Luong and Sierra, 2015), as OPIC projects are targeted at specific firms; to identify the importance of sub-national actors in causing firm losses;²⁰ or to investigate how the threats to firm profits have changed over time.

The dataset also provides a definitive historical archive to learn about the activities of OPIC, the first public provider of political risk insurance. Indeed, OPIC’s future is likely to differ systematically from its past, given the recent passage of the BUILD Act, which consolidates OPIC into a new entity, the U.S. International Development Finance Corporation.²¹ The new agency will have a larger budget and will be able to take equity stakes in projects, something that was previously forbidden to OPIC.

¹⁹Firm-level data analysis is a promising avenue for future research (see also Weymouth, 2012; Jensen, 2013), which allows for the development and assessment of more fine-grained theories.

²⁰For example, the Sandinista revolt in Nicaragua in July 1979 resulted in multiple OPIC awards.

²¹<https://www.reuters.com/article/us-usa-congress-development/congress-eying-china-votes-to-overhaul-development>.

References

- Abiad, Abdul and Ashoka Mody. 2005. "Financial Reform: What Shakes It? What Shapes It?" *American Economic Review* 95(1):66–88.
- Albertus, Michael and Victor Menaldo. 2012. "If You're Against Them You're With Us The Effect of Expropriation on Autocratic Survival." *Comparative Political Studies* 45(8):973–1003.
- Alborno, Facundo, Sebastian Galiani and Daniel Heymann. 2012. "Foreign investment and expropriation under oligarchy and democracy." *Economics & Politics* 24(1):24–46.
URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0343.2011.00391.x/abstract>
- Allee, Todd and Clint Peinhardt. 2011. "Contingent Credibility: The Impact of Investment Treaty Violations on Foreign Direct Investment." *International Organization* 65(3):401–432.
- Arel-Bundock, Vincent. 2017. "The Political Determinants of Foreign Direct Investment: A Firm-Level Analysis." *International Interactions* 43(3):424–452.
- Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer and Patrick Walsh. 2001. "New tools in comparative political economy: The Database of Political Institutions." *World Bank Economic Review* 15(1):165–176.
- Betz, Timm and Amy Pond. 2019. "Foreign Financing and the International Sources of Property Rights." *World Politics* Forthcoming.
- Biglaiser, Glen, Hoon Lee and Joseph L. Staats. 2017. "The effects of political and legal constraints on expropriation in natural resource and manufacturing sectors." *International Area Studies Review* 20(4):311–333.
- Biglaiser, Glen and Joseph L. Staats. 2012. "Finding the "Democratic Advantage" in Sovereign Bond Ratings: The Importance of Strong Courts, Property Rights Protection, and the Rule of Law." *International Organization* 66:515–535.
- Carmines, Edward G. and Richard A. Zeller. 1979. *Quantitative Applications in the Social Sciences: Reliability and validity assessment*. Thousand Oaks, CA: SAGE Publications.
- Caselli, Francesco and James Feyrer. 2007. "The Marginal Product of Capital." *Quarterly Journal of Economics* 122(2):535–568.
- Chinn, Menzie D. and Hiro Ito. 2008. "A New Measure of Financial Openness." *Journal of Comparative Policy Analysis: Research and Practice* 10(3):309–322.
- Desai, Mihir A. and Alberto Moel. 2008. "Czech Mate: Expropriation and Investor Protection in a Converging World." *Review of Finance* 12(1):221–251.
- Dinlersoz, Emin, Sebnem Kalemli-Ozcan, Henry Hyatt and Veronika Penciakova. 2018. "Leverage over the Life Cycle and Implications for Firm Growth and Shock Responsiveness." *NBER Working Paper* No. 25226.
- Elkins, Zachary, Andrew T. Guzman and Beth A. Simmons. 2006. "Competing for Capital: The Diffusion of Bilateral Investment Treaties, 1960–2000." *International Organization* 60(4):811–846.

- Freeman, Nathan W. 2013. "Domestic Institutions, Capacity Limitations, and Compliance Costs: Host Country Determinants of Investment Treaty Arbitrations, 1987–2007." *International Interactions* 39(1):54–78.
- Frieden, Jeffrey A. 1994. "International investment and colonial control: a new interpretation." *International Organization* 48(4):559–593.
- Garrett, Geoffrey. 1998. "Global Markets and National Politics: Collision Course or Virtuous Circle?" *International Organization* 52(4):787–824.
- Graham, Benjamin A. T., Noel P. Johnston and Allison F. Kingsley. 2016. "A unified model of political risk." *Advances in Strategic Management* pp. 119–160.
- Graham, Benjamin A. T., Noel P. Johnston and Allison F. Kingsley. 2018. "Even Constrained Governments Take: The Domestic Politics of Transfer and Expropriation Risks." *Journal of Conflict Resolution* 0(0):0022002717701181.
URL: <https://doi.org/10.1177/0022002717701181>
- Guriev, Sergei, Anton Kolotilin and Konstantin Sonin. 2011. "Determinants of Nationalization in the Oil Sector: A Theory and Evidence from Panel Data." *Journal of Law, Economics and Organization* 27(2):301–323.
- Hajzler, Christopher. 2012. "Expropriation of foreign direct investments: sectoral patterns from 1993 to 2006." *Review of World Economics* 148(1):119–149.
- Henisz, Witold J. 2000. "The Institutional Environment for Multinational Investment." *Journal of Law, Economics and Organization* 16(2):334–364.
- Henisz, Witold J. 2002. "The political constraint index (POLCON) dataset."
URL: <https://mgmt.wharton.upenn.edu/faculty/heniszpolcon/polcondataset/>
- Jensen, Nathan. 2003. "Democratic Governance and Multinational Corporations: The Political Economy of Foreign Direct Investment." *International Organization* 57(3):587–616.
- Jensen, Nathan. 2008. "Political Risk, Democratic Institutions, and Foreign Direct Investment." *The Journal of Politics* 70(4):1040–1052.
- Jensen, Nathan M. 2012. "Measuring Risk: Political Risk Insurance Premiums and Domestic Political Institutions." *Working Paper* available at: <http://www.sscnet.ucla.edu/polisci/cpworkshop/papers/Jensen.pdf>.
- Jensen, Nathan M. 2013. "Domestic Institutions and the Taxing of Multinational Corporations." *International Studies Quarterly* 57(3):440–448.
- Jensen, Nathan M. and Daniel J. Young. 2008. "A Violent Future? Political Risk Insurance Markets and Violence Forecasts." *Journal of Conflict Resolution* 52(4):527–547.
- 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., Noel P. Johnston, Chia yi Lee and Abdulhadi Sahin. 2012. "Economic shock, political shifts, and sovereign theft: The domestic and international determinants of investment expropriation." *Working Paper, presented at the Annual Meeting of IPES*.

