DISCUSSION PAPER SERIES

DP17838

FOREIGN DIRECT INVESTMENT AND STRUCTURAL TRANSFORMATION IN AFRICA

Bernard Hoekman, Marco Sanfilippo and Margherita Tambussi

DEVELOPMENT ECONOMICS AND INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

FOREIGN DIRECT INVESTMENT AND STRUCTURAL TRANSFORMATION IN AFRICA

Bernard Hoekman, Marco Sanfilippo and Margherita Tambussi

Discussion Paper DP17838
Published 23 January 2023
Submitted 10 January 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

FOREIGN DIRECT INVESTMENT AND STRUCTURAL TRANSFORMATION IN AFRICA

Abstract

This paper analyzes the relationship between inward FDI and structural transformation of local labour markets in Africa. We combine geolocalized information on the distribution of FDI with a novel database that provides information from 40,665,627 individuals in 2,570 subnational units over the period 1987-2019. Results are suggestive of a positive effect of FDI on structural transformation. FDI contributes to an increase in employment, and shifts of workers towards modern industries and higher-skilled occupations. No effects are found on self-employment. Results are heterogeneous, reflecting the characteristics of the foreign investor and of the business activity undertaken by foreign firms in the local market. Geospatial analysis of changes in performance of domestic firms exposed to nearby FDI projects provides evidence of horizontal spillovers and inter-industry linkages, suggesting a complementary mechanism through which FDI drives structural change.

JEL Classification: L16, F23, N17

Keywords: Jobs, Africa

Bernard Hoekman - bernard.hoekman@eui.eu
Robert Schuman Centre for Advanced Studies, EUI and CEPR

Marco Sanfilippo - marco.sanfilippo@unito.it
University of Turin and Collegio Carlo Alberto

Margherita Tambussi - margherita.tambussi@unito.it
University of Turin

Acknowledgements
We would like to thank Scott Cunningham, Giulia Malevolti, Carlo Pietrobelli, and participants to the DevEconMeet Workshop in Florence, and the UNCTAD-EUI roundtable on "Implementing the African Continental Free Trade Area" for their feedback on preliminary draft versions of the paper. Nicol`o Florenzio provided excellent research assistance. This paper has received financial support from PEDL (Project Number 8498 "The Local Effects of FDI").

Foreign Direct Investment and Structural Transformation in Africa

Bernard Hoekman Marco Sanfilippo

Margherita Tambussi*

This draft: January 10, 2023

Abstract

This paper analyzes the relationship between inward FDI and structural transformation of local labour markets in Africa. We combine geolocalized information on the distribution of FDI with a novel database that provides information from 40,665,627 individuals in 2,570 subnational units over the period 1987-2019. Results are suggestive of a positive effect of FDI on structural transformation. FDI contributes to an increase in employment, and shifts of workers towards modern industries and higher-skilled occupations. No effects are found on self-employment. Results are heterogeneous, reflecting the characteristics of the foreign investor and of the business activity undertaken by foreign firms in the local market. Geospatial analysis of changes in performance of domestic firms exposed to nearby FDI projects provides evidence of horizontal spillovers and inter-industry linkages, suggesting a complementary mechanism through which FDI drives structural change.

Keywords: FDI, Jobs, Structural Transformation, Africa

JEL Codes: L16; F23; N17

*We would like to thank Scott Cunningham, Giulia Malevolti, Carlo Pietrobelli, and participants to the DevEconMeet Workshop in Florence, and the UNCTAD-EUI roundtable on "Implementing the African Continental Free Trade Area" for their feedback on preliminary draft versions of the paper. Nicolò Florenzio provided excellent research assistance. This paper has received financial support from PEDL (Project Number 8498 "The Local Effects of FDI"). Hoekman: European University Institiute and CEPR. E-mail: bernard.hoekman@eui.eu; Sanfilippo: University of Turin and Collegio Carlo Alberto. E-mail: marco.sanfilippo@unito.it; Tambussi: University of Turin. E-mail: margherita.tambussi@unito.it.

1 Introduction

Foreign direct investment (FDI) is considered by policymakers and practitioners to be an important driver of local economic development (Javorcik, 2018). For example, foreign investors may establish supply chain linkages with domestic firms, enhancing their productive capacities (e.g. Alfaro-Urena et al., 2022, on Costa Rica), or impact the labour market by creating new employment opportunities, better jobs and the mobility of workers (e.g. Setzler and Tintelnot, 2021; Poole, 2013). The empirical literature on the effects of FDI is mixed. Systematic evidence covering a broad set of development-related outcomes is limited. While there is an extensive literature on the impact of FDI on growth and economic development more generally, very little research has been done to understand whether FDI matters for structural transformation (e.g. Alviarez et al., 2021; Muhlen and Escobar, 2019; Liu, 2022). Moreover, and with some exceptions (e.g. Toews and Vezina, 2022; Sonno, 2020; Mendola et al., 2022), empirical work generally is not granular enough to account for the heterogeneous features that can affect the "quality" of FDI projects (Javorcik, 2015), or their impact at the subnational level.

In this paper we use finely disaggregated data to evaluate the consequences of attracting FDI projects at the *local* level, conditioning on the specific activity performed by foreign investors. We look specifically at the role of FDI in driving the process of structural transformation at the sub-national level for a sample of 24 African countries over the past 30 years. By introducing new knowledge and resources into a given locality, the entry of foreign firms expands and updates the domestic knowledge pool, which is one of the drivers of structural transformation (Fu et al., 2022). FDI can influence both supply and demand forces driving structural transformation. On the supply side, foreign firms generally bring in more advanced technology to the host location, supporting productivity growth, a potential reduction in informality, or easing capital constraints (Bau and Matray, 2022). On the demand side, FDI can affect economic activity through linkages with local suppliers and income effects (Liu, 2022). Ultimately the net impact will depend on the characteristics of FDI and the host location.

Our analysis is based on geolocalized microdata on 4,918 FDI projects in 24 countries from fDi Markets, a proprietary database that provides information on the distribution of greenfield investments from 2003 onwards.[1] To recover information on the local labour market, we construct a novel database that combines population censuses from IPUMS International with information from the Demographic and Health Surveys (DHS). The information provided by these two sources is harmonized using a common administrative division identifier that is consistent over time. Our final sample spans data on 40,665,627 individuals over the period 1987-2019 in 2,567 subnational units in the 24 countries considered. We link the entry of FDI to a set of indicators related to the structural transformation of local labour markets, focusing on three dimensions of structural transformation: (1) the shift of workers from agriculture to modern sectors; (2) the shift of workers from low- to high-skilled occupations; and (3) the shift of workers out of self-employment.[2] All findings are disaggregated by the gender of workers.

For our empirical analysis, we define a geographic area as treated when it receives its first FDI project. We then compare the outcomes of interest for treated areas with the same areas before treatment and a control group of areas that did not receive FDI. We start with some descriptive evidence linking FDI to structural transformation at the administrative level in a specification that includes location

[1] fDi Markets is widely used source of information on FDI by international organizations (including UNCTAD's annual World Investment Report) and the academic community (Toews and Vezina, 2022; Brazys and Kotsadam, 2020). More information at the company's webpage.

[2] The latter is an imperfect proxy of informality. Bandiera et al. (2022) show that self-employment dominates the share of jobs in less developed countries, and Donovan et al. (2022) provide evidence showing that self-employment is likely to be associated with frequent shits into and out unemploymnet, rather than into wage work.

and country-wave fixed effects. Results are suggestive of a positive relationship between FDI and indicators of structural transformation. The presence of FDI is correlated with an increase in employment and a shift in the composition of the workforce towards modern sectors (away from agriculture) and higher-skilled occupations.

We complement the descriptive evidence with an event study analysis to help rule out any pre-trends in the performance across areas that receive FDI and those that never do so. The event study also enables us to identify if there is a persistent effect of receiving FDI on structural transformation. We follow recent econometric literature on the implementation of two way fixed effects (TWFE) and event studies (Roth et al., 2022; De Chaisemartin and D'Haultfoeuille, 2022), and adopt the approach proposed by Callaway and Sant'Anna (2021) for staggered difference-in-differences design, using the doubly-robust DiD estimator developed by Sant'Anna and Zhao (2020). This method has the advantage of correcting the potential bias due to the presence of negative weights while conditioning pre-trends to a set of covariates. The latter is important given that the presence of FDI projects tends to correlate with certain observable characteristics of the host location such as the rate of urbanization and schooling in the treated areas.

The event study analysis supports the descriptive findings. Treated locations experience an increase in employment subsequent to receiving FDI. The effect on the occupational composition of workers is limited to an increase in employment of high skilled workers in the year of the treatment, which does not persist over time. On the other hand, the role of FDI as driver of structural transformation is more evident and persistent.

To understand the implications of potential heterogeneity of FDI, we test alternative definitions of the treatment on the basis of the country of origin and the activity performed by foreign firms in the field. This reveals that most of the findings are driven by FDI projects involving the establishment of new production facilities. Conversely, we show that the entry of foreign firms in high-value added services (e.g. financial, business, R&D) drives a change in the composition of the labour force towards more skilled workers. Investment in extractive activities is not associated with structural transformation or skill upgrading, while there is some, albeit weak, long-run evidence of increases in the share of self-employment. No substantial differences emerge for FDI coming from the North (OECD countries) or the South.

We show that our results hold for a sub-sample that considers only women, who are less likely to reduce their agricultural activities but that on the other hand see some increases in self-employment, and if we exclude the most attractive areas of the country. Furthermore, we run a battery of robustness checks, which show that results continue to obtain using alternative methods. These include use of other estimators that have been proposed in the literature for the staggered setting (De Chaisemartin and d'Haultfoeuille, 2020; Borusyak, 2022).

In the final part of the paper we explore a potential demand side mechanism through which FDI may be associated with the performance of domestic firms. Specifically, for a sample that includes the 24 countries covered by our analysis, we match fDi Markets data with firm level information from the World Bank Enterprise Surveys (WBES). The data cover over 26,000 domestic firms operating in manufacturing and service sectors. We link FDI to domestic firms using both their geographic location and sector of activity. We exploit the spatial and temporal features of the FDI project and enterprise survey data by comparing the performance of domestic firms located in relative proximity to FDI projects to that of firms in locations where FDI will occur in years subsequent to when the survey data were collected. The resulting difference-in-difference controls for possible selection effects. We show that exposure to FDI (defined both in terms of horizonthal and vertical linkages) is associated with growth in sales and employment, and upgrading in domestic firms, especially in the manufacturing sector. Overall, this demand side analysis supports the idea of FDI generating new economic opportunities for neighboring domestic firms that in turn stimulates demand for workers and new skills in modern sectors of the economy.

The literature on FDI in developing countries has largely focused on economic effects, either at the macro or at the firm level (Alfaro, 2017; Javorcik, 2018; Lay and Tafese, 2020). Bringing the analysis down to the level of individuals and analyzing the local effects of FDI has several advantages compared to more aggregated studies. Following the impact evaluation literature, linking survey respondents to nearby events provides more compelling evidence on changes in the socioeconomic conditions of individuals living in proximity to an FDI project and to account for both the direct and indirect effects of local exposure to FDI projects. Toews and Vezina (2022) find large local multiplier effects of FDI on employment in Mozambique, especially for women and for the skilled.[3] Mendola et al. (2022) find a more general positive effect of FDI on local employment in a large sample of African countries.[4] We extend this literature by constructing a new geospatial dataset covering many African countries over time that permits us to analyze the relationship between FDI and structural transformation observed in African labour markets, with a focus on both the supply side of the labour market and the demand side, reflected in the performance of domestic firms that are located in the proximity of FDI projects.

Notwithstanding the presumption that the "quality" of FDI matters for devel-

[3]Also a recent work by McCaig et al. (2022) finds large employment effects due to the entry and the subsequent growth of foreign affiliates in Vietnam in respose to trade liberalization with the US.

[4]Mendez and Van Patten (2022) measure the long term impact of the United Fruit Company in Costa Rica on living standards and socioeconomic conditions of the local communities.

opment impacts (Alfaro and Charlton, 2007), evidence examining what foreign firms actually do in the field is sparse. By exploiting variation in the type of business activity performed by foreign investors, we offer a comparative overview of the differential effects of attracting different types of FDI. This type of approach has been applied in studies evaluating the impact of large mining projects by foreign firms (e.g. Tolonen, 2018), but there is little research on the effects of other activities of foreign investors, especially those related to the provision of services. Our analysis shows FDI projects that involve the establishment of local production activities and some types of value added services contribute more to structural transformation of local labour markets.

A final contribution of the paper is to add to a growing strand of research that looks at the causes of structural transformation using microdata (Lagakos and Shu, 2021; Baccini et al., 2022). Previous research has mostly focused on the role of technologies, infrastructure and trade (Bustos et al., 2016; Erten and Leight, 2019; Hjort and Poulsen, 2019) on structural transformation. We expand this literature by showing that FDI can be a driver of changes in the composition of the labour market and associated patterns of structural change. Our findings are consistent with those of Alviarez et al. (2021), who find that increasing employment in foreign affiliates pushes structural transformation in the host country.

The rest of the paper is structured as follows. Section 2 discusses the FDI and individual level data and how the spatial dimensions of the two datasets are matched. Section 3 reports the results of a descriptive exercise linking FDI to local labour market outcomes. Section 4 describes the identification strategy, based on an event study approach, and the main findings. Several extensions and robustness checks are presented in Section 5. Section 6 presents the results of an analysis of spillovers from FDI on domestic firms. Section 7 concludes.

2 Data

Individual level data. Our analysis relies on the IPUMS International Census Database, the world's largest archive of publicly available census data[5] and the Demographic and Health Survey (DHS) Program Database.[6] The DHS data are nationally representative surveys that are standardized across countries. They comprise microdata on individual and household characteristics, including employment, health and other demographic traits. Both databases are compiled through surveys, which are organized in waves. IPUMS waves are less frequent (about one per decade) but have more extensive coverage, while DHS waves are more frequent but encompass smaller sampling frames. IPUMS covers the entire population; DHS focuses on individuals (mainly women) aged 15-49 and children under five years. Another important difference relates to geographical information. The most detailed information on respondents' locations in IPUMS is generally the second-level administrative division of a country. In contrast, DHS provides information on the exact location of the individual.[7] For each country, we use the most detailed and time-consistent administrative level information available in the IPUMS International Database.[8]

We consider microdata from IPUMS International and DHS for 21 and 24 African countries, respectively, spanning 1987-2019. For consistency, these data were fil-

[5]Census data can be downloaded from the IPUMS International website.

