The International Political Economy Data Resource

Benjamin A.T. Graham
Jacob R. Tucker

September 14, 2016

Abstract:
Quantitative scholars in international relations often draw repeatedly on the same sources of country-year data across a diverse range of projects. The IPE Data Resource seeks to provide a public good to the field by standardizing and merging together 951 variables from 78 core IPE data sources into a single dataset, increasing efficiency and reducing the risk of data management errors. Easier access to data both encourages researchers to perform more robustness checks than they otherwise might and makes it easier for teachers of quantitative research methods to assign realistic exercises to their students. This resource will be updated and expanded annually and is available via the Harvard Dataverse Network.¹

¹ http://dx.doi.org/10.7910/DVN/X093TV
Introduction

Cross-national quantitative research in international relations is a data-intensive undertaking. Many researchers create novel measures of the constructs most critical to the question at hand – i.e. their variable(s) of interest. However, between control variables in the core model and a range of alternative measures employed as robustness checks, it is not uncommon for a single paper to utilize thirty or forty variables drawn from a dozen distinct datasets. In papers that use latent variable models to estimate unobservable variables of interest, the number of indicator variables used in a single paper can rise into the hundreds.\(^2\) Thus, gathering together and compiling a diverse array of variables has become a substantial part of the labor of quantitative scholarship in international relations.

Country-year data in international relations in general and international political economy (IPE) in particular is fragmented in its sources, drawn from economics, political science, sociology, international business, and other fields. Given the low readership of most journals by scholars from outside the discipline, even identifying the most appropriate existing data sources is challenging. Once identified, merging multiple datasets together requires much labor and great care, as errors at this stage can undermine all the work that follows. Scholars of international security have long utilized EUGene for a compilation of important variables.\(^3\) This project brings together a much broader range of variables, focusing primarily on those of interest to IPE scholars.

\(^2\) See, for example, Fariss (2014) assessing government respect for human rights, Hollyer, Rosendorf, and Vreeland (2014), or Fariss and Graham (2016) measuring the security of property rights.

\(^3\) EUGene refers to the Expected Utility Generation and Data Management Program (Bennett and Stam 2000).
Despite the diversity of topics studied, many projects in IPE, international security, and comparative politics employ the same standard set of core country-year data sets to measure economic, political and social conditions across countries and over time. This area of overlap has only grown as economic factors have become more central to the study of international security. This creates an opportunity for large efficiency gains through coordination. Currently, the process of merging these core datasets together is repeated scholar by scholar, project by project. This resource aims to improve and streamline this data management process.

Increased efficiency in data management not only reduces errors and saves time, it encourages researchers to perform their work more thoroughly. The cost of identifying, obtaining, and preparing additional measures deters scholars from running additional robustness tests that would increase the quality of their work. All of the measures we work with as quantitative social scientists are flawed in some way; use of more alternative measures reduces the odds that that the results we publish are driven by these flaws. A centralized source for data is particularly valuable in an interdisciplinary topic area such as IPE where information about new data resources can be slow to flow across disciplinary boundaries.

Efficient data management is valuable in teaching as well as research. Professors teaching quantitative methods courses sometimes shy away from assigning “real world” data exploration exercises due to the difficulty of pulling together the necessary datasets for their students. Similarly, one of the motivations for this resource comes from a common assignment in graduate statistics courses – to reproduce and probe the robustness of a core finding in the literature. Having a dataset like this on hand from
which students can draw additional/alternative regressors makes it easier for students to complete such a project in the span of a semester.

This project compiles 78 commonly used datasets drawn from a range of disciplines into an easy-to-merge format designed to speed the process of acquiring and preparing the necessary data for quantitative research projects.\textsuperscript{4} Political institutions, economic conditions, and political violence are all tightly intertwined, and thus, while the content area of these data is focused on IPE, we expect this resource to be of interest to scholars in a variety of related fields, including security and comparative politics. To aid security scholars in particular, version 1.5 of the IPE Data Resource is scheduled for inclusion in version 1.0 of NewGene, the successor to EUGene, a software program facilitating the creation of dyadic and directed-dyadic datasets for the study of interstate war (Poast and Stam 2015).\textsuperscript{5}

We merge component datasets together on the basis of the Gleditsch-Ward country identification system, which we argue should become the standard in IPE research (Gleditsch and Ward 1999). However, the completed dataset includes Correlates of War (COW) codes and International Financial Statistics (IFS) codes. Seventy-eight datasets fall far short of comprehensive coverage of all important IPE datasets, but it is a substantial start. Critically, the Security and Political Economy (SPEC) Lab at the University of Southern California provides institutional support for the resource.\textsuperscript{6} The

\textsuperscript{4} This count, and all details in this article, refer to version 1.5 of the data.
\textsuperscript{5} Software available at http://newgenesoftware.org [Release Pending]
\textsuperscript{6} Financial support for the lab is provided by the Dornsife College of Arts and Letters, the Center for International Studies, and the School of International Relations. Support is renewable indefinitely.
dataset will be updated and expanded annually, incorporating new versions of datasets already included and adding new relevant datasets as they become available.\textsuperscript{7}

**Methodology**

The IPE data resource incorporates component datasets covering four general topic areas: politics (37 datasets); economics (31 datasets); social conditions (8 datasets); and geography (2 datasets). See Table A1 in the Appendix for a full list. While we expect this data resource to be quite useful to security and civil conflict scholars, this resource includes only a limited number of measures of interstate and intrastate war.\textsuperscript{8} We largely leave it to scholars in these areas to merge in their dependent variable(s) of interest, which may be quite varied in form.

