

Firm-level evidence on the effects of European aid projects

Pietro Bompreszi

University of Milano Bicocca and CefES

Silvia Marchesi

University of Milano Bicocca and CefES

Tania Masi

University "G. d'Annunzio" Chieti Pescara and CefES

July, 2022

PRELIMINARY DRAFT PLEASE DO NOT CIRCULATE

Abstract: This paper introduces a new geocoded dataset of ODA project aid from 18 main European donors, and uses it to evaluate the effect of aid projects on firms' sales growth. Through the application of Natural Language Processing pipelines, we can identify geographic locations in project metadata and use it to match projects to recipient country administrative regions. Matching this data with firm survey data from the World Bank Enterprise Survey, we obtain a region-year unbalanced panel of around 100,000 firms in 121 countries from 2003 to 2016. Preliminary findings show, on average, no significant effect of European ODA projects on the growth of firm sales. The outcome changes, however, when we differentiate across geographic regions, as firms in Asia and Africa experience a positive 0.8 percent increase average sales growth for each additional project. We instead find no significant effects in Latin America or Europe. As part of a future research, we work to construct a dyadic donor country - recipient region firm panel, connecting firms in donor countries to firms in receiver regions through Orbis ownership data. This paper, aside from providing new data and initial evidence on the effectiveness of European bilateral aid projects, brings attention to the role of official intervention, an important dimension of international finance, which needs to be further investigated.

Keywords: Aid effectiveness, European ODA projects, Geocoding, Firm growth

JEL Classification: F35, O19, E24, E25

1 Introduction

With geopolitical tensions on its eastern front, Europe in the summer of 2022 seemed to have jumped back in the spirit of the post-WWII Europe transfers.¹ While the commitments discussed there may be exceptional in nature, they are a reflection of a much longer trend of European concessional lending and transfers. The EU and its member states provide more than half of global development aid (OECD-DAC 2019). The EU’s focus is on boosting the effectiveness of development assistance by increasing partner country ownership of strategies, and combining traditional financing with private-sector and domestic resources.² For example, in the specific case of Africa, given their proximity, the EU has outlined a road map to serve as the basis for negotiations on a specific new partnership between these two continents, the joint EU-African Union (AU) strategy.³

More broadly, as recently documented by Horn et al. (2020), official lending is much larger than commonly known, often surpassing total private cross-border capital flows, especially in times of global turmoil when private flows generally shrink. In an era where the global economy is running into headwinds and nationalist tendencies are on the rise, official lending in the forms of grants and loans from both multilateral development agencies and bilateral lenders will be called to fill the gap in global capital flows. Importantly, these alternative flows are determined by factors other than mere financial returns, and the question on their exact impact on local economic activity is relevant.

In this paper, we analyze the effects of European bilateral aid projects on recipients’ local economies. For the first time we track the allocation of European bilateral aid projects across the developing world, by providing sub-national information on the distribution of projects. To evaluate firms performance we use data for firms geocoded at the ADM1 level, takings sales and various firm-level indicators from the multiple rounds of the World Bank Enterprise Survey (WBES), fielded between 2003 and 2016 across a large sample of countries (mainly developing). We then match the region-wide number of aid projects from 18 major European ODA donors to study their link with reported local firm sales growth, so as to motivate the initial hypothesis of an effect on local

¹As G7 nations met for their summit in late June, the German contingent argued for a massive, long-term support of Ukraine in line with a “Marshall plan” for Ukraine.

²Current priorities for the coming years include further investments in research and innovation, climate change; health, education, sustainable growth, and security. Since the onset of the COVID-19 crisis, the EU and its member states have been adapting priorities and programs with partner countries to address the crisis, particularly in supporting efforts to guarantee equitable access to safe and effective vaccines around the world.

³These two continents share converging interests in a number of areas, including climate change, and the promotion of job creation and economic growth in Africa. European development assistance has long been seen as a strategic tool to promote these objectives.

economic growth.⁴

In the second part of the analysis reserved for future work, we want to study if the allocation of aid projects and their subsequent effect is conditional on the economic ties between donor countries and receiver regions. To do this, we plan to rely on firm ownership data provided by Orbis. In the end, exploring the impact of development policy on firm growth, the final goal would then be to assess the extent to which the EU’s strategic development policy approach is successful in achieving the stated aims (such as fostering development) or whether it may have any unintended consequences.

In our first set of results, we relate to the literature that use aid allocated at the sub-national level with measures of aid effectiveness at the firm level (Chauvet and Ehrhart 2018). In particular, we view two ways through which aid influences firm performance, the demand channel (i.e., increased demand financed by aid is met by firms’ production), and the supply channel (i.e., aid affects the productive capacity of firms). The literature on firm performance, in turn, points up three main kinds of constraints on firm growth in developing countries: the financing constraint (Beck et al. 2005; Harrison et al. 2004), lack of infrastructure, such as transport, energy, telecommunications, and water (see among others Bluhm et al. 2020; Jedwab and Moradi 2016; Rud 2012); the institutional environment (e.g., Fisman and Svensson 2007).⁵

The contributions of this work are both methodological -using tools of natural language processing in order to provide new data- and empirical, by providing insight into the channels through which aid affects growth.⁶ Recent work on the effectiveness of aid points to more than just project or recipient characteristics as determinants of positive outcomes.⁷ One hot topic of discussion has illustrated the role that donor traits have for aid effectiveness (in the case of Chinese aid, for example, see Bluhm et al. 2020, Dreher et al. 2021a, Dreher et al. 2021b, Isaksson and Kotsadam

⁴In this study, we focus on European Official Development Assistance aid, which is provided by governments and multilateral institutions to developing countries’ governments mostly with the aim to promote developmental objectives. Aid is thus not directly provided to firms.

⁵As shown by Knack (2001) and Bräutigam and Knack (2004), foreign aid may induce an institution curse and weaken economic institutions. Furthermore, from a macroeconomic point of view, aid may also adversely impact firm growth if it induces Dutch disease, that is an appreciation of the real exchange rate detrimental to outward-looking firms (Rajan and Subramanian (2011). More recently, Gehring et al. (2021) find that Chinese aid is associated with more government repression and an increased acceptance of authoritarian norms, while World Bank projects strengthen democratic values.

⁶The main advantage of using this type of data is that we are able to overcome the problems associated either with poor quality of GDP data in developing countries) or with the caveats about nightlights as a proxy for economic activity (e.g., see Henderson et al. 2012 and Chen and Nordhaus 2011). In addition, our approach allows to perform a more accurate assessment of the effect of aid at the regional level by exploiting both project and firm heterogeneity.

⁷Minasyan et al. (2017), for instance, demonstrate the importance of donor quality for aid effectiveness.

2018).⁸ Similarly, Marchesi, Masi and Paul (2021) highlight the interplay between multilateral donors such as the World Bank and Chinese project aid. What are the effects of bilateral European ODA projects, and how such effects vary across different European donors, are related questions which are of equal importance to understanding aid effectiveness, because of the sheer size and historical role of European aid donors. Using our disaggregated data on donor-recipient pairs, we aim to pin down the different channels through which distinct donors influence local economic activity in a given recipient region through aid.

The identification strategy primarily relies on a set of control variables (firm and region level) to account for the observable heterogeneity and fixed-effect estimators to control for firm-level time-invariant heterogeneity. To address the reverse causality and the existence of time-varying unobservable heterogeneity, we instrument European aid by an interaction term composed of the donor's aid budget (given by tax revenues) and the recipient-specific probability of receiving aid from Europe (Chauvet and Ehrhart 2018).

We refine the identification strategy by applying a detailed firm-sector-region mapping that distinguishes between sector-specific and region (ADM1)-specific development assistance projects provided by European donors. We categorized broadly projects which have as scopes education or healthcare spending as "social infrastructure". The remaining projects fall into one the following categories of aid: food, mineral, manufacturing, retail, infrastructure, or services. This classification allows us to hone in on the effects of these latter type of project aid, excluding from our baseline analysis those projects which cannot have a direct regional economic effect.

Our preliminary findings show, on average, no significant effect of European ODA projects on the growth of firm sales on average. The outcome changes however when we differentiate across geographic regions. We find that in Africa and Asia, one additional project corresponds to 0.8 percent increase in firms sales on average. We find no effects in Europe, which can be explained by the presence of large multilateral financing institutions within the European Union, and no effect in Latin America.

Our findings are differ from previous works such as that of Chauvet and Ehrhart (2018) who, using a panel of 4355 firms in 29 countries from the same WBES dataset, find a positive effect of both bilateral and multilateral (OECD-DAC) ODA aid on firm sales growth. There are two main reasons which we believe are responsible for these differences in the outcome. First is the

⁸Bluhm et al. find that Chinese financed infrastructure projects reduce spatial concentration of economic activity within but not across regions. Dreher et al. (2021b) find that politically directed Chinese aid does not necessarily reduce its effectiveness. Gehring et al. (2022) evaluate Chinese aid along a more holistic approach, finding no evidence that Chinese aid increases local conflict but that Chinese aid is associated with an increased acceptance of authoritarian models.

sample size. We use a sample of roughly 100,000 unique firms spread across 121 countries that is much larger than the sample used by Chauvet and Ehrhart (2018). Second, Chauvet and Ehrhart measure aid as country level flows, linking disbursements from donors at the country level to firm level outcomes through the differential responses of firms based on a series of firm-level characteristics. On the other hand, we can trace projects directly to the same geographic region of the firms, allowing for more precise estimates and an evaluation of the impact of projects on a region v. region comparison.

This paper relates to at least two broad streams of the literature. First and foremost, the vast literature on aid effectiveness. Broadly speaking, this literature converges towards either a null effect (Doucouliagos and Paldam 2009), or small positive effects (Galiani et al. 2017) of aid on growth.⁹ This effect, however, depends on whether aid was politically motivated or had a clear development focus.¹⁰ What is more, donors have also been criticized for a lack of "ownership" and underutilizing local knowledge in recipient countries (Dreher et al 2017).¹¹ More specifically, our paper closely relates to an emerging strand of literature evaluating aid-effectiveness at the project level (e.g., Denizer et al. 2013; Dreher et al. 2013, Dreher and Fuchs. 2015; Dreher et al. 2019; Dreher et al. 2020a; Feeny and Vuong 2017; Kilby 2009, 2015; Marchesi and Masi 2021; Öhler and Nunnenkamp 2014; and Shin et al. 2017) and subnational level (e.g., Bluhm et al. 2020; Chauvet and Ehrhart 2018; Cruzatti et al. 2020; Del Prete et al. 2019; Gehring et al. 2021; Dreher and Lohman 2015; Dreher et al. 2021b, Isaksson and Kotsadam 2018; Marchesi et al. 2021).

This study also relates to the literature on firm productivity growth. The need for a disaggregated level approach to obtain a deeper and complete understanding of the dynamics of productivity growth is argued by Foster et al. (2001). Firm productivity in developing countries are characterized by widespread differences in capabilities of individual firms across and within countries (Hsieh and Klenow 2009). The performance of firms is largely affected by two sets of factors: internal, i.e., within the control of a firm or business, and external aspects of operating environment (Syverson 2011). Based on the classification of aid projects used in this paper, a sector-specific project (e.g., mineral) could potentially affect the internal factors, whereas a region-specific project (e.g., infrastructure) has the potential to improve the external factors. Thus, aid projects in our study may influence firm productivity growth through reallocation of resources across firms (and sectors)

⁹For recent surveys of the aid effectiveness literature, see Doucouliagos (2009, 2019) and Dreher et al. (2018b).

¹⁰There is some empirical evidence linking a country's geopolitical proximity to DAC donors with a variety of types of preferential treatment (e.g., Alesina and Dollar 2020; Faye and Niehaus (2012), Kuziemko and Werker (2006); Kilby 2009). In turn, when aid allocation is driven by political influence aid is likely to be effective (e.g., Dreher et al. 2013, Dreher et al. 2018a, Kilby 2015).

¹¹For example, quite a few papers have argued that institutions, organizations, and policies are context-specific and that, for their successful implementation, conditional programs should suit better recipient countries' specific needs (e.g., Asmus et al. 2016, Basurto et al. 2020; Marchesi and Masi 2021).

and within firm efficiency gains (Dollar et al. 2005; Busso, et al. 2013; Macmillan et al. 2014).

In summary, this paper contributes to the current literature on aid and economic performance in several ways. First, we extend the use of the geocoded aid data by bringing together the most refined and disaggregated data on European aid. Second, distinguishing sector and region-specific aid projects provides a closer look at various channels that impinge on the levels of aid-effectiveness and contributes to the literature on the aid-effectiveness contingent on the types of aid (e.g., Clemens et al. 2012; Asmus et al. 2016). Third, we contribute to studies that examine firm performance in aid-recipient countries (e.g., Chauvet and Ehrhart (2018) by allowing aid to vary across ADM1 regions within a country, by geocoding regions where firms (in the WBES) are located, and by extending the sample to 121 developing countries. Fourth, we contribute to the literature on firm performance in developing countries. As recent studies document, firms in developing countries are typically small and unproductive and there exist very few productive firms (Hsieh and Olken 2014, Eslava et al. 2019).

We organize the rest of the paper in the following manner. In section 2, we describes the data and how we combine data from different sources; in this section, we also discuss descriptive evidence on aid and firm performance. Section 3 illustrates the empirical model and the identification strategy. Section 4 then presents the baseline results, section 5 describes the dyadic connections between donors and recipient-region firms, while section 6 shows the robustness analysis. The final section 7 concludes.

2 A geocoded dataset of European ODA donors and receivers

This paper introduces a new geocoded panel dataset of European ODA projects constructed from its raw form in the OECD DAC Creditor Reporting System (CRS). The availability of new geographic information we compile allows us to answer questions regarding the allocation of European aid as well as evaluate the mechanisms through which aid influences local economic activity. Utilizing textual information associated to projects, we identify geographic entities within the receiver country and subsequently geocode them at the ADM1 level. This, together with the geographic information for firms provided from our World Bank Enterprise Survey data, allows us to merge these data sets and explore sub-national economic outcomes. In order to visualize this overlap, Figure 1 plots administrative regions (ADM1) where firms are active (green dots) and where projects are destined (red dots). As can be seen, most aid projects are concentrated in Latin America or Africa, while firm data is also present in some regions where aid does not flow

as much, such as Eastern Europe or Central Asia.¹² In the next section, we describe in detail the construction of the ODA project aid dataset. Then, we provide a description of our firm data and the merging procedure between the two. Finally, we provide some descriptives.