- 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.
- Kantor, Mark, Michael D. Nolan and Karl P. Sauvant. 2011. *Reports on Overseas Private Investment Corporation*. Oxford, U.K.: Oxford University Press.
- Karcher, Sebastian and David A. Steinberg. 2013. "Assessing the Causes of Capital Account Liberalization: How Measurement Matters." *International Studies Quarterly* 57:128–137.
- Kerner, Andrew. 2009. "Why Should I Believe You? The Costs and Consequences of Bilateral Investment Treaties." *International Studies Quarterly* 53:73–100.
- Kerner, Andrew. 2014. "What We Talk About When We Talk About Foreign Direct Investment." *International Studies Quarterly* 58:804–815.
- Kerner, Andrew and Jane Lawrence. 2012. "What's the Risk? Bilateral Investment Treaties, Political Risk and Fixed Capital Accumulation." *British Journal of Political Science* pp. 1–15.
- Kim, Yong Kyun. 2017. "States Sued: Democracy, the Rule of Law, and Investor-State Dispute Settlement (ISDS)." *International Interactions* 43(2):300–325.
- Kobrin, Stephen J. 1979. "Political Risk: A Review and Reconsideration." *Journal of International Business Studies* 10(1):67–80.
- Kobrin, Stephen J. 1984. "Expropriation as an Attempt to Control Foreign Firms in LDCs: Trends from 1960 to 1979." *International Studies Quarterly* 28(3):329–348.
- Leeds, Brett Ashley, Michaela Mattes and Jeremy S. Vogel. 2009. "Interests, Institutions, and the Reliability of International Commitments." *American Journal of Political Science* 53(2):461–476.
- Li, Quan. 2009. "Democracy, Autocracy, and Expropriation of Foreign Direct Investment." *Comparative Political Studies* 42(8):1089–1127.
- 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.
- Li, Quan, Erica Owen and Austin Mitchell. 2018. "Why Do Democracies Attract More or Less Foreign Direct Investment? A Meta-Regression Analysis." *International Studies Quarterly* 3(1):494–504.
- Linzer, Drew A and Jeffrey K Staton. 2015. "A global measure of judicial independence, 1948–2012." *Journal of Law and Courts* 3(2):223–256.
- Lipson, Charles. 1985. *Standing guard: Protecting foreign capital in the nineteenth and twentieth centuries*. University of California Press.
- Lucas, Robert E. 1990. "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review* 80(2):92–96.
- Luong, Pauline Jones and Jazmín Sierra. 2015. "The Domestic Political Conditions for International Economic Expansion: Lessons From Latin American National Oil Companies." *Comparative Political Studies* 48(14):2010–2043.

- Mattes, Michaela, Brett Ashley Leeds and Naoko Matsumura. 2016. "Measuring change in source of leader support: The CHISOLS dataset." *Journal of Peace Research* 53(2):259–267.
- Mayer, Thierry and Soledad Zignago. 2006. "GeoDist: the CEPII distances and geographical database." bibtex: mayer_geodist_2006.
- Minor, Michael S. 1994. "The Demise of Expropriation as an Instrument of LDC Policy, 1980-1992." *Journal of International Business Studies* 25(1):177–188.
URL: <http://www.jstor.org/stable/154857>
- Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87(3):567–576.
- Peinhardt, Clint and Todd Allee. 2016. "Political Risk Insurance as Dispute Resolution." *Journal of International Dispute Settlement* 7(1):205–224.
- Pelc, Krzysztof J. 2017. "What Explains the Low Success Rate of Investor-State Disputes?" *International Organization* 71:559–583.
- Pond, Amy. 2018. "Financial Liberalization: Stable Autocracies and Constrained Democracies." *Comparative Political Studies* 51(1):105–135.
- Quinn, Dennis P. 1997. "The Correlates of Change in International Financial Regulation." *American Political Science Review* 91:531–551.
- Reinhart, Carmen M. and Kenneth Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press.
- Ross, Michael L. 1999. "The Political Economy of the Resource Curse." *World Politics* 51:297–322.
- Ross, Michael L. 2012. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton, N.J.: Princeton University Press.
- Strange, Susan. 1996. *The Retreat of the State: The Diffusion of Power in the World Economy*. Cambridge, M.A.: Cambridge University Press.
- Themnér, Lotta and Peter Wallensteen. 2011. "Armed conflict, 1946–2010." *Journal of Peace Research* 48(4):525–536.
- Vernon, Raymond. 1971. *Sovereignty at Bay*. New York, N.Y.: Basic Books.
- Wellhausen, Rachel L. 2015a. "Bondholders v. Direct Investors? Competing Responses to Expropriation." *International Studies Quarterly* 59(4):750–764.
- Wellhausen, Rachel L. 2015b. "Investor-State Disputes: When Can Governments Break Contracts?" *Journal of Conflict Resolution* 59(2):239–261.
- Weymouth, Stephen. 2012. "Firm Lobbying and Influence in Developing Countries: A Multilevel Approach." *Business and Politics* 14(4):1–26.
- Whitman, Marina Von Neumann. 1965. *Government Risk-Sharing in Foreign Investment*. Princeton, NJ: Princeton University Press.
- Wilson, Matthew Charles and Joseph Wright. 2017. "Autocratic Legislatures and Expropriation Risk." *British Journal of Political Science* 47(1):1–17.

