



OPIC Scraped Portfolio Database:
Methodological Note

Ben Leo and Jared Kalow

CENTER FOR GLOBAL DEVELOPMENT

I. Introduction

The Overseas Private Investment Corporation (OPIC) is the development finance institution of the United States government. Its mandate is to “mobilize private capital to help address critical development challenges” and to “advance U.S. foreign policy and national security priorities.” OPIC pursues this mandate by “providing investors with financing, political risk insurance, and support for private equity investment funds, when commercial funding cannot be obtained elsewhere.”¹ The agency operates on a self-sustaining basis and has provided net transfers to the US Treasury for nearly 40 consecutive years. Since its inception, it has helped to mobilize more than \$200 billion of US investment through over 4,000 development-related projects. In 2014, OPIC committed roughly \$3 billion to support 78 projects in over 40 different developing countries.²

Project-level data is available on a piecemeal basis on the OPIC website. Annual reports contain basic information for all projects approved during the respective fiscal year. OPIC also publishes annual policy reports with summary statistics on development impact. Additional information is available in the form of summary descriptions for projects approved after April 2009. While these descriptions contain the most comprehensive project-level information, they are available only in a downloadable PDF format. The OPIC website also has a database containing active projects, which contains limited information.

Despite these disparate information sources, there is not a comprehensive database that contains all public information on OPIC projects in a single, readily accessible format. This gap prevents stakeholders – including policymakers, researchers, businesses and investors, and others – from engaging with OPIC in a rigorous, data-driven manner. To address this need, we have developed the *OPIC Scraped Portfolio* database, a comprehensive source that includes all publicly available information on OPIC projects, with supplementary country-level data. This new resource is public and downloadable in multiple formats.

This brief note details the methodology, data sources, and decision points associated with the *OPIC Scraped Portfolio* database. It is organized as follows. Section II provides a general overview of publicly available project-level information and our approach for addressing any discrepancies between data sources. Section III outlines the supplemental project-level variables, data sources, and methodology. Lastly, section IV corresponds to country-level indicators.

II. Scraping the OPIC Data

A. OPIC Annual Reports

To build the *OPIC Scraped Portfolio* database, we started by gathering all available project-level information. We collected the first round of data from OPIC’s annual reports, which are available in PDF format from 2000 to 2014 on their website. Each annual report contains a table of information of all projects approved during that calendar year, with the following data fields:

¹ www.opic.gov

² Authors’ calculations.

- Project name
- US sponsors
- Short project description
- OPIC commitment amount
- Type of project (finance, investment fund, or insurance)

B. OPIC Project Summaries

For projects approved by the OPIC Board after April 2009, we scraped more information data from project summaries. These typically are 1-2 page PDF documents available on the OPIC website. These project summaries are published at least 40 days prior to formal Board consideration and typically include the following information:

- Total project size
- Specific type of finance product (loan or guaranty)
- Extended project description
- Developmental effects
- U.S economic effects
- Environmental risk category

C. Reconciling OPIC Commitment Size Discrepancies

In some instances, there are discrepancies between annual reports and project summaries in terms of the reported OPIC commitment size. This is primarily due to the advance disclosure of project summaries, which means that the OPIC Board or staff could alter the proposed commitment size during or after the formal project consideration process. In these cases, we rely upon information from the annual reports since they were published at a later date.

In several cases, the annual report's OPIC commitment exceeded both the project size and the OPIC commitment stated in the project summary. In these cases, we increased the project size to match the new OPIC commitment since the total project size must be *at least* as large as the OPIC commitment. For example, the project summary of an insurance project sponsored by Belstar Capital Limited recorded the OPIC commitment as \$180 million, with total project costs of \$180 million. The annual report reported OPIC's commitment as \$286.4 million. In this case, we increased both the OPIC commitment and the project size to \$286.4 million.

In addition, there are several projects that had project summaries but did not appear in the annual report. We assume that the Board did not approve these projects and, therefore, exclude them from the *OPIC Scraped Portfolio* database.