[Figure 1 here]

2.1 OECD ODA data

The OECD CRS provides data on roughly 3.7 million projects from all OECD donors beginning in 1973, up to when our data collection finishes in 2017. The raw data contains information on commitments, disbursements, and received amounts (in USD) of projects undertaken by OECD donors. Of particular interest to us, by collecting data at the project level the CRS provides a wealth of information aside from flows, including detailed information on the purpose of the project, descriptions, and geographic information. The first contribution of this work is to exploit textual data on project titles and descriptions in order to geocode projects at the first-order administrative (ADM1) level, allowing then for a study of the determinants of aid allocation and evaluation of the mechanisms through which impacts local economic activity. Contrary to other data on project aid, such as World Bank or Chinese aid projects, data on geocoded European aid projects are available only in a limited number. The importance of this (geocoded) lender-side microdata in the literature on aid allocation and effectiveness is evident by the breadth of recent work which explores the mechanisms through which foreign aid has sub-national level effects.

The lack of precise data on OECD ODA prevents researches from addressing these questions, despite the historically important role of OECD (and European in particular) ODA. Because we wish to focus on the relationship between European DAC donors and their respective recipient countries, we focus on 18 main European bilateral donors and construct a dataset of geocoded ODA project aid from 1980 to 2017. Subsequently, we can match this data to another dataset of firms in recipient regions, and evaluate the mechanisms of aid effectiveness and economic ties between donors and receiver regions. The remainder of this section gives an overview of the construction of the geocoded European ODA dataset. Appendix D instead outlines step by step the procedure.

Project titles and descriptions provide text data from which geographical entities can be identified. The procedure for the extraction and identification of geographical entities in use can be outlined as follows. First, we collect raw data from the OECD CRS on our 18 European donors for a total of

¹²See Figure A1 in the appendix for a decomposition of WBES firms by geographic region.

1,275,619 unique projects from 1980 to 2017.¹³ We then leverage the detailed text descriptions of projects to extract candidate geographical entities which can be matched to known cities, regions, or administrative entities within the receiver country. CRS aid data provides titles, short, and long descriptions of aid projects which are all used as sources of information. For each project we can then run Spacy's (pre-trained) Named Entity Recognition (NER) transformer model for entity identification. This particular class of algorithms use deep learning models to identify specific categories within a text, including geographic entities. The model finds least one geopolitical entity for 243,255 projects, or roughly 19% of all projects.¹⁴ In addition to this, we add the 102,848 projects for which the CRS already associates a geographic entity. The extracted entities then undergo a series of data cleaning and cross-checking.¹⁵

In summary, tests on a "golden", hand-coded data sample indicate an in-sample accuracy of over 70 percent for the NER model. Missed elements, or false negatives, include the majority of the total model errors. These include instances where the string length was too short, when the strings were only in a non-English language, and instances where the model missed the entity for no discernible reason. A lesser issue is that of false positives, or when the model identifies an entity when it is not really there. Some of these cases occur because of semantic reasons, when the model identifies political entities as geographic ones or incorrectly labels generic terms as geographic entities. We show, in the online Appendix D, that through a simple algorithmic procedure relying on a term frequency - inverse dictionary frequency and KNN fuzzy matching these false positives can be easily thrown out when matching at the ADM1 level. We also detail how the same approach can be used to cross-check for false negatives, when the NER model fails to extract entities from strings that contain them.

As a final step, we georeference the list of extracted and cleaned entities associated to each project, resulting in a dyadic donor country - recipient region panel, where within each receiver region are nested the ODA projects from the 18 European bilateral donors. In the end, we are able to obtain a total of 184,566 geocoded projects corresponding to 2,561 unique ADM1 regions in 4 broad geographic areas.

¹³1980 is the first year we have data from one of our 18 European bilateral donors, while our data collection stops in 2017.

¹⁴Note that this does not mean that the model misses 81% of geopolitical entities in the data. The vast majority of projects do not contain entities that can be geocoded. Appendix E discusses this in detail.

¹⁵In the online Appendix D, we explain in greater detail the evaluation of the NER model accuracy on our dataset, data cleaning, and how we deal with false positives and false negatives.

2.2 Firm level data

We measure firm performance using the World Bank Enterprise Survey data. Through face-to-face interviews with firm managers and owners of firms, the World Bank collected data on various indicators of the business environment, performance, and firm productivity. The survey instruments specifically include questions on (1) infrastructure and services, (2) sales and supplies, (3) degree of competition, (4) institutional capacity and access to land, (5) sources of finance, (6) business development services, (7) business-government relations, (8) labor, (9) business environment and finally (9) firm performance.

The WBES dataset has information on 146,666 firms spread across 139 countries between 2003 and 2016. We drop 35802 firms either because of the missing observations or due to multiple locations.¹⁶ Hence, we work with a final sample of 110,864 firms. Roughly 10 percent of the firms were successfully re-contacted, which produces an unbalanced panel dataset at the firm level.¹⁷ The survey is constructed to generate a representative sample of a countries manufacturing and service sectors, with the final aim of providing indicators for the investment climate in a country and the responses are harmonized across countries for comparability. The sampling methodology for each country follows a stratified random sampling according to 3 criteria (firm size, sector, geographic location). This allows a random sampling which is more representative of the economic composition of the country, since the likelihood of being selected for an interview is dependent on the individual firms' place, in the distribution of firms within a country, as well as its location with respect to geographic areas of economic activity and economically relevant sectors.

Table A1, in the online Appendix A, shows the classification of industries (coded 1 to 6) for the WBES data. Over 19 percent of the firms are in the wholesale and retail, hotels, and restaurant sector, about 11 percent in the mineral sector, 9 percent in the food sector and 5 percent in the transport, communication, information (IT), and construction sector. The rest of the firms are classified as manufacturing (about 37 percent) and "other services" (20 percent). To achieve this classification, we regroup the original 51 industrial sectors from the WBES data into the above mentioned six broad categories sectors (see Table A2 for grouping strategy). The OECD data instead provides information on about 40 categories of aid, which we reclassify to match the six WBES sectors (Table A3a and A3b in the online Appendix A). However, some European aid

¹⁶More specifically, 11,946 firms were dropped because of missing location, while we had to discard 23,552 firms that operated in more than one ADM1. Finally, 304 firms were lost because we had to drop the last year of the survey, 2018. We should emphasize here that the survey refers to the fiscal year, hence the year 2016 is in fact the year 2015.

¹⁷One of the advantages of this updated version of the WBES is the availability of multiple questionnaire waves, which gives the possibility to construct a panel with gaps for firms which participate in more than one wave.

projects (e.g., public administration or education) that are not sector-specific but could affect the performance of a firm from any sector within a region are classified as region-specific aid and coded as an additional category "0". Table A4 shows the mapping between European projects and the WBES sectors.

As a final step, we merge the WBES data with our regional European ODA project data. First we restrict the period of analysis to match data coverage from both datasets. Figure A1, in online Appendix A, visualizes the overlap; the survey coverage starts in 2003 and goes to 2016, and the majority of projects fall into this range. Following the standard practice in the aid literature (among others see Dreher et al. 2021a, 2021b), we use a two-year lag for a firm to benefit from the time an European project is committed. This allows us to evaluate any European development commitments taking place between 2001 and 2014, and the firm-level outcomes realized in the period from 2003 to 2016.

To associate WBES firms to regions, we identified the latitude and longitude of the ADM1 level using the names of the regions reported in the WBES dataset. We engage in a series of data cleaning to properly georeference survey firms: we corrected misspelled names and uncoded characters, we separate multiple locations (e.g., 5 small cities), attributed all ADM1 when "entire country" was specified, and retrieved ADM1 when different levels were specified (e.g., NUTS or North, South etc.). Finally, we geolocated each query using the Python client Geopy. We fill data gaps if the algorithm failed to find the coordinates, using Google Maps. To avoid any measurement issues, we discard firms that operates in more than one ADM1. We then map each firm into an ADM1 region following the condition that aid projects are implemented in the same ADM1 two years before the WBES interview took place. Unfortunately, this procedure, while guaranteeing the precision of the mapping between the location of the project with that of the firm, comes at the costs of losing quite a large number of observations of aid projects, which, in order to be considered, need to have been committed exactly two years before each survey date. Since a region might be under an aid project for longer periods than the ones we are able to measure, our results should then provide a lower bound for the effects of aid on firms' performance.

2.3 Descriptive Evidence

We now present some initial descriptive statistics on the sample of data we have successfully georeferenced as described in section 2.1. OECD-DAC aid is collected for different types of aid flows, including grants, loans, or equity investments. The overwhelming majority of projects however are ODA grants. Table B2, in Appendix B, provides some basic summary statistics

for our sample of projects. As mentioned, ODA grants comprise close to 97 percent of all aid projects. The average commitment value for these grants is about 400,000 U.S dollars, but with a wide distribution, as the largest of projects reaches almost 500 million in commitments.

The dyadic structure of aid data also provides important insights. Figure 2 shows the share of projects in our sample attributed to each of the 18 donors. Among all projects for which we matched to an administrative region, the majority are coming from Spain. Italy and Germany are also well represented, afterwards followed by other traditional donors such as the U.K, Norway, Belgium, and France.¹⁸ On the other side, in Figure 3 we plot the number and size of projects on the receiver side. Unsurprisingly, countries in Southern Asia like India and Pakistan receive the greatest number of projects as well as the overall largest commitments. The plot also tells us something about the nature of these projects. For example, we find cases such as Bolivia, where the difference between the number of projects and the total commitment amounts is significantly larger than its counterparts. In other words, Bolivia and in general other countries from Latin America like Peru and Colombia seem to attract many, but small, projects. We show this in greater detail in Table A6 of the online Appendix A.¹⁹

[Figures 2-3 about here]

Finally, in Figure A2 of the online Appendix A, we show the yearly variation in commitments and number of projects. The plot suggests that there was a commitment boom in projects around 2004, which materialized in later years. Interestingly there seems to be no evidence of an effect on aid activity due to the Great Financial Crises or the European Debt Crises.

Not all aid is created equally though. OECD-DAC data provides project-specific descriptions of the purpose each aid project, which can be aggregated into the sector categories we have described before. The CRS data provides 41 granular aid-sector categories which fall into 11 macro-categories of trade and tourism, energy, banking and business, industry, transport and infrastructure, environmental protection, agriculture, emergency, social infrastructure, or multisector/unspecified.²⁰ When we collapse our aid data to the ADM1 level for our analysis of firm outcomes, we identify 7 main types of aid defined as Social Infrastructure, Food, Mineral, Manufacturing, Retail, Infrastructure, and Services.

¹⁸the role of Spain as a top donor (in terms of number of projects) and the presence of Latin American countries like Bolivia, Peru and Colombia as top receivers

¹⁹The table decomposes number of unique projects and average size of commitments for each European donor. A chunk of aid is administered at the macro-regional level, which typically includes those projects defined as Social Infrastructure in our framework. The table also shows that the number and size of projects destined for Africa and the Middle East are slightly greater with respect to other receiver regions.

²⁰Table A3 in the Appendix shows the specific CRS aid sector codes.

Figures 4 and 5 provide some descriptive statistics on these categories for this collapsed dataset. The vast majority of projects are of the Social Infrastructure type, which captures things such as education, health, and basic civil-society initiatives. Among those project types which are sector specific, Food aid seems to be relatively important in terms of number of projects. We also find some results that are in line with expectations, namely that projects in the infrastructure sector are fewer relative to others, but are on average larger in size. Figure 5 then shows our final decomposition of aid projects, putting together all the information we have seen up until now. First, both the number and average size of social infrastructure projects is larger, because of the scope of these projects. We also find some interesting variation across geographic regions. Latin America for example seems to attract a large share of these broad, social infrastructure type projects, especially in comparison to projects which are more targeted. Sub-Saharan Africa, as expected is an important recipient of aid, both in terms of social infrastructure projects and sector specific ones. Furthermore, the disparity between number of projects implemented and average amount committed to these projects is lower in Sub-Saharan Africa with respect to other regions (variation between orange and green segments of the line).

[Figures 4-5 about here]

We now turn to some descriptives regarding firms. Figure B1, in the online Appendix B, plots the kdensity estimates of our main dependent variable the log of sales growth (winsorized at the 1 and 99 percentile) for the 6 main industrial sectors provided in the WBES. The distribution of firm sales do not significantly vary across sectors and follow a bell-shaped pattern. The values of log of sales growth are comparable across different industries. Figure B2 in the Appendix also looks at the distributions by sector and geographic regions. For example, in Africa it is evident how the manufacturing sector is growing less with respect to other sectors. Figure B3 in the Appendix instead speaks to the representation of the survey across different geographic regions. More than 30% of the firms in survey are from Africa and the Middle east, and again around 30% are from Asia and the Pacific. Latin America represents around 22 %, while Europe only 12%.

Finally, Table 1 provides some summary statistics on our firm and regional variables for our final, region-year level dataset of merged WBES and ODA data. The average firm is experiencing positive growth, and is medium sized (a value of 1.72 on a scale of 1 to 3). Very few firms are state owned, foreign-owned, or export-oriented. There are a total of 110,041 geocoded projects that we are able to merge with the WBES data. In each region-year unit there is on average 20 projects, with the largest regions having 153 projects in a single year.

[Table 1 about here]

3 Empirical strategy

In this section we describe the remainder of the data, including our two sources of firm level data, the World Bank Enterprise Survey (WBES) and our dyadic firm panel matching firms in donor countries with firms in receiver regions. We then outline our empirical strategies of choice. We investigate the impact of foreign aid on firm performance using the following general specification:

$$g_{i,k,h,j,(t,t-2)} = \alpha + \beta P_{h,k,j,t-2} + \gamma X_{i,k,h,j,t} + \delta Z_{h,j,t} + \tau_{k,t} + \mu_{h/i} + \varepsilon_{i,k,h,j,t}, \quad (1)$$

where g is the annual growth rate of sales, computed over three years, between year t and $t - 2$, of firm i , in industry k , region h and country j . P represents the European aid projects both at the regional and sectoral level (in terms of committed and disbursed amount) tied to region h . In section 4.3, we consider regional and sectoral projects separately. X is a set of time varying firm-level characteristics, while Z is a set of regional-level variables (logged regional population and GDP).²¹ We then include industry-year dummies $\tau_{k,t}$, in order to control for industry time-varying heterogeneity and $\varepsilon_{i,k,h,j,t}$ is the error term. We also include either region or firm fixed effects. To avoid extremely fast-growing firms driving the results, we excluded 823 firms whose sales fall into the 99th percentile.²² Finally, the standard errors are clustered at the regional level.

We control for the lagged value of *Sales*, in logarithm, which is measured at $t-2$; firm *Size*, which takes the value one for firms with fewer than 20 employees, the value two for firms with between 20 and 100 employees, and three for firms with more than 100 employees. We also control for the characteristics of firm ownership using two variables, *State* and *Foreign*. *State* is a dummy variable which is equal to one when part of (or all) the firm is owned by the state, while *Foreign* is a dummy variable which is equal to one when part of (or all) the firm is owned by a foreign individual or company. Finally, we include information on whether the firm is outward looking using *Export*, which is a dummy variable equal to one when the firm exports part of or all its sales, either directly or indirectly (as a supplier to exporting firms). The firm-level characteristics are measured in year t since we do not have their pre-determined value at year $t-2$. Definition and sources of the variables are listed in Tables B1, in the online Appendix B.

