Investor-State Dispute Settlement
Codebook

To cite:
Wellhausen, Rachel L. 2016. "Recent Trends in Investor-State Dispute Settlement."
Journal of International Dispute Settlement 7(1): 117-135. 2016.

To contact with updates or corrections:
Rachel Wellhausen
University of Texas at Austin
rwellhausen@utexas.edu

Note:
The March 2016 release contains cases through 31 December 2015. One earlier
case, from 1987, is also included.

Overview:
This database compiles information on known international investment arbitrations
filed by foreign investors against states. The time period analyzed starts with 1
January 1990. Variable codes are true as of 31 December of the final year of the data.

The data include arbitrations at venues including the International Centre for
Settlement of Investment Disputes (ICSID), the International Chamber of Commerce
(ICC), the Stockholm Chamber of Commerce (SCC), the Permanent Court of
Arbitration (PCA), the London Court of International Arbitration (LCIA), regional
arbitration centers, and ad hoc tribunals. The strategy for compiling public cases
included (1) assembling data on ICSID cases, to which the parties and often details
of the case are public knowledge; (2) collecting any public data provided by other
investment arbitration venues; (3) collecting information from news aggregators
(LexisNexis and Factiva); (4) collecting information via Google news alerts for
"expropriation," "nationalization," and "investment arbitration"; and (5) collecting
information from specialty sources. The following are particularly useful specialty
sources:
   - Luke Eric Peterson's Investment Arbitration Reporter (www.iareporter.com)
   - Andrew Newcombe's Investment Treaty Arbitration Law (www.italaw.com)
   - UNCTAD's Database of Investor-State Dispute Settlement
     (www.unctad.org/en/pages/diae/isds.aspx)
   - Global Arbitration Review (www.globalarbitrationreview.com)
   - Todd Weiler's NAFTA claims database (www.naftaclaims.com)
   - Specific company directories

The database includes any public instance in which a foreign investor sues a
sovereign state in an international tribunal. Thus, the database includes investment
treaty arbitration, which is triggered by a treaty such as a Bilateral Investment
Treaty (BIT), as well as other investment arbitrations brought under Investor-State
Dispute Settlement (ISDS) based on contractual provisions or domestic rules. The

database includes arbitrations in which the respondent is a sub-national government.

The database does not include arbitrations in which the respondent is a state-owned enterprise (SOE). The database does not include instances where foreign investors sue states for enforcement of private commercial arbitration awards. In general, no international commercial arbitration is included. Arbitrations filed by the state on behalf of foreign investors, for example at the International Court of Justice (ICJ), are not included. The database does not include arbitrations filed by states against investors in international tribunals. (This is not possible under many investment agreements, and only two such arbitrations have been filed at ICSID, ARB/76/1 and ARB/07/3.)

The database attempts to include only arbitrations that were actually filed. Thus, if a foreign investor offers a public notice of intent to file but does not actually file, it is not included. This distinction is clear when it comes to arbitrations filed at ICSID. It is less clear when arbitrations are filed at other less public venues. The coding process errors on the side of inclusion when there is evidence of intent to file but ambiguity about the subsequent process.

This is part of an ongoing data collection process. New information may be revealed in the future that resolves missing data here. Please contact Rachel Wellhausen (rwellhausen@utexas.edu) with updates or corrections.

Variable definitions:

Filed (Year filed)
This variable identifies the four-digit year in which the case was filed at an arbitration venue. The variable is coded "unknown" if the year a case is filed is not public.

End (Year end)
This variable identifies the four-digit year in which the case was concluded in the first instance. (That is, if a case is continuing into annulment proceedings, it is captured elsewhere in the data set.) Conclusion includes several possible outcomes captured elsewhere in the dataset. Ongoing arbitrations are coded "pending." The variable is coded "unknown" if public evidence shows that the arbitration process is not ongoing but the exact year in which the process stopped is not public. The variable is coded "unknown" if no more than five years have passed since the last available information on the arbitration process, but it is not clear whether the process is still pending. The "unknown" code is replaced with the four-digit year that marks five years since the last public information is available on the arbitration process. These dormant cases are assumed to be settled.

