Geocoded Official Development Aid Dataset (GODAD)

Codebook – Version 1.0, October 31, 2024

1. Introduction

The GODAD provides geo-localized information on aid projects from 18 European donors and the United States for the 1973-2020 period, based on the OECD's Creditor Reporting System (CRS), data from the Agence Française de Développement for French aid projects, as well as recipient country information for multiple donors from Aid Information Management Systems (AidData 2016, 2017). We geo-locate CRS data via a natural language processing approach and provide auxiliary information about projects along several dimensions, including donor and recipient countries, donor agencies, flow type (grants, loans, other official flows), sectors and sub-sectors of aid, and commitment and disbursement amounts. We also include data on aid projects from China (2000-2021), India (2007-2014), and the World Bank (1995-2023). Please cite the original sources of all data, as we indicate below.

Please cite the GODAD data as:

Bomprezzi, Pietro; Dreher, Axel; Fuchs, Andreas; Hailer, Teresa; Kammerlander, Andreas; Kaplan, Lennart; Marchesi, Silvia; Masi, Tania; Robert, Charlotte; Unfried, Kerstin (2024). Wedded to Prosperity? Informal Influence and Regional Favoritism. *CEPR Discussion Paper 18878.*

2. List of files included in this release

File	Details
GODAD_projectlevel.dta GODAD_projectlevel.csv	Project-location level dataset.
GODAD_adm1.dta GODAD_adm1.csv	ADM 1-year level datasets.
GODAD_adm2.dta GODAD_adm2.csv	ADM 2-year level datasets.

3. Project-level dataset: Projects from all donors

Variable name and sources	Details
project_id World Bank: AidData 2017, IATI 2023 China: Dreher et al. 2022, Custer et al. India: Asmus et al. 2024 EU/US: OECD 2023 France: Agence Française de Développement 2024	Identifier for distinct projects. The project identification key is based on "id" (World Bank), "id" (China), "aiddata_project_id" (India), "CrsID"/ "Donor"/ "Recipient" (CRS EU/US), and "id_projet" (France).

project_location_id World Bank: AidData 2017, IATI 2023 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024	Identifier for distinct project locations. Constructed as combination of *project_id* and a numerical identifier which takes values from 1 to *n*, for each unique location within a project. N/A for Agence Française de Développement data.
donor	Name of donor country.
gid_0 GADM 2023	Country string ID for country where a project is located, using the GADM version 3.6 definition. ISO 3166-1 alpha-3 country code when available (ADM0 layer).
gid_1 GADM 2023	ADM1 region string ID for the first administrative division where a project is located, using the GADM version 3.6 definition (ADM1 layer).
gid_2 GADM 2023	ADM2 region string ID for the second administrative division where a project is located, using the GADM version 3.6 definition (ADM2 layer).
name_0 GADM 2023	Name of the country where a project is located, using the GADM version 3.6 definition (ADM0 layer).
name_1 GADM 2023	Name of the ADM1 region where a project is located, using the GADM version 3.6 definition (ADM1 layer).
name_2 GADM 2023	Name of ADM2 region where a project is located, using the GADM version 3.6 definition (ADM2 layer).
latitude World Bank: AidData 2017, Kersting & Kilby 2021 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024 France: Agence Française de Développement 2024	Geographic latitude.
longitude World Bank: AidData 2017, Kersting & Kilby 2021 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024 France: Agence Française de Développement 2024	Geographic longitude.
startyear World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023 France: Agence Française de Développement 2024	Year in which the project is committed, based on "boardapprovaldate" and "Date" (World Bank), "Commitment Year" (China), "year" (India), "CommitmentDate" (CRS EU/US), "date_d_octr" (France).