The most critical decision regarding how to compile country-year data is the choice of unique country ID numbers. Over time, countries are born, conquered, merged and divided. Following every merger or division, a time-series-cross-sectional dataset must either treat the resulting units as continuations of previous countries, in which case they retain the same ID number, or as new entities with a new number. In this project, we use the ID numbers associated with the Gleditsch-Ward system (Gleditsch and Ward 1999).

The Gleditsch-Ward system asserts continuation of political units through most mergers and divisions. Thus, modern Yemen, which was formed through the merger of

\textsuperscript{7} New data will be downloaded in January and a revised version of the resource posted in March of each year.

\textsuperscript{8} The one exception is the Major Episodes of Political Violence Data (Marshall 1999), which includes economic damage within its index of conflict severity. We view this as a conflict measure particularly appropriate for use as an independent variable in analyses of economic outcomes.
People’s Democratic Republic of Yemen and the Arab Republic of Yemen in 1990, is treated as a continuation of the Arab Republic of Yemen. Imperial Russia, the Soviet Union, and modern Russia all retain the same number. When Estonia, Latvia, and Lithuania regain independence in 1991, they regain the same ID numbers they had prior to their loss of independence in 1940.

This presumption of continuity is appropriate to most of the core questions in IPE, as well as the study of intrastate violence and some (though not all) comparative politics questions. In most cases, the same firms that are operating the day before a merger or division of countries continue to operate after it; debts that were owed continue to be owed; contracts remain in place; social identities and divisions persist; and in many cases even the identity of political elites and the structure of many political institutions remains the same (e.g. Acemoglu and Robinson 2006; Alston 2008). Thus the Gleditsch-Ward system treats the economic and social unit (the country) as persisting as well, albeit in altered form. However, to facilitate use of this resource by security scholars, we make available a version of the dataset that is time-series-set for COW codes, which are the dominant set of country identifiers in the study of interstate war. It is this version of the data that is integrated into the forthcoming NewGene software.

---

9 It is notable that the Uppsala/PRIO civil conflict datasets are all based on Gleditsch-Ward numbers.

10 Time series setting refers to the process of establishing unit and time variables within Stata to identify unique time-series-cross-sectional observations. The primary version of this resource contains some observations that are duplicates by COW code and year. In particular, some component datasets include data for both Serbia (in isolation) and Yugoslavia (in total) in the same years. We can retain this information in the Gleditsch-Ward-based version of the data, where modern Serbia is treated as a continuation of the Kingdom of Serbia and as distinct from Yugoslavia, but must drop this information from the COW version, where Yugoslavia and Serbia share the same country code.
Many of the component datasets included in the IPE data resource do not include Gleditsch-Ward numbers in their native form, and no other numbering system is hegemonic. Some datasets use IFS codes, others use COW codes, etc. Thus, we append Gleditsch-Ward IDs on the basis of country names, carefully remove duplicate observations by ID and year,\(^{11}\) and then proceed to merge the datasets together.

Data cleaning involves difficult choices, and choices optimal for one research project may be inappropriate for another. Thus, it is important for scholars using this resource to be able to track and review every step that is taken from when the raw data is downloaded until the final dataset is compiled. To this end, extensive documentation is posted along with the final dataset. The codebook for this project provides the name of each dataset, a description of each variable, citations to the sources from which the data is drawn, notes regarding any data cleaning that was necessary to integrate each component dataset into the resource, and links to the relevant codebooks whenever possible. We also post the raw component data files themselves; the Stata .do files that clean these data files and append the Gleditsch-Ward Numbers (one per raw data file); and the Stata .do file that merges all of the prepped data files together. Thus, every step in the process is transparent, and scholars can review and revise any of the data cleaning choices made.

Version 1.5 of the dataset was released on August 30, 2016. All materials are available via the Harvard Dataverse Network.\(^{12}\) This resource will be updated and expanded annually.

---

\(^{11}\) In this process, we remove the observation with the least information. In most cases at least one of the duplicate observations contains exclusively missing values.

\(^{12}\) [http://dx.doi.org/10.7910/DVN/X093TV](http://dx.doi.org/10.7910/DVN/X093TV)
An Illustrative Example

In 2003, *International Organization* featured two quantitative articles, one by Nathan Jensen and one by Quan Li and Adam Resnick, both of which used country-year data to assess the relationship between domestic political institutions and inflows of foreign direct investment (FDI). In the ensuing decade, dozens more quantitative studies followed examining various political determinants of FDI inflows. The papers in this literature overlap heavily in terms of the theoretical constructs they seek to measure (e.g. democracy, government capacity, political risk), but they differ in terms of precisely how these theoretical constructs are defined and in how they are measured. While many studies employ more than one measure of key institutional variables, we argue that, across this literature, scholars do not employ alternate measures of key constructs nearly as much as they should – and nearly as much as they might if alternative measures were more readily available.

To make this case, we evaluate 21 of the most prominent articles matching two conditions: 1) Use country-year data to predict FDI inflows; 2) published between 2003 and 2013. Table 1 reports the different aspects of political institutions that are evaluated in these papers, the specific measures used to capture these institutions, and the papers that employ each measure.

---

13 Jensen (2003) and Li and Resnick (2003). While we credit these articles as seminal, they, in turn, draw on a long history of theoretical, historical, and qualitative work on the political determinants of global capital flows (e.g. Caves 1971; Kobrin; Moran; Kindleberger; Kahler 1984; Frieden 1994; Henisz 2000), and also build on prior quantitative work on the determinants of FDI flows (e.g.).