Host (Host State ISO3)

This variable identifies the state against which the foreign investor has filed the dispute. The variable is coded according to the International Standards Organization (ISO) 3-digit alphabetic codes.

Home (Home State ISO3)
This variable identifies the country of origin of the foreign investor that has filed the arbitration. The variable is coded according to the International Standards Organization (ISO) 3-digit alphabetic codes. The exception the coding rule is "BELLUX" as the code for arbitrations brought under investment treaties signed jointly by Belgium and Luxembourg.

For arbitrations under a Bilateral Investment Treaty, country of origin is determined by the state that is party to the relevant treaty. For arbitrations brought under a multilateral investment treaty or that access ISDS via other means, the country of origin is determined by: (1) the country of origin as listed in publicly available court documents; or (2) the country of origin as described in specialty press and news sources. Note that because many foreign investors have access to multiple legal means to file investment arbitration, the code need not match any popularly understood home country of the investor. The variable is coded "unknown" if public information is unavailable.

Home2 – Home5
These variables identify other countries of origin of foreign investors that are explicit parties to the arbitration. Multiple investors from the same country of origin that are parties to the same arbitration are not coded in Home2 – Home5. (The quantity of investors that are party to the arbitration is not captured in the database.) Investors from different countries of origins are coded. Additionally, at times a single investor invokes more than one bilateral treaty in filing for arbitration. In that case, Home2 – Home5 code all of the countries of origin invoked by the single investor. The variable is coded "n.a." if there are not additional parties to the arbitration from different national origins to code. The variable is coded "unknown" if there is evidence that an additional foreign investor is party to the arbitration but its country of origin could not be identified. For more information, see the description for the variable "Home."

Parties (Parties to arbitration)
This variable identifies the parties to the arbitration. Where there are many claimants, it summarizes them with the abbreviation et al. Text is taken from various sources, with the implication that it is nonsystematic.

Venue (Venue or rules adopted)
This variable identifies the venue of arbitration and/or the arbitration rules to be applied. Arbitrations at ICSID have four different possible codes. First, the code "ICSID" refers to cases that are registered for arbitration under the ICSID Convention Arbitration Rules. Second, "ICSID CR" code is for cases that are registered for conciliation under the ICSID Convention Conciliation Rules. Third,

"ICSID AF" identifies cases that are commenced under the ICSID Additional Facility, under the Additional Facility Arbitration Rules. Fourth, cases that are registered at ICSID and will follow the UNCITRAL rules are coded "ICSID UNCITRAL."[1] If it is known that arbitration is brought under UNCITRAL rules, but it is unclear in what forum that arbitration is heard, the coding is "UNCITRAL." Other codes for the arbitration tribunals in which the arbitration is brought include:

- AIC: Arab Investment Court
- ICC: International Camber of Commerce
- LCIA: London Court of International Arbitration
- MCC: Moscow Chamber of Commerce
- PCA: Permanent Court of Arbitration
- SCC: Stockholm Chamber of Commerce
- SCC Emergency Arbitration: Emergency Arbitration at the Stockholm Chamber of Commerce
- ZCC: Zurich Chamber of Commerce
- Ad hoc: Arbitration is not brought within a formal tribunal and is rather adjudicated by an ad hoc panel that does not publicly specify the applicable arbitration rules.

An investment arbitration in which it is not public information in what forum and/or under which rules it is brought is coded "unknown."

Caseno (Case no.)
This variable identifies an identifying case number for the arbitration, if one is provided by the arbitration venue. The variable is coded "n.a." if a case number is either unknown or not available.

Legal (Legal doc. invoked)
This variable identifies the instrument(s) that the claimants have invoked for purposes of consent to arbitration or conciliation when the arbitration is registered. Multiple applicable legal instruments are separated by commas.

In many arbitrations, the instrument is the Bilateral Investment Treaty (BIT) between the home country of the claimant and the host country in which the investment is located. Where possible, the particular BIT is specified.