closingyear World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023 France: Agence Française de Développement 2024	Year in which the project ends, based on "closingdate" (World Bank), "Completion Year" (China), "CompletionDate" (CRS EU/US), "date_d_ache" (France). N/A for India.
paymentyear World Bank: AidData 2017, Kersting & Kilby 2021 EU/US: OECD 2023 France: Agence Française de Développement 2024	Year in which a disbursement occurs, based on "date" or "startyear" (World Bank), "Year" (CRS EU/US), "date_d_octr" (France). N/A for China and India. In the case of France, it corresponds to the year of the first disbursement.
title World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023 France: Agence Française de Développement 2024	Title of the project, based on "project_name" (World Bank), "Title" (China), "project_title" (India), "ProjectTitle" (CRS EU/US), "nom_du_proj" (France).
description World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 EU/US: OECD 2023 France: Agence Française de Développement 2024	Detailed description of the project, based on "pdo" (World Bank), "Description" (China), "LongDescription" (CRS EU/US), "description" (France). N/A for India.
receiving_agencies World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 France: Agence Française de Développement 2024	Agencies receiving the transfer, based on "borrower" (World Bank), "Direct Receiving Agencies" (China), "libelle_ben" (France). N/A for India and EU/US.
implementing_agencies World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 EU/US: OECD 2023	Agencies responsible for implementing project, based on "impagency" (World Bank), "Implementing Agencies" (China), "ChannelReportedName" (CRS EU/US). N/A for India and France.
funding_agencies World Bank: AidData 2017, Kersting & Kilby 2021 India: Asmus et al. 2024 China: Dreher et al. 2022, Custer et al. 2023 EU/US: OECD 2023 France: Agence Française de Développement 2024	Agency issuing the transfer, based on "lendinginstr" (World Bank), "Funding Agencies" (China), "agencyname" (India), "AgencyName" (CRS EU/US), "societe" (France).
financial_type World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 EU/US: OECD 2023 France: Agence Française de Développement 2024	Type of financial transfer, based on "projectfinancialtype" (World Bank), "Flow Type" (China), "FlowName" (CRS EU/US), "groupe_de_p" (France). N/A for India.

flow_class China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023 France: Agence Française de Développement 2024	Type of financial flow, based on "Flow Class" (China), "flow_class" (India), "FlowName" (CRS EU/US), "libelle_ind" (France). N/A for World Bank.
status World Bank: AidData 2017, Kersting & Kilby 2021 China: Dreher et al. 2022, Custer et al. 2023 France: Agence Française de Développement 2024	Status of project indicating the current level of implementation, based on "status" (World Bank), "Status" (China), "etat_du_pro" (France). N/A for EU/US.
precision_code World Bank: AidData 2017, IATI 2023 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024	Geographic precision code (see 5b. for details).
data_source World Bank: AidData 2017, IATI 2023 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024 France: Agence Française de Développement 2024	Original source of project location information. Can differ from the variable *coordinates_source* when information on geographic coordinates is missing in the original source and were completed by a secondary source.
coordinates_source World Bank: AidData 2017, IATI 2023 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024 France: Agence Française de Développement 2024	Original source of geographic coordinates. It can differ from the variable *data_source* when geographic coordinates are missing in the original source and were completed by a secondary source.
location_count World Bank: AidData 2017, IATI 2023 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: Bomprezzi et al. 2024	Number of project locations, based on "project_locas" (India) and on GODAD *project_location_id* for others.
sector_codes World Bank: AidData 2017, IATI 2023 China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023	3-digit sector classification for OECD sectors, based on "sector_code" and "ad_sector_codes" (World Bank), "Sector Code" (China), "crs_sector" (India), "SectorCode" (CRS EU/US); see 5a. for details.
sector_categories China: Dreher et al. 2022, Custer et al. 2023 EU/US: OECD 2023 France: Agence Française de Développement 2024	Sector category of the project, based on the variable "sector_codes"; see 5a. for details.
sector_name China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023 France: Agence Française de Développement 2024	Sector name of the project, based on "Sector Name" (China), "crs_sector_name" (India), "SectorName" (CRS EU/US), "sector_oecd" (France); see 5a. for details.