²¹Regional GDP is measured considering the log of night-time lights (NOAA, National Geophysical Data Center, 1992-2013. Population is taken from the Socioeconomic Data and Applications Center (SEDAC), 1975, 1990, 2000, and 2014/2015.

²²Also 28,525 firms are discarded due to missing values for sales, 263 are discarded due to missing values for industry and 14,319 observations are dropped due to missing control variables.

3.1 Identification strategy for European aid projects

Equation (1) is estimated using both region and firm fixed effects, which allows us to control for firm-level time-invariant heterogeneity. To this firm fixed-effect setting, we add industry x time dummies in order to also control for industry time-varying heterogeneity. Our framework accounts for part of the observable heterogeneity -using a large set of control variables both at the firm and region level -and for the unobservable heterogeneity - using firm fixed effects and industry x year dummies. However, the estimated correlation between project aid and firm growth could still be biased by two remaining endogeneity channels: reverse causality and the existence of time varying unobservable heterogeneity. Strategies to deal with the endogeneity of aid at the macroeconomic level have evolved and improved over time.

A new strand is currently emerging in the aid effectiveness literature based on quasi-experiments, i.e., specific situations that can be taken to identify the impact of aid on growth. Early work in this area focuses on shocks affecting donor countries such as the variation in oil prices to instrument aid from Arab countries (Werker et al. 2009). Similarly, Nunn and Qian (2014) use Nunn and Qian exploit temporal variation in US wheat production, which they interact with the aid recipient’s probability to receive US food aid. Our identification strategy is then based on an instrumental variable (IV), which consists in the interaction of the donor’s aid budget with the recipient-specific probability of receiving aid from the respective donor.

More specifically, as in Chauvet and Ehrhart (2018), we calculate the European aid budget with measures of their tax revenues. The source of exogenous variation in donor economic situations is weighted by historical proximity between donors and recipient regions.²³ Our identification strategy then exploits the differential effect of changes in European donors resources. The source of exogenous variation in the European resources is then weighted by historical proximity between European countries and recipient regions. The IV equation is then the following:

$$IV_{h,t-2} = Tax\ revenues_{t-2} \times Probability_h, \quad (2)$$

where $Probability_{h,t-1}$ is the (time invariant) share of years between 1994 and 2014 that region h received any European aid, and $Tax\ revenues_{t-2}$ is the temporal variation in the tax revenues. Controlling for year fixed effects (which capture European resources) as well as for the time invariant region-specific probability component of the interaction term, the identifying assumption underlying this approach thus follows a difference-in-differences logic. Like a difference-in-differences

²³Chauvet and Ehrhart 2018 interact donors’ budget with a dummy for former colonial status and Dreher and Langlotz (2017) look at donor fractionalization interacted with the recipient probability of receiving OECD aid.

approach, we investigate the differential effect of European donors' resources on European projects to regions with high compared to low probability of receiving European projects. The identifying assumption is that firms' sales in regions with differing probabilities of receiving European aid will not be affected differently by changes in European resources, other than via the impact of aid, controlling for region and industry-year-fixed effects. For identification, we exploit the fact that European donors will be able to give more projects in years in which its liquidity is higher, so that regions with an initially higher participation probability are more likely to receive a project in these years (as displayed in Figure C3, in the online Appendix C).

If there were evidence of correlation between the two, it would only bias the results if the correlation was contingent on a country's past participation in European projects. We plot the European tax revenues over our period of estimation alongside the trend in firm sales in countries, distinguishing between different degrees of past participation. These trends (which are shown in Figure C1) are clearly parallel and not obviously correlated to European donors' budget, but most importantly, the difference between European aid and each group remains constant over time.

Given the structure of the identification strategy, the exclusion restrictions would be violated, if there was some unobservable, time-varying trend affecting sales differently across countries based on their past exposure to European projects. One might be concerned that the interacted instrumental variable violates the exclusion restriction because European resources are correlated with some omitted variables, which differentially affect firm performance in regions with low and high probabilities of receiving European aid. To address this threat, we tested for the most obvious country-level confounders, such as global GDP growth (interacted with European aid probability) and both Chinese or World Bank aid. As shown in Table C1, in the online Appendix C, the results are robust to these additional controls. However, we cannot definitively rule out the presence of omitted variable bias. The next section presents the regression results.

4 Baseline results

Table 1 presents the baseline outcomes for the European aid projects. We consider growth in sales (the difference of the amounts of firm sales, in log) as the dependent variable. In the first column of Table 1, Equation (1) is estimated using an OLS estimator without the firm fixed effects but including region (ADM1) dummies. Column 2 presents the results for the same sample when aid is instrumented. Columns (3) and (4) show the results for a panel sample with firm fixed effects and without region dummies. All specifications include industry-year dummies. The panel sample is restricted to a group of 7,807 firms, which corresponds to about 10 percent of the observations

in the OLS sample.²⁴ The variable of interest is the committed amount. The first-stage results show the coefficients for our instruments, which are always positive and significant, and with the expected sign. Kleibergen Paap tests provide further evidence in support of identification.

As our control variables are concerned, the coefficient of Sales at (t-2) is always negative and significant, at the one percent level, suggesting a catching-up effect, i.e., firms with lower sales in t-2, on average, show faster growth of sales than firms with higher sales in t-2. The coefficient of *State* is negative and significant, at the ten percent level, only in the panel specification (while is otherwise not significant) suggesting that state-owned firms perform worse on average. *Foreign*-owned firms perform better, on average, but the coefficient is positive and significant, at the one percent level, only in the OLS specifications. Export-oriented firms are associated to higher rate of growth in sales, on average. *Size* has a positive and significant coefficient, at the one percent level, suggesting that larger firms also tend to have a positive growth of sales.²⁵ Finally, the coefficients of the regional controls turn out to be not significant at conventional levels.²⁶

[Table 1 about here]

Turning to our interest variables, the coefficient of the aid amount, in the IV regression (Columns 2 and 4), turns out not to be statistically significant (while being positive and significant, at the ten percent level, in the pooled OLS specification). This result should be interpreted as the average effect of the aggregate ODA grants of all the 18 donors we included in our sample (as described in Figure 3).

Finally, we examine the effect of European aid on firm performance, by disaggregating the recipients according to their macro region. Table 2 shows the regression outcomes. It is worth mentioning that the number of observations for each sector varies: the largest sample is for African countries (24,191 firms) and the smallest is for the Central Europe (7466).

Considering the IV results, we find that European projects have a positive and significant effect on firm sales growth only in the case of Africa and Asia. Such positive effect is sizeable in quantitative terms, as one more European project, on average, increases sales in firms by about 0.8 percent in both regions. In Latin American countries and in Central Europe we do not find evidence of a causal effect of European aid on firms' performance. In the case of Europe such lack of significance

²⁴About 9,081 firms were re-contacted; among them, 2,188 firms did not match to either our classification of sectors or ADM1.

²⁵There is a strong correlation between *Size* and *Exports*, as most of the larger firms in the sample are those firms which tend to export trade (see, among others, Melitz 2003, Helpman et al. 2004).

²⁶The variable regional growth is always dropped in the panel specifications due to insufficient observations.

could be explained by the diversion of aid flows from bilateral lenders to intra-EU multilateral development institutions.

A much more detailed analysis should be undertaken in order to detect likely differences across donors, and to control for the possible bilateral donor-recipient relationship. How effects vary across different European donors is a related question which is of equal importance to understanding aid effectiveness. Using our disaggregated data on donor-recipient pairs, we can pin down the different channels through which distinct donors influence local economic activity through aid. In the next sub-section we start to provide more detailed evidence on how to explore the dyadic connections between donors and recipient-region firms.

5 Connecting donor-country and recipient-region firms

The method and results presented in this version of the paper rely on traditionally used firm data and accepted identification strategies in the aid effectiveness literature to find some initial results regarding the sub-national impact of European ODA. To do this, we collapse our CRS data to the ADM1 level, aggregating the number of projects per region and matching it with our WBES region-firm data. The advantage of this is the ability to use an identification strategy which is generally accepted in the literature, and to compare results to similar work.

This approach, however, comes with the limitations that the survey structure of the firm data limits our sample size, and we can only identify channels through the presence of differential effects based on firm-specific characteristics in the receiver region. As has been expressed throughout the paper, the final aim of this work is to explore the dyadic connections between donors and receivers, by focusing on the presence of economic ties at the firm level. Our initial data work confirms the presence of certain consistent trends in lending which are donor-recipient specific. We hypothesize that aid is not only allocated because of donor-recipient economic ties, but the effectiveness of aid projects and the regional level is also contingent on this relationship.

In order to construct a dyadic panel of donor-receiver region firms, we need to turn to the Orbis database provided by Bureau Van Dijk. Orbis provides detailed balance sheet data with a global coverage of countries and differently from other data providers captures both publicly traded and private firms. Importantly for us, Orbis also provides a comprehensive assessment of the ownership structure of firms. Figure E1, in the online Appendix E, provides an example. For each donor country, we are able to select those active firms (with consolidated balance sheets to avoid double counting) that are linked to foreign subsidiaries through the ownership structure of the

subsidiaries. The figure shows the example of a single firm in France for a given year. The heat map shows the number of subsidiaries in a given country of this firm. As can be seen, the firm is active in many countries in Latin America and Africa because of its corporate structure. If we zoom in on a certain country, for example Angola, we can retrieve information on those subsidiary firms. Figure E2 gives an example of some of the balance sheet data. Furthermore, we are provided with a number of qualitative data including the different sectors of activity and importantly, the geographic location of both owner and subsidiary down to the street level. We are currently in the process of creating a comprehensive dyadic panel based on this sample structure, linking all foreign subsidiaries f of a firm i from donor country j in every available receiver country region h , in time t . The availability of the full ownership network also allows us to check for spatial spillover effects and corporate-group specific characteristics.

The advantage of Orbis is the availability of detailed balance sheet data, which allows us to evaluate outcomes along more than one dimension, as well as precise ownership networks and location data. As with any firm-level work focusing on the global south, how nationally representative the data can be is always a question. The fact that Orbis draws from a pool of public registries as well as supplementing the data with on-the-ground surveyors reduces certain risks that the data is only drawing from one source which may vary by country. Aggregation of the final data at the industrial sector or the ADM1 level also mitigates the issues of data patchiness. With this data we will then be able to quantify the degree of economic ties between a donor-country/industry with a recipient-region/industry, and evaluate the determinants of aid allocation and its effects on local economic activity. Finally, the availability of a panel, unlike a survey, allows us to perform alternative causal identification strategies such as a stacked difference-in-difference or spatial models.

6 Robustness

This section contains a robustness analysis for our main results and the related Tables and Figures are presented in the online Appendices C We begin with issues regarding the identification strategy, in particular to address the exclusion restrictions, to then turning to discuss issues related to our survey data.

The biggest threat to identification regards the presence of underlying, time-varying heterogeneous trends, which are correlated to European tax revenues and may affect firm sales differentially, conditional on the share of years spent under an aid project (Christian and Barrett 2017). Following previous studies (e.g., Christian and Barrett, 2017; Chauvet and Ehrhart, 2018; Dreher et al., 2021a, Dreher et al 2021b), we then check the validity of our instruments by plotting trends

across regions below and above the median of the probability of receiving projects from European donors. More precisely, in Figure C1, we plot the European countries' tax revenues in tandem with the European project amount and sales growth. Since we do not have a proper panel, we construct three-year averages of project amounts and sales growth. The results show the absence of any spurious trends. We do not see any non-linear trend in the European tax revenues (a proxy for their aid budget), which are similar to the trends in project amount and sales growth for the two different groups of regions.

A similar issue is the one of alternative trends driving the first stage, which could represent a possible threat to the exclusion restriction. We explore some of these potential confounders, as we consider the presence of global GDP and both Chinese and World Bank aid. Table C1, in the online Appendix C, confirm that the results are robust to controlling for the differential effect of these alternative trends by interacting them with region-specific European aid probability.

An equally important issue to address in our model is the role of sample dependence, such as the sensitivity of the results to the inclusion of certain countries in the sample. While our country sample is vast and therefore unlikely that a given country is driving the results, issues of sample dependence could arise from the firm sample within countries. The stratified random sampling methodology for the WBES explained in section 2.2, at least theoretically, guarantees that the patterns for firm sales growth are not being driven by a particular set of firms more exposed to European aid.²⁷

Another quite evident limitation to survey data is the problem of recontacting firms. Beside promising best practices and efforts to create panel data in their survey, the WBES provides no guarantee that firms which can be recontacted will be. And there is no way to know why some firms don't appear in future waves of the survey. The biggest limitation which would affect our results on firm sales growth is firms dropping out because they go bust, what we call the survivor bias. If this were the case however, we would expect that the distribution of firms with repeated interviews versus the distribution of single-presence (no repeated interviews) firms would be significantly different. Figure C3, in the online Appendix C, shows that the two distributions are rather similar.

²⁷Besides, firm-level controls should also control for these potential channels.

7 Conclusions

Over the past two decades, the concept and very nature of aid has changed with the emergence of new donors like China. At the same time, a more accurate evaluation of traditional European DAC donors is also needed, especially considering the recent debate on the relevance of European intervention to assist Ukraine or to discourage migration from developing countries.

This paper evaluates the effect of development project aid from European donors on firms' sales growth, using a large dataset of almost 200000 European ODA projects and a dataset of 86,000 unique firms, spanning 121 countries between 2001 and 2016. Our identification strategy exploits the differential effect of changes in donors' aid budget on project participation (Chauvet and Ehrhart 2018). Preliminary findings show, on average, no significant effect of European ODA projects on the growth of firm sales. The outcome changes, however, when we differentiate across geographical regions, as countries in Africa and Asia seem to benefit from European intervention, while Latin American and Central European countries do not. For the latter, such lack of significance could be explained by the diversion of aid flows from bilateral lenders to intra-EU multilateral development institutions.

As part of a future research, we work to construct a dyadic donor country - recipient region firm panel, connecting firms in donor countries to firms in receiver regions through Orbis ownership data. This paper, aside from providing new data and initial evidence on the effectiveness of European bilateral aid projects, brings attention to the role of official intervention, an important dimension of international finance, which needs to be further investigated.

References

- [1] Alesina, Alberto and David Dollar. 2000. Who gives foreign aid to whom and why? *Journal of Economic Growth* 5: 33–63.
- [2] Asmus Gerda, Axel Dreher and Peter Nunnenkamp, 2016, Is Targeted Aid More Effective? Sector-specific Needs, the Composition of Aid and its Effects on Growth, unpublished Manuscript
- [3] Basurto, Pia, Pascaline Dupas and Jonathan Robinson. 2020. Decentralization and Efficiency of Subsidy Targeting: Evidence from Chiefs in Rural Malawi. *Journal of Public Economics* 185.