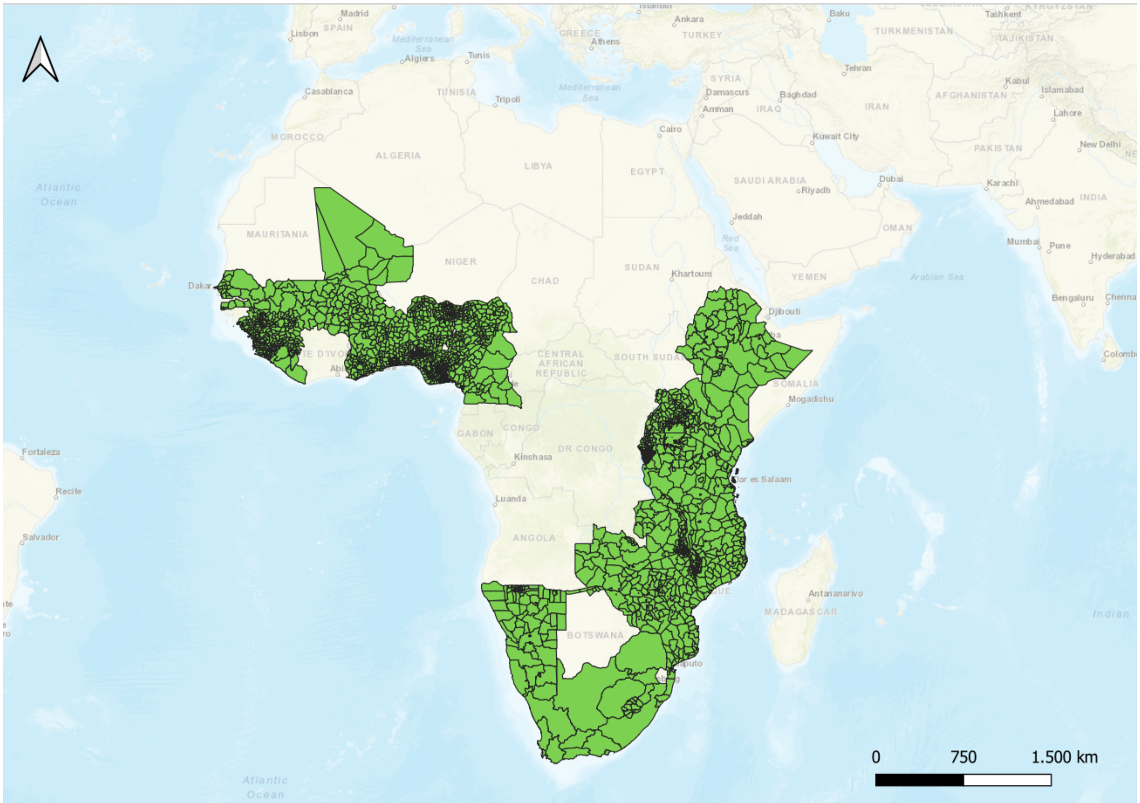
[6]Survey data can be downloaded from the DHS Program website.

[7]The DHS Program provides the coordinates for georeferencing the respondents' position. These coordinates identify geographical points that are called clusters, which are randomly displaced in order to maintain the confidentiality of the survey. Urban clusters contain an error that ranges from 0 to 2 kilometers, while for rural clusters the range widens to 5 kilometers. Most DHS surveys consider 25-30 households per cluster. For further details on the random displacement of clusters, see the Program's webpage at GPS data collection.

[8]To solve the time-consistency problems in the administrative designations we follow the procedures discussed by Alesina et al. (2021). Since the DHS clusters change over time, we combine the DHS locations with the IPUMS polygonal shapefiles. For most of the countries, it was possible to use the consistent shapefiles provided by IPUMS international. However, as IPUMS does not provide time-consistent shapefiles for Burundi, Namibia and Nigeria, for these countries we use shapefiles provided by the Humanitarian Data Exchange (HDX). More information at the organization's webpage.

tered, selecting information for individuals aged between 15 and 49 years and excluding DHS waves with female-only respondents.[9] Our sample includes data on 40,665,627 individuals from 82 different DHS waves and 49 IPUMS waves, localized in 24 countries and 2,570 subnational units (see Table A1 in the appendix). Figure 1 reports the African countries included in our sample and the associated subnational units.

Figure 1: African countries disaggregated by subnational units



Source: authors' elaboration on IPUMS shapefiles and Humanitarian Data Exchange (HDX) shapefiles.

To harmonize the information provided by the two sources of individual level data, we follow the procedure proposed by Bandiera et al. (2022) to combine IPUMS and DHS data to produce nationally representative indicators of labour

[9]Egypt, Madagascar and Morocco do not include data for men in DHS and are therefore dropped from the sample. We include this information when running some extensions of our main analysis for samples including only women respondents.

market activities and workforce characteristics.[10] Our dataset goes one step further as we aggregate at the sub-national level, creating labour market indicators at the local level. A limitation of this approach is that we can consider only those African countries where DHS data and GPS information allow us to geocode respondents.[11] An additional issue is related to the lack of consistent shapefiles for a long time span (due to changes in the definition of the administrative divisions) which prevent us from considering some additional IPUMS waves.[12]

We focus on the following outcomes: employed population; high skilled workers; blue collar and white collar occupations, activities within or outside the agricultural sector; and self-employment. Employed population indicates the individual's employment status both in IPUMS and DHS[13] High-skilled workers include legislators, senior officials, managers, professionals, technicians and associate professionals in IPUMS; professionals, technicians and managers in DHS.[14] Blue collar occupations encompass skilled agricultural and fishery workers, crafts

[10]Diao et al. (2017) use DHS data to explore labor market dynamics.

[11]We exclude countries for which DHS data are not reported (Botswana, South Sudan), countries with only one DHS wave (Angola, Chad) and countries without GPS data (Mauritius, Sudan). Although they have only one DHS available, Mozambique and South Africa are included because there are at least two IPUMS waves for these countries. Moreover, we do not consider Niger, Ivory Coast and the Democratic Republic of Congo because we do not have data on FDI projects for these countries.

[12]For example, Burkina Faso 1985, Ethiopia 1984, Ghana 1984, Guinea 1983, Malawi 1987, South Africa 1996 have not any geographical variable attached to them and therefore they were dropped from the sample.

[13]In DHS, the respondent is employed if he or she worked in the past seven days, while in IPUMS the individual's employment status indicates whether the respondent is part of the labor force in the week prior the census. In Mali 1987-2009, Rwanda 2002, Sierra Leone 2004 the reference period is the past month while in Liberia 2008, Senegal 1988-2013, Sierra Leone 2015 and Zimbabwe 2012 the reference period is the past year. In Benin 1992-2013 the reference period is three months and in Ethiopia 2007 the reference period is 2 months.

[14]IPUMS codes the individual's primary occupation according to the major categories in the International Standard Classification of Occupations (ISCO) scheme for 1988. For someone with more than a job, the primary occupation is typically the one in which the individual had spent the most time or earned the most money. The 11 categories are: (1) legislators, senior officials and managers; (2) professionals; (3) technicians and associate professionals; (4) clerks; (5) service workers and shop and market sales; (6) skilled agricultural and fishery workers; (7) crafts and related trades workers; (8) plant and machine operators and assemblers; (9) elementary occupations; (10) armed forces; (11) other occupations, unspecified or not elsewhere classified. DHS follows the ISCO scheme for 2008. The 9 categories are: (1) professionals, technicians or managers; (2) clerical; (3,7) sales and services; (4,5) agriculture; (6) domestic service; (8) skilled manual; (9) unskilled manual.

and related trades workers, plant and machine operators, assemblers and workers in elementary occupations in IPUMS, and agricultural, domestic, skilled and unskilled manual workers in DHS. White collar occupations refer to clerks, services and shops workers in IPUMS; clerks, sales and services workers in DHS. Employment in the agricultural sector include workers in agriculture, fishing and forestry, mining and extraction for IPUMS, and agriculture for DHS.[15] Activities outside the agricultural sector comprise all the residual occupations defined by ISCO 2008 in DHS and all the other sectorial categories defined by the ISIC classification in IPUMS. Finally, self-employment includes individuals employed by a family member or self-employed in DHS and individuals self-employed in IPUMS.[16]

We calculate location-wave level shares for the outcomes of interest, weighting each individual-level observation using the weights provided by IPUMS and by DHS to ensure that these indicators are a more precise approximation of the population of interest.[17]

FDI. The foreign investment data come from the Financial Times fDi Markets database. Data are gathered from various sources, including the media and national investment promotion agencies. This database includes the location of each

[15] To create this variable, we use the ISCO scheme for 2008 for DHS. IPUMS records the disaggregated sectoral classification, distinguishing between 17 levels roughly conforming to the International Standard Industrial Classification (ISIC). The categories are: (1) agriculture, fishing and forestry; (2) mining and extraction; (3) manufacturing; (4) electricity, gas, water and waste management; (5) construction; (6) wholesale and retail trade; (7) hotels and restaurants; (8) transportation, storage and communications; (9) financial services and insurance; (10) public administration and defense; (11) services, not specified; (12) business services and real estate; (13) education; (14) health and social work; (15) other services; (16) private household services; (17) other industry, not elsewhere classified.

[16] IPUMS indicates the status of an economically active person with respect to his or her employment. There are three categories: (1) self-employed; (2) wage/salaried worker; and (3) unpaid worker. DHS provides information on type of employer. The categories are: (1) employed by family member; (2) employed by non-family member; (3) self-employed.

[17] Unlike IPUMS, we need to correct for lower sampling rates of men compared to women in DHS before calculating the weighted average. More information on the weighting procedure in the DHS data can be found at the DHS forum webpage.

project, the name and country of origin of the investor.[18] It also provides detailed information on the sector (corresponding to the NAICS 2007 classification) and activity of the FDI project in the host country. This variable covers 17 categories, including business services, sales, manufacturing, extraction, construction, R&D, and retail. This feature of the data is important because it permits assessment of the potentially heterogeneous effects of FDI on structural transformation. Our sample includes 4,918 greenfield FDI projects in 24 countries, covering the period 2003-2020. During this period, South Africa, Kenya and Nigeria were the top three recipients; the US, UK and South Africa the top three investors (see Table A2). Investments in financial and business services and in software and ICT services accounted for more than 30% of the total number of projects received by African countries in our sample during the period for which the data are available. Business and sales activities and manufacturing activities attracted 69% of the projects (see Table A3).

We have added geographical coordinates to each project for which information on location (city, region, country) was available. This was possible for 72% of all projects in the sample countries.[19] Figure A1 in the Appendix reports the geographic distribution of FDI projects. Based on this information, we created a panel version of the fDi Markets database for each country, counting the number of projects in each administrative unit before each IPUMS/DHS wave. Similar information was created by grouping FDI projects by their main characteristics,

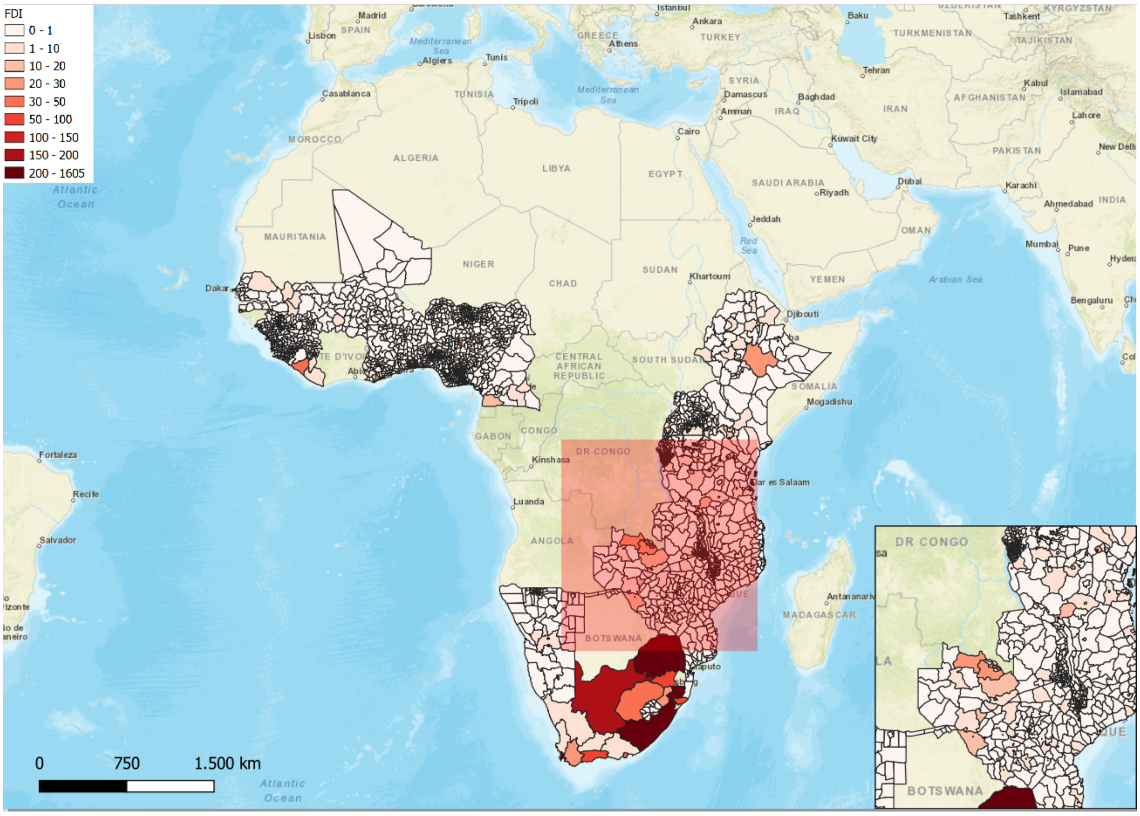
[18] fDi Markets provides as well information on capital expenditure and jobs created at the project level. However, for a large majority of the projects included in the sample (over 70%) this is an estimated data, which is generated using a proprietary econometric model of the data provider. For this reason, we prefer not to use these variables in our analysis.

[19] The total number of projects in the African countries considered in our sample was 6,787. Of these, only 4,918 projects were georeferenced (3,261 at the city-level and 1,657 at the state-level). Tables A4 and A5 in the Appendix summarize the characteristics of the projects that we were not able to geolocalize. Overall, their geographic and sectoral composition is not too far from the one of the sample, a slight exception being perhaps a higher relevance of manufacturing projects among those not geolocalized.

including their origin (i.e. OECD, non-OECD, African countries, China), sector and the type of activity performed by the foreign firm in the field.

Combining the data at geographic level. For each country, we combine the DHS and IPUMS data with the FDI data at the level of each administrative unit using the unit identifier. Figure 2 shows the African countries considered in our sample disaggregated by administrative units and the number of FDI projects at the sub-national level.[20] Figure 3 provides an example of how the dataset is structured. It reports data for Mozambique across the three available waves (1997 IPUMS, 2007 IPUMS, 2011 DHS). Each province in each wave describes the share of workers in the agricultural sector and the locations of the FDI projects (black dots) received up to the year of the wave.

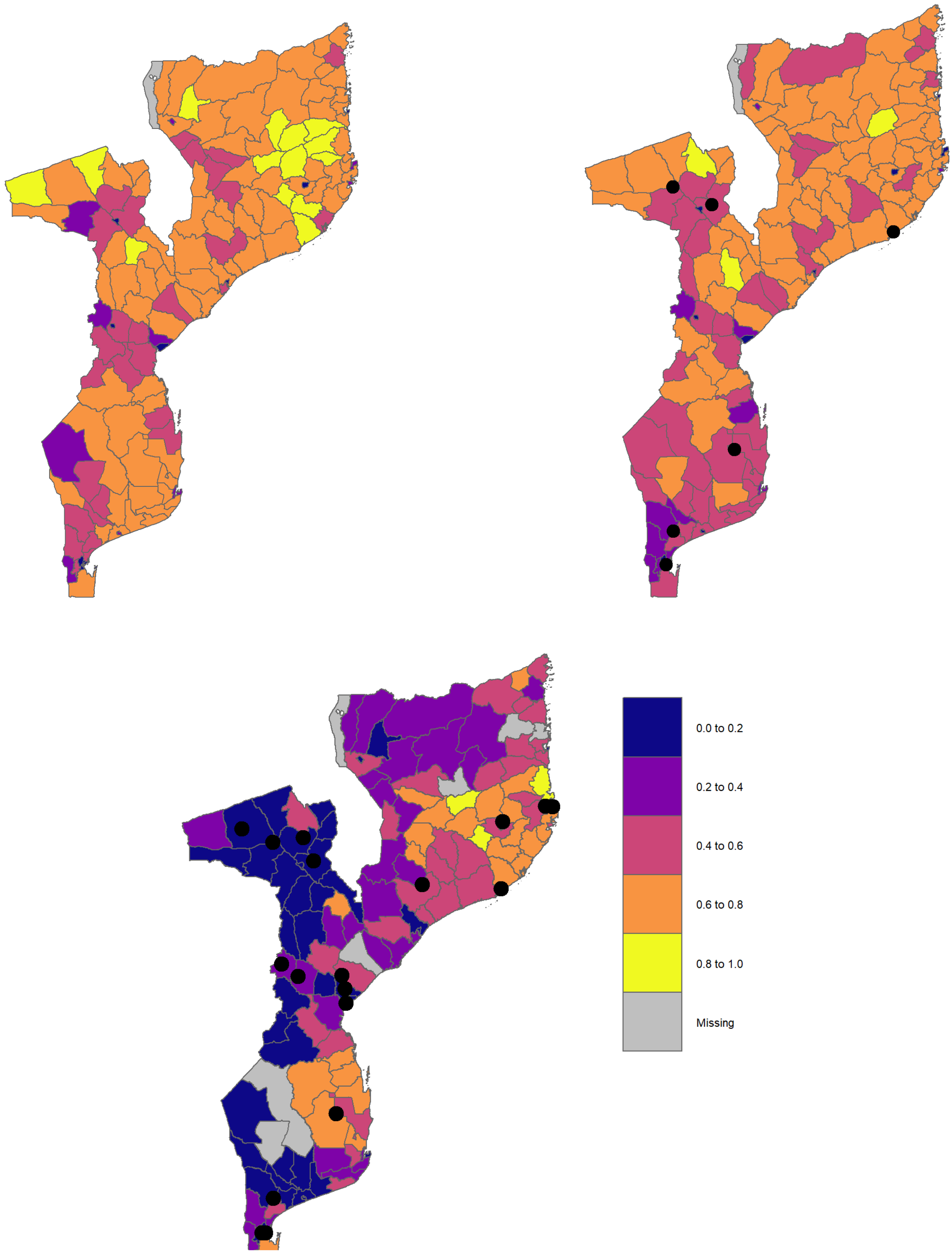
Figure 2: Subnational units and number of FDI projects



Source: authors' elaboration on fDi Markets data, on IPUMS shapefiles and Humanitarian Data Exchange (HDX) shapefiles.

[20]Table A6 reports some descriptive statistics of our sample.

Figure 3: Share of workers in the agricultural sector in Mozambique and the FDI locations



Note: authors' elaboration on the final dataset combining fDi Markets, DHS and IPUMS data. The maps report administrative units' share of agricultural occupations and projects' locations. In the first row, the map on the left reports the share in the 1997 IPUMS wave while the map on the right reports the share in the 2007 IPUMS wave and the projects' locations received by 2007. In the second row, the map reports the share in the 2011 DHS wave and location of FDI projects that occurred up to 2011.

3 Descriptive Evidence

In our analysis, we link the entry of FDI to changes in the local labour markets in recipient destinations. Throughout the analysis, we will consider the treatment as an absorbing state, i.e. once receiving its first FDI a location remains treated for the rest of the period considered.[21]

We begin by providing some descriptive evidence based on the following regression:

$$y_{ict} = \beta_0 + \beta_1 (FDI)_{ict} + \beta_2 (X)_{ict} + \gamma_i + \theta_{ct} + \epsilon_{ict} \tag{1}$$

where y_{ict} is one of the outcomes of interest, as defined in Section 2. They include: (1) the share of employed population in location i; (2) the share of population within or outside the agricultural sector; (3) the share of population in skilled or unskilled activities; (4) the share of population in self-employment. γ and θ are location and country-wave fixed effects. All the regressions are weighted using the total population of the area. We control for location specific factors (X_{ict}), including the average age of the population, the share of women and urban residents, as well as the share of individuals with education beyond secondary school. Standard errors are clustered at the level of the location.

Results of the descriptive analysis are summarized in Table 1. They provide suggestive evidence linking the presence of FDI with indicators of the relative size and composition of the respective local labour markets. We find evidence for a positive association between receiving FDI and an increase in the share of employed population in the area (column 1). This is accompanied with a positive correlation between FDI and indicators of structural transformation. The pres-

[21] This is plausible in the case of FDI projects, whose nature is more sticky compared to other type of investments. Still, fDi Markets does not record divestment, so that it is not possible to know whether a project that was registered in a given year is still active a few years later.

ence of FDI is associated with higher shares of (i) high-skilled workers and workers in white-collar occupations (columns 2 and 3); and (ii) workers employed in non-agricultural activities (manufacturing and services, column 6). We find no evidence of an association between FDI and informal work.

Table 1: Results of the TWFE model

VARIABLES	(1) Employment	(2) High Skilled	(3) White collar	(4) Blue collar	(5) Agriculture	(6) Non Agriculture	(7) Self Employment
FDI	0.0184**	0.00561**	0.00863*	0.00651	0.00434	0.0143**	0.0205
	(0.00852)	(0.00261)	(0.00503)	(0.00901)	(0.00932)	(0.00709)	(0.0132)
Constant	0.227***	-0.165***	-0.0373	0.508***	0.456***	-0.231***	0.180***
	(0.0411)	(0.0223)	(0.0368)	(0.0510)	(0.0525)	(0.0510)	(0.0567)
Observations	10,725	10,367	10,367	10,367	9,758	9,758	9,959
R-squared	0.790	0.772	0.852	0.836	0.866	0.873	0.817
Controls	Y	Y	Y	Y	Y	Y	Y
ADM FE	Y	Y	Y	Y	Y	Y	Y
Country*wave FE	Y	Y	Y	Y	Y	Y	Y
Mean DV	0.688	0.0393	0.151	0.502	0.377	0.300	0.559

The unit of observation is the province. The variable *FDI* is a dummy variable equal to 1 if the province has received at least one project by year t. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. Controls variables are the share of female, the share of people who live in the urban areas, the share of individuals with at least secondary education and average age in each province. We include location and country-wave fixed effects and the total population of the area as a weight. Standard errors clustered at the provincial level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

4 Identification strategy and results

The relationship estimated in equation (1) rests on the assumption that, absent the treatment, areas receiving FDI would have performed similarly to areas that never do so. But FDI location choices are not random. Socio-economic characteristics of local communities may influence FDI decisions. This may not be addressed by estimating equation (1) using standard TWFE methods, even in the absence of pre-trends. Specifically, when the treatment – as in our case – is heterogeneous over time across units, the coefficients estimated through TWFE may not provide the correct weighted average of treatment effects across units (and time).[22] This is because with heterogeneous treatment one ends up comparing treated units both with never treated (and not yet treated) units, which is correct, as well as with already-treated units, which is not. The introduction of the latter type of comparison results in negative weights for some of the estimated coefficients, which can bias both the size and sign of the coefficient of interest. Several alternative methods have been proposed in order to deal with these issues, provided that conditions related to the presence of parallel trends and absence of anticipation hold.

We adopt the approach proposed by Callaway and Sant'Anna (2021), which provides estimators for staggered difference-in-differences that are robust to the presence of heterogeneous treatment. This estimates group- and time-specific average treatment effects on the treated using two-period/two-group difference-in-differences estimators and then aggregates them to produce summary treatment effect estimates, weighting by the size of each treatment group. More specifically, we use the doubly robust difference-in-differences estimator proposed by Sant'Anna and Zhao (2020) as generalized by Callaway and Sant'Anna (2021) to

[22]Recent reviews by Roth et al. (2022); De Chaisemartin and D'Haultfoeuille (2022)) provide an overview of the issues and proposed solutions.

a setting with multiple periods and multiple groups. This method combines the outcome regression approach and the inverse probability weighting using pre-treatment characteristics to either condition parallel trends or weighting the control group on the probability of being treated. This is an appropriate approach in our setting given that it is unlikely that parallel trends hold unconditionally. If we compare the characteristics of locations one period before being treated with control locations, there are notable differences in terms of some outcomes and characteristics, including the share of urban population and the degree of secondary education (see Table 2).

Table 2: Summary Statistics: Treated and Control areas

Variable	Treated (at t-1)	Controls
Employed	62.45%	69.06%
High skilled	4.87%	3.77%
White collar	14.82%	14.77%
Blue collar	41.78%	50.86%
Agriculture	28.79%	38.69%
Non Agriculture	32.18%	29.17%
Self Employment	41.99%	56.68%
Age	28.26	28.54
Female	52.72%	54.52%
Urban	43.8%	23.27%
Secondary	30.63%	25.02%

Note: The table reports information on the characteristics of locations included in our sample. Information on the locations that we identify as treated in our analysis refer to one period before treatment.

To address these concerns we include initial characteristics in terms of average age, gender composition, urban population and the share of the population with at least secondary education. We use both the never treated units and those not-yet treated as control group.

We first report some diagnostics on the relevance of negative weights in our sample. This is done following the approach proposed by De Chaisemartin and d'Haultfoeuille (2020), which identifies the proportion of ATTs that receive a negative weight due to the staggered treatment and therefore could bias the sign of the TWFE estimates. Table A7 in the Appendix reports the weights attached to each coefficient of the outcomes of interest in regression 1. Considering the outcomes of interest, slightly more than 90 percent of the weights are strictly positive, while the sum of the negative weights is equal to -0.01. Column 4 reports the ratio between the absolute expected value of the coefficients and the standard deviation of the weights. If this ratio is close to zero, treatment effect heterogeneity would be a serious concern for the validity of $\hat{\beta}_{twfe}$. In our case, it ranges between 0.0038 and 0.0266, which indicates a plausible amount of heterogeneity for all the outcomes of interest.

4.1 Results

In this Section, we present results based on the event study estimates. Note that the structure of our data poses some limitations to the number of leads and lags that we can effectively include. The median (mean) area is observed 3 times (3.2), with similar values if we focus on the treated areas only. Based on this, we report estimates that use 3 periods before and 2 periods after the treatment. When interpreting the results, it should be noted that this is indicative of a relatively long time span as the median (mean) distance across the waves considered in the analysis is 5 (4.5) years.

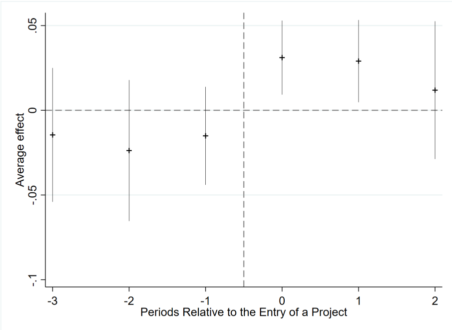
Figure 4 summarizes the findings of the event study on the outcomes of interest, reporting the estimated coefficients together with their 95% confidence interval.[23] The main message emerging from Figure 4 is that for most of the outcomes there

[23] Table A8 provides additional details by reporting the event study estimates of the average treatment effect on the treated of the main specification.

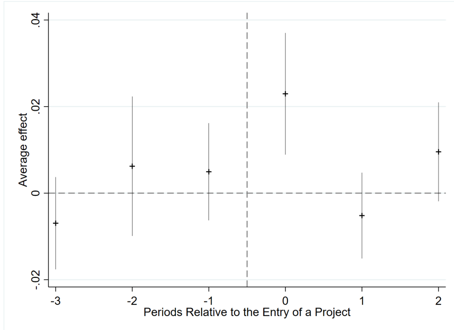
is no evidence of violation of parallel trends. This is reassuring in view of the consistency with results of the descriptive exercise in table 1, permitting a more straightforward interpretation of the results. In addition, the event study estimates help to assess whether the effects of FDI are likely to persist over time.

Results of the event study show that employment increases following entry of a project in a given geographic location for about two periods. Thereafter it becomes insignificant. The most striking result is the evidence of structural transformation away from agriculture, with non-agricultural shares of employment rising consistently over time. Changes in the composition of occupations by skill level are less clear. There is a jump in the share of skilled employment in the year of the treatment, which disappears subsequently. No evidence emerges for other occupations, with results for white collar occupations clearly affected by pre-trends. Similarly, there is no evidence that FDI leads to changes in the share of self-employment.

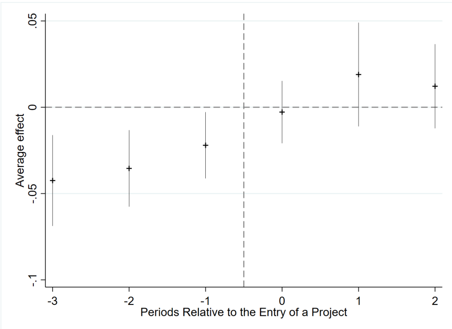
Figure 4: Relationship between FDI projects and the outcomes of interest



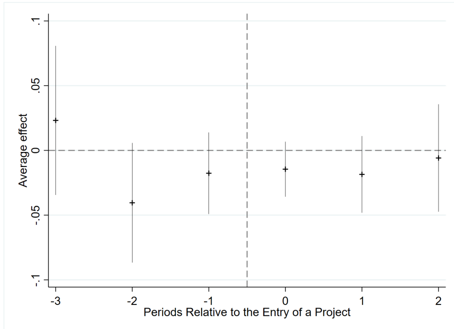
(a) Employment



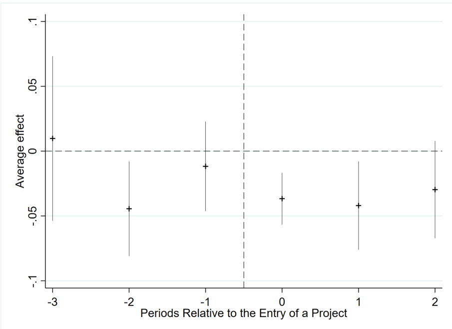
(b) High Skilled



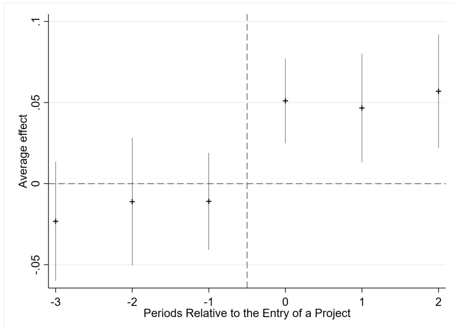
(c) White collar



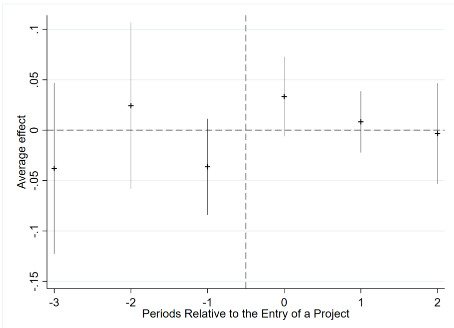
(d) Blue Collar



(e) Agriculture



(f) Non Agriculture



(g) Self-Employment

Note: this figure plots the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is a location. As controls we include the share of females, the share of people who lives in urban areas, the share of individuals with at least secondary education and average age in each location, while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after an FDI project entry.

5 Heterogeneity and Robustness

5.1 Heterogeneity

Activity and origin of FDI. To account for the potential heterogeneity of FDI, we differentiate projects both according to the country of origin of the investment and what foreign firms actually do in the field. Regarding the former, we distinguish between FDI coming from OECD member countries and those originating in the South (all the other countries).[24] There is substantial evidence for developing countries that the origin of FDI may be associated with differences in potential effects, e.g., FDI from large emerging economies such as China can have a positive impact because of a smaller technological distance between source and host country (e.g. Amighini and Sanfilippo, 2014; Gold et al., 2017).

As far as the activity of foreign investors is concerned, we distinguish projects whose core activities involve physical production in the location of the investment from projects that mostly generate value added through intangible activities (e.g. financial or business-related services).[25] We also consider projects in the extractive industry. The type of activity matters since it helps qualify the nature of the local spillovers of FDI. While some activities (e.g. production and, to a lesser extent, extraction of natural resources) are labour intensive and more likely to stimulate jobs and local demand (Toews and Vezina, 2022), others (e.g. business or financial services) can foster private sector development by easing existing frictions (e.g. access to credit). In doing this we define the event as the first time a given location receives a project in a given sector, activity or from a

[24] In our sample the OECD countries are Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, South Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

[25] More precisely, this includes projects in financial, business, and ICT related services.

specific source.

Results of this exercise are summarized in Appendix Figures A2, which plot ATTs for FDI across by sector and Figure A3, which does so by origin of FDI.[26] The results suggest that the heterogeneous characteristics of foreign investors can affect the relationship between FDI and structural transformation. FDI projects involving production (manufacturing) facilities are associated with structural change, a finding that is consistent with the body of evidence on the direct and indirect benefits of inward FDI in industry in developing countries (e.g. Javorcik, 2018; Alfaro, 2017; Javorcik, 2015). Locations that receive FDI projects in extractive activities show an increase in the share of the population that is employed that is less likely to persist over time, a reduction in the share of skilled workers, and a lack of structural transformation towards non-agricultural employment. There is also some evidence of an increase in the share of self-employment. The results do not reveal large differences for FDI originating in OECD countries relative to non-OECD countries (Figure A3).

Female employment. Appendix Figure A4 reports event study results for a subsample in which information is aggregated using only data on female respondents for the same countries and waves used in the foregoing analysis. As mentioned in Section 2, Egypt, Madagascar and Morocco include only DHS data for women and were not included in our main sample. These data can be used for an analysis focusing only on women. Overall, results based on the two samples are similar and in line with those obtained in figure 4 but we observe some differences. The treated areas show an increase in employed women, characterized by shifts in the female workforce to high-skilled jobs and jobs outside agriculture, but we find no strong evidence of women moving out of agriculture in response

[26] Tables A9 and A10 in the Appendix provide additional details by reporting the event study coefficients of the average treatment effect on the treated according to the type of activity performed and the country of origin.

to the entry of FDI. We find a small, contemporaneous to the treatment year, increase in the share of self-employment for women living in locations that received FDI.[27]

5.2 Robustness checks

Alternative estimators. We compare results obtained with the estimator by Callaway and Sant'Anna (2021) with results obtained using two alternative estimators proposed by De Chaisemartin and d'Haultfoeuille (2020) and Borusyak et al. (2021) to address similar settings like ours.[28] Appendix Figure A5 reports the results of the three estimators combined. Overall, both the direction and the size of the estimated coefficients are very similar.

Dropping the capital city. A potential source of concern for the analysis is that FDI projects are more likely to occur in areas within a country that are more developed, which may drive some of our findings.[29] To address this potential con-

[27]Table A11 in the Appendix reports the event study estimates of the average treatment effect on the treated considering only female respondents in the same countries and waves used in the main analysis and in the sample that includes also Egypt, Madagascar and Morocco.

[28]De Chaisemartin and d'Haultfoeuille (2020) and Borusyak et al. (2021) propose two different estimators for staggered difference-in-differences design. On the one hand, De Chaisemartin and d'Haultfoeuille (2020) propose an estimator that rules out dynamic effects and therefore the treatment effect is instantaneous. They use a weighted average across t of two type of the difference-in-differences: the first difference-in-difference compares the outcomes evolution of untreated groups in the two periods and groups becoming treated while the second difference-in-difference compares the outcome evolution of groups treated in the two periods. Moreover, it considers the unconditional difference-in-difference designs. On the other hand, Borusyak et al. (2021) propose an imputation estimator which allows for dynamic effects, meaning that it assumes that group's current outcome does not depend only on its current treatment and the outcome at time t is allowed to depend on its past treatments. It estimates the counterfactual using a TWFE model that is fit using only pre-treatment data. If the outcomes are not too serially correlated, it can be more efficient but it also relies on a stronger parallel trends assumption that may be more susceptible to bias.

[29]It can also be argued that information gathered from fDi Markets is likely to be biased towards certain locations within a country insofar as the coverage of media outlets and official reporting on FDI projects, the main sources of information that feed fDi Markets, are more likely to cover the areas where most economic activity occurs.

cern, we replicate the analysis *excluding* areas hosting the capital city.[30] Results of this exercise, summarized in Appendix Figure A6 shows that results mostly do not change. An important exception concerns employment, for which the coefficient remains positive but is no longer statistically significant in post-treatment periods.[31]

6 Mechanisms: Firm level responses to FDI

The main result of the analysis is to find evidence that FDI is a driver of structural transformation within local labor markets in the sample of African countries. This leaves open the mechanisms through which FDI projects drive such structural transformation. In this Section we analyze one potential demand side channel: spillover effects from FDI on proximate domestic firms that may impact local labour demand.

To do this we exploit geolocalized information on each FDI project and match it to firm-level data from the World Bank Enterprise Surveys (WBES) for all African countries for which this is possible.[32] The WBES provide nationally representative[33] firm-level information for many countries around the world, including most African countries. For the analysis, we use a harmonized version of the

[30]For Ethiopia, Ghana, Kenya, Liberia, Mozambique, Namibia, Rwanda, Sierra Leone and South Africa, the capital city includes more than one sub-national unit. For these countries we therefore drop more than one area.

[31]Table A12 shows the event study coefficients of the average treatment effect on the treated for the sample without the provinces with the capital city.

[32]This was done by computing the the geographical distance between each FDI project and each WBES firm using the R function geosphere::distm and appending each firm's ID to the distance matrix. editing the format of the distance matrix to a long version and merging the distance matrix with the FDI data. This algorithm was applied to each country/wave sub-sample. Each country/wave data set includes firm-level information (firm ID, ISIC code, geographical coordinates) and project-level information (project ID, distance from the firm, company data, ISIC codes, geographical coordinates).

[33]WBES are based on stratified random samples of companies extracted from public registries. Stratification is by size, location, and sector.

data set that provides standardized variables for all the surveys run from 2006 to 2020. Our sample includes 26,351 firms from all the 24 countries covered in our main analysis (Appendix Table A13 provides summary information on the sample size for each country and wave). Around half of the firms are in the manufacturing sector, with services covering the other half. The outcome of this matching exercise for the countries covered in the analysis is reported in Figure A7 in the Appendix.

In our empirical specification, we link exposure of domestic companies to FDI projects to firm-level indicators measuring various dimensions of economic performance and activities that may be affected by firm exposure to FDI. We use the following indicators: labour productivity (total sales per employee); total sales; number of employees; share of skilled workers in total; the ratio of total cost of labor on the number of employees; and investment in fixed assets (a dummy taking one if a firm has invested in fixed assets over the past 3 years).

To identify the implications of FDI exposure for domestic firms we employ a method that exploits spatial and temporal variation in the entry of new FDI projects.[34] We compare areas in which a FDI project has already occurred and those where a project has not yet been implemented at the time of the survey, but that will take place in a subsequent period.[35] To implement this approach, we first define a buffer around the centroid of each of the places in which a firm included in the WBES is located and then group the firms as follows:

1. those within a certain cut-off distance from an FDI project that was received before the year in which the survey occurred (which we label as *Active*);

2. those within a certain cut-off distance from an FDI project that has not yet

[34] This method has been applied to evaluate the impact of development projects, as well as the effects of FDI (e.g. Brazys and Kotsadam, 2020)

[35] As fDi Markets span a longer time period than most of the firm surveys we can determine instances where FDI will occur in time periods not covered in the WBES.

occurred but will take place in a period following the survey year (*Inactive*);[36] and

3. those outside the cut-off distance from either an active or an inactive project (the control group).

We then estimate the following regression:

$$y_{it} = \beta_1(Active)_{jrt} + \beta_2(Inactive)_{jrt} + \beta_3(X)_{it} + \theta_r + \phi_j + \omega_{ct} + \epsilon_{ijrt} \tag{2}$$

Where Y is an outcome of interest for a firm i in industry j, location (city or region) r, and time t, and X is a vector of firm characteristics (including their age and size group).[37] Location θ_r, industry ϕ_j and country-year ω_{ct} fixed effects account for common spatial and temporal trends, spatial clustering across firms, and country-specific time-contingent factors, such as regulations, that may influence the relationship examined. Standard errors are clustered at the region-industry level.

We use a buffer extending 50km around each firm's location in our main specifications This distance has been adopted in studies using similar methods in the literature (Tolonen, 2018; Brazys and Kotsadam, 2020). As a robustness check, we also consider the sensitivity of our results to different buffer sizes.[38] Figure 5 illustrates the identification strategy, using the buffer around Umuahia, in Nigeria, as an example. The red dots in the middle of the circle represents a domestic firm. The blue dots are FDI projects that are located in the neighbourhood of the firm. The blue dots inside the circle are considered in the definition of the treatment

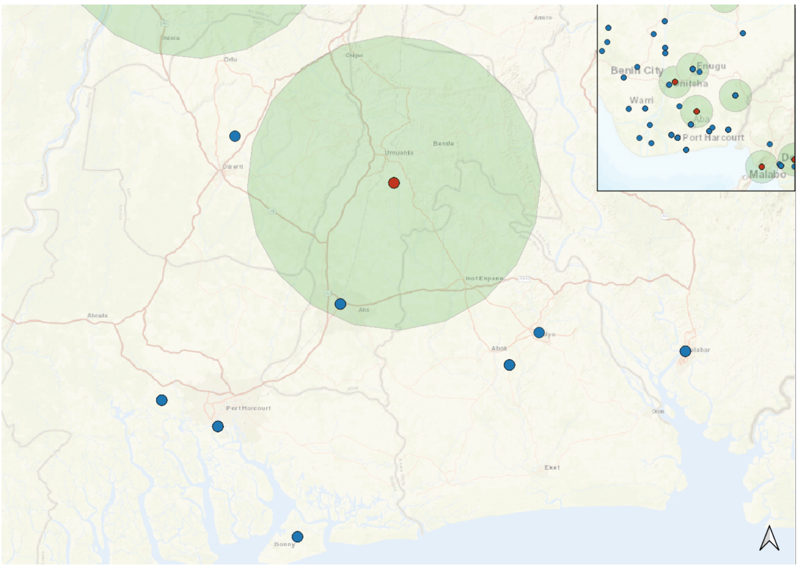
[36] When creating this group, locations in which there are already active projects are excluded.

[37] This variable is provided by the WBES and takes the value of 1 if a firm is small (less than 20 employees), 2 if it is medium-sized (20-99); and 3 if large (100 and over).

[38] Results of robustness checks using the different buffers are reported in Table A15 in the Appendix.

(either as active or inactive, depending on when the projects are undertaken relative to a given WBES wave). The ones outside the circle will be included in the control group.

Figure 5: Example of the buffer around Umuahia in Nigeria



Note: authors' elaboration on WBES and fDi Markets data.

This identification strategy relies on the estimation of two differences. The first difference β_1 captures the impact on a given outcome Y of FDI *inclusive* of any selection effect; the second difference β_2 is meant to capture only the selection effect. The variable of interest for the analysis is the difference between these two coefficients (i.e., $\beta_1 - \beta_2$). Including the "inactive" coefficient allows us to compare the outcome for firms in the proximity of current FDI projects with those of firms that will receive a project in the future. This provides us with an estimated coefficient (the difference) that accounts for unobservable time-invariant characteristics that may affect both firm outcomes and FDI location choice.

We employ a binary definition of treatment. Whether a location is active or inactive depends on whether it hosts (or will host) at least one FDI project over the time span considered. Given the granularity of the data, this does not represent a major issue since the median (average) number of FDI projects around each individual firm in our sample, conditional on being treated, is exactly 1 (2.6).

We measure the spillover effect of FDI by considering as treated those firms that (a) operate in the same 4-digit (ISIC Rev. 3) industry, and that (b) have FDI projects located within the buffer considered in our analysis. We expect this measure to capture competition effects as well as knowledge and technological spillovers due to the fact that foreign firms operate in the same narrowly defined industry, and thus may share similar production techniques (Fons-Rosen et al., 2017), and are in relatively close proximity, allowing more frequent exchanges of ideas and workers (Kee, 2015; Newman et al., 2020).

Results are summarized in Table 3. Recall that our coefficient of interest is the difference $\beta_1 - \beta_2$. Proximity to FDI is associated with several dimensions of domestic firm performance. First, domestic firms expand their economic activity following the entry of proximate FDI projects. This is reflected in the positive and significant impact of FDI on labour productivity and sales, as well as in a slightly significant positive impact on the probability to invest in fixed assets. Similarly, we find that firms increase their employment. We do not find, however, evidence on shifts towards more skilled workers (although there are a many missing values) or changes in the cost of labour. Results are mainly driven by the sub-sample of firms (and hence FDI projects) in the manufacturing sector.

In a related exercise, we repeat the analysis considering firms as treated if exposed to FDI in sectors to which they are related through input-output linkages (AppendixTable A14).[39] Although our data do not allow us to determine whether

[39]To account for vertical spillovers, we rely on I/O coefficients, which are available from Eora for most countries in our sample. We construct weights using the national I/O tables for 2010. For each country, we consider the (26x26) matrix of sectors included in Eora. After extracting this matrix, we calculate the gross value of domestic output for each of the 26 sectors, and for each sector, the share of the sector's output used by other sectors (backward linkage). These coefficients are used to calculate measures of exposure to FDI weighted by their cross-sectoral dependence. To do this, we construct a concordance table that links the 26 Eora sectors to the sectors defined by WBES and fDi Markets, using the 2-digit ISIC classification. Considering the number of foreign projects, we define backward linkages as the weighted sum of the number of foreign projects in

a domestic firm sells directly to a foreign firm locally (in contrast to, e.g., Alfaro-Urena et al., 2022), this is the best proxy for the foreign demand effect that we can use. We again find evidence of increases in productivity and sales for firms exposed to the demand of FDI firms. On the other hand, we also find evidence of a decrease in employment. This could be a consequence of firms' upgrading, given that in this case we find a more marked evidence of increased investments in capital as well as a rise in labour costs (a proxy for wages) following exposure to new potential demand from foreign buyers.

each domestic firm's geographic buffer, the weights being the share of output sold by the sector of firm i and the sector of the FDI project. These measures provide a proxy for the probability that domestic firms enter the supply chain of foreign investors and participate in GVC-related activities.

Table 3: Mechanisms: the effect of FDI on firms outcomes

VARIABLES	(1) Productivity	(2) Total Sales	(3) Investment	(4) Number of employees	(5) Skilled employees	(6) Labor cost on employment
Active (50 km)	0.190**	0.272***	0.0209	0.0665**	0.0361	0.0337
	(0.0749)	(0.0800)	(0.0215)	(0.0283)	(0.0374)	(0.0795)
Inactive (50 km)	-0.00273	-0.0143	-0.0229	-0.00589	-0.00769	-0.00705
	(0.0628)	(0.0698)	(0.0166)	(0.0224)	(0.0125)	(0.0597)
Constant	13.28***	13.84***	0.226***	0.525***	0.620***	11.43***
	(0.0407)	(0.0440)	(0.00932)	(0.0150)	(0.0101)	(0.0380)
Observations	18,981	19,117	21,008	20,976	7,631	18,211
R-squared	0.677	0.694	0.126	0.804	0.219	0.692
Difference	0.192	0.287	0.0438	0.0724	0.0438	0.0407
p-value difference	0.0335	0.00296	0.0970	0.0395	0.253	0.664

Note: The analysis is based on the estimation of equation 2. The unit of observation is a domestic firm. Firms who report foreign ownership in the WBES sample are dropped. The table reports at the bottom the coefficient of interest, the difference between the coefficients Active and Inactive, and its p-value. In this Table, the treatment is defined as the proximity (within a 50km buffer) to at least an FDI in the same sector. All regressions include a dummy for firm size (small, medium, large), the age of the firm, city, industry (2-digit ISIC Rev 3.1) and country-year fixed effects. Standard errors clustered at the city-industry level in parentheses*** $p<0.01$, ** $p<0.05$, * $p<0.1$

7 Conclusions

Using a novel database that combines population census data from IPUMS International and Demographic Health Surves with foreign investment data from fDi Markets, we link the entry of FDI projects to a set of indicators related to the structural transformation of local labor markets in Africa. Descriptive analysis reveals that the entry of a FDI project is correlated with an increase in employment and a shift of the workforce towards non-agricultural sectors and more skilled occupations. Most these findings are driven by FDI in the manufacturing sectors.

To go beyond correlations, we apply an event-study analysis following the approach proposed by Callaway and Sant'Anna (2021) for staggered difference-in-difference settings to rule out any pre-trends. The results confirm the descriptive analysis. We find that treated areas show evidence of structural transformation, with an increase in employment characterized by shifts of workers to higher-skilled jobs and occupations outside the agricultural sector. Our analysis of a potential demand side mechanism through which FDI inflows may impact on local labor markets – effects of FDI on neighboring domestic firms – reveals that FDI is associated with changes in the performance of domestic firms and their demand for workers.

References

ALESINA, A., S. HOHMANN, S. MICHALOPOULOS, AND E. PAPAIOANNOU (2021): "Intergenerational mobility in Africa," *Econometrica*, 89, 1–35.

ALFARO, L. (2017): "Gains from Foreign Direct Investment: Macro and Micro Approaches," *The World Bank Economic Review*, 30, 2–15.

ALFARO, L. AND A. CHARLTON (2007): "Growth and the Quality of Foreign Direct Investment: Is All FDI Equal?" CEP Discussion Papers dp0830, Centre for Economic Performance, LSE.

ALFARO-URENA, A., I. MANELICI, AND J. P. VASQUEZ (2022): "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," *The Quarterly Journal of Economics*.

ALVIAREZ, V., C. CHEN, N. PANDALAI-NAYAR, L. VARELA, K.-M. YI, AND H. ZHANG (2021): "Multinationals and Structural Transformation," RIETI Discussion Paper Series 21-E-100.

AMIGHINI, A. AND M. SANFILIPPO (2014): "Impact of South–South FDI and Trade on the Export Upgrading of African Economies," *World Development*, 64, 1–17.

BACCINI, L., M. FIORINI, B. HOEKMAN, AND M. SANFILIPPO (2022): "Services, Jobs, and Economic Development in Africa," *The World Bank Research Observer*, lkac006.

BANDIERA, O., A. ELSAYED, A. SMURRA, AND C. ZIPFEL (2022): "Young Adults and Labor Markets in Africa," *Journal of Economic Perspectives*, 36, 81–100.

BAU, N. AND A. MATRAY (2022): "Misallocation and Capital Market Integration: Evidence from India," *Econometrica*.

BORUSYAK, K. (2022): "DID_IMPUTATION: Stata module to perform treatment effect estimation and pre-trend testing in event studies," .

BORUSYAK, K., X. JARAVEL, AND J. SPIESS (2021): "Revisiting event study designs: Robust and efficient estimation," *arXiv preprint arXiv:2108.12419*.

BRAZYS, S. AND A. KOTSADAM (2020): "Sunshine or Curse? Foreign Direct Investment, the OECD Anti-Bribery Convention, and Individual Corruption Experiences in Africa," *International Studies Quarterly*, 1–12.

BUSTOS, P., B. CAPRETTINI, AND J. PONTICELLI (2016): "Agricultural Productivity and Structural Transformation. Evidence from Brazil," *American Economic Review*, 106, 1320–1365.

CALLAWAY, B. AND P. H. C. SANT'ANNA (2021): "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 225, 200–230.

DE CHAISEMARTIN, C. AND X. D'HAULTFOEUILLE (2020): "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 110, 2964–96.

DE CHAISEMARTIN, C. AND X. D'HAULTFOEUILLE (2022): "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey," *The Econometrics Journal*.

DIAO, X., K. HARTTGEN, AND M. MCMILLAN (2017): "The Changing Structure of Africa's Economies," *The World Bank Economic Review*, 31, 412–433.

DONOVAN, K., W. J. LU, AND T. SCHOELLMAN (2022): "Labor Market Dynamics and Development," Tech. rep.

ERTEN, B. AND J. LEIGHT (2019): "Exporting out of Agriculture: The Impact of WTO Accession on Structural Transformation in China," *The Review of Economics and Statistics*, forthcoming.

Fons-Rosen, C., S. Kalemli-Ozcan, B. E. Sorensen, C. Villegas-Sanchez, and V. Volosovych (2017): "Foreign Investment and Domestic Productivity: Identifying Knowledge Spillovers and Competition Effects," Working Paper 23643, National Bureau of Economic Research.

Fu, X., D. Emes, and J. Hou (2022): "Multinational enterprises and structural transformation in emerging and developing countries: A survey of the literature," *International Business Review*.

Gold, R., H. Gorg, A. Hanley, and A. Seric (2017): "South-South FDI: is it really different?" *Review of World Economics*, 153, 657–673.

Hjort, J. and J. Poulsen (2019): "The Arrival of Fast Internet and Employment in Africa," *American Economic Review*, 109, 1032–1079.

Javorcik, B. (2015): "Does FDI Bring Good Jobs to Host Countries?" *The World Bank Research Observer*, 30, 74–94.

——— (2018): "Eight Things Development Professionals Should Know about Foreign Direct Investment," PEDL Synthesis Paper N.3.

Kee, H. L. (2015): "Local intermediate inputs and the shared supplier spillovers of foreign direct investment," *Journal of Development Economics*, 112, 56–71.

Lagakos, D. and M. Shu (2021): "The Role of Micro Data in Understanding Structural Transformation," STEG Pathfinding Paper.

Lay, J. and T. Tafese (2020): "Promoting private investment to create jobs: A review of the evidence," PEGNET Policy Studies 02/2020.

Liu, Y. (2022): "Does Foreign Direct Investment Catalyze Local Structural Transformation and Human Capital Accumulation? Evidence from China," World Bank Policy Research Working Paper N. 9952.

McCaig, B., N. Pavcnik, and W. F. Wong (2022): "FDI Inflows and Domestic Firms: Adjustments to New Export Opportunities," Working Paper 30729, National Bureau of Economic Research.

Mendez, E. and D. Van Patten (2022): "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company," *Econometrica*.

Mendola, M., G. Prarolo, and T. Sonno (2022): "Curse or Blessing? Multinational Corporations and Labor Supply in Africa," CEPR Discussion Paper N. 16964.

Muhlen, H. and O. Escobar (2019): "The role of FDI in structural change: Evidence from Mexico," *The World Economy*.

Newman, C., J. Page, J. Rand, A. Shimeles, M. Söderbom, and F. Tarp (2020): "Linked-in by FDI: The Role of Firm-Level Relationships for Knowledge Transfers in Africa and Asia," *Journal of Development Studies*, 56, 451–468.

Poole, J. P. (2013): "Knowledge Transfers from Multinational to Domestic Firms: Eivdence from Worker Mobility," *Review of Economics and Statistics*, 95, 393–406.

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2022): "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *arXiv preprint arXiv:2201.01194*.

Sant'Anna, P. H. and J. Zhao (2020): "Doubly robust difference-in-differences estimators," *Journal of Econometrics*, 219, 101–122.

Setzler, B. and F. Tintelnot (2021): "The Effects of Foreign Multinationals on Workers and Firms in the United States," *The Quarterly Journal of Economics*, 136, 1943–1991.

Sonno, T. (2020): "Globalization and Conflicts: the Good, the Bad, and the Ugly of Corporations in Africa," mimeo.

TOEWS, G. AND P.-L. VEZINA (2022): "Resource Discoveries, FDI bonanzas, and local Multipliers: Evidence from Mozambique," *The Review of Economics and Statistics*, 104, 1046–1058.

TOLONEN, A. (2018): "Local Industrial Shocks and Infant Mortality," *The Economic Journal*, 129, 1561–1592.

Appendix

Table A1: Coverage of IPUMS data and DHS data by countries and years

Country	DHS Waves	IPUMS Waves
Benin	1996, 2001, 2012, 2017	1992, 2002, 2013
Burkina Faso	1993, 1999, 2003, 2010	1996, 2006
Burundi	2010, 2016	NA
Cameroon	1991, 2004, 2011, 2018	1987, 2005
Ethiopia	2000, 2005, 2010, 2016	1994, 2007
Ghana	1993, 1998, 2003, 2008, 2014	2000, 2010
Guinea	1999, 2005, 2012, 2018	1996, 2014
Kenya	2003, 2008 2014	1989, 1999, 2009
Lesotho	2004, 2009, 2014	1996, 2006
Liberia	2007, 2013, 2019	2008
Malawi	2000, 2004, 2010, 2015	1998, 2008
Mali	1996, 2001, 2006, 2012, 2018	1987, 1998, 2009
Mozambique	2011	1997, 2007
Namibia	2000, 2006, 2013	NA
Nigeria	2003, 2008, 2013, 2018	NA
Rwanda	2005, 2010, 2014, 2019	1991, 2002, 2012
Senegal	1993, 1997, 2005, 2010, 2019	1988, 2002, 2013
Sierra Leone	2008, 2013, 2019	2004, 2015
South Africa	2017	2001, 2007, 2011, 2016
Tanzania	1999, 2010, 2015	1988, 2002, 2012
Togo	1998, 2013	2010
Uganda	2000, 2006, 2011, 2016	1991, 2002, 2014
Zambia	2007, 2013, 2018	1990, 2000, 2010
Zimbabwe	1999, 2005, 2010, 2015	2012

Note: For DHS, we adopt the year of data collection from the survey documentation given that some datapoints might have been collected in the previous or in the following year. This is not the case for IPUMS which collects data during a single year.

Table A2: Countries per number of investments received and made

Investors	Freq	Recipients	Freq
United States	14.46%	South Africa	34.30%
United Kingdom	13.89%	Kenya	12.67%
South Africa	6.28%	Nigeria	11.64%
Germany	5.98%	Ghana	7.63%
France	4.98%	Mozambique	5.06%
China	4.55%	Tanzania	4.21%
India	3.84%	Ethiopia	3.50%
Switzerland	3.58%	Uganda	3.29%
Japan	2.83%	Zambia	3.27%
UAE	2.83%	Rwanda	2.40%

Note: authors' elaboration on fDi Markets data. The sample includes 4918 greenfield FDI projects in 24 countries from 2003 to 2020.

Table A3: Sectors and activities per number of projects

Sectors	Freq.	Activities	Freq.
Financial services	17.14%	Business Services	26.01%
Business services	10.94%	Sales, Marketing & Support	24.26%
Software & IT services	8.70%	Manufacturing	19.38%
Communications	7.77%	Logistics, Distribution & Transportation	4.53%
Food & Beverages	6.73%	Electricity	3.97%
Transportation & Warehousing	5.31%	Extraction	3.13%
Metals	4.94%	Construction	3.05%
Industrial equipment	4.45%	Headquarters	2.91%
Renewable energy	3.92%	Research & Development	2.70%
Coal, oil & gas	3.54%	ICT & Internet Infrastructure	2.50%
Real estate	2.95%	Retail	2.09%
Chemicals	2.89%	Education & Training	2.03%
Automotive OEM	2.52%	Maintenance & Servicing	1.79%
Hotels & tourism	2.05%	Customer Contact Centre	0.98%
Building materials	1.93%	Technical Support Centre	0.35%
Electronic components	1.77%	Recycling	0.24%
Textiles	1.53%	Shared Services Centre	0.08%

Note: authors' elaboration on fDi Markets data. The sample includes 4918 greenfiled FDI projects in 24 countries from 2003 to 2020.

Table A4: Countries per number of non-geolocalized investments received and made

Investors	Freq	Recipients	Freq
United States	11.08%	South Africa	20.01%
United Kingdom	9.52%	Nigeria	11.66%
India	9.31%	Kenya	10.38%
China	7.98%	Ghana	9.63%
South Africa	7.28%	Tanzania	6.58%
UAE	3.64%	Ethiopia	5.78%
Kenya	3.53%	Uganda	4.92%
France	3.42%	Zambia	4.49%
Canada	3.26%	Mozambique	4.01%
Japan	3.21%	Senegal	2.73%

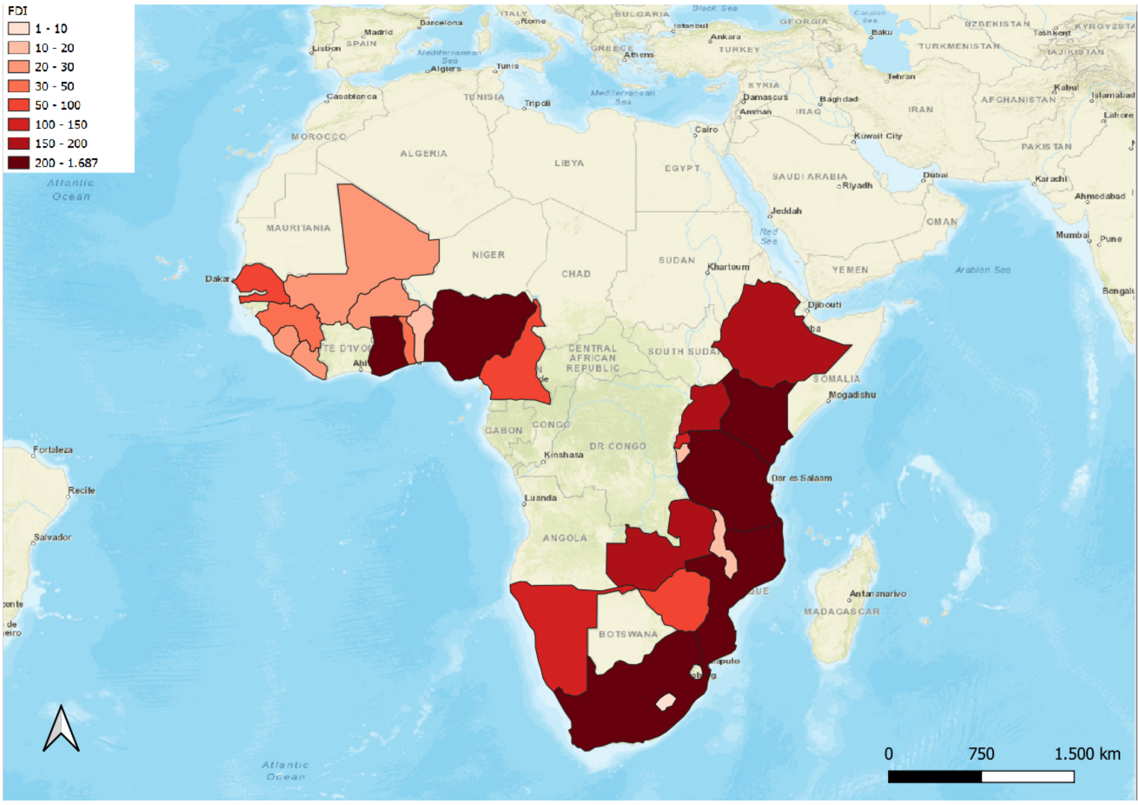
Note: authors' elaboration on fDi Markets data. The sample includes 1869 greenfield FDI projects in 24 countries from 2003 to 2020. These projects were not georeferenced and therefore they were excluded from our sample.

Table A5: Sectors and activities per number of non-geolocalized projects

Sectors	Freq.	Activities	Freq.
Financial services	13.06%	Manufacturing	26.32%
Communications	11.82%	Sales, Marketing &Support	22.26%
Metals	8.08%	Business Services	18.30%
Business Services	7.28%	Extraction	8.08%
Food & Beverages	7.28%	ICT & Internet Infrastructure	6.58%
Software & IT Services	6.15%	Logistics, Distribution & Transportation	3.75%
Coal, oil & gas	5.62%	Electricity	3.48%
Automotive OEM	4.12%	Retail	2.68%
Chemicals	3.37%	Construction	2.30%
Transportation & Warehousing	3.32%	Research & Development	1.71%
Industrial Equipment	3.21%	Education & Training	1.39%
Renewable Energy	3.00%	Headquarters	0.96%
Consumer Products	2.41%	Customer Contact Centre	0.80%
Electronic components	2.19%	Maintenance & Servicing	0.70%
Building materials	2.03%	Recycling	0.97%
Textiles	1.93%	Technical Support Centre	0.32%

Note: authors' elaboration on fDi Markets data. The sample includes 1869 greenfield FDI projects in 24 countries from 2003 to 2020. These projects were not georeferenced and therefore they were excluded from our sample.

Figure A1: Geographic distribution of FDI projects across Africa



Source: authors' elaboration on fDi Markets data.

Table A6: Descriptive Statistics - Entire Sample

	Mean	SD	Median	Num. of Obs.
Female	0.544	0.073	0.537	11206
Urban	0.249	0.343	0.070	10886
Age	28.543	1.616	28.293	11206
Secondary Educ. +	0.261	0.272	0.149	11176
Employment	0.688	0.165	0.704	11140
Self Employment	0.559	0.222	0.574	10516
Employee	0.091	0.099	0.059	10516
Agriculture	0.376	0.252	0.371	10035
Non Agriculture	0.299	0.2	0.262	10035
High skilled	0.039	0.051	0.023	10751
White collar	0.150	0.137	0.107	10751
Blue collar	0.501	0.223	0.499	10751

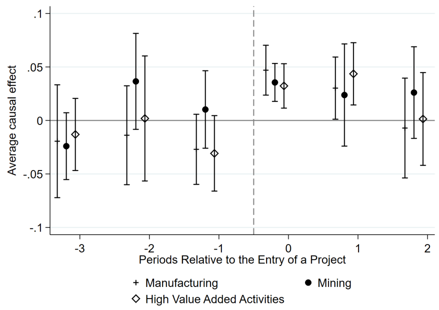
Note: authors' elaboration on the final dataset which combines fDi Markets data, IPUMS data and DHS data.

Table A7: Negative weights

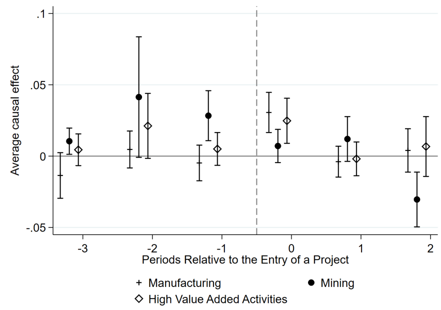
	Positive	Negative	Sum Negative	$\lvert E[\hat{\beta}_{twfe}]\rvert / SD_w$
Employment	587	52	-0.007	0.0183
High Skilled	537	53	-0.01	0.0096
White collar	537	53	-0.01	0.0051
Blue collar	537	53	-0.01	0.0095
Agriculture	519	51	-0.01	0.0038
Non agriculture	519	51	-0.01	0.0266
Self Employment	530	48	-0.01	0.0105

Note: this table reports the positive and negative weights and the sum of the latter attached to each coefficient of the outcomes of interest. Moreover it reports the standard deviation of the average treatment effects which indicate how serious the problem of heterogeneity is.

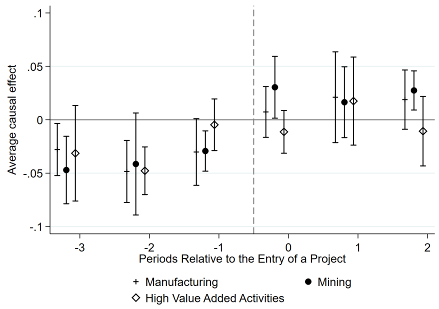
Figure A2: Relationship between FDI using other definitions and the outcomes



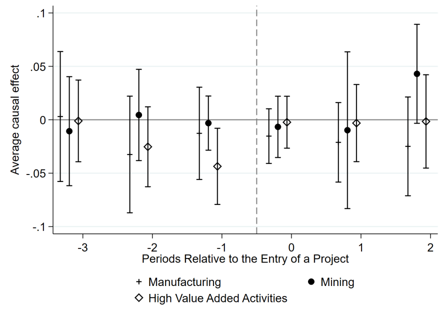
(a) Employment



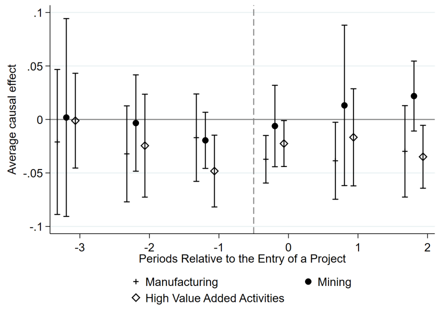
(b) High Skilled



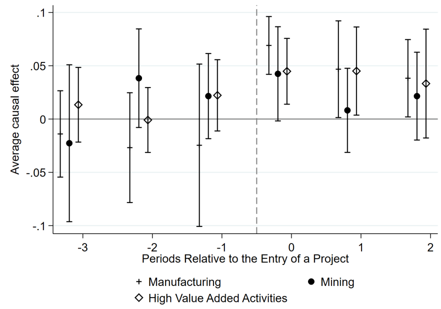
(c) White collar



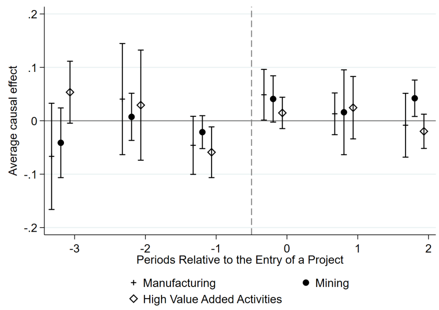
(d) Blue Collar



(e) Agriculture



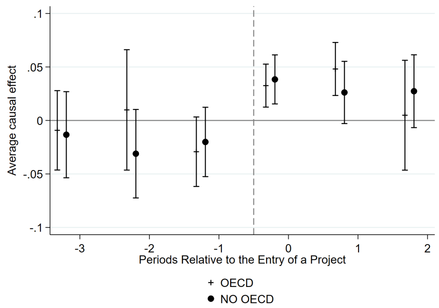
(f) Non Agriculture



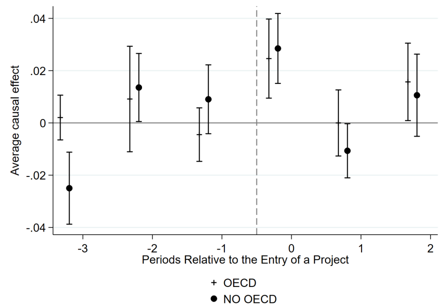
(g) Self-Employment

Note: this figure plots the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is a location. As controls we include the share of females, the share of people who lives in urban areas, the share of individuals with at least secondary education and average age in each location, while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after an FDI project entry.

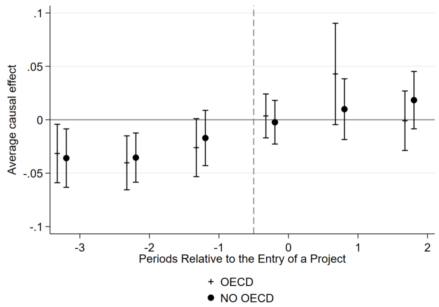
Figure A3: Relationship between FDI considering the donor and the outcomes



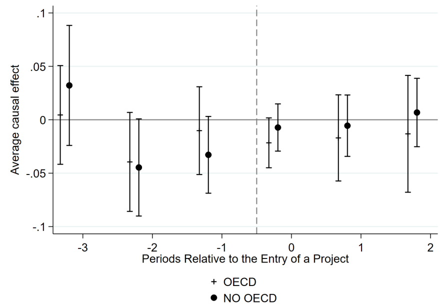
(a) Employment



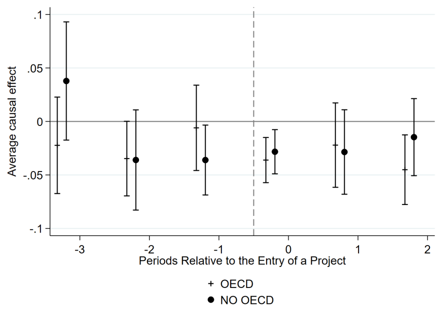
(b) High Skilled



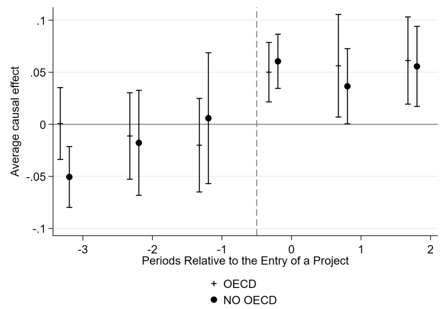
(c) White collar



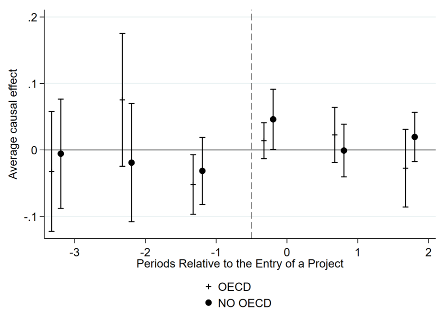
(d) Blue Collar



(e) Agriculture



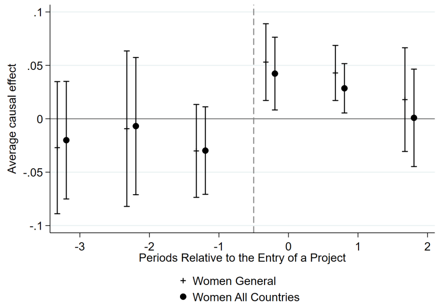
(f) Non Agriculture



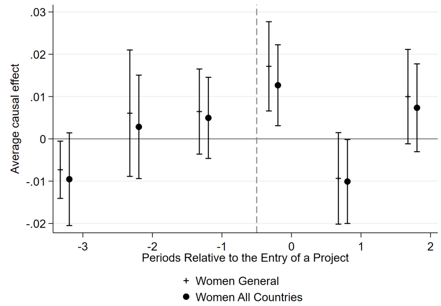
(g) Self-Employment

Note: this figure plots the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is a location. As controls we include the share of females, the share of people who lives in urban areas, the share of individuals with at least secondary education and average age in each location, while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after an FDI project entry.

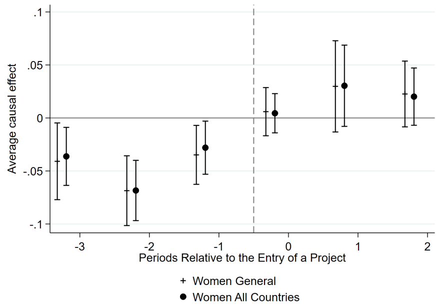
Figure A4: Relationship between FDI and the outcomes disaggregated by gender



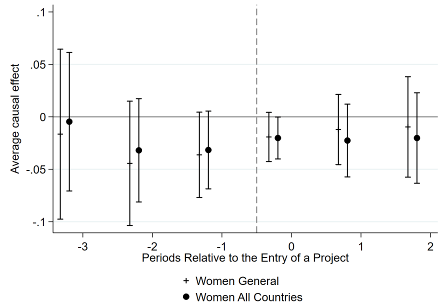
(a) Employment



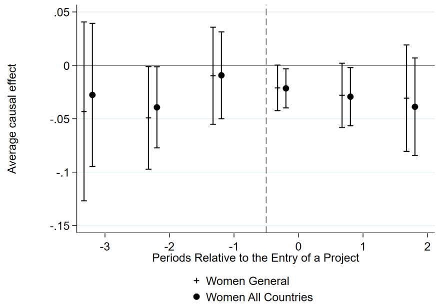
(b) High Skilled



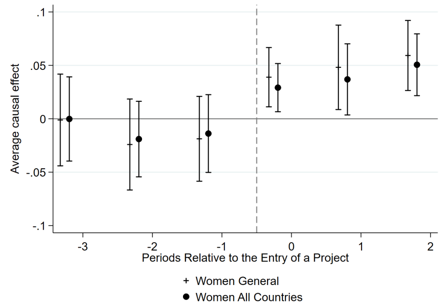
(c) White collar



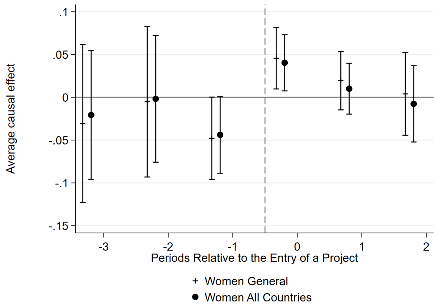
(d) Blue Collar



(e) Agriculture



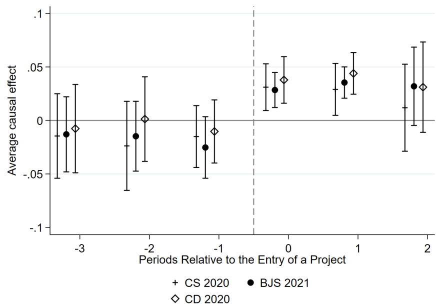
(f) Non Agriculture



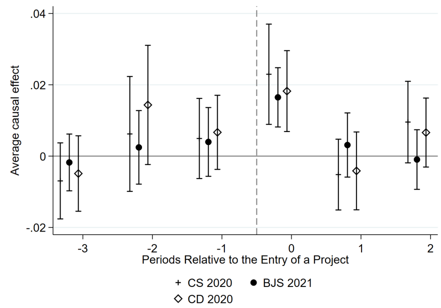
(g) Self-Employment

Note: this figure plots the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is a location. As controls we include the share of females, the share of people who lives in urban areas, the share of individuals with at least secondary education and average age in each location, while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after an FDI project entry.

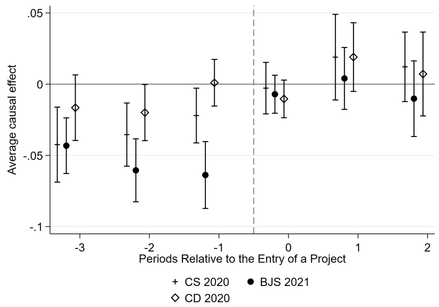
Figure A5: Relationship between FDI and the outcomes using other estimators



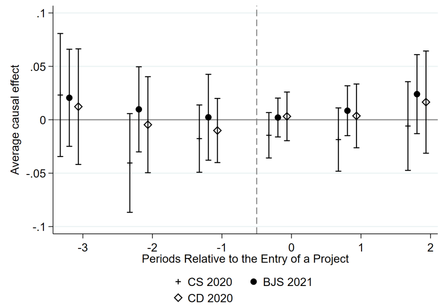
(a) Employment



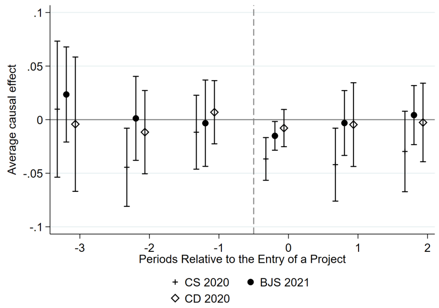
(b) High Skilled



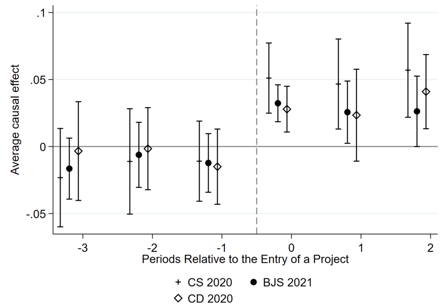
(c) White collar



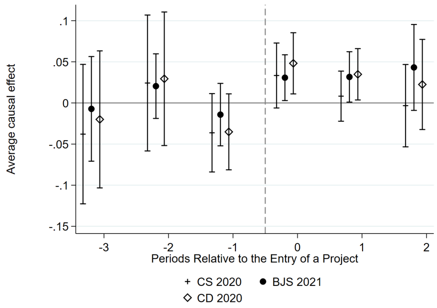
(d) Blue Collar



(e) Agriculture



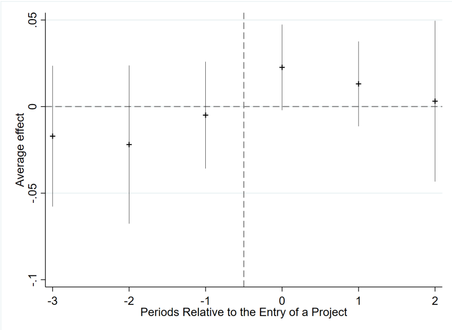
(f) Non Agriculture



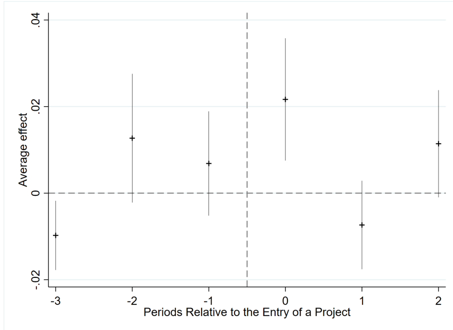
(g) Self-Employment

Note: this figure plots the event-study estimates of the average treatment effect on the treated using the following three methods: (1) the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021) (CS 2020); (2) the method proposed by De Chaisemartin and d'Haultfoeuille (2020) (CD 2020); and (3) the method proposed by Borusyak et al. (2021) (BJS 2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. We control for the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include total population of the area as a weight. We plot estimates 3 periods before and 2 periods after an FDI project entry.

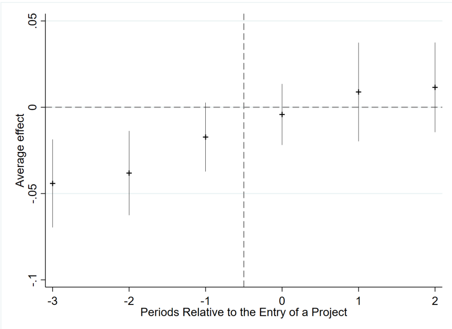
Figure A6: Relationship between FDI and the main outcomes dropping capital



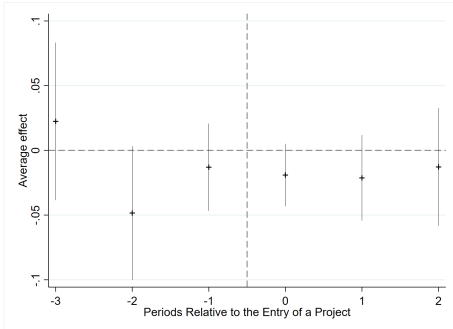
(a) Employment



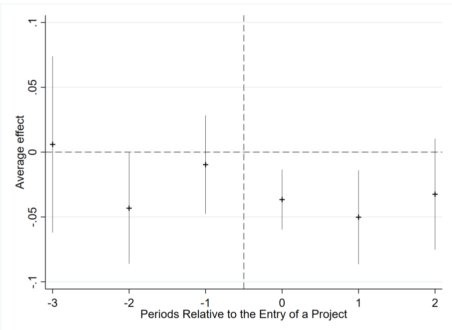
(b) High Skilled



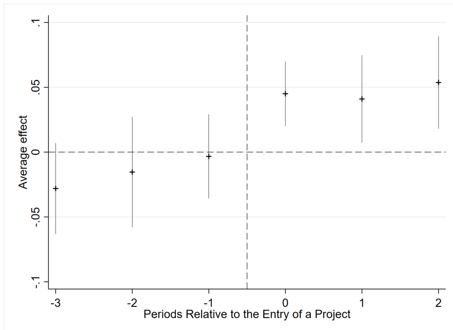
(c) White collar



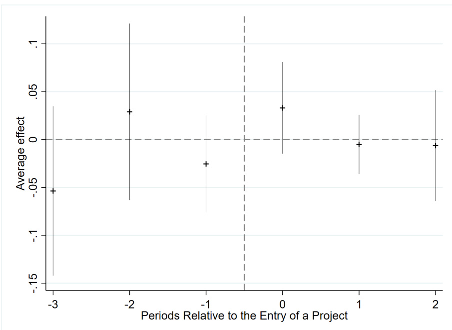
(d) Blue Collar



(e) Agriculture



(f) Non Agriculture



(g) Self-Employment

Note: this figure plots the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). The sample excludes all the administrative area in which the capital city of a country is located. All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. We control for the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include total population of the area as a weight. We plot estimates 3 periods before and 2 periods after an FDI project entry.

Table A8: Event Study Coefficients - Relationship between FDI projects and the outcomes of interest

	Employment	High Skilled	White	Blue	Agri	No Agri	Self
3 periods before event	-0.0145	-0.00695	-0.0425**	0.0232	0.00979	-0.0232	-0.0379
	(0.0202)	(0.00544)	(0.0134)	(0.0294)	(0.0324)	(0.0187)	(0.0433)
2 periods before event	-0.0238	0.00622	-0.0354**	-0.0404	-0.0444*	-0.0111	0.0242
	(0.0212)	(0.00822)	(0.0113)	(0.0236)	(0.0186)	(0.0201)	(0.0422)
1 period before event	-0.0151	0.00494	-0.0220*	-0.0176	-0.0117	-0.0109	-0.0363
	(0.0148)	(0.00573)	(0.00980)	(0.0161)	(0.0176)	(0.0153)	(0.0243)
Period of event	0.0311**	0.0230**	-0.00280	-0.0145	-0.0367***	0.0511***	0.0334
	(0.0112)	(0.00717)	(0.00922)	(0.0108)	(0.0102)	(0.0134)	(0.0202)
1 period after event	0.0290*	-0.00519	0.0190	-0.0185	-0.0420*	0.0466**	0.00825
	(0.0124)	(0.00506)	(0.0153)	(0.0151)	(0.0174)	(0.0171)	(0.0155)
2 periods after event	0.0119	0.00954	0.0122	-0.00587	-0.0297	0.0569**	-0.00333
	(0.0208)	(0.00583)	(0.0124)	(0.0212)	(0.0192)	(0.0179)	(0.0255)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Table A9: Event Study Coefficients - Relationship between FDI projects and the outcomes of interest considering the type of activity performed

	Emp(Man)	Emp(Min)	Emp(HIVA)	High Skill(Man)	High Skill(Min)	High Skill(HIVA)	White(Man)
3 periods before event	-0.0194	-0.0240	-0.0131	-0.0136	0.0105*	0.00445	-0.0279*
	(0.0269)	(0.0159)	(0.0172)	(0.00816)	(0.00469)	(0.00568)	(0.0125)
2 periods before event	-0.0138	0.0365	0.00185	0.00465	0.0414	0.0212	-0.0484**
	(0.0236)	(0.0229)	(0.0298)	(0.00661)	(0.0216)	(0.0116)	(0.0148)
1 period before event	-0.0270	0.0103	-0.0308	-0.00479	0.0283**	0.00503	-0.0302
	(0.0167)	(0.0185)	(0.0180)	(0.00637)	(0.00894)	(0.00587)	(0.0159)
Period of event	0.0470***	0.0355***	0.0323**	0.0306***	0.00710	0.0248**	0.00732
	(0.0119)	(0.00905)	(0.0106)	(0.00718)	(0.00594)	(0.00807)	(0.0121)
1 period after event	0.0302*	0.0238	0.0436**	-0.00390	0.0120	-0.00192	0.0211
	(0.0148)	(0.0244)	(0.0148)	(0.00552)	(0.00801)	(0.00603)	(0.0217)
2 periods after event	-0.00709	0.0261	0.00138	0.00399	-0.0304**	0.00673	0.0189
	(0.0238)	(0.0218)	(0.0221)	(0.00776)	(0.00979)	(0.0107)	(0.0142)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Man" refer to the projects in manufacturing, the columns indicated with "Min" refer to the projects in mining and the columns indicated with "HIVA" refer to the projects in high-value added activities. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Event Study Coefficients - Relationship between FDI projects and the outcomes of interest considering the type of activity performed *(Cont.)*

	White(Min)	White(HIVA)	Blue(Man)	Blue(Min)	Blue(HIVA)	Agri(Man)	Agri(Min)
3 periods before event	-0.0471**	-0.0314	0.00304	-0.0107	-0.00107	-0.0210	0.00183
	(0.0161)	(0.0228)	(0.0310)	(0.0260)	(0.0195)	(0.0346)	(0.0472)
2 periods before event	-0.0414	-0.0478***	-0.0325	0.00448	-0.0253	-0.0322	-0.00336
	(0.0244)	(0.0114)	(0.0279)	(0.0218)	(0.0191)	(0.0229)	(0.0230)
1 period before event	-0.0293**	-0.00469	-0.0127	-0.00315	-0.0436*	-0.0170	-0.0195
	(0.00962)	(0.0123)	(0.0220)	(0.0129)	(0.0182)	(0.0208)	(0.0134)
Period of event	0.0304*	-0.0113	-0.0153	-0.00665	-0.00230	-0.0372**	-0.00613
	(0.0148)	(0.0102)	(0.0131)	(0.0146)	(0.0124)	(0.0113)	(0.0194)
1 period after event	0.0164	0.0175	-0.0211	-0.00977	-0.00311	-0.0386*	0.0131
	(0.0169)	(0.0210)	(0.0190)	(0.0374)	(0.0184)	(0.0184)	(0.0382)
2 periods after event	0.0274**	-0.0107	-0.0249	0.0430	-0.00155	-0.0298	0.0219
	(0.00935)	(0.0166)	(0.0236)	(0.0236)	(0.0223)	(0.0218)	(0.0167)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Man" refer to the projects in manufacturing, the columns indicated with "Min" refer to the projects in mining and the columns indicated with "HIVA" refer to the projects in high-value added activities. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Event Study Coefficients - Relationship between FDI projects and the outcomes of interest considering the type of activity performed *(Cont.)*

	Agri(HIVA)	No Agri(Man)	No Agri(Min)	No Agri(HIVA)	Self(Man)	Self(Min)	Self(HIVA)
3 periods before event	-0.00114	-0.0140	-0.0227	0.0134	-0.0667	-0.0412	0.0535
	(0.0226)	(0.0207)	(0.0376)	(0.0179)	(0.0507)	(0.0333)	(0.0296)
2 periods before event	-0.0245	-0.0268	0.0383	-0.000890	0.0406	0.00733	0.0292
	(0.0245)	(0.0263)	(0.0236)	(0.0155)	(0.0531)	(0.0225)	(0.0526)
1 period before event	-0.0483**	-0.0246	0.0215	0.0223	-0.0460	-0.0213	-0.0590*
	(0.0171)	(0.0389)	(0.0204)	(0.0171)	(0.0277)	(0.0157)	(0.0242)
Period of event	-0.0225*	0.0691***	0.0424	0.0448**	0.0487*	0.0409	0.0147
	(0.0109)	(0.0138)	(0.0225)	(0.0158)	(0.0242)	(0.0221)	(0.0150)
1 period after event	-0.0167	0.0468*	0.00817	0.0450*	0.0130	0.0159	0.0245
	(0.0232)	(0.0232)	(0.0201)	(0.0211)	(0.0199)	(0.0405)	(0.0298)
2 periods after event	-0.0349*	0.0383*	0.0215	0.0333	-0.00836	0.0421*	-0.0198
	(0.0150)	(0.0185)	(0.0210)	(0.0260)	(0.0305)	(0.0174)	(0.0164)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Man" refer to the projects in manufacturing, the columns indicated with "Min" refer to the projects in mining and the columns indicated with "HIVA" refer to the projects in high-value added activities. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Table A10: Event Study Coefficients - Relationship between FDI projects and the outcomes of interest considering the donors

	Empl(Oecd)	Empl(no Oecd)	Skilled(Oecd)	Skilled(no Oecd)	White(Oecd)	White(no Oecd)	Blue(Oecd)
3 periods before event	-0.00922	-0.0133	0.00205	-0.0250***	-0.0316*	-0.0359*	0.00452
	(0.0189)	(0.0205)	(0.00437)	(0.00703)	(0.0140)	(0.0140)	(0.0236)
2 periods before event	0.00987	-0.0311	0.00914	0.0136*	-0.0403**	-0.0354**	-0.0395
	(0.0287)	(0.0211)	(0.0103)	(0.00664)	(0.0129)	(0.0118)	(0.0236)
1 period before event	-0.0292	-0.0201	-0.00448	0.00904	-0.0261	-0.0171	-0.0102
	(0.0166)	(0.0165)	(0.00522)	(0.00672)	(0.0139)	(0.0132)	(0.0210)
Period of event	0.0326**	0.0384**	0.0246**	0.0285***	0.00359	-0.00232	-0.0216
	(0.0103)	(0.0117)	(0.00772)	(0.00683)	(0.0105)	(0.0104)	(0.0119)
1 period after event	0.0481***	0.0262	-0.0000180	-0.0107*	0.0429	0.00993	-0.0170
	(0.0126)	(0.0148)	(0.00645)	(0.00528)	(0.0242)	(0.0145)	(0.0206)
2 periods after event	0.00487	0.0273	0.0157*	0.0106	-0.000902	0.0184	-0.0132
	(0.0262)	(0.0174)	(0.00756)	(0.00802)	(0.0142)	(0.0137)	(0.0279)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Oecd" refer to the projects received by an oecd country while the columns indicated with "No oecd" refer to the projects received by a non-oecd country. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Event Study Coefficients - Relationship between FDI projects and the outcomes of interest considering the donors *(Cont.)*

	Blue(no Oecd)	Agri(Oecd)	Agri(no Oecd)	No Agri(Oecd)	No Agri(no Oecd)	Self(Oecd)	Self(no Oecd)
3 periods before event	0.0321	-0.0223	0.0378	0.000777	-0.0506***	-0.0324	-0.00565
	(0.0287)	(0.0230)	(0.0282)	(0.0176)	(0.0149)	(0.0460)	(0.0419)
2 periods before event	-0.0446	-0.0347	-0.0360	-0.0112	-0.0178	0.0753	-0.0192
	(0.0232)	(0.0178)	(0.0239)	(0.0212)	(0.0257)	(0.0509)	(0.0453)
1 period before event	-0.0328	-0.00597	-0.0361*	-0.0200	0.00588	-0.0521*	-0.0316
	(0.0183)	(0.0204)	(0.0167)	(0.0229)	(0.0321)	(0.0228)	(0.0257)
Period of event	-0.00725	-0.0361***	-0.0283**	0.0501***	0.0605***	0.0138	0.0461*
	(0.0113)	(0.0108)	(0.0105)	(0.0146)	(0.0133)	(0.0138)	(0.0231)
1 period after event	-0.00555	-0.0221	-0.0286	0.0562*	0.0365*	0.0226	-0.000916
	(0.0147)	(0.0201)	(0.0201)	(0.0251)	(0.0184)	(0.0212)	(0.0202)
2 periods after event	0.00679	-0.0451**	-0.0147	0.0613**	0.0556**	-0.0275	0.0195
	(0.0163)	(0.0166)	(0.0184)	(0.0214)	(0.0196)	(0.0298)	(0.0189)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Oecd" refer to the projects received by an oecd country while the columns indicated with "No oecd" refer to the projects received by a non-oecd country. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Table A11: Event Study Coefficients - Relationship between FDI projects and the outcomes of interest disaggregated by gender

	Emp(Gen)	Emp(All)	High Skilled(Gen)	High Skilled(All)	White(Gen)	White(All)	Blue(Gen)
3 periods before event	-0.0271	-0.0200	-0.00731*	-0.00954	-0.0409*	-0.0363**	-0.0165
	(0.0316)	(0.0281)	(0.00345)	(0.00559)	(0.0185)	(0.0140)	(0.0413)
2 periods before event	-0.00930	-0.00687	0.00606	0.00283	-0.0686***	-0.0684***	-0.0444
	(0.0371)	(0.0328)	(0.00762)	(0.00624)	(0.0168)	(0.0145)	(0.0303)
1 period before event	-0.0301	-0.0298	0.00645	0.00495	-0.0348*	-0.0280*	-0.0362
	(0.0222)	(0.0209)	(0.00513)	(0.00489)	(0.0142)	(0.0128)	(0.0208)
Period of event	0.0530**	0.0423*	0.0171**	0.0127**	0.00594	0.00445	-0.0192
	(0.0183)	(0.0174)	(0.00538)	(0.00488)	(0.0116)	(0.00943)	(0.0120)
1 period after event	0.0429**	0.0285*	-0.00935	-0.0101*	0.0298	0.0304	-0.0121
	(0.0132)	(0.0118)	(0.00552)	(0.00506)	(0.0220)	(0.0195)	(0.0171)
2 periods after event	0.0179	0.000865	0.00998	0.00735	0.0226	0.0201	-0.00963
	(0.0248)	(0.0233)	(0.00569)	(0.00530)	(0.0158)	(0.0138)	(0.0244)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed women, the share of women in high skill jobs, in white collar jobs and in blue collar jobs, the share of women employed in agriculture and outside agriculture, and the share of women self-employed or working in a family business. As controls we include the share of women who lives in the urban areas, the share of female individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Gen" refer to the sample of female respondents for the same countries and waves used in the main analysis. The columns indicated with "All" refer to the sample of women which includes also Egypt, Madagascar and Morocco. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Event Study Coefficients - Relationship between FDI projects and the outcomes of interest disaggregated by gender *(Cont.)*

	Blue (All)	Agri (Gen)	Agri (All)	No Agri (Gen)	No Agri (All)	Self (Gen)	Self (All)
3 periods before event	-0.00463	-0.0431	-0.0277	-0.00115	-0.000195	-0.0307	-0.0207
	(0.0337)	(0.0427)	(0.0342)	(0.0219)	(0.0201)	(0.0471)	(0.0383)
2 periods before event	-0.0319	-0.0492*	-0.0393*	-0.0241	-0.0190	-0.00507	-0.00185
	(0.0251)	(0.0245)	(0.0194)	(0.0217)	(0.0180)	(0.0449)	(0.0377)
1 period before event	-0.0316	-0.00972	-0.00935	-0.0188	-0.0139	-0.0479	-0.0438
	(0.0189)	(0.0232)	(0.0208)	(0.0203)	(0.0186)	(0.0246)	(0.0229)
Period of event	-0.0202*	-0.0211	-0.0216*	0.0389**	0.0291*	0.0455*	0.0404*
	(0.0102)	(0.0109)	(0.00933)	(0.0142)	(0.0115)	(0.0183)	(0.0168)
1 period after event	-0.0226	-0.0280	-0.0293*	0.0481*	0.0369*	0.0195	0.0101
	(0.0177)	(0.0153)	(0.0139)	(0.0202)	(0.0170)	(0.0174)	(0.0151)
2 periods after event	-0.0202	-0.0307	-0.0388	0.0592***	0.0506***	0.00395	-0.00762
	(0.0220)	(0.0254)	(0.0233)	(0.0167)	(0.0148)	(0.0247)	(0.0227)

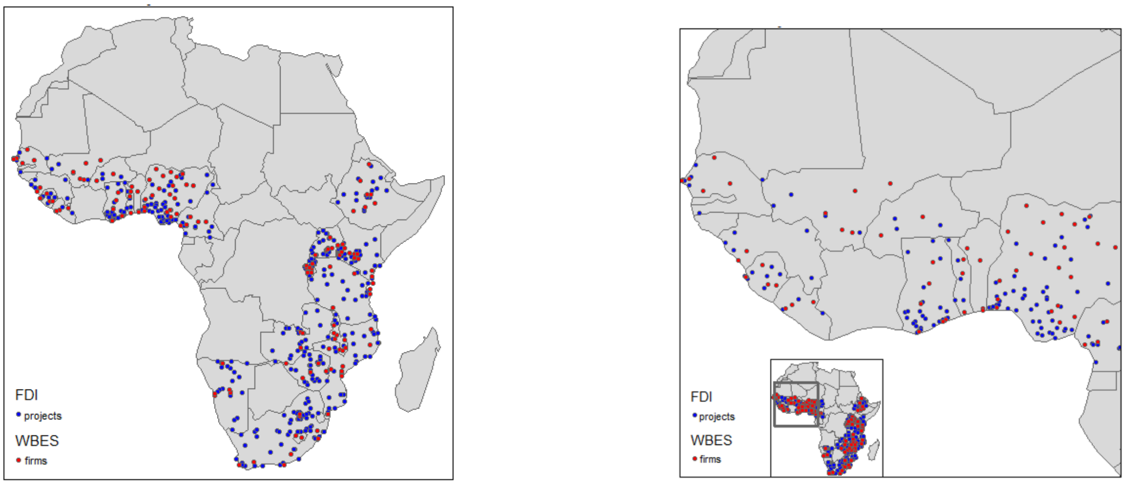
Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed women, the share of women in high skill jobs, in white collar jobs and in blue collar jobs, the share of women employed in agriculture and outside agriculture, and the share of women self-employed or working in a family business. As controls we include the share of women who lives in the urban areas, the share of female individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. The columns indicated with "Gen" refer to the sample of female respondents for the same countries and waves used in the main analysis. The columns indicated with "All" refer to the sample of women which includes also Egypt, Madagascar and Morocco. Standard errors clustered at the province level in parentheses. *** p<0.001, ** p<0.01, * p<0.05.

Table A12: Event Study Coefficients - Relationship between FDI projects and the outcomes of interest without capital cities

	Employment	High Skilled	White	Blue	Agri	No Agri	Self
3 periods before event	-0.0171	-0.00977*	-0.0441***	0.0224	0.00597	-0.0281	-0.0537
	(0.0207)	(0.00408)	(0.0130)	(0.0311)	(0.0347)	(0.0179)	(0.0451)
2 periods before event	-0.0219	0.0127	-0.0381**	-0.0484	-0.0433*	-0.0154	0.0290
	(0.0233)	(0.00759)	(0.0125)	(0.0264)	(0.0219)	(0.0217)	(0.0470)
1 period before event	-0.00495	0.00685	-0.0173	-0.0130	-0.00966	-0.00332	-0.0255
	(0.0157)	(0.00615)	(0.0102)	(0.0172)	(0.0194)	(0.0166)	(0.0259)
Period of event	0.0226	0.0217**	-0.00418	-0.0191	-0.0366**	0.0450***	0.0331
	(0.0126)	(0.00721)	(0.00903)	(0.0123)	(0.0118)	(0.0127)	(0.0244)
1 period after event	0.0131	-0.00736	0.00884	-0.0213	-0.0502**	0.0410*	-0.00513
	(0.0125)	(0.00521)	(0.0146)	(0.0169)	(0.0185)	(0.0172)	(0.0158)
2 periods after event	0.00311	0.0114	0.0115	-0.0128	-0.0325	0.0537**	-0.00630
	(0.0237)	(0.00632)	(0.0133)	(0.0232)	(0.0218)	(0.0182)	(0.0295)

Note: this table reports the event-study estimates of the average treatment effect on the treated using the doubly-robust estimator in Sant'Anna and Zhao (2020) and Callaway and Sant'Anna (2021). All the coefficients are estimated using the 95% confidence interval. The unit of observation is the province. The outcomes of interest are the share of employed population, the share of population in high skill jobs, in white collar jobs and in blue collar jobs, the share of population employed in agriculture and outside agriculture, and the share of population self-employed or working in a family business. As controls we include the share of female, the share of people who lives in the urban areas, the share of individuals with at least secondary education and average age in each province while we include the total population of the area as a weight. We plot the estimates 3 periods before and 2 periods after the treatment. Standard errors clustered at the province level in parentheses. *** $p<0.001$, ** $p<0.01$, * $p<0.05$

Figure A7: Geographic location of WBES firms (red dots) and FDI projects (blue dots)



Note: authors' elaboration on WBES and fDiMarkets data.

Table A13: WBES firms and FDI projects

Country	Waves of WBES	WBES firms	FDI Projects
Benin	2009, 2016	300	12
Burkina Faso	2009	394	28
Burundi	2006, 2014	427	12
Cameroon	2009, 2016	724	89
Ethiopia	2011, 2015	1492	174
Ghana	2007, 2013	1214	375
Guinea	2006, 2016	373	34
Kenya	2007, 2013, 2018	2439	624
Lesotho	2009, 2016	301	6
Liberia	2009, 2017	301	27
Malawi	2009, 2014	673	10
Mali	2007, 2010, 2016	1035	23
Mozambique	2007, 2018	1080	249
Namibia	2006, 2014	909	105
Nigeria	2007, 2014	4567	561
Rwanda	2006, 2011, 2019	813	119
Senegal	2007, 2014	1107	99
Sierra Leone	2009, 2017	302	21
South Africa	2007, 2020	2034	1731
Tanzania	2006, 2013	1232	207
Togo	2009, 2016	305	31
Tunisia	2013, 2020	1207	155
Uganda	2006, 2013	1325	162
Zambia	2007, 2013, 2019	1805	161
Zimbabwe	2011, 2016	1199	93

Note: The sample includes all firms surveyed for the countries included in the sample used in the individual level analysis. The number WBES firms include as well foreign firms (which are excluded from the empirical analysis). FDI projects refer to the number of projects recorded before the respective WBES waves.

Table A14: Mechanisms: the effect of FDI on firms outcomes–Backward Linkages

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Productivity	Total Sales	Investment	Number of employees	Skilled employees	Labor cost on employment
active_back_50	0.408**	0.343*	0.0265	-0.0981*	0.0743	0.616***
	(0.165)	(0.187)	(0.0372)	(0.0542)	(0.0900)	(0.178)
inactive_back_50	-0.152	-0.186	-0.0492	-0.00659	0.0924	0.239***
	(0.109)	(0.116)	(0.0304)	(0.0344)	(0.152)	(0.0906)
Constant	13.23***	13.78***	0.258***	0.528***	0.551***	11.23***
	(0.0762)	(0.0827)	(0.0188)	(0.0252)	(0.0921)	(0.0673)
Observations	10,525	10,579	11,334	11,339	2,654	10,082
R-squared	0.643	0.654	0.157	0.792	0.279	0.655
Difference	0.560	0.528	0.0758	-0.0915	-0.0181	0.377
p-value difference	0.000805	0.00501	0.0441	0.0590	0.891	0.0213

Note: The analysis is based on the estimation of equation 2. The unit of observation is a domestic firm. Firms who report foreign ownership in the WBES sample are dropped. The table reports at the bottom the coefficient of interest, the difference between the coefficients Active and Inactive, and its p-value. In this Table, the treatment is defined as the proximity (within a 50km buffer) to at least an FDI in a 2-digit sector that, according to the country I/O table, is buying inputs from the sector in which the domestic firm operates. All regressions include a dummy for firm size (small, medium, large), the age of the firm, city, industry (2-digit ISIC Rev 3.1) and country-year fixed effects. Standard errors clustered at the city-industry level in parentheses.*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A15: Alternative Buffers

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Productivity	Total Sales	Investment	Number of employees	Skilled employees	Labor cost on employment
Panel A: 25 km buffer						
Difference	0.194	0.279	0.0356	0.0598	0.0498	0.0440
p-value difference	0.0389	0.00605	0.177	0.135	0.243	0.651
Panel B: 50 km buffer						
Difference	0.192	0.287	0.0438	0.0724	0.0438	0.0407
p-value difference	0.0335	0.00296	0.0970	0.0395	0.253	0.664
Panel C: 100 km buffer						
Difference	0.118	0.186	0.0495	0.0599	0.0533	0.0162
p-value difference	0.166	0.0430	0.0465	0.0551	0.111	0.852
Panel D: 200 km buffer						
Difference	0.0939	0.119	0.0181	0.0264	0.0294	0.0331
p-value difference	0.256	0.192	0.407	0.373	0.369	0.690

Note: The analysis is based on the estimation of equation 2. We only report the coefficient of the difference between Active and Inactive projects at different buffers, from 25 to 200km. The unit of observation is a domestic firm. Firms who report foreign ownership in the WBES sample are dropped. All regressions include a dummy for firm size (small, medium, large), the age of the firm, city, industry (2-digit ISIC Rev 3.1) and country-year fixed effects. Standard errors clustered at the city-industry level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$