url China: Goodman et al. 2024 France: Agence Française de Développement 2024	Project URL under which project details can be found, "Source URLs" (China), "url" (France). N/A for India, World Bank and EU/US.
comm World Bank: World Bank 2023; deflated with DAC deflators China: Dreher et al. 2022, Custer et al. 2023 India: Asmus et al. 2024 EU/US: OECD 2023; deflated with DAC deflators France: Agence Française de Développement 2024; deflated with DAC deflators	Monetary value of the commitment for a project in constant 2014 USD, based on "curr_total_commitment" (World Bank), "Amount (Constant USD 2021)" (China), "usd_commitment_pt_con" (India), "UDS_Commitment" (CRS EU/US), "engagements" (France), and deflated to the year 2014.
comm_nominal World Bank: World Bank 2023 China: Dreher et al. 2022, Custer et al. 2023 EU/US: OECD 2023 France: Agence Française de Développement 2024	Monetary value of the commitment for a project in current USD, based on "curr_total_commitment" (World Bank), "Amount (Nominal USD)" (China), "UDS_Commitment" (CRS EU/US), "engagements" (France).
comm_ibrd World Bank: World Bank 2023; deflated with DAC deflators	Monetary value of the commitment by the IBRD for a project in constant 2014 USD. The value is calculated based on "curr_ibrd_commitment" and is deflated to the year 2014.
comm_ida World Bank: World Bank 2023; deflated with DAC deflators	Monetary value of the commitment by the IDA for a project in constant 2014 USD. The value is calculated based on "*idacommamt*" and is deflated to the year 2014.
disb World Bank: Kersting & Kilby 2021; deflated with DAC deflators India: Asmus et al. 2024 EU/US: OECD 2023; deflated with DAC deflators France: Agence Française de Développement 2024; deflated with DAC deflators	Monetary value of the disbursement for a project in constant 2014 USD. The value is calculated based on "disbursement" (World Bank), "usd_disbursment_pt_con" (India), "USD_Disbursement" (CRS EU/US), "versements_" (France) and deflated to the year 2014. Gross disbursements can be negative in case ineligible expenditures or voluntary refunds from previous years are recorded as negative disbursements.
disb_nominal World Bank: Kersting & Kilby 2021 EU/US: OECD 2023 France: Agence Française de Développement 2024; converted to USD using OECD exchange rate	Monetary value of the disbursement for a project in current USD. The value is calculated based on "disbursement" (World Bank), "USD_Disbursement" (CRS EU/US), "versements_" (France). Gross disbursements can be negative in case ineligible expenditures or voluntary refunds from previous years are recorded as negative disbursements.

disb_ibrd World Bank: Kersting & Kilby 2021; deflated with DAC deflators	Monetary value of the disbursement by the IBRD for a project in constant 2014 USD. The value is calculated based on "disbursement" (World Bank) and deflated to the year 2014. Gross disbursements can be negative in case ineligible expenditures or voluntary refunds from previous years are recorded as negative disbursements.
disb_ida World Bank: Kersting & Kilby 2021; deflated with DAC deflators	Monetary value of the disbursement by the IDA for a project in constant 2014 USD. The value is based on "disbursement" (World Bank) and deflated to the year 2014. Gross disbursements can be negative in case ineligible expenditures or voluntary refunds from previous years are recorded as negative disbursements.
comm_loc_evensplit World Bank: World Bank 2023 China: Goodman et al. 2024 India: Asmus et al. 2024 EU/US: OECD 2023, Bomprezzi et al. 2024	Estimation of the total monetary value in constant 2014 USD of commitments assigned to the location of a project (based on the variable *comm*). Values are calculated by dividing total amounts committed (*comm)* by the number of locations (*location_count*).
comm_ibrd_loc_evensplit World Bank: World Bank 2023	Estimation of the total monetary value in constant 2014 USD of IBRD commitments assigned to the location of a project (based on the variable *comm*). Location amounts are estimated by dividing total amounts committed for a project by the number of locations corresponding to this project.
comm_ida_loc_evensplit World Bank: World Bank 2023	Estimation of the total monetary value in constant 2014 USD of IDA commitments assigned to the location of a project (based on the variable *comm*). Location amounts are estimated by dividing total amounts committed for a project by the number of locations corresponding to this project.
disb_loc_evensplit World Bank: Kersting & Kilby 2021 India: Asmus et al. 2024 EU/US: OECD 2023, Bomprezzi et al. 2024	Estimation of the total monetary value in constant 2014 USD of disbursements assigned to the location of a project (based on the variable *disb*). Values are calculated by dividing total amounts disbursed (*disb)* by the number of locations (*location_count*).
disb_ibrd_loc_evensplit World Bank: Kersting & Kilby 2021	Estimation of the total monetary value in constant 2014 USD of IBRD disbursements assigned to the location of a project (based on the variable *disb_ibrd*). Location amounts are estimated by dividing total amounts disbursed for a project by the number of locations corresponding to this project. Gross disbursements can be negative in case ineligible expenditures or voluntary refunds from previous years are recorded as negative disbursements.
disb_ida_loc_evensplit World Bank: Kersting & Kilby 2021	Estimation of the total monetary value in constant 2014 USD of IDA disbursements assigned to the location of a project (based on the variable *disb_ida*). Location amounts are estimated by