14 These articles are: Adam and Filippaios (2007); Ali, Fiess, and MacDonald (2010); Asiedu and Lien (2011); Biglaiser and DeRouen (2006); Blanton and Blanton (2007); Busse and Hefeker (2007); Büthe and Milner (2008); Dutta and Roy (2011); Egger and Pfaffermayr (2004); Haftel (2010); Jensen (2003, 2004, 2008); Kern (2009); Li (2006); Mathur and Singh (2013); Mosley and Uno (2007); Ramasamy and Yeung (2010).
Table 1: Political Determinants of Foreign Direct Investment

| Concept                                      | Measure                                                                 | Articles Using                                      |
|----------------------------------------------|-------------------------------------------------------------------------|-----------------------------------------------------|
| Democracy/Executive Constraints              | Democracy (Polity IV)\(^{15}\)                                        | Jensen (2008)                                       |
|                                              |                                                                         | Blanton and Blanton (2007)                          |
|                                              |                                                                         | Büthe and Milner (2008)                             |
|                                              |                                                                         | Mosley and Uno (2007)                               |
|                                              |                                                                         | Biglaiser and DeRouen (2006)                        |
|                                              |                                                                         | Haftel (2010)                                       |
|                                              |                                                                         | Li (2006)                                           |
|                                              |                                                                         | Asiedu and Lien (2011)                              |
|                                              |                                                                         | Kerner (2009)                                       |
|                                              | Democracy (Alvarez, Cheibub, Limongi, Przeworski)\(^{16}\)            | Buthe and Milner (2008)                             |
|                                              |                                                                         | Jensen (2003)                                       |
|                                              | Electoral Democracy (Freedom House)\(^{17}\)                          | Asiedu and Lien (2011)                              |
|                                              |                                                                         | Mathur and Singh (2013)                             |
|                                              |                                                                         | Büthe and Milner (2008)                             |
|                                              | Democratic Accountability (ICRG)\(^{18}\)                             | Asiedu and Lien (2011)                              |
|                                              |                                                                         | Busse and Hefeker (2006)                            |
|                                              | Checks (DPI)\(^{19}\)                                                | Jensen (2008)                                       |
|                                              | Domestic political constraints (polcon)\(^{20}\)                     | Buthe and Milner 2008                               |
|                                              |                                                                         | Jensen (2008)                                       |
|                                              |                                                                         | Haftel (2010)                                       |
| Rule of Law                                  | Law and Order (ICRG)                                                  | Ali, Fiess, MacDonald (2010)                        |
|                                              |                                                                         | Dutta and Roy (2011)                                |
|                                              |                                                                         | Busse and Hefeker (2006)                            |
|                                              |                                                                         | Asiedu and Lien (2011)                              |
|                                              |                                                                         | Jensen (2003)\(^{21}\)                             |
|                                              | Rule of law (WGI)\(^{22}\)                                          | Li (2006)                                           |
|                                              |                                                                         | Morrissey and Udomkerdmongkol (2012)                |
|                                              | Control of Corruption (WGI)                                           | Morrissey and Udomkerdmongkol (2012)                |
|                                              | Corruption (ICRG)                                                    | Ali, Fiess, MacDonald (2010)                        |
|                                              |                                                                         | Ramasamy and Yeung (2010)                           |
|                                              |                                                                         | Biglaiser and DeRouen (2006)                        |

\(^{15}\) Marshall and Gurr (2012).
\(^{16}\) Alvarez et al (1996).
\(^{17}\) Freedom House (2015).
\(^{18}\) Political Risk Services Group (2011).
\(^{19}\) Keefer (2010); Thorsten et al (2001).
\(^{20}\) Henisz (2010); Henisz (2000); Henisz (2002).
\(^{21}\) Jensen’s citation is to Easterly (1999) but the underlying data seems to be from ICRG.
\(^{22}\) Kaufmann, Kraay and Mastruzzi (2009).
| Category | Variable | Source |
|----------|----------|--------|
| Political Violence | Corruption Perceptions (Transparency International) | Mathur and Singh (2013) |
| | Internal Conflict and/or External Conflict (ICRG) | Ali, Fiess, MacDonald (2010) |
| | Internal Conflict and/or External Conflict (UCDP/PRIO) | Blanton and Blanton (2007) |
| | Political Stability and Absence of Violence (WGI) | Morrissey and Udomkerdmongkol (2012) |
| | Socioeconomic Conditions, Ethnic Tensions and/or Religion in Politics (ICRG) | Dutta and Roy (2011) |
| | Major Episodes of Political Violence (Center for Systemic Peace) | Biglaiser and DeRouen (2006) |
| | Terrorist incidents (ITERATE) | Li (2006) |
| Government Capacity/Effectiveness | Bureaucratic Quality (ICRG) | Asiedu and Lien (2011) |
| | Government Reputation (Easterly) | Jensen (2003) |
| | Military in Politics (ICRG) | Dutta and Roy (2011) |
| | Property Rights Protection (Economic Freedom in the World) | Mathur and Singh (2013) |
| Risk of Adverse Government Actions | Transfer Risk (Credendo Group) | Jensen (2008) |
| | Expropriation Risk (Credendo Group) | Jensen (2008) |
| | Expropriation | Jensen (2003) |
| | Political Risk (ICRG) | Asiedu and Lien (2011) |

---

23 Transparency International (2012).
24 Gleditsch et al (2002).
25 Marshall (2012). Biglaiser and DeRouen refer to the variable from this dataset as “Societal Conflict.”
26 Mickolus (1982).
27 Easterly (1999).
28 Gwartney, Lawson, and Hall (2014).
29 For up to date panel data, see Graham, Kingsley and Johnston (2015).
30 Easterly (1999).
What Table 1 makes clear is that, even when different studies converge on a relatively narrow and well-defined aspect of political institutions, scholars still have multiple measures to choose from. Different scholars select different measures, and often the choice between different possible measures is not discussed.