In other arbitrations, the instrument invoked is the contract between the parties, which specifies the ISDS arbitration procedure to be followed in case of dispute. These cases are coded as "Contract." Where arbitrations are brought at ICSID, we rely on ICSID's coding of "Contract"; when brought elsewhere, we code "Contract" if it is specified as the legal basis of the arbitration in available sources. Additionally, if the parties asked the tribunal to arbitrate the dispute in an arbitration venue with a submission agreement, the variable is coded "Submission Agreement." If both a

---

[1] For details on the different type of arbitration rules at ICSID, see
https://icsid.worldbank.org/apps/ICSIDWEB/icsiddocs/Pages/Overview.aspx.

contract/agreement and an international treaty are invoked in filing, we give preference to the international treaty code.

In some instances the instrument invoked is a domestic investment law in the host country. Where arbitrations are brought at ICSID, we rely on ICSID's coding of domestic investment law; when brought elsewhere, we code appropriately based on available sources. If both a domestic investment law and an international treaty are invoked in filing, we give preference to the international treaty code.

If a preferential trade agreement (PTA) is the instrument invoked at filing, we code as "PTA" or with the name of the agreement, including but not limited to "NAFTA" and "CAFTA-DR." If the Energy Charter Treaty is invoked at filing, we code appropriately.

Finally, some arbitrations are coded as "Other (as coded by ICSID)." This replicates the ICSID coding. These can be thought of as equivalent to "unknown." For arbitrations at any venue (ICSID or elsewhere), if the legal instruments invoked are not specified and/or documents on the arbitration process are not public, this variable is coded "unknown."

Invsttype (Text on investment type)
This variable uses descriptive text to characterize the type of investment at stake in the arbitration. For ICSID arbitrations, the source for the variable is the subject of the dispute provided by ICSID. For other arbitrations, the characterization is based on documents in the arbitration process (such as the Notice of Arbitration or Final Award) or on other sources. If the type of the investment at stake is not specified, the variable is coded "unknown."

Sector (Sector 3-letter code)
This variable identifies the economic sector of the investment with which the arbitration is concerned. The possible options are agriculture (AGR), mining (MIN), petroleum and natural gas (OIL), manufacturing (MFG), utilities (UTL), construction (CON), telecommunications (TEL), financial sector (FIN), and other services (SVC). For ICSID cases the source of information is the economic sector that ICSID provides on its website. For cases in other arbitration venues, the characterization is based on sources as described above. Note that because the sector is coded based on investment characteristics, the code may be in an industry other than the one typically associated with the claimant's business. If the sector of the investment is not specified, this variable is coded "unknown."

A few codes that may not be obvious to the user are specified here. In other cases, the user can draw on the "Invsttype" variable to gain more information.
- Utilities: Types of investments include highways, penitentiaries, railroads
- Services: Types of investments include airport concessions, port concessions, tourism such as construction/operation of hotels and resorts

- Real estate: Types of investment include purely construction (and not operation) projects, property management
- If the type of investment could be coded as manufacturing as well as another sector, manufacturing is the privileged code. For example, this is applicable to some investments that could be thought of as manufacturing and services.

Asought_text (Text on award sought)
This variable uses descriptive text to specify the award that the investor seeks, in the currency in which the investor seeks it. It takes this descriptive text directly from available sources. The text may or may not include a range of compensation sought or claims of seeking "up to" or "over" a certain value. (In general, the data does not code whether the investor seeks interest on the final award and arbitration costs; this is true in nearly all instances. The norm is to include penal interest that has accrued to the point of the claim in the total compensation sought.) The variable is coded "unknown" if public information is not available.

Asought_usm (Award sought in US$m, exchange rate at time of filing, less interest/costs)
This variable quantifies the text in "Asought_text" in millions of US dollars at the exchange rate at the time the arbitration is registered if necessary.[2] If "Asought_text" contains a range, the variable codes the minimum requested by the investor. In general, the goal of this variable is to capture the absolute minimum the investor sees at stake in the arbitration. It does not quantify any additional claims for interest on the final award or arbitration costs. If "Asought_text" is "unknown," this variable is also coded "unknown."

Concluded (Arbitration concluded)
This is a dummy variable that captures whether the entire arbitration process, including any possible annulment or set-aside proceedings, has concluded. The variable is coded "yes" if the initial arbitration is concluded and no additional proceedings occurred, or if additional proceedings are over. The variable is coded "no" if either the initial arbitration or additional proceedings are ongoing.

Winner (Status/winner pre-annulment)
This variable codes the manner in which the arbitration ended in the first instance. The variable is coded "pending" if the arbitration process is ongoing (as of 31 December 2014). It is coded "settlement" if the case was discontinued and/or there is public information that the parties reached an agreement and withdrew the case before a final ruling. The variable is coded "investor" if the tribunal reaches a final ruling and determines that the state was at fault. (This coding is not affected by the size of the award won by the investor.) This variable is coded "state" if the tribunal's final ruling does not determine that the state was at fault. Reasons for a "state" code include, but are not limited to, the tribunal finding that the investor does not

---

[2] Values are rounded to one decimal point. For exchange rates, see xe.com or fxtop.com.

deserve compensation, that the tribunal does not have jurisdiction, or the that tribunal finding that the investor did not have adequate standing to file the arbitration. The variable is coded "unknown" if the arbitration process is known to have ended, but the outcome of the process is not public.

Awardinvst_text (Text on award to investor)
This column uses descriptive text to specify the amount if any that is awarded to the investor at the close of the arbitration. The amount is provided in the currency in which the award is rendered. The column is coded "settlement if "Winner" is coded "settlement." (The terms of the settlement are included if publicly available.) The column is coded "n.a." if "Winner" is coded "pending." If the investor is known to have won an award but the amount of that award is unknown, or if "Winner" is coded "unknown," then the variable is coded "unknown." Otherwise, the variable is coded "0" if "Winner" is not coded "investor."

The descriptive text may also include notes on how the panel decided to rule on the legal costs incurred by the parties in the course of the arbitration and any costs of the arbitration itself, but this data has not been systematically collected.

Awardinvst_usm (Award to investor in US$m, not including costs/interest)
This variable quantifies the text in "Awardinvest_text" and is coded with a numeric value in millions of US dollars at the exchange rate at the time the arbitration is concluded. [3] The value coded is the base award and does not include additional sums that may include interest and/or awards for legal costs. In general, the goal of this variable is to capture the absolute minimum, base award the investor won in arbitration. In other respects the coding matches "Awardinvst_text."

Annul (Annulment filed)
This variable identifies whether either party to the arbitration submitted an application for annulment of an award rendered by an ICSID tribunal. The variable is coded "yes" if an annulment process was initiated. The variable is coded "no" if tge arbitration was heard at ICSID and no party initiated an annulment process. The variable is "n.a." if the arbitration was not heard at ICSID. It is also coded "n.a." if an ICSID arbitration is still pending.

In these cases, the subsequent columns are also coded "no" in order distinguish them clearly from the pending arbitration cases.

Annul_filed (Year annulment filed)
This variable identifies the four-digit year that the annulment process was initiated if the variable "Annul" is coded "yes." The variable is coded "n.a." otherwise.

Annul_end (Year annulment ends)

---

[3] Values are rounded to one decimal point. For exchange rates, see xe.com or fxtop.com.

This variable identifies the four-digit year that the annulment process concluded if the column "Annul" is coded "yes." The variable is coded "pending" if the annulment process is ongoing. The variable is coded "n.a." otherwise.

Annulstatus (Status of annulment)
This variable identifies the status of the annulment process, if "Annul" is coded "yes." The variable is coded "upheld" if the annulment process concluded either in a ruling or a dismissal and the original ruling was upheld. The variable is coded "overturned" if the annulment process concluded and the original ruling was overturned. The variable is coded "partially overturned" if the annulment process concluded and the original ruling was partially overturned. The variable is coded "discontinued" if one or both parties requested the discontinuation of the annulment process. The variable is coded "pending" if the annulment process is ongoing. The variable is coded "n.a." otherwise.