	dividing total amounts disbursed for a project by the number of locations corresponding to this project. Gross disbursements can be negative in case ineligible expenditures or voluntary refunds from previous years are recorded as negative disbursements.
comm_loc_ratio China: Goodman et al. 2024	Estimation of the total monetary value in constant 2014 USD assigned to the location of a project, multiplying the project's commitment value with the share of the project that is located in a sub-national region (based on the variable *intersection_ratio*).
project_area China: Goodman et al. 2024	The total surface of the project in km².
intersection_ratio China: Goodman et al. 2024	For locations with precision levels 1 to 3, this value represents the percentage of the project area within the ADM2 unit. For locations with precision level 4, it represents the percentage within the ADM1 unit. Expressed as a value between 0 and 1.
purpose_code EU/US: OECD 2023	Based on "PurposeCode" (EU/US), which identifies a contribution's sector of destination. N/A for other donors.
place_name EU/US: Bomprezzi et al. 2024	Geographic place name extracted by the Named Entity Recognition pipeline.
aims_project_location_id AIMS: AidData 2016/2017	Identifier provided by AIMS for unique project locations.
aims_precision_code AIMS: AidData 2016/2017	Precision codes as provided by AIMS.
aims_place_name AIMS: AidData 2016/2017	Geographic place name as provided by AIMS.
aims_latitude AIMS: AidData 2016/2017	Latitude as provided by AIMS.
aims_longitude AIMS: AidData 2016/2017	Longitude as provided by AIMS.
honig_id Honig 2019	Honig (2019) project ID for merging with his Project Performance Database.
afd_latitude France: Agence Française de Développement 2024	Latitude as provided in the data from the Agence Française de Développement (based on their variable "latitude"). See 5e. for details.
afd_longitude France: Agence Française de Développement 2024	Longitude as provided in the data from the Agence Française de Développement (based on their variable "longitude"). See 5e. for details.
afd_concours_id France: Agence Française de Développement 2024	ID of the "concours" in the data from the Agence Française de Développement (based on their variable "id_concours"). See 5e. for details.

4. ADM1/ADM2 datasets

Variable	Description	Value for placeholder
gid_0 GADM 2023	Country string ID for country where a project is located using the GADM version 3.6 definition. ISO 3166-1 alpha-3 country code when available (ADM0 layer).	
gid_1 GADM 2023	ADM1 region string ID for the first administrative division where a project is located, using the GADM version 3.6 definition (ADM1 layer).	
gid_2 GADM 2023	ADM2 string ID for the second administrative division where a project is located, using the GADM version 3.6 definition (ADM2 layer).	
name_0 GADM 2023	Name of country where a project is located, using the GADM version 3.6 definition (ADM0 layer).	
name_1 GADM 2023	Name of ADM1 region where a project is located, using the GADM version 3.6 definition (ADM1 layer).	
name_2 GADM 2023	Name of ADM2 region where a project is located, using the GADM version 3.6 definition (ADM2 layer).	
year	Year of observation.	
{country}_{type}_{sector} OECD 2023, Bomprezzi et al. 2024; deflated with DAC deflators	Sum of monetary value, in current or constant 2014 USD, given to project locations within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets financed by a certain country.	- {country}: AUT, BEL, CHE, DEN, ESP, FIN, FRA, GER, GRE, ICE, IRE, ITA, LUX; NED, NOR, POR, SWE, UK, USA, EUR - {type}: "comm" (for commitments in constant USD), "comm_nominal" (for commitments in current USD), "disb" (for disbursements in constant USD), "disb_nominal" (for disbursements in current USD) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), "prod" (for "Production Sectors"), or "other" for other sectors