To use the concept of veto-player-type constraints on the executive as an example, there are three prominent options: the executive constraints ($XCONST$) component of the Polity IV democracy score (Marshall and Jaggers 2002); the count of

| Civil and Human Rights | Labor Rights (Collective Labor Rights Dataset)\(^{31}\) | Adam and Filippaios (2007) |
|-----------------------|-------------------------------------------------|---------------------------|
|                       | Political Terror Scale\(^{32}\)                | Mosley and Uno (2007)    |
|                       | Civil Liberties (Freedom House)                | Adam and Filippaios (2007) |
| Government Stability  | Government Stability (ICRG)                    | Busse and Hefeker (2006)  |
|                       | Democratic Accountability & Military in Politics (ICRG) | Dutta and Roy (2011) |
|                       | Political Instability (Banks and Wilson)\(^{34}\) | Asiedu and Lien (2011)\(^{33}\) |
|                       | Voice and Accountability (WGI)                 | Ali, Fiess, MacDonald (2010) |
|                       | Regulatory Quality (WGI)                       | Asiedu and Lien (2011)\(^{33}\) |
| International Agreements | Number of Preferential Trade Agreements | Büthe and Milner (2008) |
|                       | GATT/WTO membership                            | Büthe and Milner (2008) |
|                       | Bilateral Investment Treaties                   | Too numerous to list    |
|                       | IMF Agreements                                  | Jensen (2004)           |

---

\(^{31}\) Mosley and Uno (2007).
\(^{32}\) Gibney et al. (2015).
\(^{33}\) Asiedu and Lien reverse this variable and call it “instability of government.”
\(^{34}\) Banks and Wilson (2015).
\(^{35}\) Because these articles vary in their year of publication, they sometimes draw on different versions of the measures listed, some of which change names between articles. We have denoted articles as employing the same measure when they employ a different edition of the same dataset.
veto players (checks) in the Database of Political Institutions (Beck et al. 2001); and the political constraints (polcon) measure designed by Witold Henisz (Henisz 2000). Though they measure the same underlying theoretical construct, each is coded differently, making each appropriate in different circumstances. Each of the three measures hails from a different discipline: xconst was created by political scientists; checks by economists; and polcon by a business scholar, thus many scholars are not even aware of all three. The pairwise correlations between them range from .64 to .79 – correlations that are strong, but not sufficiently so to render the choice between them moot.

| Table 2: Pairwise Correlations of Constraints Measures |
|-----------------------------------------------|
| Variables          | Xconst | Checks | Polcon |
|---------------------|--------|--------|--------|
| Xconst (Polity)     | 1.00   |        |        |
| Checks (DPI)        | 0.68   | 1.00   |        |
| Polcon (Henisz)     | 0.79   | 0.64   | 1.00   |

Similarly, Figure 2 shows substantial differences in distribution across these measures. Both polcon and checks feature a large cluster of countries in which they assess the chief executive to be entirely unconstrained – 43% of the observations for checks and 39% of the observations for polcon. Conversely, only 17% of observations receive the lowest score on the seven-point xconst scale; 31% receive the top score.

---

36 Polcon is omitted from version 1.5 of the resource because Henisz could not be reached to secure his permission. If permission is granted, the Polcon data will be added to the next update of the resource.
In light of these important differences, no scholar should use any of these measures, even as a control variable, without fully understanding the substantive differences between each and assessing, at least for their own information, the robustness of their findings to each of the three. However, many scholars fail to test the robustness of their results to alternative measures of key variables simply because of the time required to track down, prepare, and merge together the necessary data. The more readily available alternative measures are, the higher-quality analysis scholars are likely to perform.

A fundamental contribution of the IPE data resource is to make a wide range of alternate measures easily available to scholars. The ambition of the resource is to include all or most of the variables used in statistical models of the core outcomes in international political economy – from currency crises to trade policy to human rights. With 951 variables drawn from 78 component data sources, version 1.5 already places us quite far along that path.

Existing Resources

A number of scholars and institutions have taken the step of providing catalogs of links to frequently used datasets. For example, Paul Hensel maintains the International Relations Data Site, which is essentially a catalog of links to individual datasets created
by various scholars. The Quality of Governance Institute (QoG) provides the resource most similar to that which we offer here, but with more of a comparative politics focus than an IPE focuses (Teorell et al 2016). Thus, the largest difference between these two resources is in the substantive topics they cover: 106 datasets are compiled by QoG and 78 datasets here, of which only 28 overlap. However, these resources also differ in their approach to country codes. The QoG (rightly) takes an approach to country codes that is tailored to the study of domestic institutions and processes. Every time countries merge or divide in the QoG data, the resulting units are given new country IDs. Thus, in the QoG data, a unified Yemen is treated as neither a continuation of South Yemen or of North Yemen, but rather as an entirely new entity. This is appropriate in many comparative politics contexts because the nature of domestic politics has been fundamentally altered, but as discussed in the methodology section, it is less ideal for IPE. Thus, the IPE data resource uses Gleditsch-Ward and Correlates of War IDs instead.