III. Supplemental Project-Level Data

Drawing upon information from the OPIC annual reports and project summaries, we added several additional data fields. These include: (1) sector classification; (2) US sponsor type; (3) Fortune 500 status; (4) leverage ratio; (5) development effect descriptor.

A. Sector

We developed a sector tree methodology (see appendix II) to classify OPIC projects into 16 distinct categories. The publicly disclosed project descriptions underpin these classification decisions. When appropriate, we classified projects into sub-sectors and secondary sub-sectors. There were several cases where the sector classification was unclear, so we developed rules for each of these cases:

- Sector-Based Lending Projects: These projects (e.g., agriculture lending facilities) are coded as a financial services sector project, with the thematic area listed as the respective sub-sector. By illustration, a social investment facility that provides lending to SME-operated healthcare facilities is coded as a ‘financial services’ project in the ‘healthcare’ sub-sector.
- Sales Projects: These projects are coded as retail sector projects, and then the specific sector in which they operate is coded in the sub-sector. For example, an irrigation sales project is coded in the ‘financial services’ sector, in the ‘water’ sub-sector, and in the ‘equipment’ sub-sector 2.
- Banking Projects: Banking sector projects are classified into small- and medium-sized enterprises (SMEs), microfinance institutions (MFIs), housing, leasing, or consumer loans in the sub-sector 2 field. Projects that explicitly cover more than one of these lending segments (e.g., a project that offers microfinance and SME loans) are coded as ‘general’ in the sub-sector 2 field.
- Investment Funds: Large investment funds with no discernible sector-based focus are typically coded as ‘general’.

B. US Sponsor Type

We classified US sponsors into three categories: individual, NGO, and corporation. When we were unable to make a sponsor name-based decision, we ran an internet query to identify whether the related sponsor has a .com or .org URL address. Overall, the US sponsor type field entails a relatively high room for error. There is little information on many US sponsors, and US sponsors may not fit neatly into these categories. By default, the type is corporation, and we only classify sponsors as NGOs or individuals when it is obvious. There are also several cases of mixed sponsor categories (for example, one corporation and two individuals).

C. Fortune 500 Status

We determined whether a Fortune 500 company sponsored a project using a set of historical Fortune 500 lists dating back to 2000.³ There are some cases where the Fortune 500 company is mentioned more than once in the respective annual list. In these cases, we adjusted the name of the Fortune 500 companies to match the name of the U.S. sponsor. For example, we ensured that Citibank projects matched to Citibank as a Fortune 500 company, and used a matching function in Stata to mark all Fortune 500-sponsored project as such.

D. Leverage Ratios

Project leverage rates are calculated as the ratio of the total project size to the OPIC commitment. In general terms, there is no uniform methodology for calculating leverage ratios; this is one of several potential approaches.⁴ However, the ratio of total funding to public funding is one of the common ways of calculating leverage ratios. Since OPIC has only disclosed total project sizes in the project summaries, we were unable to calculate leverage ratios for projects approved before 2009.

E. Developmental Effect Descriptor

Development effect category classifications are based on the descriptive adjective used in the project summaries. In each document, there is common language for describing the developmental effect, usually along the lines of, “this project is expected to have [adjective] development impact.” Based on discussions with OPIC staff, these descriptive adjectives likely suggest the following categories:

- “Highly Developmental”: *high, highly, significant, strong, and upper end.*
- “Developmental”: *moderate, positive, some, and substantial.* We also assume that the lack of descriptive adjectives suggests that the project is in the “developmental” category.
- “Indeterminate”: *minimal.*⁵

F. Detailed Project-Level Information

We also included three text categories from the project summaries. These include: extended project descriptions, developmental effects, and U.S. economic effects. Each of these categories provides deeper information on each project.

³ “Fortune 500 Companies - Archived List of Best Companies from 1995”; “Downloads.” Fortune 500 provides historical lists from 1955 to 2005 on their website. Lists of Fortune 500 companies for all years are available on another site (TopForeignStocks.com) and match up to the Fortune 500 lists from 2000-2005.

⁴ Brown and Jacobs, “Leveraging Private Investment: The Role of Public Sector Climate Finance.”

⁵ There are two other projects coded as indeterminate with clearance in progress, one project with multiple downstream investments to be scored separately, and one project that was not scored on the developmental effect matrix.

IV. Country-Level Data

To supplement OPIC project-level data, we have incorporated a range of country-level data on commercial and political risks, private credit depth, and other ‘additionality’ indicators. This information is gathered from a range of sources, including: the World Bank, the Bank for International Settlements, and Delcredere Ducroire (Belgian credit agency). The source for the indicators include:

- World Bank: GNI per capita, income classification⁶, population, private credit depth, GDP, FDI
- Bank for International Settlements: US bank exposure
- Bureau for Economic Analysis: US FDI stock
- Delcredere Ducroire: War Risk, expropriation risk, transfer risk, commercial risk

A. Country Level Risk

Country-level risk data comes from Delcredere Ducroire, the Belgian public credit insurer. Delcredere evaluates all developed and developing countries for commercial risk, war risk, expropriation risk, and transfer risk. A previous review of political risk indicators found that Delcredere was the best public source for country-level risk data. Moreover, plant location consultants often use their ratings to evaluate risks even if their clients ultimately do not purchase risk insurance.⁷ In addition, Delcredere prices for insurance reflect other agencies’ insurance prices, implying that their risk rankings may be in line with other risk measures.⁸ The current risk ratings are available on their database. We were able to obtain historical data by directly contacting Delcredere Ducroire.

B. Bilateral Investment Treaty and OECD Country Status

We include two binary variables for whether the developing country has a bilateral investment treaty (BIT) with the United States and whether it is an OECD member. The BIT coding reflects whether the agreement had entered into force prior to the respective OPIC project being approved. The same approach was applied to the OECD variable. There were no projects in OECD countries the same year that the country joined the organization, meaning that all projects were definitively approved when the partner country was an OECD member.

⁶ We utilized the World Bank’s historical “analytical classifications” for all countries. In some countries, the World Bank has made ex-post adjustments to GNI per capita figures due to GDP rebasing exercises or statistical adjustment reasons. The historical World Bank analytical categories correspond to a given country’s classification at the time of OPIC board approval. For additional details, see [the World Bank’s note on analytical classifications of countries](#).

⁷ Jensen, “Measuring Risk.”

⁸ Ibid.

C. Presidential Initiatives: Power Africa and Feed the Future

Similarly, we coded binary variables for Power Africa and Feed the Future based on whether the project occurred in a focus country of either US development initiative. For Power Africa, we coded power projects in sub-Saharan Africa in 2013 and later. For Feed the Future, we coded agriculture projects in Feed the Future focus countries in 2009 and later. As a disclaimer, there were multiple power- or agriculture-related projects that were approved during the first year of both initiatives. It was not clear whether these projects were approved before or after the initiatives were announced. Despite this, we coded all qualifying projects in that year as part of the initiative.

Metadata Chart

Data	Source	Notes
Year	OPIC Annual Report	
Region	OPIC Annual Report	The annual reports divide projects into 6 regions (Africa and the Middle East, Asia and the Pacific, Europe and Eurasia, Latin America and the Caribbean) and a separate category for Global. We further separate 1) Africa and the Middle East into sub-Saharan Africa and MENA, and 2) Europe and Eurasia into Europe and NIS.
Country	Annual Report	
GNI Per Capita	World Bank	
Income Classification	World Bank	
Income Classification (2)	Internal classification	LIC = 1, LMIC = 2, UMIC = 3, HIC = 4
OECD	OECD/internal classification	There were no “overlap” projects – projects approved in the same year the country joined OECD.
Private Credit Depth	World Bank	
US FDI Stock	Bank for International Settlements	
US Bank Exposure	Bureau of Economic Analysis	
US BIT	State Department/internal classification	We consider first of year for the BIT treaty the year that the BIT entered into force, rather than the year it was signed.
Doing Business Rank	Doing Business	Rankings are adjusted to be on a 1-189 scale for each year.
DD War Risk	Delcredere Ducreire	Data requested from Delcredere Ducreire.
DD Expropriation Risk	Delcredere Ducreire	Data requested from Delcredere Ducreire.
DD Transfers Risk	Delcredere Ducreire	Data requested from Delcredere Ducreire.
DD Commercial Risk	Delcredere Ducreire	Data requested from Delcredere Ducreire.
Population	World Bank	
GDP	World Bank	
Total FDI	World Bank	
OPIC Commitment as a % of FDI	Internal classification	
Type	OPIC Annual Report	
Type (Details)	OPIC Project Summary	
Sector	Internal classification	Based on the project description, we classified each project into a sector using our sector tree.
Sub-Sector/Sub-Sector 2	Internal classification	Based on the project description, we classified each project into sub-sectors when appropriate using our sector tree.
US Sponsor Type	Internal classification	Based on the US sponsors, we assessed whether the US sponsor was a corporation, NGO, individual, or some combination of the three. The default type is corporation.
US Sponsor (1-6)	OPIC Annual Report/OPIC Project Summary	In some cases, the annual reports only provide 1-2 sponsors and the project summaries provide additional sponsors.
Project Name	OPIC Annual Report	
Project Description (Annual Report)	OPIC Annual Report	The annual reports provide a concise project description.
Project Description (Project Summary)	OPIC Project Summary	The project summaries usually provide an extended project description. Text is copied directly from project summary.
OPIC Commitment, Unadjusted	OPIC Annual Report	

Total Project Size, Unadjusted	OPIC Project Summary	
OPIC Commitment	CPI index/internal calculation	
Total Project Size	CPI index/internal calculation	
OPIC Leverage Ratio	Internal classification	Calculated as the proportion of the total project size to the OPIC commitment.
% of Annual Commitments	Internal classification	
Loan Term	OPIC Project Summary	Only available for some projects.
Environmental Risk Category	OPIC Project Summary	There are 71 projects between 2009 and 2014 with no environmental risk category in the project summary.
Developmental Effects	OPIC Project Summary	Text is copied directly from project summary.
US Impact	OPIC Project Summary	Text is copied directly from project summary.
Developmental Effect Descriptor	OPIC Project Summary	Determined by the adjective preceding developmental effect in the “Developmental Effect” category of the project summary.
Feed the Future	Feed the Future/ internal classification	Projects coded 1 are agriculture (sector) projects in focus countries in 2009 or later. Projects coded 2 are agriculture projects in focus countries before 2009.
Power Africa	Power Africa/ internal classification	Projects coded 1 projects are power (sub-sector) projects in sub-Saharan Africa in 2014 or later. Projects coded 2 are power projects in sub-Saharan Africa before 2014.
Project Summary	OPIC Project Summary	Link to project summary
Annual Report Year	OPIC Annual Report	
Project Summary Year	OPIC Project Summary	In some cases, there was a discrepancy in the year between the annual report and the project summary; we always rely on the annual report.
Fortune 500	Internal classification	Fortune 500 status determined by US sponsor placement on the Fortune 500 list in the given year
Credit Depth Quartile	Internal classification	Calculated based on each country’s credit depth percentile in the universe of countries for a given year
US Bank Exposure Quartile	Internal classification	Calculated based on each country’s US Bank Exposure percentile in the universe of countries for a given year

Sector Tree Classifications



Works Cited

Brown, Jessica, and Michael Jacobs. "Leveraging Private Investment: The Role of Public Sector Climate Finance." Background Note. Overseas Development Institute, April 2011.
<http://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7082.pdf>.

"Downloads." TopForeignStocks.com. Accessed February 9, 2016.
<http://topforeignstocks.com/downloads/>.

"Fortune 500 Companies - Archived List of Best Companies from 1995." *Fortune 500*. Accessed February 9, 2016.
http://archive.fortune.com/magazines/fortune/fortune500_archive/full/1955/.

Jensen, Nathan M. "Measuring Risk: Political Risk Insurance Premiums and Domestic Political Institutions." *International Finance*, 2005.
<https://ideas.repec.org/p/wpa/wuwpif/0512002.html>.