projectscount_{country}_{se ctor} OECD 2023, Bomprezzi et al. 2024	Number of projects within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets for projects financed by a certain country. Authors' computations.	- {country}: AUT, BEL, CHE, DEN, ESP, FIN, FRA, GER, GRE, ICE, IRE, ITA, LUX; NED, NOR, POR, SWE, UK, USA, EUR - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), "prod" (for "Production Sectors"), or "other" for other sectors
IND_{type} Asmus et al. 2024	Sum of monetary value, in constant USD2014, given to Indian project locations within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets.	- {type}: "comm" (for commitments) or "disb" (for disbursements)
IND_projectscount Asmus et al. 2024	Number of Indian projects within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets.	
IND_dummy _{type} Asmus et al. 2024	Dummy variable equal to 1 if the ADM1 (ADM2) region in the ADM1 (ADM2)-level dataset received at least one Indian project location.	- {type}: "comm" (for commitments) or "disb" (for disbursements)
CHN_{type}_{flow}_{sector} Dreher et al. 2022, Custer et al. 2023, Goodman et al. 2024	Sum of monetary value, in current or constant 2014 USD, committed to Chinese Official Finance project locations within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets.	- {type}: "comm" (for commitments in constant USD), "comm_nominal" (for commitments in current USD) - {flow}: "oda" (for ODA-like) or "oof" (for OOF-like) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors")
CHN_loccount_{type}_{flow }_{sector} Dreher et al. 2022, Custer et al. 2023, Goodman et al. 2024	Number of Chinese Official Finance project locations within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets.	- {type}: "comm" (for commitments) - {flow}: "oda" (for ODA-like) or "oof" (for OOF-like) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors")
CHN_dummy_ {type}_{flow}_{sector} Dreher et al. 2022, Custer et al. 2023, Goodman et al. 2024	Dummy variable equal to 1 if the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets received at least one Chinese Official Finance project location.	- {type}: "comm" (for commitments) - {flow}: "oda" (for ODA-like) or "oof" (for OOF-like) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors")

WB_{type}_{sector}_{financing} Commitments: World Bank 2023 Disbursements: Kersting & Kilby 2021 Geographic location: AidData 2017 (1995-2014), IATI 2023, Kersting & Kilby 2021 (1998-2023) Deflator: deflated with DAC deflators	Sum of monetary value, in current or constant 2014 USD, given to World Bank project locations within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets.	- {type}: "comm" (for commitments in constant USD), "comm_nominal" (for commitments in current USD), "disb" (for disbursements in constant USD), "disb_nominal" (for disbursements in current USD) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors") - {financing}: "ibrd" or "ida"
WB_loccount _{type}_{sector}_{financing} Commitments: World Bank 2023 Disbursements: AidData 2017 (1995-2014), IATI 2023, Kersting & Kilby 2021 (1998-2023)	Number of World Bank project locations contained in the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets.	- {type}: "comm" (for commitments) or "disb" (for disbursements) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors") - {financing}: "ibrd" or "ida"
WB_dummy_ {type}_{sector}_{financing} Commitments: World Bank 2023 Disbursements: Kersting & Kilby 2021 Geographic location: AidData 2017 (1995-2014), IATI 2023, Kersting & Kilby 2021 (1998-2023)	Dummy variable equal to 1 if the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets received at least one World Bank project.	- {type}: "comm" (for commitments) or "disb" (for disbursements) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors") - {financing}: "ibrd" or "ida"
AFD_{type}_{flow}_{sector} Agence Française de Développement 2024 ; deflated with DAC deflators	Sum of monetary value, in constant 2014 USD, committed to additional French projects within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets. See 5e. for details.	- {type}: "comm" (for commitments) - {flow}: "oda" (for ODA) or "oof" (for OOF) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors")
AFD_projectscount_ {type}_{flow}_{sector} Agence Française de Développement 2024	Number of additional French projects within the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets. See 5e. for details.	- {type}: "comm" (for commitments) - {flow}: "oda" (for ODA) or "oof" (for OOF) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors")

AFD_dummy_ {type}_{flow}_{sector} Agence Française de Développement 2024	Dummy variable equal to 1 if the ADM1 (ADM2) region in the ADM1 (ADM2)-level datasets received at least one additional French project. See 5e. for details.	- {type}: "comm" (for commitments) - {flow}: "oda" (for ODA) or "oof" (for OOF) - {sector}: "eco" (for "Economic Infrastructure and Services"), "soc" (for "Social Infrastructure and Services"), or "prod" (for "Production Sectors")

5. Additional comments

a. Project sectors

The variable *sector_category* categorizes projects into three broad sector categories as defined by the OECD: "Economic Infrastructure and Services," "Social Infrastructure and Services," and "Production Sectors." We code this variable using information from sector variables as reported in the source data: *"Sector Name"* in the Chinese Official Finance dataset, "SectorName" in the CRS, *"sector_code"* in IATI, and *"ad_sector_codes"* for World Bank projects. We define sector categories as follows:

Category	Sector	Sector code
Social Infrastructure and Services	Education, Health, Water Supply and Sanitation, Government and Civil Society, Other Social Infrastructure and Services, Population Policies/Programs and Reproductive Health	110, 111, 112, 113, 114, 120, 121, 122, 123, 130, 140, 150, 151, 152, 160
Economic Infrastructure and Services	Transport and Storage, Communications, Energy, Banking and Financial Services, Business and Other Services	210, 220, 230, 231, 232, 233, 234, 235, 236, 240, 250
Production Sector	Agriculture, Forestry, Fishing, Tourism, Industry, Mining, Construction, Trade Policies and Regulations	310, 311, 312, 313, 320, 321, 322, 323, 330, 331, 332

b. Geographic precision

Geographic precision for World Bank, Chinese, and Indian data

The assigned precision codes range from 1 to 9 for each project where locations are identified at the following spatial level:

Code	Name	Description
1	Exact location	The coordinates correspond to an exact location, such as a populated place or a hill. The code is also used for locations that join a location which is a line (such as a road or railroad). Lines are not coded – only the points that connect lines. All points that are mentioned in the source are coded.
2	Near exact location	The location is mentioned in the source as being "near," in the "area" of, or up to 25 km away from an exact location. The coordinates refer to that adjacent, exact location.
3	Second order administrative division	The location is, or lies in, a second order administrative division (ADM2), such as a district, municipality, or commune.
4	First order administrative division	The location is, or lies in, a first order administrative division (ADM1), such as a province, state, or governorate.
5	Estimated coordinates	The location can only be related to estimated coordinates, such as when a location lies between populated places; along rivers, roads and borders; more than 25 km away from a specific location; or when sources refer to parts of a country greater than ADM1 (e.g., "northern Uganda").
6	Independent political entity	The location can only be related to an independent political entity, meaning the pair of coordinates that represent a country.
7	Unclear – capital	Unclear. The capital is assumed to be one of two possible locations. (The other option is the country level, with precision 9.)

8	Local or national capital	The location is estimated to be a seat of an administrative division (local capital) or the national capital. For example, if aid goes to Luanda without further specification on the location, and there is an ADM1 and a capital called Luanda, then code the coordinates of the capital with precision 8. If it is not spelled out that aid goes to the capital, but if it is clear that it goes to a government ministry or to government financial institutions, and if those institutions are most likely located in the capital, then the coordinates of the capital are coded with precision 8. (However, if it can be verified that the recipient institution is located in the capital then precision 1 is used.)
9	Unclear – country	Unclear. The locations are estimated to be the country level (often paired with the capital, with precision 7).

The description can be accessed via the IATI website (https://iatistandard.org/en/iati-standard/203/codelists/geographicalprecision/).

Geographic precision for CRS data
The CRS data are geocoded using a Named Entity Recognition pipeline to extract geographic entities, combined with the Google Maps API to assign coordinates. Precision is rated on a scale from 1 to 13, with 1 representing the highest level of precision. This scale reflects the detail of the address provided by the geocoder; for example, "Sector 8 Dwarka, Dwarka, New Delhi, Delhi 110077, India" is considered more precise than "Delhi 110077, India." In the current version of the dataset, additional ad-hoc data cleaning may be necessary for coordinate-level spatial analysis.

c. Chinese Official Finance project locations
AidData's Geospatial Global Chinese Development Finance Dataset (Goodman et al. 2024) provides precise geospatial boundaries (called features) of those Global Chinese Development Finance Dataset 3.0 projects (Dreher et al. 2022, Custer et al. 2023) that could be identified on OpenStreetMaps (OSM). These projects have precision level 1 or 2. For each project that could be identified in OSM, the original data provide the exact geospatial boundaries of the project. In case that the boundaries of a project span across multiple regions, we partition projects into as many locations as the number of ADM units covered. For example, if the boundaries of a project cover two regions, the project is partitioned into two locations. If the boundaries of a project are contained within only one region, the project is considered to have only one location. The locations are defined at two regional levels: the ADM1 level if precision level is 4, and the ADM2 level if precision level is 3 or lower. Separating projects into project locations allows aggregation at the regional level.