World Bank. 2011. “World Investment and Political Risk 2011.”. Washington DC: The World Bank.

World Bank. 2017. “World Development Indicators.”.

Online appendix

Negative binomial models

TABLE 1 – Negative Binomial regression models with multiple imputation (pol-con3 models).

	All	Expropriation	Inconvertibility	Violence
main				
Veto players	-3.27*** (0.82)	-2.15** (1.03)	-3.58*** (1.25)	-4.45*** (0.81)
Distance	-0.07* (0.04)	-0.01 (0.05)	-0.04 (0.06)	-0.17** (0.08)
Capital-intensive	0.33 (0.29)	0.20 (0.26)	0.31 (0.35)	0.41 (0.55)
Resource Rent	0.18 (0.13)	0.49** (0.21)	0.05 (0.19)	0.18 (0.20)
GDP	-0.34*** (0.08)	-0.02 (0.13)	-0.47*** (0.10)	-0.44*** (0.17)
Coal.Turnover	2.25** (0.91)	2.84** (1.36)	1.36 (1.15)	3.29*** (0.83)
Conflict	1.21*** (0.24)	1.21*** (0.46)	0.86** (0.40)	2.08*** (0.43)
KA Openness	-0.62*** (0.11)	-0.49** (0.23)	-1.13*** (0.24)	-0.10 (0.13)
Econ.Crisis	0.67 (0.44)	0.11 (0.60)	1.35** (0.55)	-0.20 (0.64)
Contract duration	-0.12*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.12*** (0.04)
Constant	2.92 (1.92)	-7.54** (3.43)	4.66* (2.39)	4.78 (3.75)
lnalpha				
Constant	3.80*** (0.34)	2.00 (2.24)	4.75*** (0.33)	3.67*** (1.17)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 2 – Negative Binomial regression models with multiple imputation (xconst models).

	All	Expropriation	Inconvertibility	Violence
main				
Executive constraints	-0.25*** (0.06)	-0.13 (0.08)	-0.35*** (0.10)	-0.21** (0.09)
Distance	-0.06 (0.04)	-0.01 (0.05)	-0.04 (0.06)	-0.15** (0.07)
Capital-intensive	0.31 (0.28)	0.18 (0.26)	0.27 (0.35)	0.36 (0.53)
Resource Rent	0.22 (0.13)	0.53** (0.23)	0.06 (0.18)	0.25 (0.21)
GDP	-0.35*** (0.08)	-0.04 (0.13)	-0.48*** (0.10)	-0.50*** (0.16)
Coal.Turnover	2.50*** (0.82)	2.85** (1.39)	1.92* (1.02)	3.47*** (0.74)
Conflict	1.27*** (0.25)	1.25*** (0.47)	0.95** (0.40)	2.14*** (0.38)
KA Openness	-0.64*** (0.12)	-0.50** (0.23)	-1.16*** (0.26)	-0.10 (0.14)
Econ.Crisis	0.65 (0.44)	0.14 (0.60)	1.35** (0.55)	-0.20 (0.64)
Contract duration	-0.12*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.13*** (0.03)
Constant	3.36* (1.88)	-7.24** (3.44)	5.26** (2.42)	5.96* (3.48)
lnalpha				
Constant	3.88*** (0.32)	2.18 (2.17)	4.77*** (0.29)	3.79*** (1.08)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 3 – Negative Binomial regression models with multiple imputation (polyarchy models).

	All	Expropriation	Inconvertibility	Violence
main				
Polyarchy	-2.19*** (0.67)	-0.78 (1.21)	-2.95*** (1.00)	-2.50** (0.98)
Distance	-0.06 (0.04)	-0.00 (0.05)	-0.03 (0.06)	-0.15* (0.08)
Capital-intensive	0.32 (0.28)	0.17 (0.26)	0.32 (0.35)	0.36 (0.54)
Resource Rent	0.24* (0.14)	0.55** (0.23)	0.13 (0.19)	0.25 (0.21)
GDP	-0.33*** (0.08)	-0.04 (0.14)	-0.43*** (0.11)	-0.46*** (0.16)
Coal.Turnover	2.35*** (0.83)	2.70** (1.30)	1.60 (1.06)	3.46*** (0.78)
Conflict	1.15*** (0.25)	1.21*** (0.46)	0.79* (0.42)	2.05*** (0.42)
KA Openness	-0.62*** (0.12)	-0.49** (0.23)	-1.11*** (0.24)	-0.07 (0.14)
Econ.Crisis	0.68 (0.43)	0.14 (0.58)	1.38** (0.52)	-0.21 (0.64)
Contract duration	-0.12*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.12*** (0.04)
Constant	2.60 (1.90)	-7.32** (3.53)	3.84 (2.56)	5.05 (3.55)
lnalpha				
Constant	3.94*** (0.33)	2.17 (2.31)	4.83*** (0.32)	3.89*** (1.11)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 4 – Negative Binomial regression models with multiple imputation (liberaldem models).

	All	Expropriation	Inconvertibility	Violence
main				
liberaldem	-2.83*** (0.86)	-0.78 (1.23)	-4.32*** (1.44)	-2.35* (1.36)
Distance	-0.05 (0.04)	-0.00 (0.05)	-0.03 (0.06)	-0.14* (0.08)
Capital-intensive	0.32 (0.28)	0.18 (0.26)	0.31 (0.35)	0.37 (0.53)
Resource Rent	0.24* (0.14)	0.55** (0.23)	0.12 (0.18)	0.26 (0.21)
GDP	-0.32*** (0.08)	-0.05 (0.13)	-0.41*** (0.11)	-0.48*** (0.17)
Coal.Turnover	2.40*** (0.81)	2.66** (1.29)	1.80* (1.03)	3.32*** (0.78)
Conflict	1.11*** (0.25)	1.19*** (0.45)	0.70* (0.41)	2.04*** (0.44)
KA Openness	-0.62*** (0.12)	-0.49** (0.23)	-1.11*** (0.24)	-0.07 (0.14)
Econ.Crisis	0.65 (0.43)	0.14 (0.59)	1.35** (0.51)	-0.24 (0.64)
Contract duration	-0.12*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.12*** (0.03)
Constant	2.23 (1.91)	-7.31** (3.51)	3.21 (2.50)	5.15 (3.63)
lnalpha				
Constant	3.91*** (0.32)	2.22 (2.25)	4.75*** (0.30)	3.89*** (1.12)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 5 – Negative Binomial regression models with multiple imputation (judicial_independence models).

	All	Expropriation	Inconvertibility	Violence
main				
Judicial independence	-3.45*** (0.73)	-2.43* (1.24)	-3.85*** (1.11)	-4.08*** (0.85)
Distance	-0.06* (0.04)	-0.01 (0.05)	-0.04 (0.06)	-0.15* (0.08)
Capital-intensive	0.28 (0.27)	0.16 (0.26)	0.27 (0.34)	0.35 (0.56)
Resource Rent	0.15 (0.12)	0.45** (0.20)	0.01 (0.17)	0.18 (0.20)
GDP	-0.30*** (0.09)	0.01 (0.14)	-0.42*** (0.10)	-0.40** (0.18)
Coal.Turnover	2.28*** (0.81)	2.99** (1.42)	1.26 (0.98)	3.62*** (0.82)
Conflict	1.22*** (0.26)	1.15** (0.46)	0.96** (0.43)	1.98*** (0.40)
KA Openness	-0.63*** (0.11)	-0.50** (0.23)	-1.09*** (0.25)	-0.12 (0.14)
Econ.Crisis	0.63 (0.43)	0.11 (0.59)	1.31** (0.52)	-0.25 (0.64)
Contract duration	-0.12*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.12*** (0.04)
Constant	2.53 (1.96)	-7.91** (3.62)	4.27* (2.45)	4.16 (3.87)
lnalpha				
Constant	3.77*** (0.33)	1.89 (2.30)	4.70*** (0.33)	3.69*** (1.16)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 6 – Negative Binomial regression models with multiple imputation (checks models).

	All	Expropriation	Inconvertibility	Violence
main				
Check	-0.26*** (0.09)	-0.15 (0.15)	-0.41** (0.16)	-0.12 (0.14)
Distance	-0.04 (0.04)	0.00 (0.05)	-0.01 (0.06)	-0.13* (0.07)
Capital-intensive	0.33 (0.28)	0.16 (0.26)	0.34 (0.36)	0.36 (0.52)
Resource Rent	0.28** (0.14)	0.55** (0.22)	0.18 (0.19)	0.28 (0.21)
GDP	-0.34*** (0.08)	-0.02 (0.14)	-0.43*** (0.10)	-0.54*** (0.15)
Coal.Turnover	2.22*** (0.84)	2.79** (1.39)	1.45 (1.09)	3.21*** (0.67)
Conflict	1.25*** (0.24)	1.23*** (0.48)	0.89** (0.39)	2.20*** (0.38)
KA Openness	-0.61*** (0.11)	-0.49** (0.23)	-1.10*** (0.24)	-0.07 (0.15)
Econ.Crisis	0.61 (0.44)	0.10 (0.59)	1.30** (0.54)	-0.25 (0.62)
Contract duration	-0.13*** (0.02)	-0.12*** (0.03)	-0.12*** (0.02)	-0.13*** (0.03)
Constant	2.57 (1.84)	-7.76** (3.71)	3.40 (2.45)	6.23* (3.38)
lnalpha				
Constant	3.95*** (0.35)	2.08 (2.33)	4.81*** (0.32)	3.82*** (1.08)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Logit models

TABLE 7 – Logistic regression models with multiple imputation (polcon3 models).

	All	Expropriation	Inconvertibility	Violence
main				
Veto players	-3.27*** (0.72)	-2.42** (1.03)	-3.21*** (1.19)	-5.27*** (0.95)
Distance	-0.07** (0.03)	-0.02 (0.05)	-0.08* (0.05)	-0.12* (0.06)
Capital-intensive	0.17 (0.26)	0.19 (0.26)	0.17 (0.36)	0.27 (0.49)
Resource Rent	0.18 (0.12)	0.52** (0.21)	-0.00 (0.17)	0.11 (0.18)
GDP	-0.30*** (0.07)	-0.03 (0.13)	-0.34*** (0.08)	-0.55*** (0.14)
Coal.Turnover	2.12*** (0.77)	2.84** (1.36)	1.36 (1.01)	3.02*** (0.57)
Conflict	1.16*** (0.23)	1.17** (0.47)	0.83** (0.33)	1.85*** (0.33)
KA Openness	-0.53*** (0.11)	-0.50** (0.24)	-0.93*** (0.18)	-0.01 (0.13)
Econ.Crisis	0.64* (0.37)	0.26 (0.61)	1.10** (0.50)	0.30 (0.49)
Contract duration	-0.10*** (0.01)	-0.12*** (0.03)	-0.09*** (0.01)	-0.11*** (0.03)
Constant	1.65 (1.72)	-7.31** (3.33)	1.88 (1.80)	6.89** (3.12)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 8 – Logistic regression models with multiple imputation (xconst models).

	All	Expropriation	Inconvertibility	Violence
main				
Executive constraints	-0.24*** (0.05)	-0.15* (0.08)	-0.29*** (0.08)	-0.27*** (0.09)
Distance	-0.06** (0.03)	-0.01 (0.05)	-0.08 (0.05)	-0.10 (0.06)
Capital-intensive	0.15 (0.26)	0.16 (0.26)	0.14 (0.35)	0.28 (0.49)
Resource Rent	0.22* (0.13)	0.56** (0.23)	0.02 (0.17)	0.19 (0.19)
GDP	-0.31*** (0.07)	-0.05 (0.13)	-0.34*** (0.08)	-0.60*** (0.14)
Coal.Turnover	2.40*** (0.74)	2.90** (1.39)	1.82* (0.97)	3.41*** (0.60)
Conflict	1.21*** (0.24)	1.22** (0.49)	0.88** (0.34)	1.96*** (0.32)
KA Openness	-0.54*** (0.11)	-0.51** (0.24)	-0.96*** (0.18)	-0.02 (0.13)
Econ.Crisis	0.65* (0.38)	0.28 (0.61)	1.12** (0.52)	0.32 (0.49)
Contract duration	-0.10*** (0.01)	-0.12*** (0.03)	-0.09*** (0.02)	-0.11*** (0.03)
Constant	2.08 (1.70)	-7.04** (3.39)	2.25 (1.75)	7.89** (3.05)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 9 – Logistic regression models with multiple imputation (polyarchy models).

	All	Expropriation	Inconvertibility	Violence
main				
Polyarchy	-2.35*** (0.71)	-1.15 (1.26)	-2.88*** (0.98)	-3.30*** (1.12)
Distance	-0.06* (0.03)	-0.01 (0.05)	-0.07 (0.05)	-0.10 (0.07)
Capital-intensive	0.13 (0.26)	0.16 (0.26)	0.12 (0.36)	0.25 (0.49)
Resource Rent	0.24* (0.13)	0.58** (0.23)	0.05 (0.17)	0.21 (0.19)
GDP	-0.28*** (0.07)	-0.04 (0.13)	-0.30*** (0.08)	-0.54*** (0.14)
Coal.Turnover	2.28*** (0.72)	2.77** (1.30)	1.66* (1.00)	3.40*** (0.64)
Conflict	1.08*** (0.23)	1.16** (0.47)	0.69** (0.33)	1.81*** (0.37)
KA Openness	-0.51*** (0.12)	-0.50** (0.24)	-0.91*** (0.19)	0.01 (0.14)
Econ.Crisis	0.64* (0.36)	0.29 (0.60)	1.13** (0.49)	0.28 (0.51)
Contract duration	-0.10*** (0.01)	-0.12*** (0.03)	-0.09*** (0.02)	-0.10*** (0.03)
Constant	1.21 (1.69)	-7.30** (3.43)	1.06 (1.77)	6.55** (3.13)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 10 – Logistic regression models with multiple imputation (liberaldem models).

	All	Expropriation	Inconvertibility	Violence
main				
liberaldem	-3.06*** (0.92)	-1.19 (1.33)	-4.13*** (1.46)	-3.99** (1.80)
Distance	-0.05* (0.03)	-0.01 (0.05)	-0.06 (0.05)	-0.09 (0.07)
Capital-intensive	0.14 (0.26)	0.16 (0.26)	0.12 (0.35)	0.25 (0.49)
Resource Rent	0.23* (0.13)	0.58** (0.23)	0.05 (0.17)	0.20 (0.19)
GDP	-0.27*** (0.07)	-0.05 (0.13)	-0.28*** (0.08)	-0.54*** (0.15)
Coal.Turnover	2.20*** (0.72)	2.71** (1.29)	1.57 (0.97)	3.23*** (0.64)
Conflict	1.04*** (0.23)	1.14** (0.46)	0.64** (0.32)	1.75*** (0.38)
KA Openness	-0.51*** (0.12)	-0.50** (0.24)	-0.91*** (0.20)	0.01 (0.14)
Econ.Crisis	0.63* (0.36)	0.28 (0.61)	1.11** (0.48)	0.26 (0.50)
Contract duration	-0.10*** (0.01)	-0.12*** (0.03)	-0.09*** (0.02)	-0.10*** (0.03)
Constant	0.84 (1.72)	-7.32** (3.46)	0.49 (1.77)	6.33* (3.22)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 11 – Logistic regression models with multiple imputation (judicial_independence models).

	All	Expropriation	Inconvertibility	Violence
main				
Judicial independence	-3.33*** (0.61)	-2.85** (1.19)	-3.22*** (0.94)	-4.96*** (1.09)
Distance	-0.07** (0.03)	-0.02 (0.05)	-0.08 (0.05)	-0.10 (0.07)
Capital-intensive	0.13 (0.26)	0.15 (0.26)	0.12 (0.35)	0.28 (0.52)
Resource Rent	0.16 (0.11)	0.47** (0.20)	-0.00 (0.16)	0.10 (0.17)
GDP	-0.26*** (0.07)	0.01 (0.14)	-0.31*** (0.08)	-0.52*** (0.15)
Coal.Turnover	2.35*** (0.75)	3.04** (1.42)	1.59* (0.97)	3.45*** (0.62)
Conflict	1.08*** (0.24)	1.11** (0.47)	0.74** (0.35)	1.77*** (0.32)
KA Openness	-0.55*** (0.11)	-0.51** (0.24)	-0.96*** (0.19)	-0.03 (0.13)
Econ.Crisis	0.62* (0.36)	0.26 (0.60)	1.09** (0.49)	0.27 (0.44)
Contract duration	-0.10*** (0.01)	-0.12*** (0.03)	-0.09*** (0.02)	-0.11*** (0.03)
Constant	1.26 (1.72)	-7.73** (3.52)	1.53 (1.79)	6.59** (3.10)
Observations	82536	82536	82536	82536
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 12 – Logistic regression models with multiple imputation (checks models).

	All	Expropriation	Inconvertibility	Violence
main				
Check	-0.33*** (0.10)	-0.27 (0.18)	-0.43*** (0.15)	-0.06 (0.20)
Distance	-0.05 (0.03)	-0.01 (0.05)	-0.06 (0.05)	-0.09 (0.09)
Capital-intensive	0.14 (0.26)	0.13 (0.26)	0.14 (0.37)	0.54 (0.57)
Resource Rent	0.27** (0.12)	0.59*** (0.22)	0.10 (0.16)	0.36* (0.22)
GDP	-0.29*** (0.08)	-0.02 (0.14)	-0.28*** (0.09)	-0.72*** (0.17)
Coal.Turnover	2.31*** (0.80)	3.04** (1.40)	1.57 (1.07)	2.99*** (0.80)
Conflict	1.21*** (0.24)	1.21** (0.49)	0.89*** (0.34)	2.27*** (0.32)
KA Openness	-0.47*** (0.10)	-0.45** (0.21)	-0.83*** (0.17)	0.28* (0.15)
Econ.Crisis	0.63* (0.35)	0.29 (0.58)	1.06** (0.48)	0.63 (0.72)
Contract duration	-0.09*** (0.01)	-0.12*** (0.02)	-0.08*** (0.02)	-0.08* (0.05)
Constant	1.13 (1.88)	-7.72** (3.61)	0.30 (2.19)	8.66** (3.85)
Observations	67970	67970	67970	67970
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

Cox Proportional Hazard models

TABLE 13 – Cox proportion hazard regression models with multiple imputation (polcon3 models).

	All	Expropriation	Inconvertibility	Violence
Veto players	-2.39*** (0.73)	-2.17** (1.09)	-1.95* (1.14)	-4.13*** (1.20)
Distance	-0.08** (0.03)	0.02 (0.05)	-0.09* (0.06)	-0.13** (0.05)
Capital-intensive	0.20 (0.24)	0.24 (0.27)	0.09 (0.40)	0.22 (0.48)
Resource Rent	0.34*** (0.13)	0.57*** (0.20)	0.25 (0.18)	0.07 (0.19)
GDP	-0.27*** (0.07)	-0.07 (0.12)	-0.25*** (0.08)	-0.61*** (0.12)
Coal.Turnover	2.29** (0.93)	3.20** (1.38)	1.57 (1.19)	2.96*** (0.86)
Conflict	1.10*** (0.23)	0.96** (0.43)	0.61* (0.32)	2.03*** (0.33)
KA Openness	-0.42*** (0.13)	-0.52** (0.25)	-0.81*** (0.22)	0.03 (0.15)
Econ.Crisis	0.55 (0.34)	0.35 (0.55)	0.86 (0.51)	0.47 (0.56)
Observations	72312	73821	73611	74389
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

TABLE 14 – Cox proportion hazard regression models with multiple imputation (xconst models).

	All	Expropriation	Inconvertibility	Violence
Executive constraints	-0.23*** (0.05)	-0.20*** (0.08)	-0.25*** (0.09)	-0.27*** (0.10)
Distance	-0.08*** (0.03)	0.01 (0.06)	-0.11* (0.06)	-0.12** (0.05)
Capital-intensive	0.17 (0.24)	0.21 (0.27)	0.05 (0.39)	0.22 (0.49)
Resource Rent	0.35*** (0.14)	0.59*** (0.21)	0.25 (0.17)	0.12 (0.20)
GDP	-0.26*** (0.07)	-0.06 (0.13)	-0.24*** (0.08)	-0.64*** (0.13)
Coal.Turnover	2.60*** (0.90)	3.44** (1.41)	2.00* (1.15)	3.31*** (0.82)
Conflict	1.15*** (0.23)	1.02** (0.43)	0.65** (0.32)	2.11*** (0.32)
KA Openness	-0.44*** (0.13)	-0.53** (0.25)	-0.83*** (0.23)	0.01 (0.16)
Econ.Crisis	0.58 (0.35)	0.37 (0.56)	0.89 (0.53)	0.50 (0.54)
Observations	72312	73821	73611	74389
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 15 – Cox proportion hazard regression models with multiple imputation (polyarchy models).

	All	Expropriation	Inconvertibility	Violence
Polyarchy	-1.79** (0.82)	-1.15 (1.14)	-1.85** (0.87)	-3.59*** (1.33)
Distance	-0.07** (0.03)	0.02 (0.06)	-0.10* (0.06)	-0.12** (0.06)
Capital-intensive	0.17 (0.25)	0.22 (0.28)	0.06 (0.40)	0.22 (0.48)
Resource Rent	0.37*** (0.14)	0.62*** (0.22)	0.27 (0.18)	0.14 (0.20)
GDP	-0.25*** (0.08)	-0.07 (0.12)	-0.22** (0.09)	-0.57*** (0.14)
Coal.Turnover	2.40*** (0.86)	3.14** (1.32)	1.79 (1.14)	3.27*** (0.88)
Conflict	1.05*** (0.22)	0.97** (0.43)	0.52 (0.33)	1.92*** (0.37)
KA Openness	-0.41*** (0.14)	-0.52* (0.26)	-0.80*** (0.23)	0.04 (0.16)
Econ.Crisis	0.57* (0.33)	0.38 (0.54)	0.88* (0.50)	0.46 (0.54)
Observations	72312	73821	73611	74389
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

TABLE 16 – Cox proportion hazard regression models with multiple imputation (liberaldem models).

	All	Expropriation	Inconvertibility	Violence
liberaldem	-2.35** (0.96)	-1.26 (1.24)	-2.76** (1.12)	-4.57** (1.91)
Distance	-0.07** (0.03)	0.02 (0.05)	-0.09* (0.05)	-0.10* (0.05)
Capital-intensive	0.17 (0.24)	0.22 (0.28)	0.04 (0.40)	0.21 (0.48)
Resource Rent	0.36*** (0.14)	0.62*** (0.22)	0.26 (0.18)	0.13 (0.20)
GDP	-0.24*** (0.08)	-0.07 (0.12)	-0.20** (0.09)	-0.57*** (0.15)
Coal.Turnover	2.37*** (0.86)	3.11** (1.32)	1.78 (1.16)	3.13*** (0.85)
Conflict	1.01*** (0.22)	0.94** (0.42)	0.47 (0.31)	1.86*** (0.38)
KA Openness	-0.41*** (0.14)	-0.52* (0.26)	-0.79*** (0.23)	0.04 (0.16)
Econ.Crisis	0.56 (0.33)	0.38 (0.55)	0.87* (0.50)	0.43 (0.53)
Observations	72312	73821	73611	74389
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

TABLE 17 – Cox proportion hazard regression models with multiple imputation (judicial_independence models).

	All	Expropriation	Inconvertibility	Violence
Judicial independence	-3.07*** (0.57)	-3.11*** (1.07)	-2.34*** (0.82)	-5.20*** (1.33)
Distance	-0.08*** (0.03)	0.01 (0.06)	-0.10* (0.06)	-0.11** (0.05)
Capital-intensive	0.13 (0.24)	0.17 (0.27)	0.04 (0.38)	0.22 (0.53)
Resource Rent	0.30** (0.12)	0.50** (0.19)	0.23 (0.17)	0.02 (0.18)
GDP	-0.21*** (0.08)	-0.02 (0.13)	-0.21** (0.09)	-0.55*** (0.14)
Coal.Turnover	2.49*** (0.92)	3.43** (1.42)	1.72 (1.20)	3.20*** (0.86)
Conflict	1.01*** (0.22)	0.90** (0.43)	0.52* (0.31)	1.90*** (0.33)
KA Openness	-0.44*** (0.13)	-0.52** (0.25)	-0.83*** (0.23)	0.00 (0.15)
Econ.Crisis	0.55 (0.33)	0.33 (0.54)	0.86 (0.51)	0.42 (0.50)
Observations	72312	73821	73611	74389
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.010

TABLE 18 – Cox proportion hazard regression models with multiple imputation (checks models).

	All	Expropriation	Inconvertibility	Violence
Check	-0.18* (0.10)	-0.19 (0.14)	-0.24 (0.16)	-0.16 (0.21)
Distance	-0.06 (0.03)	0.03 (0.05)	-0.07 (0.06)	-0.11** (0.05)
Capital-intensive	0.19 (0.24)	0.20 (0.28)	0.08 (0.41)	0.23 (0.46)
Resource Rent	0.40*** (0.13)	0.63*** (0.21)	0.30 (0.18)	0.18 (0.21)
GDP	-0.27*** (0.08)	-0.06 (0.13)	-0.23*** (0.09)	-0.67*** (0.14)
Coal.Turnover	2.24** (0.91)	3.17** (1.38)	1.65 (1.17)	2.94*** (0.83)
Conflict	1.16*** (0.24)	1.00** (0.44)	0.64* (0.33)	2.23*** (0.34)
KA Openness	-0.41*** (0.14)	-0.53** (0.26)	-0.79*** (0.23)	0.04 (0.17)
Econ.Crisis	0.54 (0.33)	0.35 (0.55)	0.86 (0.51)	0.41 (0.55)
Observations	72312	73821	73611	74389
Pseudo R^2	.	.	.	.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Selection bias in the OPIC dataset

In this note, we use Directed Acyclic Graphs (DAGs) to discuss some of the threats to inference that researchers should keep in mind when they use the OPIC dataset. We consider a simple (but somewhat realistic) theoretical model, and discuss the conditions under which key quantities of interest are identifiable. Our goal is to illuminate the additional empirical challenges that arise in this specific observational context, where analysts can only observe firms who buy political risk insurance.

Theoretical model

Figure 1 shows one idealized model. When using the OPIC dataset, researchers can control for the sector in which each investment is made. They can also control for the dollar amount of the insurance coverage.¹ With Figure 1, we want to study how **Unmeasured firm characteristics** – beyond sector and project size – affect inference.

Based on subject matter knowledge, we posit the following relationships :

- **Unmeasured firm characteristics** can affect the probability of **Investment**, the likelihood of buying **Insurance**, and the chances of experiencing a **Loss** at the hands of the host government.
- **Democracy** in the host country affects both the attractiveness of the host for new **Investment**, and the probability that the host will inflict a **Loss** on the foreign investors.
- **Insurance** is purchased by firms with an **Investment** in place, and with certain **Unmeasured firm characteristics** that increase their vulnerability to host government intervention.
- **Claims** can only be filed by investors who suffered a **Loss** and who bought **Insurance**

Each of these relationships is encoded as an arrow in Figure 1. The boxes around **Insurance** and **Investment** indicate that, in the OPIC dataset, we only observe firms who have both invested and contracted insurance. By construction, any model estimated using these data conditions on both variables.

Simulated data

If we assume that every relationship in Figure 1 is linear, we can simulate 1,000,000 data points as follows :

1. The coverage amount variable is incomplete. The sector categories were produced by OPIC and are probably less than ideal.

FIGURE 1 – A simplified theoretical model.

```

N = 1e6
unmeasured = rnorm(N) # continuous
democracy = rnorm(N) # continuous
investment = democracy + unmeasured + rnorm(N) > 0 # binary
insurance = investment + unmeasured + rnorm(N) > 0 # binary
loss = -1.5 * democracy + investment + unmeasured + rnorm(N) # continuous
claim = loss + insurance + rnorm(N) # continuous
dat = data.frame(unmeasured, democracy, investment, insurance, loss, claim)

```

In this simulation process, we *know* that the (direct) causal effect of democracy on loss equals -1.5.

No selection

Consider the ideal case where information on *all* firms is available, that is, the ideal case where we are *not* forced to condition on **Investment** and **Insurance**.

The first thing to note is that in order to identify the direct effect of **Democracy** on **Loss**, we need to control for **Investment**. The intuition is straightforward : If democracies attract more investment, our dataset might show larger numbers of loss events in democracies simply because there is more “at risk” investments there. Our quantity of interest is the direct effect of **Democracy** on **Loss**, *conditional* on the level of **Investment**.

Unfortunately, since **Investment** is a collider, controlling for that variable opens a new path between our dependent and independent variables :

$$\text{Democracy} \rightarrow \boxed{\text{Investment}} \leftarrow \text{Unmeasured} \rightarrow \text{Loss} \quad (1)$$

TABLE 19 – No selection

	Total effect	Collider bias	Direct effect
(Intercept)	0.50*** (0.00)	−0.58*** (0.00)	0.00 (0.00)
democracy	−1.27*** (0.00)	−1.77*** (0.00)	−1.50*** (0.00)
investmentTRUE		2.16*** (0.00)	1.00*** (0.00)
unmeasured			1.00*** (0.00)
R ²	0.38	0.59	0.77
Adj. R ²	0.38	0.59	0.77
Num. obs.	1000000	1000000	1000000
RMSE	1.63	1.32	1.00

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

If we do not close this path, we get what Judea Pearl calls “collider bias.” To close the path, we need to control for *both* **Investment** and for the **Unmeasured firm characteristics**. Obviously, this is non-trivial given that these characteristics are unmeasured...

We illustrate this by estimating three models using the data simulated above :

```
models = list()
models[['Total effect']] = lm(loss ~ democracy, data = dat)
models[['Collider bias']] = lm(loss ~ democracy + investment, data = dat)
models[['Direct effect']] = lm(loss ~ democracy + investment + unmeasured, data = dat)
```

Table 19 prints the results for those three models. The first column shows that if we do not control for the level of investment, our estimate of the effect of **Democracy** on **Loss** is biased toward zero. The second columns shows that controlling for **Investment** does not solve the problem, because it produces collider bias. The third column shows that analysts must control for the **Unmeasured firm characteristics** in order to obtain unbiased estimates of the direct effect.

The main lesson to draw from this exercise is that the threat to inference posed by unmeasured firm characteristics is not specific to the OPIC dataset. When unmeasured factors are related to both the probability of investment and the probability of loss, we are unlikely to recover a true causal estimate from the observational data.

TABLE 20 – Selection

	Biased	Unbiased
(Intercept)	2.70*** (0.00)	2.00*** (0.00)
democracy	−1.72*** (0.00)	−1.50*** (0.00)
unmeasured		1.00*** (0.00)
R ²	0.47	0.60
Adj. R ²	0.47	0.60
Num. obs.	431449	431449
RMSE	1.63	1.42

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Selection

Now, we consider data that more closely resemble those in the OPIC dataset. We repeat the previous experiment, but truncate the sample based on values of the **Insurance** and **Investment** variables. This is meant to reflect the fact that our OPIC dataset only includes firms who have decided to both invest and buy insurance. We also add a new twist : Our dependent variable is now insurance **Claims** rather than **Losses**.

```
selection = (dat$investment == TRUE) & (dat$insurance == TRUE)
models = list()
models[['Biased']] = lm(claim ~ democracy, data = dat[selection, ])
models[['Unbiased']] = lm(claim ~ democracy + unmeasured, data = dat[selection, ])
```

Interestingly, the fact that unmeasured firm characteristics are related to the probability of buying insurance does not pose additional difficulties (in this simple model), because we are already conditioning on **Insurance** (by excluding uninsured projects from the database), thus blocking the problematic path.

Also interesting is that in this simple simulation, the two left-most models of Table 19 are *more* biased than the first model in Table 20, where data are affected by OPIC-like selection.

When does the problem disappear ?

In essence, the selection problem that afflicts the OPIC data can be thought of as a straightforward failure to meet the selection on observables assumption : Unmeasured firm characteristics may be related to both the probability of investing and the likelihood of experiencing a loss. If researchers are willing to assume that one of those two links is inexistant, they can treat their estimates

as unbiased (again, in the circumscribed context of this idealized model). To see this, consider to new simulated datasets.

Here, the effect of Unmeasured firm characteristics on Investment is zero :

```
# simulate
N = 1e7
unmeasured = rnorm(N) # continuous
democracy = rnorm(N) # continuous
investment = democracy + 0 * unmeasured + rnorm(N) > 0 # binary
insurance = investment + unmeasured + rnorm(N) > 0 # binary
loss = -1.5 * democracy + investment + unmeasured + rnorm(N) # continuous
claim = loss + insurance + rnorm(N) # continuous
dat = data.frame(unmeasured, democracy, investment, insurance, loss, claim)

# selection
selection = (dat$investment == TRUE) & (dat$insurance == TRUE)

# estimate
models = list()
models[['Investment']] = lm(claim ~ democracy, data = dat[selection, ])
```

Here, the effect of Unmeasured firm characteristics on Losses is zero :

```
# simulate
N = 1e7
unmeasured = rnorm(N) # continuous
democracy = rnorm(N) # continuous
investment = democracy + unmeasured + rnorm(N) > 0 # binary
insurance = investment + unmeasured + rnorm(N) > 0 # binary
loss = -1.5 * democracy + investment + 0 * unmeasured + rnorm(N) # continuous
claim = loss + insurance + rnorm(N) # continuous
dat = data.frame(unmeasured, democracy, investment, insurance, loss, claim)

# selection
selection = (dat$investment == TRUE) & (dat$insurance == TRUE)

# estimate
models[['Losses']] = lm(claim ~ democracy, data = dat[selection, ])
```

Table 21 shows that estimates of the Democracy coefficient in those two datasets are unbiased, even if we face the same double selection problem as before (i.e., we only observe realized investment projects by firms who buy insurance).

TABLE 21 – No effect of unmeasured firm characteristics on investment or losses

	Investment	Losses
(Intercept)	2.29*** (0.00)	2.00*** (0.00)
democracy	-1.50*** (0.00)	-1.50*** (0.00)
R ²	0.36	0.47
Adj. R ²	0.36	0.47
Num. obs.	3803154	4315566
RMSE	1.66	1.41

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Conclusion

None of what we wrote above demonstrates that the statistical models reported in the main paper actually meet the conditions required for causal identification. Moreover, the models we looked at above were necessarily quite simple ; one could invent more complicated models where the same conclusions do not hold exactly. Still, we hope that this discussion was helpful in illuminating the special challenges that arise when working with the kind of data introduced in our paper.

In the end, researchers who analyze the OPIC data face a choice between two paths. The first option is to make modest claims about generalizability : The OPIC data allows us to study the determinants of losses by the subset of firms who would buy political risk insurance. The second path that researchers can take is a more ambitious one : Claiming that the insights obtained from OPIC data analysis can generalize to a broader set of firms. In order to make such general claims, however, analysts must be willing to make the necessary theoretical leap, that is, to claim that there are no red arrows in Figure 1.