The IPE Data Resource also has similarities to the data management components of the Expected Utility Generation and Data Management Program (EUGene) developed by Bennett and Stam (2000), to the World Development Indicators (WDI), which are compiled by the World Bank, the GROWup project created by Girardin et al (2015), and to the PRIO-GRID dataset (Tollefsen, Strand, and Buhaug 2012). EUGene and its forthcoming successor, NewGene, cover a range of variables often used in models of interstate conflict, while WDI covers measures of economic and human development, largely eschewing measures of political institutions. Thus a key distinction is again the substantive area of focus of each data aggregation effort. PRIO-GRID brings together a

---

37 http://www.paulhensel.org/data.html
38 This count was produced using the January 2016 version of the QoG Standard Data.
range of economic, social, and political information on the basis of geographic space, rather than country boundaries. PRIO-GRID covers some of the same variables as this project, but it brings them together with regard to a different set of geographic units – quadratic grid cells rather than countries.

While the IPE Data Resource is a stand-alone resource, the coming integration of this data with NewGene offers an opportunity to push forward the integration of data resources between security and IPE. While the most obvious initial benefit of this integration is to make political economy data easily available in appropriate formats to scholars of interstate conflict, the potential to build dyadic and directed-dyadic economic information (e.g. trade, investment, and migration flows) into future generations of NewGene offers the potential to, in turn, facilitate the use of detailed security data (e.g. disputes and alliances) by IPE scholars.

A Note on Citation

The component datasets compiled in this project were each constructed, through great cost and effort, by scholars and institutions that deserve credit for their work.39 The codebook to the IPE Data Resource provides the appropriate citation for each component dataset. Any publication that draws on data from this resource must cite to the original datasets from which the variables in question are drawn. Citations to this paper (i.e. to the IPE Data Resource itself) are necessary only if the aggregation and data management efforts represented by the resource contribute to the research produced.

39 Permission to repost data was obtained from the authors of each component dataset.
Conclusion

This data resource serves the field by bringing together 78 datasets from political science, economics, international business, and sociology and carefully merging them together in a theoretically informed manner. Datasets are merged together on the basis of the Gleditsch-Ward system of country IDs, but an alternate version of the dataset employing the COW numbering system is also available. Because the project has the institutional support necessary to sustain annual updates, it will be able to remain relevant and up to date, updating existing data and incorporating new component datasets as they become available. The value of this project comes in the form of increased efficiency, increased research and teaching quality, and the diffusion of new data resources across disciplinary boundaries. As the costs of doing high quality work fall, we can all do better work and more of it.

Bibliography

Adam, Antonis, and Fragkiskos Filippaios. 2007. "Foreign direct investment and civil liberties: A new perspective." European Journal of Political Economy 23, no. 4: 1038-1052.

Ali, Fathi A., Norbert Fiess, and Ronald MacDonald. 2010. "Do institutions matter for foreign direct investment?" Open Economies Review 21, no. 2: 201-219.

Allee, Todd, and Clint Peinhardt. 2010. "Delegating differences: Bilateral investment treaties and bargaining over dispute resolution provisions." International Studies Quarterly 54, no. 1: 1-26.

Alvarez, Mike, JoséAntonio Cheibub, Fernando Limongi, and Adam Przeworski. 1996. "Classifying Political Regimes." Studies In Comparative International Development 31: 3-36.

Asiedu, Elizabeth, and Donald Lien. 2011. "Democracy, foreign direct investment and natural resources." Journal of International Economics 84, no. 1: 99-111.
Banks, Arthur S., Wilson, Kenneth A. 2015. Cross-National Time-Series Data Archive. Databanks International. Jerusalem, Israel.

Bauer, Molly, Cesi Cruz, and Benjamin Graham. 2012. "Democracies Only: When Do IMF Agreements Serve as a Seal of Approval?" *The Review of International Organizations* 7: 33-58.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *The World Bank Economic Review* 15: 165–76.

Bennett, D. Scott, and Allan C. Stam. 2000. "EUGene: A Conceptual Manual." *International Interactions* 26: 179-204.

Bennett, D. Scott and Poast, Paul and Stam, Allan C. 2015. “NewGene: A Conceptual Manual.” Available at SSRN: http://ssrn.com/abstract_id=2596297

Biglaiser, Glen, and Karl R. DeRouen. 2006. "Economic reforms and inflows of foreign direct investment in Latin America." *Latin American Research Review* 41, no. 1: 51-75.

Blanton, Shannon Lindsey, and Robert G. Blanton. 2007. "What attracts foreign investors? An examination of human rights and foreign direct investment." *Journal of Politics* 69, no. 1: 143-155.

Busse, Matthias, and Carsten Hefeker. 2007. "Political risk, institutions and foreign direct investment." *European Journal of Political Economy* 23, no. 2: 397-415.

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, no. 4: 741-762.

Dutta, Nabamita, and Sanjukta Roy. 2011. "Foreign direct investment, financial development and political risks." *The Journal of Developing Areas* 44, no. 2: 303-327.

Egger, Peter, and Michael Pfaffermayr. 2004. "The Impact of Bilateral Investment Treaties on Foreign Direct Investment." *Journal of Comparative Economics* 32: 788-804.

Fariss, Christopher J. "Respect for human rights has improved over time: Modeling the changing standard of accountability." *American Political Science Review* 108(02): 297-318.

Freedom House. 2015. Freedom in the World. https://freedomhouse.org/report-types/freedom-world#.VZV8p_lVikp
Gibney, Mark, Linda Cornett, Reed Wood, Peter Haschke, and Daniel Arnon. 2015. The Political Terror Scale 1976-2015. 20 June 2015, from the Political Terror Scale website: http://www.politicalterrorscale.org.

Girardin, Luc, Philipp Hunziker, Lars-Erik Cederman, Nils-Christian Bormann, et al. 2015. GROWtp - Geographical Research On War, Unified Platform. ETH Zurich. http://growup.ethz.ch/.

Gleditsch, Kristian S, and Michael D Ward. 1999. "Interstate System Membership: A Revised List of the Independent States since 1816." International Interactions 25: 393-413.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. "Armed conflict 1946-2001: A new dataset." Journal of Peace Research 39, no. 5: 615-637.

Graham, Benjamin A.T., Allison F. Kingsley, and Noel P. Johnston. 2015. "Even Constrained Governments Steal: The Domestic Politics of Transfer and Expropriation Risks." Working paper.

Gwartney, James, Robert Lawson, Joshua Hall, James Gwartney, Robert Lawson, and Joshua Hall. 2014. "2014 Economic Freedom Dataset, published in Economic Freedom of the World: 2014 Annual Report."

Haftel, Yoram Z. 2010. "Ratification counts: US investment treaties and FDI flows into developing countries." Review of International Political Economy 17.2: 348-377.

Henisz, Witold J. 2000. "The Institutional Environment for Economic Growth." Economics and Politics 12 (1): 1-31.

Henisz, Witold J. 2000. "The Institutional Environment for Multinational Investment." Journal of Law, Economics, and Organization 16: 334–64.

Henisz, Witold J. 2002. "The Institutional Environment for Infrastructure Investment.” Industrial and Corporate Change 11 (2): 355-389.

Henisz, Witold J. 2010. "Political Constraint Index."

Hicks, Raymond and Kristina Johnson. 2011. "The Politics of Globalizing Production: Why we See Investment Chapters in Preferential Trade Agreements." Working paper.

Hollyer, James R., B. Peter Rosendorff, and James Raymond Vreeland. 2014. "Measuring Transparency." Political Analysis.
Jensen, Nathan M. 2003. "Democratic governance and multinational corporations: Political regimes and inflows of foreign direct investment." International Organization 57, no. 03: 587-616.

Jensen, Nathan. 2008. "Political risk, democratic institutions, and foreign direct investment." The Journal of Politics 70, no. 04: 1040-1052.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2009. "Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2008." World Bank Policy Research Working Paper 4978.

Kaufmann, Daniel and Kraay, Aart and Zoido, Pablo. October 1999. “Aggregating Governance Indicators.” World Bank Policy Research Working Paper No. 2195.

Keefer, Philip. 2010. DPI 2010: Database of Political Institutions: Changes and Variable Definitions. Development Research Group, World Bank.

Kerner, Andrew. 2009. "Why should I believe you? The costs and consequences of bilateral investment treaties." International Studies Quarterly 53, no. 1: 73-102.

Li, Quan. 2006. "Democracy, Autocracy, and Tax Incentives to Foreign Direct Investors: A Cross-National Analysis." Journal of Politics 68, no. 1: 62-74

Marshall, Monty G., Ted Robert Gurr. 2012. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2012." http://www.systemicpeace.org/polity/polity4.htm

Marshall, Monty G. and Keith Jaggers. 2002. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002."

Marshall, Monty G. 2012. Center for Systemic Peace (CSP) Major Episodes of Political Violence, 1946-2012, URL: www.systemicpeace.org/warlist.htm.

Marshall, Monty G. 1999. Third World War. Lanham, MD: Rowman & Littlefield.

Mickolus, Edward F. 1982. International Terrorism: Attributes of Terrorist Events, 1968-1977 [ITERATE 2]. ICPSR07947-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. http://doi.org/10.3886/ICPSR07947.v1

Morrissey, Oliver, and Manop Udomkerdmongkol. 2012. "Governance, private investment and foreign direct investment in developing countries." World development 40, no. 3: 437-445.

Mosley, Layna and Saika Uno, 2007. Collective Labor Rights Dataset, University of North Carolina and University of Notre Dame.
Mosley, Layna, and Saika Uno. 2007. "Racing to the bottom or climbing to the top? Economic globalization and collective labor rights." Comparative Political Studies 40, no. 8: 923-948.

Mathur, Aparna, and Kartikeya Singh. 2013. "Foreign direct investment, corruption and democracy." Applied Economics 45, no. 8: 991-1002.

Political Risk Services Group. 2011. “International Country Risk Guide.”

Ramasamy, Bala, and Matthew Yeung. 2010. "The determinants of foreign direct investment in services." The World Economy 33, no. 4: 573-596.

Teorell, Jan, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Anna Khomenko & Richard Svensson. 2016. The Quality of Government Standard Dataset, version Jan16. University of Gothenburg: The Quality of Government Institute, http://www.qog.pol.gu.se doi:10.18157/QoGStdJan16

Thorsten Beck, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. "New tools in comparative political economy: The Database of Political Institutions." 15:1, 165-176 (September), World Bank Economic Review.

Treier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." American Journal of Political Science 52: 201-17.

Tollefsen, Andreas Forø, Håvard Strand, and Halvard Buhaug. 2012. “PRIO-GRID: A Unified Spatial Data Structure.” Journal of Peace Research 49(2): 363-374.

Tsebelis, George. 1995. "Decision Making in Political Systems: Veto Players in Presidentialism, Parliamentarism, Multicameralism and Multipartyism." British Journal of Political Science 25: 289-325.

Tsebelis, George. 2002. Veto Players: How Political Institutions Work. Princeton, NJ: Princeton.
## Appendix

### Table A1: List of Component Datasets

| Institution/Authors | Dataset Name | # of countries and years | Topic                                                                 |
|---------------------|--------------|--------------------------|----------------------------------------------------------------------|
| **Economic Datasets** |              |                          |                                                                      |
| Gwartney, Lawson, and Hall | Economic Freedom of the World | 158 | 1970-2013 | Indicators of economic freedom                                     |
| Freedom House       | Freedom in the World | 195 | 1972-2014 | Measures of political and civil liberties                          |
| Heritage Foundation | Index of Economic Freedom | 187 | 1995-2016 | Index of economic freedom                                         |
| Aizenman, Chinn, and Ito | Open Economy Trilemma | 188 | 1960-2014 | Monetary independence, exchange rate stability, and financial openness |
| Singer, Bremer, and Stuckey | COW Material Capabilities | 212 | 1816-2001 | CINC scores                                                        |
| Stockholm International Peace Research Institute | SIPRI Military Expenditures | 172 | 1988-2015 | Military expenditures                                               |
| Claessens and van Horen | Bank Ownership | 137 | 1995-2009 | Annual values for number of foreign banks in each country          |
| Bodea and Hicks     | Central Bank Independence | 82 | 1972-2010 | Central bank independence score for each country                    |
| IMF                 | Contract Intensive Money | 184 | 1948-2012 | Contract intensity of the money supply from the International Financial Statistics |
| Bauer, Cruz and Graham | IMF Agreement | 188 | 1960-2008 | Whether countries are under an IMF agreement in a given year       |
| Beck, Demirgüc-Kunt, and Levine | Financial Development and Structure | 175 | 1960-2011 | Size, activity, and efficiency of financial intermediaries and markets |
| Levy-Yeyati and Sturzenegger | De facto exchange rate regime | 183 | 1974-2004 | De facto exchange rate regimes                                     |
| Graham, Johnston, and Kingsley | LexisNexis-based information Measures | 171 | 1992-2012 | Number of business and investment relative articles written about a country in a given year (English) |
| World Bank          | Newspaper Circulation Data | 192 | 1961-2010 | Data on newspaper circulation                                       |
| Graham, Johnston, and Kingsley | I/B/E/S Analyst Coverage [IB] | 94 | 1993-2015 | Number of unique I/B/E/S analysts covering firms based in a country in a given year. |
| Forbes and Warnock  | Surges, Stops, Flight and | 58 | 1973-2009 | Surges, stops, flights and retrenchments in capital flows.         |
| Source | Data Description | Start Year | End Year | Notes |
|--------|------------------|------------|----------|-------|
| UNCTAD (UN Conference on Trade and Development) | FDI flows & stocks | 197 | 2014 | Inward FDI flows and stocks as reported to the UNCTAD |
| U.S. Bureau of Economic Analysis | Outward FDI Stocks, Historical Cost Basis | 191 | 2014 | Outward FDI stocks from the U.S. |
| U.S. Bureau of Economic Analysis | Outward FDI Flows, by industry | 59 | 2014 | Outward FDI flows from the U.S., separated by industry. |
| U.S. Bureau of Economic Analysis | Outward FDI flows for PPE | 196 | 2013 | Outward FDI flows from the US for net property, plant and equipment |
| US Department of the Treasury | List of Tax Havens | 211 | 2015 | List of know tax havens |
| Ashford | Oil Exports Data | 188 | 2001 | Data on oil revenues generated from the WDI, BP Statistical Index, and the Energy Information Administration |
| Feenstra, Inklaar and Timmer | PENN World Tables 8.0 | 166 | 2011 | A range of economic indicators |
| World Bank | World Development Indicators | 195 | 2014 | Information about national economic development such as food exports, interest payments on external debts, armed forces personnel. |
| Graham | Augmented WDI Data | 161 | 2011 | GDP Total, GDP/Capita, GDP Growth, Population, labor/capital ratio |
| OECD | OECD Statistics | 34 | 2014 | Long-term interest rates |
| International Labor Organization | ILOSTAT | 201 | 2013 | Unit labor cost |
| Li | Tax incentives | 202 | | Measures of certain types of tax incentives |
| World Bank | Doing Business Indicators | 181 | 2013 | Variables that measure barriers to starting a business within each country |
| World Bank | Global Financial Development Database | 197 | 2014 | Characteristics of the banking sector within each country. |
| World Bank | MFN Tariff Data | 183 | 2010 | MFN tariff rates |

**Political Datasets**

| Source | Dataset | Start Year | End Year | Notes |
|--------|---------|------------|----------|-------|
| Alvarez, Cheibub, Limongi, and Przeworski | Regime Type | 135 | 1990 | Regime variables that classify each country's government type |
| Marshall and Gurr | Polity IV | 184 | 2014 | Information about political regimes |
| Data Source | Indicator Description | Start Year - End Year | Notes |
|-------------|-----------------------|-----------------------|-------|
| Coppedge et al | Varieties of Democracy | 170 | 1900 - 2015 | Indicators of democracy |
| Beck, Clarke, Groff, Keefer, et al. | Database of Political Institutions | 176 | 1975 - 2015 | Information about the party system, political parties, and political orientations within each country. |
| Geddes, Wright, and Frantz | Autocratic Regimes | 120 | 1946 - 2010 | Information about political regimes and regime failures |
| Boix, Miller, and Rosato | Regime Type | 211 | 1800 - 2010 | Binary democracy measure |
| Boix, Miller, and Rosato. | System Level Democracy | 194 | 1800 - 2010 | These system-level democracy measures are built from the Boix et al. data. |
| Strom, Gates, Graham and Strand | Powersharing | 175 | 1975 - 2010 | Underlying political structures and their ability to support credible policy commitments |
| Linzer and Staton | Measure of Judicial Independence | 198 | 1948 - 2012 | Latent measure of judicial independence. |
| Kaufmann, Kraay, and Mastruzzi | Worldwide Governance Indicators | 198 | 1996 - 2014 | Measures of political stability, government accountability, and rule of law within a country |
| Marshall and Cole | State Fragility Index | 166 | 1995 - 2011 | Stability of each country based on social, economic, and political factors. |
| Bertelsmann Stiftung | Sustainable Governance Indicators | 41 | 2014 - 2015 | Measures of sustainable governance practices. |
| World Justice Project | Rule of Law Index | 102 | 2013 - 2015 | Measurements on how the rule of law is experienced by the general public. |
| Bertelsmann Stiftung | Transformation Index | 130 | 2006 - 2016 | Indicators of the quality of democracy, a market economy and political management in developing and transitioning countries. |
| World Bank | CPIA Database | 197 | 2005 - 2015 | Measures of the extent to which private economic activity is facilitated by an effective legal system. |
| Graham | Bureaucratic Risk and Policy Risk | 205 | 1988 - 2013 | Political risk indices |
| Cingarelli and Richards | CIRI Human Rights dataset | 200 | 1981 - 2007 | Government adherence to 13 internationally recognized human rights |
| Fariss | Latent Human Rights Protection | 180 | 1949 - 2013 | Information about several different measures of human rights protection |
| Mosley and Uno | Collective Labor Rights | 90 | 1985 - 2002 | Indicators of labor rights for developing nations |
| Reporters Without Borders | Freedom of the Press | 177 | 2012 - 2014 | Alternate freedom of the press measure |
| Freedom House | Freedom of the Press | 198 | 1993 - 2012 | Measures of freedom of the press |
| Hollyer, Rosendorff, and Vreeland | Transparency | 125 | 1980 - 2010 | HRV index value for each country in a given year |
| Dataset Provider                        | Dataset Name                          | Year Range | Description                                                                 |
|----------------------------------------|---------------------------------------|------------|-----------------------------------------------------------------------------|
| Transparency International             | Corruption Perceptions                | 216        | Transparency International corruption index                                 |
| Monty G. Marshall, Oneal, and Russett  | Major Episodes of Political Violence   | 176        | Information on major wars and episodes of political violence in a country   |
| Nordhaus, Marshall                     | MID Propensity                        | 165        | Level of military spending in each country and its effect on democracy in that country. |
| Fortna                                 | UN Peacekeeping                       | 30         | Information on the presence of UN Peacekeeping Forces                       |
| Cornett, Wood and Haschke              | Political Terror Scale                | 192        | Measures of political violence and terror                                  |
| Cornett, Gibney and Haschke            | Societal Violence Scale               | 194        | Measures of societal violence                                               |
| Gleditsch, Wallensteen, Eriksson,      | UCDP/PRIO Armed Conflict              | 118        | Measures of conflict intensity and involvement                             |
| Gleditsch                              | PRIO Battle Deaths Dataset            | 117        | Annual battle deaths                                                        |
| UCDP                                   | UCDP Battle Deaths Dataset            | 94         | Annual battle deaths                                                        |
| Strand                                 | Onset of Armed Conflict               | 176        | List of armed conflict onsets                                               |
| Park                                   | Patent protection index               | 122        | Index of strength of patent protection and patent rights in each country    |
| Zhao                                   | Strength of Property Protection       | 47         | Cross-sectional measure of strength of property rights protection           |
| Schwab, World Economic Forum           | Global Competitiveness Dataset        | 39         | Government transparency                                                     |
| Multiple Sources                       | Bilateral Investment Treaties         | 190        | Number of bilateral investment treaties signed by each country              |
| Graham, IPE Data Resource              | Membership in WTO, IMF, EU, and NATO  | 222        | Dummy variable specifying which international organization each country belongs to |

**Social and Cultural Datasets**

| Dataset Provider                        | Dataset Name                          | Year Range | Description                                                                 |
|----------------------------------------|---------------------------------------|------------|-----------------------------------------------------------------------------|
| Barro and Lee                          | Educational Attainment                | 146        | Information about the years of schooling, primary, secondary, tertiary for population 25 and over |
| OECD                                   | Program for International Student Assessment Scores | 70        | Scores from an international assessment of 15-year-old students’ literacy in reading, mathematics, and science. |
| World Bank                             | Education Statistics                  | 197        | General education statistics                                                |
| Source                                      | Dataset Type                  | Dataset ID | Time Period                  | Description                                                                 |
|---------------------------------------------|-------------------------------|------------|------------------------------|-----------------------------------------------------------------------------|
| Wimmer, Cederman, and Min                   | Ethnic Power Relations       | 165        | 1946-2013                    | Ethnic groups and their political sway in each country                      |
| UN Office on Drugs and Crime                | Crime Data                   | 196        | 2000-2014                    | Measures of the levels of crime.                                            |
| NationsOnline                               | Languages                     | 250        | 1800-2013                    | First two official languages in spoken in each country                      |
| Brown and James                             | Religious Characteristics of States | 198     | 1815-2010                    | Estimates of religious demographics                                         |
| United Nations                              | UN Emigrant Stock Data       | 188        | 1990, 2000, 2010, 2013       | Global stock of emigrants from each country                                |

### Geographic Datasets

| Source                                      | Dataset Type                  | Dataset ID | Time Period                  | Description                                                                 |
|---------------------------------------------|-------------------------------|------------|------------------------------|-----------------------------------------------------------------------------|
| Graham                                      | Archipelagos Data             | 58         | N/A                          | Number of archipelagos                                                     |
| Mayer and Zignago (CEPII)                   | Bilateral distance to U.S.    | 222        | 1800-2015                    | Geographic distance, linguistic ties and historical ties to U.S.           |