- [4] Beck, Thorsten, Asli Demirguc-Kunt and Vojislav Maksimovic. 2005. Financial and legal constraints to growth: does firm size matter? *Journal of Finance* 60: 137–177.
- [5] Bluhm, Richard, Axel Dreher, Andreas Fuchs, Bradley C. Parks, Austin Strange and Michael J. Tierney. 2020. Connective Financing: Chinese Infrastructure Projects and the Diffusion of Economic Activity in Developing Countries. CEPR Discussion Paper 14818.
- [6] Bompreszi, Pietro and Silvia Marchesi. 2021. A firm level approach on the effects of IMF intervention. CefES Working Paper 476.
- [7] Bräutigam, Deborah A., Stephen Knack. 2004. Foreign aid, institutions, and governance in Sub-Saharan Africa. *Economic Development and Cultural Change* 52: 255–285.
- [8] Busso, Matias, Lucia Madrigal, and Carmen Pagés. 2013. Productivity and Resource Misallocation in Latin America. *The BE Journal of Macroeconomics* 13: 903–932.
- [9] Chauvet, Lisa and Hélène Ehrhart. 2018. Aid and growth: evidence from firm-level data. *Journal of Development Economics* 135: 461–477.
- [10] Chen, Xi and William D. Nordhaus. 2011. Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences* 108: 8589–8594.
- [11] Clemens, Clemens, Steven Radelet, Rikhil Bhavnani and Samuel Bazzi, 2012, Counting chickens when they hatch: Timing and the Effects of Aid on Growth. *Economic Journal* 122, 561: 590-617.
- [12] Clemens Michael A and Mariapia Mendola. 2020. Migration from Developing Countries: Selection, Income Elasticity, and Simpson’s Paradox. Center for Global Development .Working Paper 539
- [13] Cuzatti, John C., Axel Dreher and Johannes Matzat. 2020. Chinese Aid and Health at the Country and Local Level. AidData Working Paper 97.
- [14] Del Prete Davide, Michele Di Maio. Aminur Rahman. 2019. Firms amidst the War. Mimeo.
- [15] Dennison Susi, Shoshana Fine and Richard Gowan. 2019. False moves: Migration and development aid. European Council Foreign Relation. Policy Brief.
- [16] Denizer, Cevdet, Daniel Kaufmann and Art Kraay. 2013. Good countries or good projects? Macro and micro correlates of World Bank project performance. *Journal of Development Economics* 105: 288-302.

- [17] Dollar, David, Mary Hallward-Driemeier, and Taye Mengistae. 2005. Investment Climate and Firm Performance in Developing Economies. *Economic Development and Cultural Change* 54(1): 1–31.
- [18] Doucouliagos, Hristos and Martin Paldam. 2009. The aid effectiveness literature: The sad results of 40 years of research. *Journal of Economic Surveys* 23(3):433–461.
- [19] Doucouliagos, Hristos. 2019. The Politics of International Aid. In Roger Congleton, Bernard Grofman, and Stefan Voigt (eds.), *Oxford Handbook of Public Choice*.
- [20] Dreher, Axel, Stephan Klasen, James Raymond Vreeland and Eric Werker. 2013. The costs of favoritism: is politically driven aid less effective? *Economic Development and Cultural Change* 62: 157-191.
- [21] Dreher, Axel and Steffen Lohmann. 2015. Aid and growth at the regional level. *Oxford Review of Economic Policy* 31: 420—446.
- [22] Dreher, Axel and Andreas Fuchs. 2015. Rogue Aid? An Empirical Analysis of China’s Aid Allocation. *Canadian Journal of Economics* 48: 988-1023.
- [23] Dreher, Axel, Sarah Langlotz and Silvia Marchesi. 2017. Information transmission and ownership consolidation in aid programs. *Economic Inquiry* 55: 1671-1688.
- [24] Dreher, Axel, Vera Eichenauer and Kai Gehring. 2018a. Geopolitics, Aid and Growth: The Impact of UN Security Council Membership on the Effectiveness of Aid. *World Bank Economic Review* 32: 268-286.
- [25] Dreher, Axel, Valentin F. Lang and Sebastian Ziaja. 2018b. Foreign aid, in: T. Risse, T. Börzel and A. Draude (eds.), *The Oxford Handbook of Governance and Limited Statehood*, pp. 394-415, Oxford University Press.
- [26] Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C. Parks, Paul A. Raschky, Michael J. Tierney. 2019. African Leaders and the Geography of China’s Foreign Assistance. *Journal of Development Economics* 140: 44-71.
- [27] Dreher, Axel, Andreas Fuchs, Bradley C. Parks, Austin Strange, Michael J. Tierney. 2021a. Aid, China, and Growth: Evidence from a New Global Development Finance Dataset. *American Economic Journal: Economic Policy* 13: 135-174.

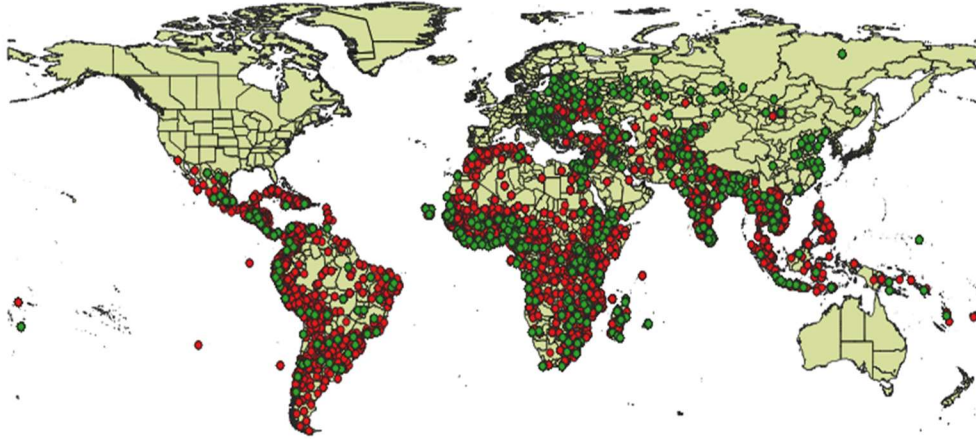
- [28] Dreher, Axel, Andreas Fuchs, Roland Hodler, Bradley C. Parks, Paul A. Raschky, Michael J. Tierney. 2021b. Is Favoritism a Threat to Chinese Aid Effectiveness? A Subnational Analysis of Chinese Development Projects. *World Development* 139: 105291.
- [29] Eslava, Marcela, John Haltiwanger and Alvaro Pinzón. 2019. Job creation in Colombia vs. the US: ‘up or out dynamics’ meets ‘the life cycle of plants’. NBER Working Paper 25550.
- [30] Faye Michael and Paul Niehaus. 2012. Political Aid Cycles. *American Economic Review*, 102(7): 3516–3530
- [31] Feeny, Simon and Vu Vuong. 2017. Explaining aid project and program success: Findings from Asian Development Bank Interventions. *World Development* 90: 329-34.
- [32] Fisman, Raymond and Jakob Svensson. 2007. Are corruption and taxation really harmful to growth? Firm level evidence. *Journal of Development Economics* 83: 63–75.
- [33] Foster, L., Haltiwanger, J. and C.J. Krizan (2001) Aggregate Productivity Growth: Lessons from Microeconomic Evidence, in Edward Dean, Michael Harper and Charles Hulten (eds.) *New Developments in Productivity Analysis*, Chicago: University of Chicago Press.
- [34] Galiani, Sebastian, Stephen Knack, Lixin Colin Xu and Ben Zou. 2017. The effect of aid on growth: Evidence from a quasi-experiment. *Journal of Economic Growth* 22.
- [35] Gehring, Kai and Valentin F. Lang. 2020. Stigma or Cushion? IMF Programs and Sovereign Creditworthiness. *Journal of Development Economics* 146:1-14.
- [36] Gehring, Kai, Lennart Kaplan and Melvin Wong. 2021. China and the World Bank: How Contrasting Development Approaches Affect the Stability of African States, AidData Working Paper 87
- [37] Harrison, Ann E., Inessa Love and Margaret S. McMillan. 2004. Global capital flows and financing constraints. *Journal of Development Economics* 75: 269–301.
- [38] Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. 2004. Export versus FDI with Heterogeneous Firms. *American Economic Review* 94: 300-316.
- [39] Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. Measuring economic growth from outerspace. *American Economic Review* 102: 994–1028.
- [40] Hsieh, Chang-Tai and Peter J. Klenow. 2009. Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics* 124: 1403-1448.

- [41] Hsieh, Chang-Tai and Benjamin A. Olken. 2014. The Missing "Missing Middle. *Journal of Economic Perspectives* 28.
- [42] Isaksson, Ann-Sofie and Andreas Kotsadam. 2018. Chinese aid and local corruption. *Journal of Public Economics* 159: 146-159.
- [43] Jedwab, Remi and Alexander Moradi. 2016. The permanent effects of transportation revolutions in poor countries: evidence from Africa. *The Review of Economic and Statistics* 98: 268–284.
- [44] Kaja, Ashwin and Eric Werker. 2010. Corporate Misgovernance at the World Bank and the Dilemma of Global Governance. *World Bank Economic Review* 24: 171-198.
- [45] Kilby, Christopher. 2009. The political economy of conditionality: An empirical analysis of World Bank loan disbursements. *Journal of Development Economics* 89: 51-61.
- [46] Kilby, Christopher. 2015. Assessing the impact of World Bank preparation on project outcomes. *Journal of Development Economics*.115: 111-123.
- [47] Knack, Stephen. 2001. Aid dependence and the quality of governance: cross-country empirical tests. *Southern Economic Journal* 68: 310–329.
- [48] Kuziemko, Ilyana and Eric Werker. 2006. How much is a Seat on the Security Council Worth? Foreign Aid and Bribery at the United Nations. *Journal of Political Economy* 114: 905-930.
- [49] Marchesi, Silvia, Laura Sabani and Axel Dreher. 2011. Read my Lips: The Role of Information Transmission in Multilateral Reform Design. *Journal of International Economics* 84: 86-98.
- [50] Marchesi, Silvia and Tania Masi. 2021. Delegation of Implementation in project aid. *The Review of International Organizations* 16: 655-687.
- [51] Marchesi, Silvia, Tania Masi. and Saumik Paul. 2021. Project aid.and firm performance. IZA Discussion Paper 14705..
- [52] Melitz, M. J. 2003. The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71: 1695–1725.
- [53] Minasyan, Anna, Peter Nunnenkamp and K. Richert. 2017. Does aid effectiveness depend on the quality of donors? *World Development* 100:16–30.
- [54] Nunn, Nathan and Qian, Nancy. 2014. US Food Aid and Civil Conflict. *American Economic Review* 104:1630–66.

- [55] OECD-DAC (2019) <http://stats.oecd.org/qwids/> (last accessed November 2019)
- [56] Öhler, Hannes and Peter Nunnenkamp. 2014. Needs based Targeting or Favoritism? The Regional Allocation of Multilateral Aid within Recipient Countries. *Kyklos* 67: 420-446.
- [57] Paul, Saumik and Hironobu Isaka. 2019. Labor Income Share at the Firm Level: Global Trends, IZA Discussion Paper No. 12852.
- [58] Rajan, Raghuram and Arvind Subramanian. 2007. Does aid affect governance? *American Economic Review* 97: 322–327.
- [59] Rajan, Raghuram and Arvind Subramanian. 2011. Aid, Dutch disease, and manufacturing growth. *Journal of Development Economics* 94: 106–118.
- [60] Rajan, Raghuram and Luigi Zingales, 1998. Financial dependence and growth. *American Economic Review* 88: 559–586.
- [61] Rud, Juan Pablo. 2012. Electricity provision and industrial development: evidence from India. *Journal of Development Economics* 97: 352–367.
- [62] Syverson, Chad. 2011. What Determines Productivity? *Journal of Economic Literature* 49: 326–365.
- [63] Werker, Eric, Faisal Z. Ahmed, Charles Cohen. 2009. How is Foreign Aid Spent? Evidence from a Natural Experiment. *American Economic Journal: Macroeconomics* 1: 225–244.
- [64] Wooldridge, Jeffrey M. 2015. Control Function Methods in Applied Econometrics. *Journal of human resource*, 15: 420-445.
- [65] World Bank Global Spread Of Conflict By Country And Population. 2018. World Bank Group.
- [66] World Development Indicators (2021). World Bank.
- [67] World Enterprise Analysis Unit. 2018. Firm Level Productivity Estimates. World Bank.

Figures

Figure 1: European project and firm distribution across countries



Note: Green dots refer to WBES firms, while red dots are the European projects.