As financial information is only available at the overall project level and not at the locations level, the location amounts are estimated in two ways:

1. *comm_evensplit_ratio*: The project amount is equally divided across all identified locations for the project, in the same way as for other donors.
2. *comm_loc_ratio*: The project amount is distributed proportionally based on the intersection ratio, which reflects the share of the project's "location" area relative to the total project area. We calculate it as: $comm\_loc\_ratio = comm \times intersection\_ratio$.

In cases where the project location is coded at precision levels 3 or 4 (aligned with administrative boundaries ADM1 or ADM2), the latitude and longitude provided in the dataset represent the centroid of the administrative unit, as defined by AidData 3.0 (Dreher et al. 2022, Custer et al. 2023). For locations with precision levels 1 or 2, the latitude and longitude correspond to the centroid of the project location, which is the centroid of the portion of the project within the ADM2 unit if the project

spans multiple ADM2 regions. If the project does not span multiple ADM2 regions, the latitude and longitude represent the centroid of the project polygon.

Note that projects with precision level 1 that could not be matched to an ADM2 region in GADM are not included in the dataset. A project may not be matched to an ADM2 region for several reasons: the project could be located on water, or the receiving country may not have ADM2 boundaries available in GADM. This is the case for: Antigua and Barbuda, Armenia, Bahamas, Barbados, Cabo Verde, Comoros, Cook Islands, Curaçao, Dominica, French Polynesia, Grenada, Guam, Israel, Jamaica, Kiribati, Lesotho, Libya, Maldives, Marshall Islands, Mauritius, Micronesia (Federated States of), Moldova, Montenegro, Nauru, Niue, North Macedonia, Palau, Saint Lucia, Seychelles, Sint Maarten, Tonga, Trinidad and Tobago, and Turkmenistan.

d. World Bank Operations

We build the geocoded dataset for World Bank projects from the official list containing all projects since 1947. We match these projects with two datasets to retrieve locations. The World Bank Geocoded Research Release from AidData contains projects from 1995 until 2014 but is incomplete. We merge these data with geocodes provided by the International Aid Transparency Initiative (IATI), available since 1998.

e. Project data from the Agence Française de Développement

We derive information on projects funded by France from the Agence Française de Développement (AFD) Open Data platform: https://opendata.afd.fr/pages/accueil/ (Permanent link). The AFD provides geolocated information on projects financed by three agencies: AFD, Proparco, and the Ministry of Europe and Foreign Affairs (MEAE). We accessed the data included in this dataset in September 2024. We include data from the AFD at the "concours"-level, which refers to a financial contribution made by one of these agencies toward a development project. This means that one project can correspond to several rows if it resulted in several concours. We do not have information on projects locations.

Some projects funded by France appear in both the AFD data and the CRS data. When geolocation information was not extracted from the CRS data, it is supplemented with geographic coordinates from the AFD data, as indicated in the *coordinates_source* variable. For projects where geographic information was extracted from CRS, we provide both CRS and AFD coordinates for comparison. French-funded projects that could not be identified within the CRS data are also included in the dataset, but should be treated with caution, as they may potentially be duplicates of projects in CRS that were not identifiable. Users can filter the dataset using *data_source* == "Agence Francaise de Developpement" to focus specifically on data from the AFD that are not present in CRS or to exclude these projects.

The ADM1 (ADM2)-level datasets include the aggregation of French-funded projects whose IDs could not be matched in the CRS data, stored in separate variables from other French-funded projects. These variables are identifiable by the inclusion of 'AFD' in their names.

f. Project data from the Aid Information Management Systems

We include the geographic information provided on aid projects registered in the Aid Information Management System (AIMS) for a set of aid recipient countries (AidData 2016, 2017). These recipient countries are Burundi, Colombia, the Democratic Republic of Congo, Honduras, Iraq, Nepal, Nigeria, Senegal, Sierra Leone, Somalia, Timor-Leste, and Uganda. We match projects with the CRS based on mutual characteristics such as project titles and descriptions and include in the project-level data the corresponding information on location and coordinates as registered in AIMS. When geolocation information was not extracted from the CRS data, it is supplemented with geographic coordinates from the AIMS data, as indicated in the *coordinates_source* variable. For projects where geographic information was extracted from CRS, we provide both CRS and AIMS coordinates for comparison.