Figure 2: European share of projects by donors

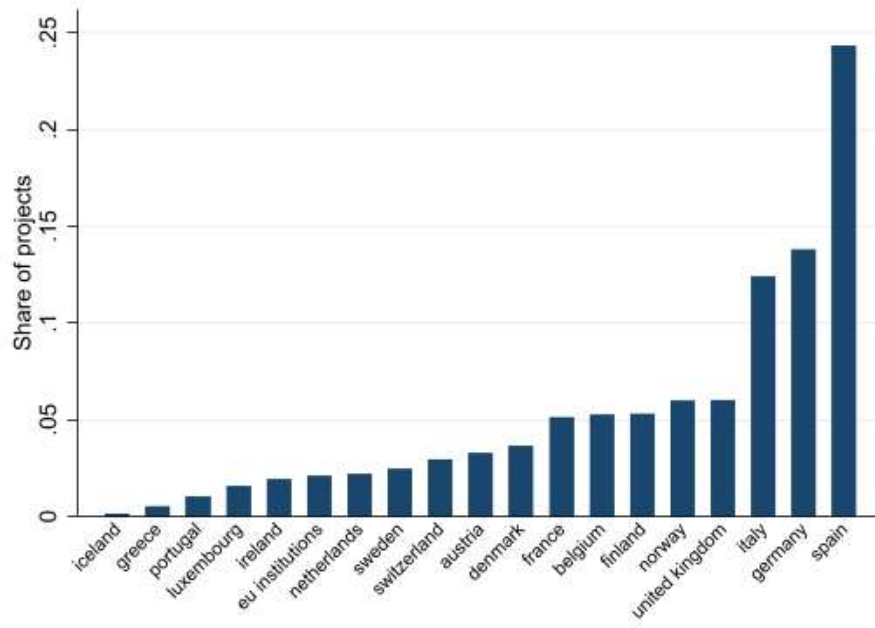


Figure 3: Number and size of projects to receiver countries, top 30th percentile

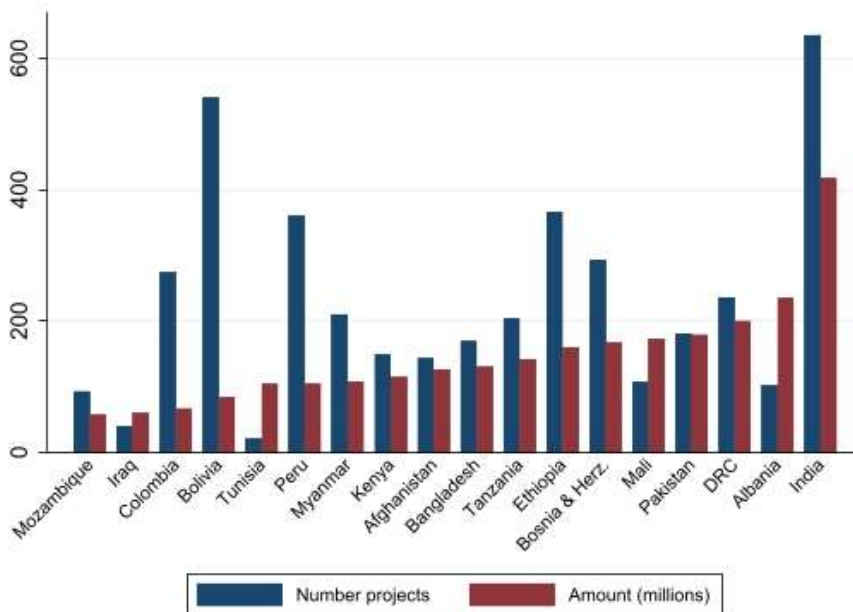


Figure 4: Number of projects by strategic aid sectors

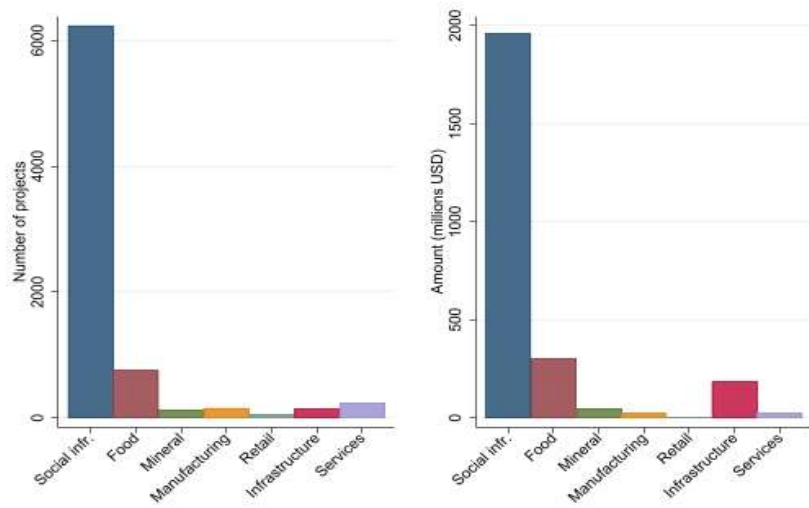
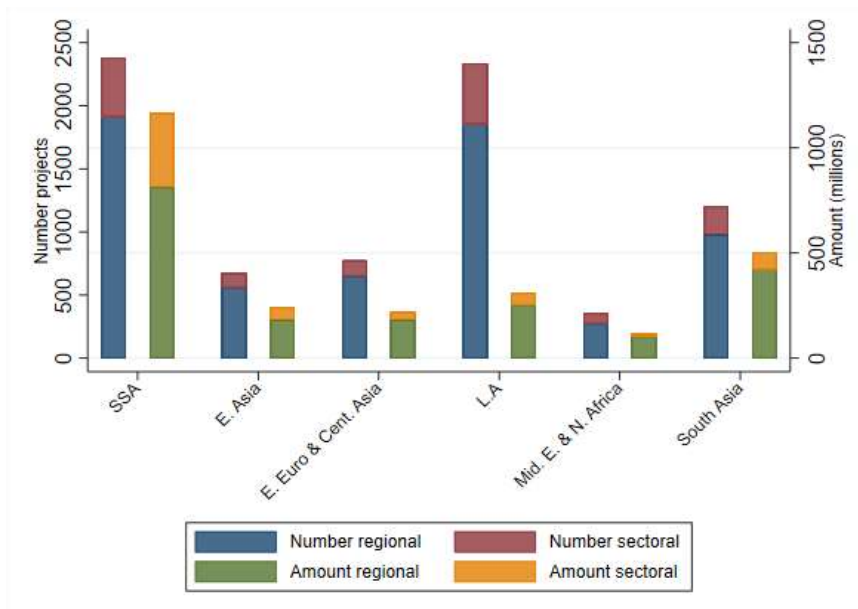


Figure 5: Aid projects by geographic regions and sectors



Tables

Table 1: Firm sales growth and European ODA projects

Dependent variable: Annual growth rate of sales	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Number of European ODA projects	0.001 (0.850)	0.001 (0.621)	0.000 (0.108)	-0.000 (-0.009)
Log Sales (base year)	-0.095*** (-16.732)	-0.095*** (-16.659)	-0.134*** (-9.601)	-0.134*** (-9.658)
State ownership (Yes/No)	-0.001 (-0.028)	-0.001 (-0.045)	-0.149* (-1.930)	-0.149* (-1.942)
Foreign ownership (Yes/No)	0.064*** (7.920)	0.064*** (7.907)	0.033 (1.125)	0.033 (1.138)
Exports goods (Yes/No)	0.043*** (6.642)	0.043*** (6.638)	0.020 (1.013)	0.020 (1.012)
Firm Size	0.172*** (18.341)	0.172*** (18.405)	0.125*** (5.839)	0.125*** (5.771)
Log regional population	-0.092 (-0.391)	-0.094 (-0.399)	-0.087 (-0.167)	-0.086 (-0.170)
First stage				
Tax Revenues x Probability		61.24***		54.81*
p-value		(0.003)		(0.081)
Observations	67,202	67,204	8,035	8,039
R-squared	0.221	0.125	0.635	0.175
Region FE	YES	NO	YES	NO
Firm FE	NO	YES	NO	YES
Industry x year FE	YES	YES	YES	YES
Kleibergen-Paap LM stat (p-value)		0.0085		0.1113
Kleibergen-Paap LM stat (F stat)		8.767		3.043
Panel observations		3938		3938

Notes: Column 1 uses an OLS estimator with region dummies. Column 2 uses an IV estimator with region dummies. Column 3 uses the within estimator with firm fixed effects. Column 4 uses an IV estimator with firm fixed effect. The variable of interest is log of committed aid. All models include industry-year dummies and firm and regional level controls. Kleibergen-Paap p-values are for the underidentification LM test. Standard errors are clustered at the regional level. t-statistics in parenthesis, ***p<0.01, **<p0.05, *p<0.1.

Table 2: Firm sales growth and aid projects by macro region

	(1)	(2)	(3)	(4)
Dependent variable: Annual growth rate of sales	OLS	IV	OLS	IV
<i>Africa</i>				
Aid	0.003** (2.413)	0.006** (2.030)	-0.000 (-0.088)	0.010 (1.071)
Observations	24,191	24,191	2,101	2,101
R-squared	0.267	0.127	0.680	0.139
Kleibergen-Paap LM stat (p-value)		0.0702		0.1021
Kleibergen-Paap LM stat (F-stat)		4.369		1.580
<i>Asia</i>				
Aid	0.004** (2.554)	0.007** (2.141)	0.008*** (2.822)	0.007 (0.821)
Observations	15,557	15,557	2,397	2,401
R-squared	0.233	0.115	0.621	0.173
Kleibergen-Paap LM stat (p-value)		0.062		0.018
Kleibergen-Paap LM stat (F-stat)		5.215		7.308
<i>Latin America</i>				
Aid	-0.001* (-1.699)	-0.001 (-1.088)	-0.000 (-0.423)	-0.000 (-0.076)
Observations	19,988	19,989	3,167	3,167
R-squared	0.189	0.146	0.678	0.282
Kleibergen-Paap LM stat (p-value)		0.006		0.051
Kleibergen-Paap LM stat (F-stat)		12.840		6.891
<i>Europe</i>				
Aid	0.000 (0.458)	0.063 (0.254)	0.001 (0.285)	-0.014 (-1.355)
Observations	7,466	7,467	368	370
R-squared	0.214	0.238	0.668	0.218
Kleibergen-Paap LM stat (p-value)		0.797		0.159
Kleibergen-Paap LM stat (F-stat)		0.061		6.654

Notes: Column 1 uses an OLS estimator with region dummies. Column 2 uses an IV estimator with region dummies. The variable of interest is log of committed amount. All models include industry-year dummies and firm and regional level controls. Kleibergen-Paap p-values are for the underidentification LM test. Standard errors are clustered at the regional level. t-statistics in parenthesis, ***p<0.01, **<p0.05, *p<0.1.

ONLINE APPENDICES

ONLINE APPENDIX A: European Aid Data and the World Bank Enterprise Survey

p. 2

Table A1: Distribution of firms across broad sectors in the WBES data

Table A2: The WBES sectoral classification (disaggregated level)

Table A3: Mapping of sectors between European projects and WBES data

Table A4: Typologies of WBP and WBES sector codes

Table A5: World Bank Enterprise survey years (2003 – 2016), by country

Table A6: European Project data (2001 – 2014), by country

Table A7: European Project, by donor and recipient macro region

Figure A1: WBES firms by geographic region

ONLINE APPENDIX B: Descriptive

p. 10

Table B1: Definition and Sources

Table B2: Summary statistics

Table B3: European project types

Table B4: Description of the Channels

Figure B1: European Commitments and Disbursements, by aid sector

Figure B2: European projects, by aid sector

Figure B3: Distribution of sales by sectors

Figure B4: Distribution of sales by sectors and regions

ONLINE APPENDIX C: Identification

p. 13

Table C1: Underlying trends (European projects): Global GDP growth and Chinese/WB # project aid

Figure C1: Parallel trend Europe

Figure C2: First stage marginal effects, Grants

Figure C3: Survival bias

ONLINE APPENDIX D: Geographical entity extraction

p. 16

Figure D1: Raw CRS data

Figure D2: Geo-entity extraction example

Figure D3: Model error decomposition

Figure D4: Distribution of projects by sector, pre and post geocoding

ONLINE APPENDIX E: Orbis ownership data

p. 20

Figure E1: Corporate ownership structures

Figure E2: Subsidiary firm

ONLINE APPENDIX A: World Bank Aid Data and the World Bank Enterprise Survey

A.1 Mapping European projects to World Bank Economic Survey data

This section describes the methodology that we follow to locate the firms from World Bank Enterprise Survey (WBES) data into each the region to which the European development projects are located. Since geocoded data allows us to identify the specific region of each European project within a country, as a first step we match regions from both data sets using the names of the regions. However, geocodes are not available from the WBES data. As a second-best approximation, we identify the latitude and longitude of the regions using the names of the regions that could not be directly matched from the names of the regions available in the European project data. Once the regions from both datasets are fully matched, we then follow three steps to identify and allocate each European project to specific firms within a region.

Step 1

First, we re-organize the World Bank economic survey sectors into six broad categories of industries: food, mineral, other manufacturing, wholesale and retail, transport, communication and hotel and other services. The distribution of 110,864 firms across this broad classification of industries is given in Table A1 below.

Table A1: Distribution of firms across broad sectors in the WBES data

WBES broad sector categories	Number	Percent
Firms in food sector (=1)	9,534	8.6%
Firms in mineral sector (=2)	11,749	10.6%
Firms in manufacturing sectors (=3)	40,511	36.6%
Firms in wholesale and retail sectors, hotels and restaurant (=4)	21,307	19.3%
Firms in transport, communications (IT) and construction (=5)	5,285	4.8%
Firms in other services sectors (=6)	22,211	20%

Notes: The total number of firms is 110597 as 267 observations are lost due to missing industry code. *Source:* authors' calculations based on World Bank Enterprise Survey (WBES) data.

We compute this table using the mapping presented in Table A2 of a more disaggregated classification of sectors into these six broad categories.

Step 2

As a second step we use the description of the European k projects to identify whether they are region-specific (related to all firms) or sector-specific (related to firms in a sector). As shown in Table A3, there are in total X European project sectors, which are again regrouped into seven WBES categories (we add region-specific projects as the seventh category that affects firms from all sectors in a region). The seven WBES categories are summarized in Table A4.

Step 3

Next, we apply the above mapping to a feasible period of analysis. The World Bank Enterprise Survey data is available for the period from 2003 to 2016, whereas the information of European projects is available from 1995 to 2014. Following the literature, we use two-year lag assuming it takes about two years for a firm to potentially benefit since a European project is committed. This allows us to evaluate any development commitments taking place between 2001 and 2014, and the firm-level outcomes realized in the period from 2003 to 2016. The two tables A5 and A6 below then show the year in which European projects were undertaken and the number of firms surveyed in each round.

Table A2: The WBES sectoral classification (disaggregated level)

Code	Sector	WBES Broad categories
1	Basic Metals & Metal Products	2 Mineral
2	Basic Metals/Fabricated Metals/Machinery	2 Mineral
3	Chemicals & Chemical Products	2 Mineral
4	Chemicals, Non-Metallic Mineral, Plastic	2 Mineral
5	Chemicals, Plastics & Rubber	2 Mineral
6	Construction	6 Other services
7	Electronics	3 Other manufacturing
8	Electronics & Communications Equip.	3 Other manufacturing
9	Fabricated Metal Products	2 Mineral
10	Food	1 Food
11	Food/Leather/Wood/Tobacco/Rubber Product	3 Other manufacturing
12	Furniture	3 Other manufacturing
13	Garments	3 Other manufacturing
14	Hospitality & Tourism	5 Transport, communication and hotels
15	Hotels & Restaurants	5 Transport, communication and hotels
16	IT & IT Services	5 Transport, communication and hotels
17	Leather Products	3 Other manufacturing
18	Machinery & Equipment	3 Other manufacturing
19	Machinery & Equipment & Electronics	3 Other manufacturing
20	Machinery & Equipment, Electronics	3 Other manufacturing
21	Manufacturing	3 Other manufacturing
22	Manufacturing Panel	3 Other manufacturing
23	Minerals, Metals, Machinery & Equipment	2 Mineral
24	Mining Related Manufacturing	2 Mineral
25	Motor Vehicles	3 Other manufacturing
26	Motor Vehicles & Transport Equip.	3 Other manufacturing
27	Non-Metallic Mineral Products	2 Mineral
28	Other Manufacturing	3 Other manufacturing
29	Other Services	6 Other services
30	Other Services Panel	6 Other services
31	Petroleum products, Plastics & Rubber	2 Mineral
32	Printing & Publishing	3 Other manufacturing
33	Rest of Universe	3 Other manufacturing
34	Retail	4 Wholesale and retail trade
35	Retail & IT	4 Wholesale and retail trade
36	Retail Panel	4 Wholesale and retail trade
37	Rubber & Plastics Products	2 Mineral
38	Services	2 Mineral
39	Services of Motor Vehicles	4 Wholesale and retail trade
40	Services of Motor Vehicles/Wholesale/Re	4 Wholesale and retail trade
41	Textiles	3 Other manufacturing
42	Textiles & Garments	3 Other manufacturing
43	Textiles, Garments, Leather & Paper	3 Other manufacturing
44	Tourism	5 Transport, communication and hotels
45	Transport	5 Transport, communication and hotels
46	Transport, Storage, & Communications	5 Transport, communication and hotels
47	Wholesale	4 Wholesale and retail trade
48	Wholesale & Retail	4 Wholesale and retail trade
49	Wood Products	3 Other manufacturing
50	Wood Products & Furniture	3 Other manufacturing
51	Wood products, Furniture, Paper & Public	3 Other manufacturing

Table A3a: OECD provided aid sectors

Broad sector	Sector names and codes
Agriculture, Forestry and Fishing	Agriculture (311), Forestry (312), Fishing (313)
Banking, Business, Financial Services	Banking and Financial services (240), Business and other services (250)
Budget Support	General budget support (510), Development food assistance (520), Other commodity assistance (530), Actions related to debt (600)
Energy and Mineral	Energy Policy (231), renewable energy generation (232), non-renewable energy generation (233), Hybrid energy (234), Nuclear energy (235), Energy distribution (236), Mineral resources and mining (322)
Emergency and Disaster	Emergency response (720), Reconstruction relief (730), Disaster prevention (740)
Environmental protection	General environment protection (410)
Industry	Industry (321)
Multisector/Unspecified	Multisector (430), Unspecified (998)
Trade and tourism	Trade policies (331), Tourism (332)
Transport Communication and Construction	Transport and storage (210), Communications (220), Construction (323)
Social infrastructure	Education (111), Basic education (112), Secondary education (113), Post-secondary education (114), Health (121), Basic health (122), Population policies/programmes (130), Water supply & sanitation (140), Government & civil society (151), Conflict peace & security (152), Other social infrastructure (160)

Table A3b: Mapping of sectors between European projects and WBES data

European projects	WBES Categorise
Agriculture, Forestry and Fishing	1
Banking, Business, Financial Services	0
Transport Communication and Construction	5
Emergency and Disaster	0
Energy and Mineral	2
General Budget Support	0
Environmental Protection	0
Industry	3
Social infrastructure and services	0
Trade and Tourism	4
Multisector/Unspecified	

Notes: The 11 European sector are mapped into 7 WBES sector classification groups as shown in Table A4

Table A4: Typologies of European and WBES sector codes

European project categories	WBES sector categories	Code
<i>Region-specific (social infrastructure)</i>	<i>Firms in all sector</i>	0
	<i>Firms in food sector</i>	1
Sector-specific	<i>Firms in mineral sector</i>	2
	<i>Firms in other manufacturing sectors</i>	3
	<i>Firms in wholesale and retail sectors, hotels and restaurant (=4)</i>	4
	<i>Firms in transport, communications (IT) and construction (=5)</i>	5
	<i>Firms in other services sectors</i>	6

Table A5: World Bank Enterprise survey years (2003 – 2016), by country

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Afghanistan	647	.	526
Albania	304	.	175	.	.	.	360	.	.	.
Angola	.	.	.	425	.	.	.	360
Antigua and Barbuda	151
Argentina	.	.	.	1,063	.	.	.	1,054	985.
Armenia	374	.	.	.	360	.	.	.
Azerbaijan	380	.	.	.	390	.	.	.
Bahamas	150
Bangladesh	1,504	.	.	.	250	.	1,442	.	.	.
Barbados	150
Belarus	273	360	.	.	.
Belize	150
Benin	.	197	150	150
Bhutan	250	253	.
Bolivia	.	.	.	613	.	.	.	362
Bosnia and Herzegovina	361	.	.	.	360	.	.	.
Botswana	.	.	.	342	.	.	.	268
Brazil	1,642	1,802
Bulgaria	1,015	.	288	.	.	.	293	.	.	.
Burkina Faso	.	.	.	139	.	.	394
Burundi	.	.	.	270	157	.	.
Cambodia	472	.	.	373
Cameroon	.	.	.	207	.	.	363	361
Cape Verde	.	.	.	98	.	.	156
Central African Republic	150
Chad	150
Chile	.	.	.	1,017	.	.	.	1,033
People's Republic of China	2,700
Colombia	.	.	.	1,000	.	.	.	942
Congo	151
Costa Rica	538
Croatia	633	.	159	.	.	.	360	.	.	.
Czech Republic	250	.	.	.	254	.	.	.
Côte d'Ivoire	526	361
DRC	.	.	.	340	.	.	.	359	.	.	529	.	.	.
Djibouti	266	.	.	.
Dominica	150
Dominican Republic	360	359
Ecuador	453	.	.	658	.	.	.	366
Egypt	2,897	.	.	1,814
El Salvador	.	.	.	693	.	.	.	360	719
Eritrea	179
Estonia	273	.	.	.	273	.	.	.
Eswatini	.	.	.	307	150
Ethiopia	644	.	.	.	848	.
Fiji	164
FYR Macedonia	366	.	.	.	360	.	.	.
Gabon	179
Gambia	.	.	.	174
Georgia	373	360	.	.	.
Ghana	494	720	.	.	.
Grenada	153
Guatemala	.	.	.	522	.	.	.	590
Guinea	.	.	.	223	150
Guinea Bissau	.	.	.	159
Guyana	165
Honduras	450	.	.	436	.	.	.	360	332

Hungary	291	.	.	.	310	.	.
India	9,281	.	.
Indonesia	1,444	1,320	.
Iraq	756
Israel	483	.	.
Jamaica	376
Jordan	573	.	.
Kazakhstan	544	.	.	.	600	.	.
Kenya	657	781	.	.
Kosovo	270	.	.	.	202	.	.
Kyrgyz Republic	235	.	.	.	270	.	.
Lao PDR	360	.	.	379	.	.	368
Latvia	271	.	.	.	336	.	.
Lebanon	561	.	.
Lesotho	151	150
Liberia	150
Lithuania	276	.	.	.	270	.	.
Madagascar	445	.	.	.	532	.	.
Malawi	150	523	.
Malaysia	1,000	.
Mali	155	.	.	.	490	.	360	185
Mauritania	.	.	.	237	150	.
Mauritius	398
Mexico	.	.	.	1,480	.	.	1,480
Micronesia	68
Moldova	363	.	.	.	360	.	.
Mongolia	362	.	.	.	360	.	.
Montenegro	116	.	.	.	150	.	.
Morocco	407	.	.
Mozambique	479
Myanmar	632	607
Namibia	.	.	.	329	580	.
Nepal	368	.	.	.	482	.	.
Nicaragua	452	.	.	478	.	.	336	333
Niger	.	.	125	.	.	150
Nigeria	1,891	.	3,157	.	.	.	2,676	.
Pakistan	935	1,247	.	.
Panama	.	.	.	604	.	.	365
Papua New Guinea	65	.
Paraguay	.	.	.	613	.	.	361
Peru	.	.	.	632	.	.	1,000
Philippines	1,326	1,335
Poland	455	.	.	.	542	.	.
Romania	541	.	.	.	540	.	.
Russian Federation	1,004	.	.	4,220	.	.	.
Rwanda	.	.	.	212	.	.	.	241
Samoa	109
Senegal	506	601	.	.
Serbia	388	.	.	.	360	.	.
Sierra Leone	150
Slovak Republic	275	.	.	.	268	.	.
Slovenia	276	.	.	.	270	.	.
Solomon Islands	151	.
South Africa	603	.	.	.	937
South Sudan	738	.
Sri Lanka	610
St. Kitts and Nevis	150
St. Lucia	150
St. Vincent and the Grenadines	154
Sudan	662	.
Suriname	152
Sweden	600	.	.

Tajikistan	360	359	.	.	.
Tanzania	.	.	.	419	813	.	.	.
Thailand	1,000
Timor-Leste	150	126	.
Togo	155	150
Tonga	150
Trinidad and Tobago	370
Tunisia	592	.	.	.
Turkey	1,152	1,344	.	.	.
Uganda	.	.	.	563	762	.	.	.
Ukraine	851	1,002	.	.	.
Uruguay	.	.	.	621	.	.	607
Uzbekistan	366	390	.	.	.
Vanuatu	128
Venezuela	.	.	.	120	.	.	320
Viet Nam	.	.	1,150	.	.	1,053	996	.
West Bank and Gaza	434	.	.	.
Yemen	477	.	.	353	.	.	.
Zambia	484	720	.	.	.
Zimbabwe	599	600

Table A6: European Project, by donor and recipient macro region

Donor country	Receiver region									
	Africa & Mid. East		Asia & Pacific		Europe		Latin America		Macro-regional or Unspecified	
Austria	1329	0.1349	596	0.0880	641	0.2653	533	0.1421	139	0.386
Belgium	2825	0.4140	717	0.2534	31	0.3471	1006	0.1719	69	0.2408
Denmark	1143	0.9441	770	0.8251	35	0.7352	145	0.5605	87	0.6999
EU Institutions	948	.	236	0.0143	294	.	152	.	195	.
Finland	1217	0.4280	845	0.2929	212	0.1205	317	0.1551	97	0.9096
France	1022	0.5086	345	0.3084	64	0.2579	287	0.2597	77	0.0395
Germany	6078	0.4891	6010	0.3959	762	0.4037	4030	0.1424	1	
Greece	133	0.1133	114	0.1400	334	0.3625	7	0.0561	20	0.6924
Iceland	52	0.3789	8	0.2315	4	0.0885	1	0.1655	15	0.1893
Ireland	1229	0.2125	200	0.1416	15	0.4783	57	0.0813		
Italy	5310	0.2553	1494	0.1626	784	0.0874	2312	0.0980	152	0.3116
Luxembourg	452	0.1843	828	0.1147	51	0.2677	602	0.1744	24	0.3897
Netherlands	913	0.8232	604	1.1269	206	0.8201	210	0.7008	70	1.6275
Norway	2138	0.6628	1599	0.5981	559	0.2815	559	0.3104	1	1.8687
Portugal	782	0.2783	154	0.9473	14	1.5973	35	0.0674	14	0.1830
Spain	6455	0.1295	1206	0.2135	348	1.1931	12708	0.1115	493	0.1796
Sweden	1445	1.0436	820	0.6533	581	0.3403	252	0.5858	44	1.0287
Switzerland	1276	0.5772	993	0.6913	491	0.7336	284	0.5082	67	2.3478
United Kingdom	1883	1.2906	1646	1.5430	364	0.2922	554	0.3752	9	0.2226

Notes: OECD aid decomposed by donor country and receiver region. For each receiver (macro) region, first column represents number of unique projects, and second column shows average commitment size (in million USD). Regional and unspecified are project occurring at the multi-country level, i.e., not attributable to the country or sub-country level. Source: OECD CRS (2019).

Figure A1: Data coverage of WBES vs. European project aid

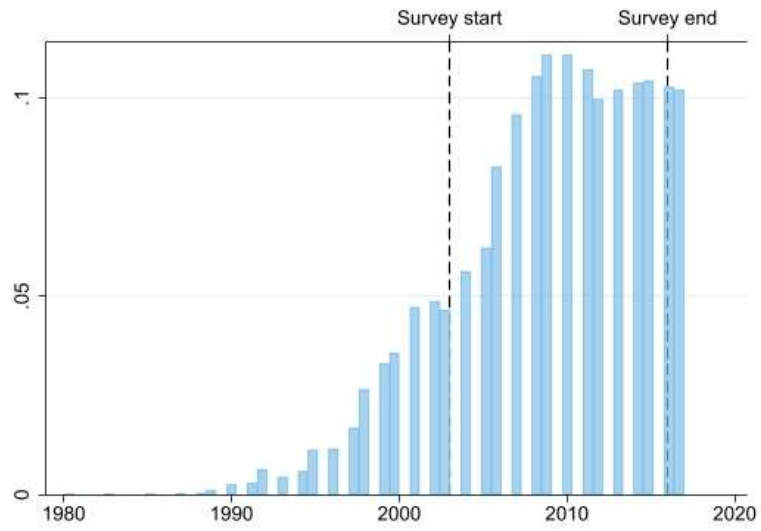
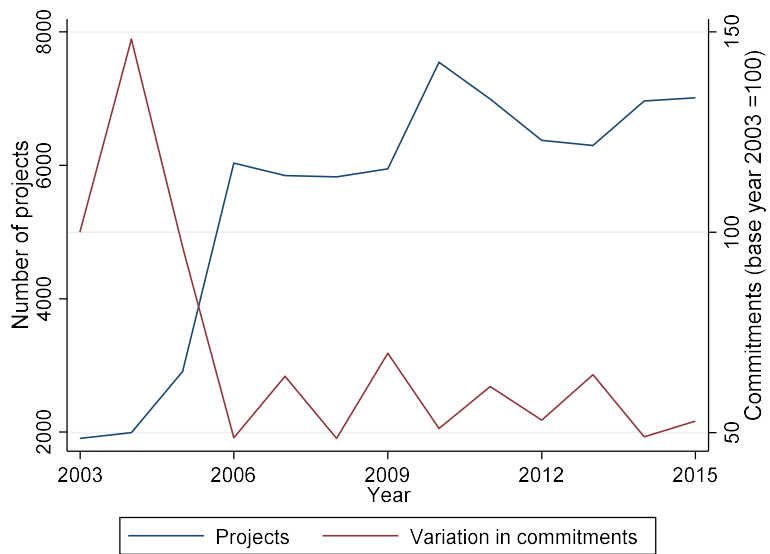


Figure A2: Projects and commitments over time



APPENDIX B: Summary statistics and variable definition

Table B1: Definition and Sources

Variable	Description	Source	Unit
FIRM			
Log Sales (base year)	Establishment Sales 3 Years Ago	World Bank Enterprise Survey	log
Sales growth	Average annual growth rate of sales	Own elaboration from WBES	%
Foreign	Dummy=1 if owned by private foreign individuals, companies or organizations	World Bank Enterprise Survey	Binary
Export	Dummy=1 if sales from indirect exports>0	World Bank Enterprise Survey	Binary
Size	Small, Medium, And Large Firm Categories Based On No. Of Employees	World Bank Enterprise Survey	1 Small(<20) 2 Medium(20-99) 3 Large(100 & over)
Firm has an overdraft facility	Dummy=1 if firms have an overdraft facility	World Bank Enterprise Survey	Binary
No obstacle with access to finance	No obstacle with access to finance, categorical variable (1-5)	World Bank Enterprise Survey	
Firm has internationally-recognized quality certification	Dummy=1 if firm has internationally-recognized quality certification	World Bank Enterprise Survey	Binary
Firm has a checking/saving account	Dummy=1 if firm has a checking/saving account	World Bank Enterprise Survey	Binary
Financial statements certified by external auditor	Dummy=1 if firm has financial statements certified by external auditor	World Bank Enterprise Survey	Binary
No obstacle with electricity	No obstacle with electricity, categorical variable (1-5)	World Bank Enterprise Survey	
No obstacle with transport	No obstacle with transport, categorical variable (1-5)	World Bank Enterprise Survey	
No obstacle with political instability	No obstacle with political instability, categorical variable (1-5)	World Bank Enterprise Survey	
No obstacle with crime, theft and disorder	No obstacle with crime, theft and disorder, categorical variable (1-5)	World Bank Enterprise Survey	
No obstacle with corruption	No obstacle with corruption, categorical variable (1-5)	World Bank Enterprise Survey	
REGIONAL			
Log regional population	Gridded population of the World (ADM1), log values	Hosted by CIESIN, at Columbia University (2000, 2005, 2010, 2015, 2020)	Log
Log regional GDP	Night-time lights	NOAA, National Geophysical Data Centre (1992-2013)	Log
Total # European project	Number of total European projects	Own elaboration from OECD CRS (2019)	
Tot Amount European project	Amount of total European projects	Own elaboration from OECD CRS (2019)	Log
p(Grant)	Number of Grant/Loan projects in the ADM1, two years before	Own elaboration from OECD CRS (2019)	
pa(Grant)	Amount of the Grant/Loan projects in the ADM1, two years before, millions of US\$	Own elaboration from OECD CRS (2019)	Log
IV			
European budget	Tax revenue (% of GDP). Following Chauvet and Ehrhart (2018) is a measure of the European donor capacity to commit to new financing of loans and grants at any point in time.	World Development Indicators (2021)	
Probability to receive European aid, Grants	Measures the share of years in the sample in which an ADM1 region received at least one European aid project.	Own calculation	

Table B2: Summary statistics

Variable	N	Mean	SD	Min	Max
FIRM CHARACTERISTICS					
Sales growth	67,202	0.11	0.45	-8.53	2.57
Log Sales (base year)	67,202	16.83	3.27	5.50	37.24
State	67,202	0.01	0.12	0	1
Foreign	67,202	0.12	0.33	0	1
Export	67,202	0.22	0.41	0	1
Size	67,202	1.72	0.76	1	3
REGIONAL VARIABLES					
Log regional population	67,202	14.68	1.45	7.81	18.35
European projects (ODA Grants)					
Number Total	67,202	19.79	25.16	0	153
Amount of European projects (mill. US\$)	67,202	3.08	11.80	0	158.64

Notes: These summary statistics refer to the OLS specification with region fixed effects

Table B3: Distribution of project commitments by flow type

	Number	Mean	S.d	Min	Max	Kurtosis
Flow Type						
ODA Grants	96182	.3787908	2.805383	0	392.36	5377.132
ODA Loans	3508	4.797094	21.33473	0	518.376	145.2091
Equity Investments	88	2.304792	6.008075	0	48.8305	42.80079

Notes: Distribution of project commitments, in millions of USD, for European donors over the matched geocoded sample (WBES and European ODA).

Figure B1: Distribution of firm sales by industrial sectors

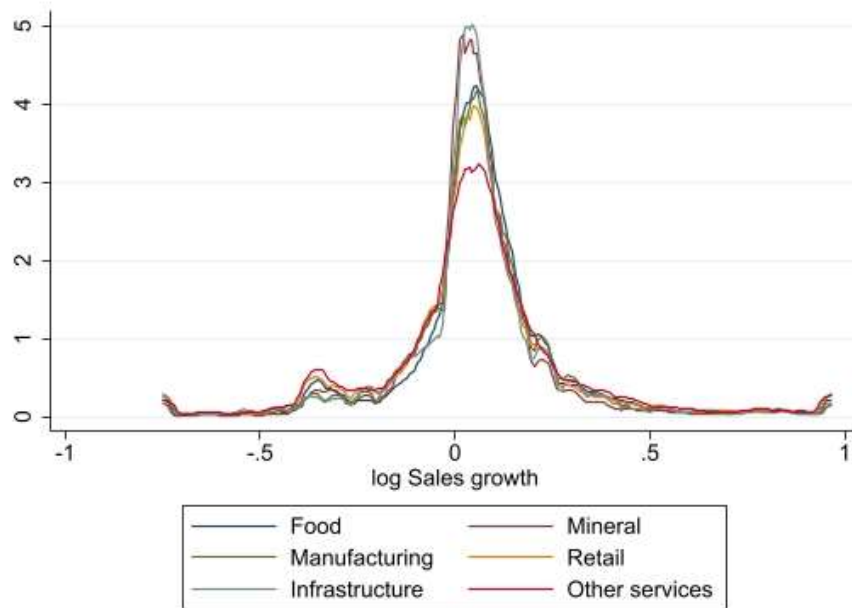


Figure B2: Distribution of sales by sectors and regions

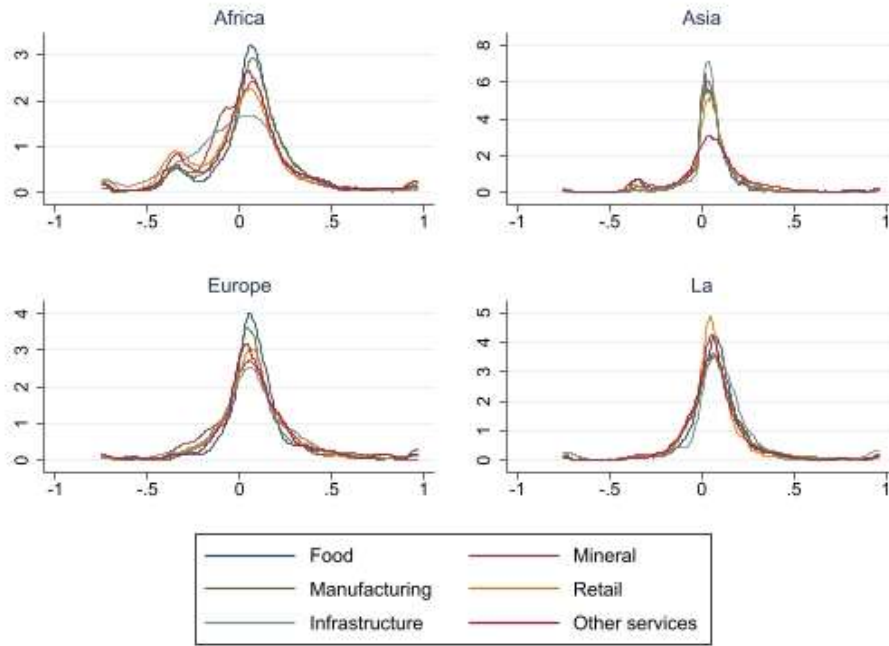
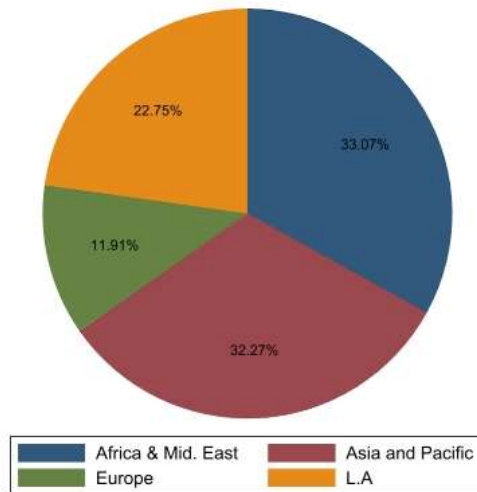


Figure B3: Firm representativeness by geographic region



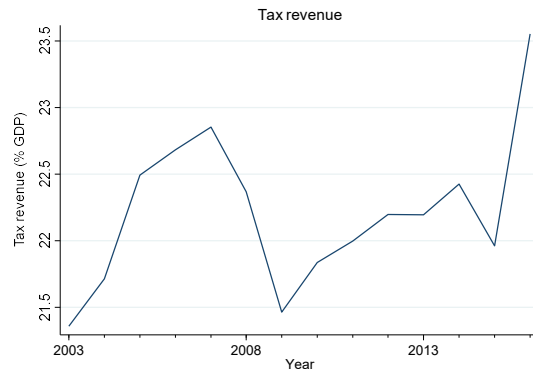
ONLINE APPENDIX C: Identification

Table C1: Underlying trends: Global GDP growth and Chinese and WB aid

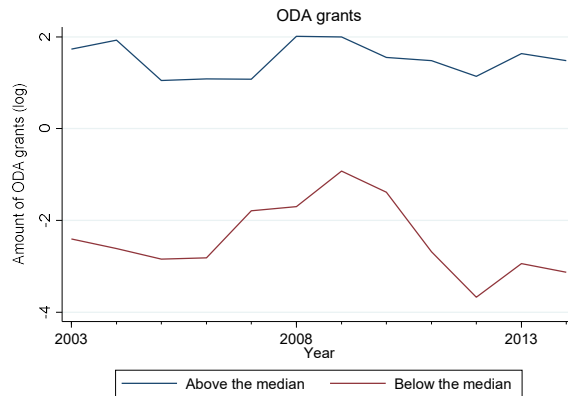
	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS	2SLS FE	2SLS	2SLS FE	2SLS	2SLS FE
Total Amount of European ODA Commitments	-0.038 (-1.003)	-0.019 (-1.024)	-0.044 (-1.058)	-0.025 (-1.336)	-0.044 (-1.139)	-0.027 (-1.410)
Global GDP x European Probability	0.036 (0.498)	0.019 (0.363)			0.003 (0.032)	-0.026 (-0.453)
Total Amount of Chinese ODA Commitments			0.005 (1.095)	0.004 (1.376)	0.005 (1.043)	0.005 (1.377)
Total Amount of World Bank ODA Commitments			-0.007 (-1.105)	-0.004 (-0.491)	-0.007 (-1.104)	-0.004 (-0.536)
Observations	67,204	7,807	67,204	7,807	67,204	7,807
R-squared	0.115	0.172	0.113	0.170	0.113	0.169
Region FE	YES	NO	YES	NO	YES	NO
Firm FE	NO	YES	NO	YES	NO	YES
Industry x year FE	YES	YES	YES	YES	YES	YES
Kleibergen-Paap LM stat (p-value)	0.0663	0.0439	0.0497	0.0065	0.0406	0.0069
Panel observations		3822		3822		3822

Notes: Standard errors are clustered at the regional level. t-statistics are reported in parentheses; *, ** and *** denote significance levels at 10%, 5% and 1%, respectively. Columns (1)-(2) control for the global GDP growth, interacted with the share of years, between 2000 and 2014, that region h received European aid. Column (3)-(4) control for both World Bank and Chinese ODA commitments, while columns (5)-(6) control for all.

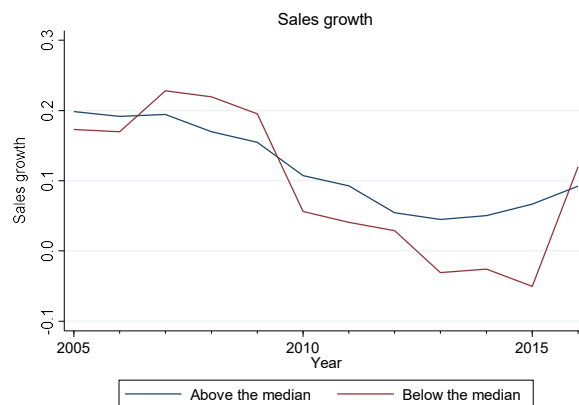
Figure C1: Parallel trends, Donor budgets and European ODA projects



Panel A



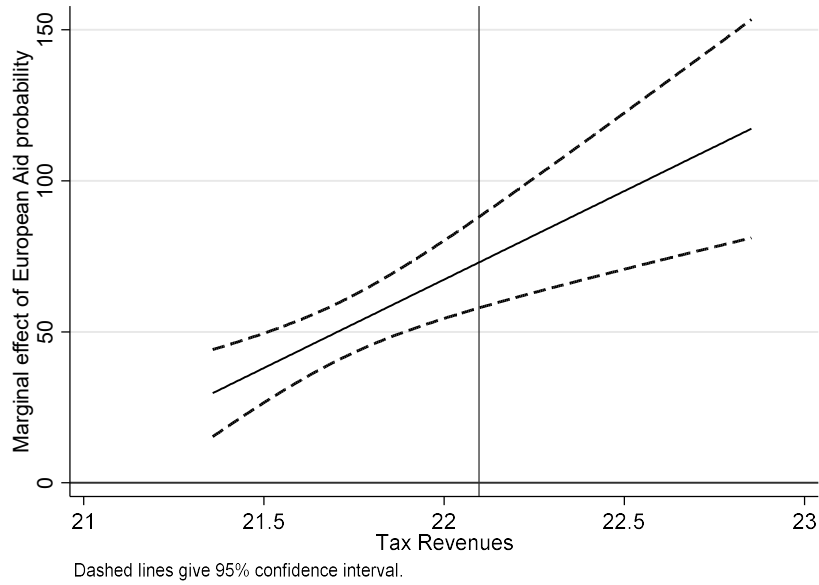
Panel B



Panel C

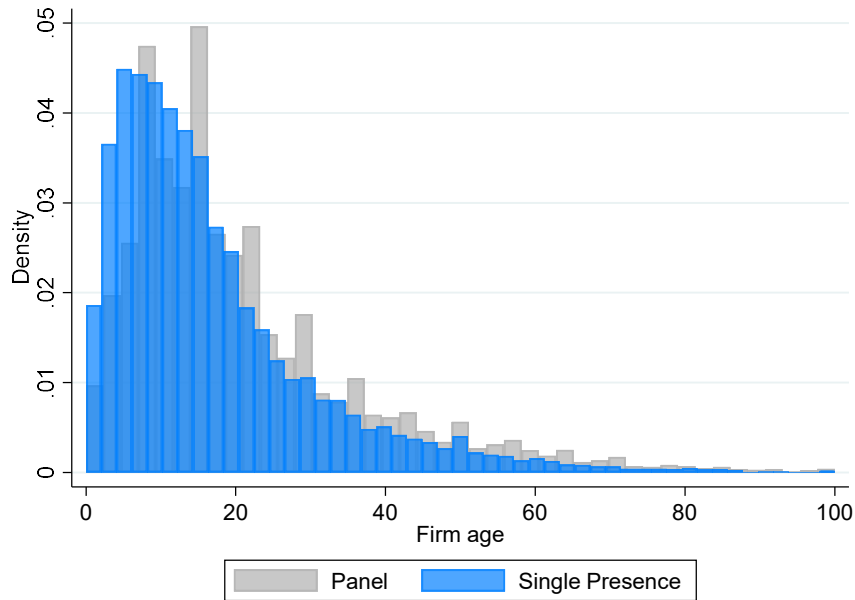
Notes: Panel A shows European donors' tax revenues. Panel B shows the three-year average amount of European ODA projects within the group of regions (ADM1) which is above the median of the probability of receiving projects, and the group that is below the median over time. Panel C shows the three-year average sales growth rate within these two groups over time. The averages are calculated using observations from the sample of Column 2, in Table 1.

Figure C2: First stage marginal effects, Grants



Notes: Marginal effects of share of past years under European ODA projects on current participation in European ODA projects, for differing levels of tax revenues in a given year. Corresponds to regression of Column 1, in Table 4. Dotted lines show 95% confidence interval.

Figure C3: Survival bias



Notes: Distributions of firm age in full sample (excluding conflict countries) for firms which only appear in one wave of the survey (single presence) versus firms that are recontacted at least once over different waves.

ONLINE APPENDIX D: Geographical entity extraction

The geocoding of CRS projects requires first the extraction of appropriate geographical information from CRS projects. Textual data associated to individual projects is entered in a free-form manner, so that the quality of descriptive information varies. This section outlines first the nature of the raw CRS data, then the steps taken in cleaning and evaluating this raw data, before explaining the geographical entity extraction model and supplementary text matching algorithms used. Finally, we present a series of model accuracy evaluation metrics and robustness checks and cleaning of the final data.

Raw data description

The OECD collects and publishes ODA data in 3 phases: i) aggregate level preliminary ODA data for the prior calendar year and forward spending plans for the next 3 years, ii) final detailed data including all project level data (CRS) for the prior calendar year, iii) update and revisions in June and September. For our work we rely on the Creditor Reporting System (CRS) data which provides detailed information on individual aid activities from which the aggregate data is derived. We bulk download the data under text file formats, as this is the only way to obtain commitments data prior to 1995 or disbursements prior to 2002. The CRS database provides information about projects along several dimensions: donor and recipient countries and regions, income group of receivers, donor agencies, channel of delivery (government, NGO, institutes), flow type (grants, loans, other official flows), sectors and sub-sectors of aid, commitment and disbursement amounts, expected start and completion dates, effective year, and project titles and descriptions. The data we are interested in for the purpose of identifying geographic locations comes primarily from the project title and description. However, we also rely on non-ambiguous sources of information like the receiver country in our procedure to cross-check results. Figure D1 shows a sample of our raw data with the text data.

Figure D1: Raw CRS data

crsid	donorcountry	receivercountry	projecttitle	longdescription	year
2010010962	Germany	Afghanistan	Stabilisation of German led region	Stabilisation program for Kunduz, Takhar, Badakhshan and Northern Baghlan. AKF/KfW	2010
1999001509z	Sweden	Afghanistan	UNICEF/AFG/99/EPI		1999
2009001595	United Kingdom	Afghanistan	Helmand Alternative Livelihoods Programme (HALP): Management Consultant	Helmandi farmers within the Food Zone increase wheat production and are supported by government institutions more able to implement rural livelihoods and counter-narcotics programmes.	2010
2011001477	Denmark	Afghanistan	Region of Origin Phase IIB 2012-2013	NSP is a national programme and an implementing partner for the ROI programme.	2011
2011000493	Netherlands	Afghanistan	BBC Kunduz	Civic Education	2014
2003015457z	Netherlands	Afghanistan			2005
2004004006z	France	Afghanistan	APPUI AU MINISTERE AFGHAN DE L'AGRICULTURE		2004
2009000802	Netherlands	Afghanistan	KAB Dihzak irrigation - SADA	KAB-URU UTR project irrigatie en stabiliteit	2009
2009060903	France	Afghanistan	Actions dans le domaine de la gouvernance		2009
2008010266	Germany	Afghanistan	Skateboarding hall Kabul	Build a skateboarding facility in Kabul to engage youth throughout Afghanistan, building technical skills, confidence and life opportunities	2008
2009001452	Sweden	Afghanistan	Afghanistan elec obs EU09	Val i Afganistan augusti 2009. 1 LTO rekryterad.	2010

The following sections provide an overview of our data compiling procedure.

Initial data cleaning

A first step consists in cleaning the data from duplicate entries. The provided CRS identifier fails to identify unique projects, and the presence of many projects lacking an ID requires some pre-analysis cleaning. Furthermore, a share of duplicate observations arises from the presence of large multi-sectoral projects, encompassing more than one type of aid. First, we identify all duplicate entries by sorting along all provided dimensions of the data and dropping duplicates. From our raw data of 18 selected European donors, we are left with 1,275,619 unique projects. The second part of the data pre-processing consists in identifying non-English language text. This is relevant for several reasons regarding the accuracy of the geo-entity extraction, as will be explained in greater detail in the following sections. We rely on a free Python library for language identification in this initial phase, but we are in the processing of refining this strategy.

Running model

To run the model, we rely on the Spacy library for natural language processing tools. We use the (pre-trained) Spacy core English transformer pipeline and leverage the Named Entity Recognition (NER) model. These models are typically used to identify within text pieces of information such as names, actions, or geopolitical entities. The advantage of this specific pipeline is in its speed, flexibility, and method of processing text data. Transformer models process all inputs bidirectionally, unlike traditional recurrent neural networks which process sequentially. This allows first for greater parallelization in computations and hence speed, and improved accuracy because the model learns to interpret sentences, or pieces of string, from multiple directions. Furthermore, this described parallelization has allowed for these models to be trained on massive datasets, thereby resulting in more accurate models. Specifically, we use the RoBERTa-base model trained on the entire English language Wikipedia and the online book corpus, a large online collection of digitalized books. When running the model on our data, the different components of the pipeline, such as the NER, all interact with the transformer component simultaneously, and different components not required can be switched off, allowing for gains in speed in our processing of the data. We run apply the NER feature of this pipeline to our three sources of text information for each project: the project title, the short description, and the long description. We obtain as an output then for each of these input strings a list of extracted entities by the model. Figure D2 shows a stylized example of what the model would identify in our text data.

Figure D2: Geo-entity extraction example

Project title:
Skateboarding hall **Kabul GPE**

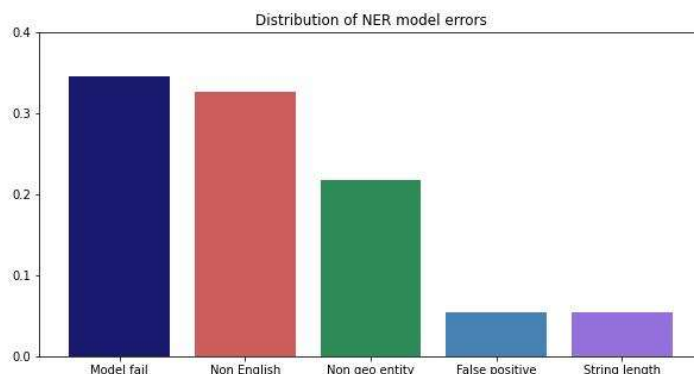
Project description:
Build a skateboarding facility in **Kabul GPE** to engage youth throughout **Afghanistan GPE**, building technical skills, confidence and life opportunities

The primary caveat regards the presence of non-English language texts. Model accuracy, as well be defined in the next section, cannot be evaluated accordingly on a non-English text because the model was pre-trained on English language text. Future iterations of this work will consider the use of non-English language transformer models, which have improved drastically in accuracy. Section 5 of this Appendix also presents a simple algorithm that has been implemented to extract entities where the NER fails. For now, the data presented in this version reflects the entities extracted from this NER model.

Evaluating model accuracy

In the end the model finds a geographic entity in at least one of the strings provided for 243,255 projects, or about 19%. We note that this does not mean that it missed information in 81% of projects, as many projects simply do not have geographic information contained in the text or do not have text at all. We evaluate our model with a use of a golden dataset. This is a random sample of 200 unique projects which are then hand-coded with the correct outcome the model should predict. This can either be the name of a geographical entity if it exists in the string, or nothing if not entity exists. We then run the same model on this dataset and confront the model outcomes to the true outcomes. We find that the model correctly identifies the outcome for 72% of projects. Figure D3 shows the decomposition of the remaining 28% of model errors on this golden dataset.

Figure D3: Model error decomposition



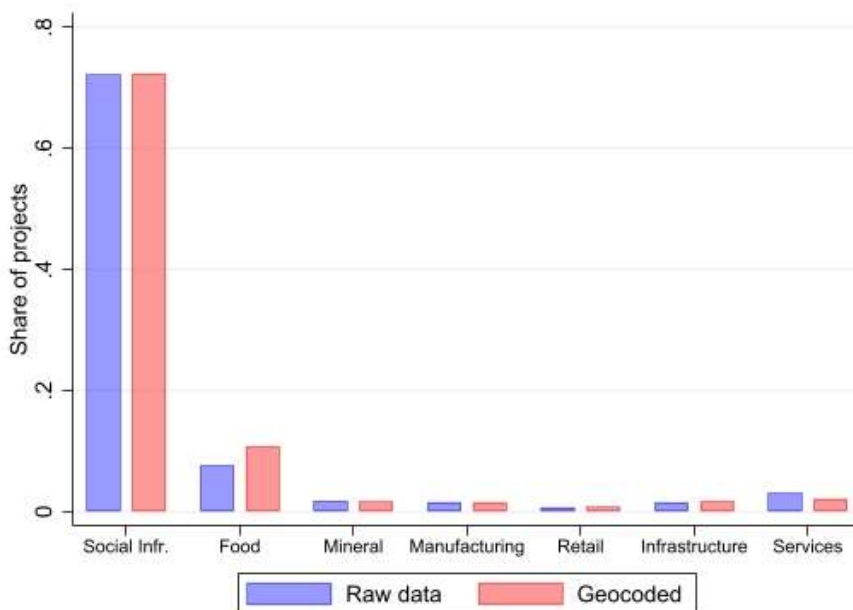
We can classify errors along 5 broad categories and by 2 error types. False positives represent the instances where the model classifies some part of the string a geopolitical entity when it is not. False negatives are the cases where the model misses an entity. First, for around 35% of the cases the model fails, either picking up false positives or reporting false negatives, for non-discernible reasons. This can be due to tricky syntax in the sentence for example. Around 25% of errors instead are due to the model reporting entities which are not geopolitical and hence not geo-referenceable. The most common example is proper names like People’s Bank. We report as a separate category of “False Positives” those cases where for example the “Swedish” was tagged, but it provides no relevant information to the location of the project. There are also a small number of false positives deriving from erroneous input of the raw data. These make up less than 10% of cases. Also, less than 10% of cases are those false negatives where the model fails to pick up an entity due to the string length being too short. Finally, a significant portion of errors derive from the non-English language texts, as we have before mentioned. We will mention in the next sections possible solutions to this.

Final data cleaning

As a final robustness check, before proceeding to geocode the extracted entities, we can perform a series of data cleaning and cross-checking procedures. To do this, we rely on the non-ambiguous geographic information in the raw data and a hierarchical application of a fuzzy matching algorithm. Specifically, we apply a term frequency - inverse dictionary frequency algorithm combined with a K-nearest neighbour (KNN) approach. In the first part, we split texts into chunks and filter out “noisy” words based on the frequency of words in the full dataset. The KNN algorithm then matches the candidate words with a hierarchical dataset of country administrative region names. This dataset consists in a set of organized text files, where for each country we have lists of ADM1 names and cities within these regions. The procedure is essentially a record-linkage approach, which returns a closeness score for each matched candidate word. We then only keep the match ranked as most precise. The use of this additional information for each project title and description is as a robustness. The advantage is that the algorithm always extracts at least one match for each string. The fact that we only match within a list of receiver country-specific regions and cities mitigates the issue of random matches. Furthermore, the availability of a precision score associated to each match, unlike with the NER output, allows us to quantitatively evaluate each match. We can use this additional information in the following way to deal with false positives and false negatives in the NER output. Identifying false positives is rather straightforward. First, we can run a simple string matching between the NER output and the KNN output in the instances when the KNN output precision metric corresponds to certainty (close to 100% matching). If in turn the NER output and the KNN are sufficiently close, we are more confident in the NER output. Similarly, we can run our record-linkage algorithm directly between the NER output and the country-specific list of geographic entities. Finally, it should also be noted that false positives are also thrown out in the geocoding procedure, when the GeoPy library is not able to identify the input as a geographic location. As a final cross-check, the georeferenced entities from GeoPy are cross checked to ensure they are of the receiver country. Dealing with false negatives is trickier. As we showed in Figure E3, the majority of missed cases stem from the presence of non-English language text. We can credibly fill in some of these gaps for the cases where the KNN output has precision close to 100%, and the NER or CRS provided geography

data is missing. However, relying on the KNN output without a cross-reference when the precision metric is not very high results in too many errors. For this reason, the next iteration of the dataset requires the use of language specific NER models. To finish, Figure D4 shows the distribution of the share of total projects for the raw European ODA data, before the ge-entity extraction and geocoding, and on the final dataset with only the geocoded and collapsed data. As can be seen, the distribution is largely the same, providing evidence that the procedure outline in this section did not introduce excessive biases in the data through sample selection.

Figure D4: Distribution of projects by sector, pre and post geocoding



ONLINE APPENDIX E: Orbis ownership data

Figure E1: Corporate ownership structures



PRIDE FORAMER
LESCAR, France

Active

Private

BvD ID: FR689800886 Orbis ID: 003870862 The Global Ultimate Owner of this controlled subsidiary is VALARIS LTD

Geographic footprint

Number of companies in the corporate group per country



The map represents the geographic location of the companies in the corporate group

All 247 companies in the corporate group are represented

Legend

Number of entities by country

- More than 40
- From 19 to 40
- From 7 to 19
- Less than 7

- 📍 Country of the company
- 📍 Country of the Global Ultimate Owner

Ultimate Owner definition

The path from the company to its Ultimate Owner is minimum 50.01%

I consider a company the Ultimate Owner if there are no identified shareholders or if the shareholder percentage is not known.

Figure E2: Subsidiary firm



PRIDE FORAMER

LUANDA, Angola

Active

Private

BvD ID: AO100005268 Orbis ID: 231305439 This entity is a foreign company of PRIDE FORAMER.

Key information

RUA DO FUTUNGO, 54
 LUANDA
 Angola
 Phone: +244 222 332 231
 Website: www.pride.com

Activity: Corporate, Mining & Extraction

Operating revenue (Turnover) for 2019
\$ 16.0 m

P/L [=Net Income] for 2019
 Not available

Ownership
 1 shareholder
 0 subsidiary
 247 companies in the corporate group

PEPs and sanctions
 This company is not the same or similar to a risk relevant name

Financial profile
 Limited financials, Local registry filing

	2019
	USD
	12 months
Operating revenue (Turnover)	16,040,000
Number of employees	1,450