g. Coding missing

In the aggregated ADM1/ADM2-level datasets, we code commitments and disbursements as missing if we know there is at least one project in the respective region and year based on the original datasets, but we have no information on their monetary value. Variables are also coded as missing for years outside the time period of the original datasets. We include CRS data over the 1973-2020 period. AidData's Global Chinese Development Finance Dataset (3.0) covers the period from 2000 to 2021. Data on India (the World Bank) are available for the 2007-2014 (1995-2023) periods.

h. Mistakes in original data
We do not drop data that may include mistakes in the original source. For example, CRS data includes 500 projects that Italy seem to have been committed in 1900. We include these data nevertheless.

6. Bibliography

Agence Française de Développement (AFD), 2024, AFD Open Data: Données aide au développement. Paris, France: AFD. Accessed on 09.2024. https://opendata.afd.fr/explore/.

AidData, 2016, Colombia AIMS Geocoded Research Release Level 1 v1.1.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, DRC AIMS Geocoded Research Release Level 1 v1.3.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Honduras AIMS Geocoded Research Release Level 1 v1.3.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Iraq AIMS Geocoded Research Release Level 1 v1.3.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Nepal AIMS Geocoded Research Release Level 1 v1.4.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Nigeria AIMS Geocoded Research Release Level 1 v1.3.2 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Senegal AIMS Geocoded Research Release Level 1 v1.5.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Somalia AIMS Geocoded Research Release Level 1 v1.1.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Timor-Leste AIMS Geocoded Research Release Level 1 v1.4.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2016, Uganda AIMS Geocoded Research Release Level 1 v1.4.1 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2017, Burundi AIMS Geocoded Research Release_Level 1 v1.0 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2017, Sierra Leone AIMS Geocoded Research Release Level 1 v1.0 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 03.2024. http://aiddata.org/research-datasets.

AidData, 2017, World Bank Geocoded Research Release Level 1 v1.4.2 Geocoded Dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on 12.05.2022. http://aiddata.org/research-datasets.

Asmus, G., Eichenauer, V.Z., Fuchs, A., and Parks, B., 2024, Does India Use Development Finance to Compete with China? A Subnational Analysis. *Journal of Conflict Resolution*, forthcoming.

Custer, S., Dreher, A., Elston, T.-B., Escobar, B., Fedorochko, R., Fuchs, A., Ghose, S., Lin, J.J., Malik, A.A., Parks, B.C., Solomon, K., Strange, A., Tierney, M.J., Vlasto, L., Walsh, K., Wang, F., Zaleski, L., and Sheng Zhang, 2023, Tracking Chinese Development Finance: An Application of AidData's TUFF 3.0 Methodology. Williamsburg, VA: AidData at William & Mary.

Dreher, A., Fuchs, A., Parks, B. C., Strange, A., & Tierney, M.J., 2022, Banking on Beijing: The Aims and Impacts of China's Overseas Development Program. Cambridge, UK: Cambridge University Press.

Goodman, S., Zhang, S., Malik, A.A., Parks, B.C., Hall, J., 2024, AidData's Geospatial Global Chinese Development Finance Dataset. *Scientific Data* 11, 529.

GADM, 2023. Global Administrative Areas (Version 3.6). https://gadm.org/.

Honig, Daniel, 2019, Project Performance Database February 2019, Harvard Dataverse, V1, https://doi.org/10.7910/DVN/444GNW.

International Aid Transparency Initiative, 2023, Accessed on 02.08.2023. http://www.iatiregistry.org/api/.

Kersting, E. and C. Kilby, 2021, Do Domestic Politics Shape U.S. Influence in the World Bank? *Review of International Organizations* 16(1): 29-58.

OECD 2024, Geographical Distribution of Financial Flows: Deflators (Edition 2023), OECD International Development Statistics (database), Accessed on 30 October 2024. https://doi.org/10.1787/2b6a144b-en.

OECD Creditor Reporting System (CRS), 2023, International Development Statistics (IDS) online databases. Accessed on 01.02.2023. https://stats.oecd.org/Index.aspx?DataSetCode=crs1.

World Bank, 2023, World Bank Projects & Operations, Accessed on 06.09.2023. https://datacatalog.worldbank.org/search/dataset/0037800.

World Bank, 2024, World Development Indicators, Accessed on 30.10.2024. https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS.