ASP Course on Microeconometric Methods (2021) - Take Home Assignment -

Project A

Susann Adloff, Linda Maokomatanda, Saskia Meuchelböck

Section 1

We review three papers that provide a seamless synergy into our project that estimates the impact of foreign direct investment on firm performance by taking a holistic approach to the literature. We first take a look at a paper by Borin and Mancini (2016) which displays the effects on firm performance in the firm's country of origin when they decide to invest abroad, thus becoming a multi-national enterprise. In an intermediate stage, we review a paper by Bajgar and Javorcik (2020) that investigates analyze spillover effects from FDI on domestic firm performance resulting from links along the supply chain. To conclude our review and introduce section 2, we review Chen (2011) studies the role of the origin of FDIs on target firm performance changes.

Borin and Mancini (2016) investigate the ex-post effects of foreign direct investment (FDI) on firm performance using data from the Bank of Italy's annual Survey of Industrial and Service Firms (the Invind survey) as well as balance sheet data from the Company Accounts Data Service (henceforth CADS). They extend the discussion from empirical and theoretical literature that shows how multinational firms (MNE) exhibit a competitive advantage before investing abroad, by conducting microeconomic data analysis to evaluate the policy implications of firm heterogeneity. The authors specify the best-case scenario to be the implementation of policy measures and internationalization strategies that are capable of enhancing both firm performance and employment. The ex-ante causal relationship (from performance to internationalization) in Borin and Mancini (2016)'s paper introduces a severe form of endogeneity, in that ex-post performance might reflect not only foreign investment, but also pre-existent advantages in terms of managerial ability, know-how and technology. To accurately evaluate the ex-post effects of FDI that take into account the inherent self-selection problem, the authors use a propensity score matching procedure to analyze the causal relationship between firm performance and foreign direct investment, looking in particular at the potential gains in terms of productivity and potential losses in terms of employment in the parent firm due to the acquisition of multinational status; as well as to evaluate whether these effects are evenly spread across new MNEs or concentrated among certain groups of investors. To implement the propensity matching estimation, the authors choose a company, ex-ante very similar to the first one, but that does not choose to invest abroad to act as a proxy of the unobservable counterfactual, enabling them to compare the evolution of new MNEs with the performance of the exact same company with no investment; since it is not possible to observe the same company in these two different scenarios. They find that firms investing abroad for the very first time, especially in advanced economies, show higher productivity and employment dynamics in the years following the investment: the average positive effect on TFP is driven by new multinationals operating in specialized and high-tech sectors, while the positive employment gains are explained by an increase of the white collar component. On average, the authors do not find a negative effects on the parent firm's blue collar component.

Bajgar and Javorcik (2020) analyze spillover effects from FDI on domestic firm performance resulting from links along the supply chain. More specifically, they investigate the relationship between quality upgrading by Romanian exporters and the presence of foreign affiliates in upstream and downstream industries, respectively. They use customs data merged with firm-level data for the years 2005 to 2011 and find a positive and robust link between the product quality of domestic exporters – proxied by two measures established in the literature, namely unit values and a measure proposed by Khandelwal et al. (2013) – and the presence of foreign-owned firms in upstream (i.e. input-supplying) industries. The results indicate that a one percentage point increase in FDI presence upstream is related to a roughly 0.5% increase in product quality of domestic exporters, on average and ceteris paribus. The positive relationship between export quality and the presence of foreign-owned firms in downstream industries is less robust. To estimate these effects, the authors employ OLS regressions with product quality as the dependent variable and a measure of FDI presence in upstream and downstream industries, respectively, as the main independent variable of interest. They control for firm-product-destination fixed effects, region-time fixed effects and linear time trends at the region-industry level. However, the authors acknowledge that reverse causality could be a problem, for example if foreign firms primarily invest in industries where high-quality domestic inputs become increasingly available, and these developments are not captured by the linear time trends. To mitigate these concerns, they lag their FDI variables by one year in all specifications. In addition, they perform a "strict exogeneity" test suggested by Wooldridge (2010), including also contemporaneous and lead FDI values into their model. They argue that if foreign firms enter Romania as a consequence of quality upgrading rather than the other way around, the coefficients on the lead values of FDI should also be statistically significant, which is not the case. They conclude that the increase in foreign presence in the upstream industries over the studied period corresponded to a circa 4% increase in the quality of exports by local firms.

Chen (2011) studies the role of the origin of FDIs on target firm performance changes. It uses data on acquisitions of U.S firms between 1979 and 2006, comparing firm level performance indicators before and after acquisition, focusing on the difference between firms acquired by domestic firms (USFs), firms from industrialized countries (ICFs) and firms from developing countries (DCFs). The identification strategy is based on a diff-in-diff analysis which gives rise to a twofold selection problem. The empirical set up defines a control group which is domestically acquired firms and two treatment groups. As trans-boundary firm acquisition is more challenging than domestic acquisitions (Helpman et al 2004), firms that engage in the former are likely to be different from those that invest domestically, which might translate into "more skillful" target firm selection on part of foreign acquirers. Hence, the firms chosen for foreign investments are likely to structurally differ pre-treatment from those firms that are chosen for domestic acquisitions. Secondly, selection criteria for target firms might also differ based on the origin of the acquiring firm. The author argues that systematic differences can be expected in the aimed at restructuring process depending on the origin of the acquirer, due to structural differences between these groups in technological progress and relative input costs between target and acquirer. Consequently, selection bias might be found in the comparison of control and the treatment groups, as well as in the comparison of the two treatment groups. To solve this issue the author implements a propensity score matching procedure. The scoring is achieved through a multinomial logit model estimation of observables available to acquires before the acquisition that provide information on present and potential future target firm performance, as well as year, industry and state fixed effects. The matching is done using a kernel matching procedure. This step allows to impose pre-treatment homogeneity in firm performance across comparison groups and thus to interpret any difference in performance between matched pairs to result from the difference in treatment. The results of this estimation show that firms acquired by ICFs show 13% higher improvements in terms of labour productivity, a 10% higher profit increase and a 19% higher increase in sales in the five years post acquisition relative to the year before the acquisition, compared to firms acquired by USFs. Between those two groups no difference in overall employment changes was found. For firms acquired by DCFs the author finds that labour productivity gains and changes in sales and employment are 1% lower than those of domestic acquisitions, albeit higher profit gains. Comparing firm-level effects of FDIs originating from industrialized countries to those from developing countries, gains in sales, labour productivity and employment are significantly higher for target firms acquired by ICFs.

In Section 2, we will provide a discussion of the most important features of the data we will be working with, including some interesting patterns and correlations that we have discovered in the data. We also provide some summary statistics and graphs of the variables of interest.

Section 2

The aim of our research project is to estimate the impact of foreign direct investment (FDI) on labor market outcomes. More specifically, we analyze the effect of FDI on employment and wages. We use firm-level data for the years 2015 to 2017. In 2016, some of the firms receive FDI. The receipt of FDI is referred to as treatment in the following. Table 1 shows summary statistics for the main variables included in our dataset, both for 2015 – i.e. our pre-treatment period – and 2017 – i.e. the post-treatment period. We have information on wages, employment, total factor productivity (TFP), debt, export intensity as well as a dummy variable indicating whether a firm engages in research and development activities. In total, 11,323 firms are included in our dataset. On average, wages decreased from 2015 to 2017, while TFP, employment, export intensity, and the share of firms engageing in R&D increased. We do not have information on debt in 2017.

Table 2 reports our main variables by treatment, i.e. for two subsamples: firms receiving FDI in 2016 (4,460 or 39%) and firms not receiving FDI (6,884). In addition, we report the differences in means as well as the p-values of a t-test of the null hypothesis that means are equal across groups. For most of the pre-treatment variables the null hypothesis can be rejected, providing first evidence that treatment is not randomly assigned but depends on firm-level characteristics. Firms receiving FDI, on average, pay lower wages, employ more workers, have a lower level of TFP and export more of their production. There are no significant differences regarding R&D activities.

In addition, our dataset provides information on several firm-level characteristics that do not change over the time period studied: the firm's level of technology, its ownership struc-

Table 1: Summary statistics

	Mean	Std. Dev.	Min.	Max.	Obs.
Wages in 2015 (logs)	7.333	3.839	-7.332	22.432	11323
Total factor productivity in 2015	3.041	2.047	-5.359	11.357	11323
Employment in 2015 (logs)	4.411	3.040	-6.229	15.993	11323
Debt in 2015 (logs)	0.504	0.353	-0.200	1.300	11323
Export intensity in 2015	0.159	0.080	0.010	0.483	11323
RD in 2015 (dummy)	0.121	0.326	0.000	1.000	11323
Wages in 2017 (logs)	5.010	3.083	-6.185	17.042	11323
Total factor productivity in 2017	3.656	2.056	-4.701	11.811	11323
Emplyoment in 2017 (logs)	5.030	3.095	-6.218	16.388	11323
Export intensity in 2017	0.270	0.108	0.019	0.950	11323
RD in 2017 (dummy)	0.407	0.491	0.000	1.000	11323

Table 2: Summary statistics by treatment status

	No FDI	FDI	Mean diff.	P-value
Wages in 2015 (logs)	7.529	7.031	0.498***	0.000
	(3.849)	(3.804)		
Total factor productivity in 2015	3.185	2.821	0.364***	0.000
	(2.060)	(2.005)		
Employment in 2015 (logs)	3.766	5.405	-1.639***	0.000
	(3.054)	(2.737)		
Export intensity in 2015	0.131	0.204	-0.073***	0.000
	(0.068)	(0.076)		
Debt in 2015 (logs)	0.511	0.493	0.019**	0.006
	(0.349)	(0.358)		
RD in 2015 (dummy)	0.117	0.128	-0.012	0.063
	(0.321)	(0.334)		
Observations	6863	4460		

Means reported in columns 1 and 2, standard deviations in parenthesis.

ture, and its access to the world market, proxied by a dummy variable indicating whether there is a port within 500km of the firm's location. Roughly one third of the firms in our dataset are in the low-tech and medium-high tech industries each, and roughly one sixth each is in the low- and high-tech sectory, respectively. At 40%, the largest share of firms in our dataset is independent, followed by state-owned firms (28%) and subsidiaries (23%). Only 8% of firms are listed companies. One third of the firms has access to a port within 500km.

Table 3 gives an overview of the frequency distribution of treatment across these characteristics and provides further evidence that treatment is non-random. Thus, the share of firms receiving FDI is comparatively high in the low- and medium-low-technology sector. Among the firms receiving FDI, roughly 70% are in those sectors, compared with only 40% of the firms not receiving FDI. Access to ports is also unbalanced across treatment groups, and more prevalent among those firms receiving FDI. Regarding the ownership structure, the share of listed companies is lower among treated than non-treated firms, while the shares of independent and state-owned firms are somewhat higher.

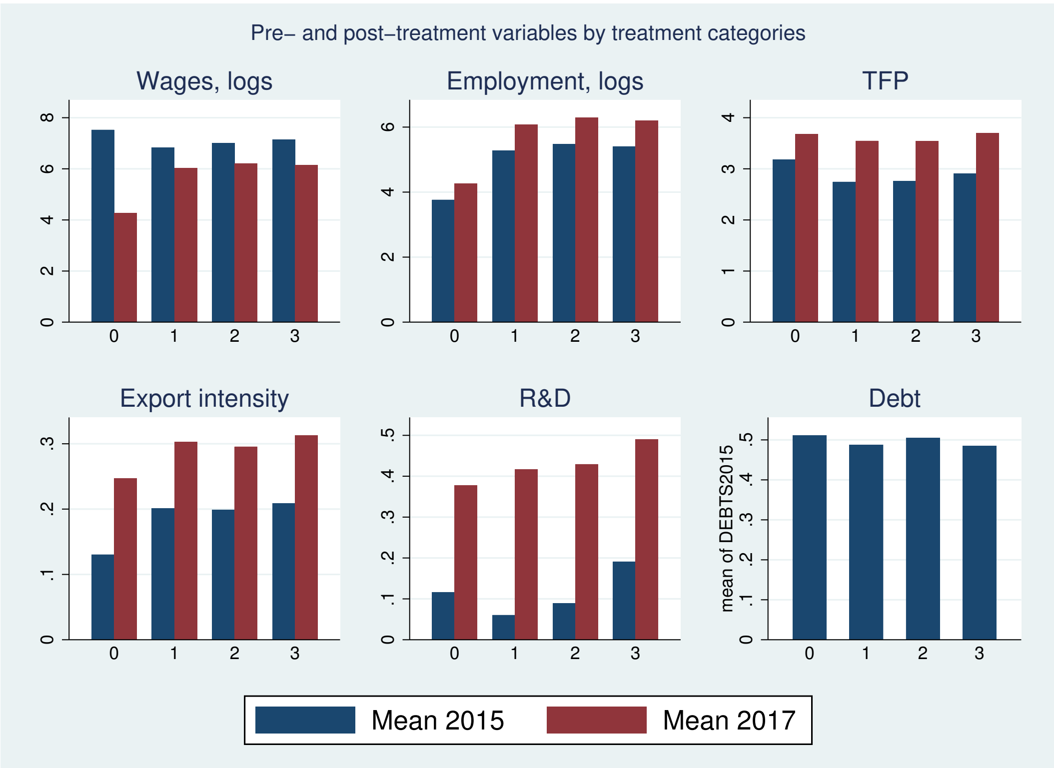
Table 3: Characteristics of treated and non-treated firms

	Treatment					
Characteristics	*No FDI*		*FDI*		*Total*	
	No.	%	No.	%	No.	%
Low-tech	1,869	27.2	2,325	52.1	4,194	37.0
Medium low	904	13.2	781	17.5	1,685	14.9
Medium high	2,432	35.4	1,107	24.8	3,539	31.3
High-tech	1,658	24.2	247	5.5	1,905	16.8
Total	6,863	100.0	4,460	100.0	11,323	100.0
No ports within 500km	4,988	72.7	2,378	53.3	7,366	65.1
Ports within 500km	1,875	27.3	2,082	46.7	3,957	34.9
Total	6,863	100.0	4,460	100.0	11,323	100.0
Listed companies	736	10.7	173	3.9	909	8.0
Subsidiaries	1,615	23.5	1,015	22.8	2,630	23.2
Independent	2,702	39.4	1,891	42.4	4,593	40.6
State	1,810	26.4	1,381	31.0	3,191	28.2
Total	6,863	100.0	4,460	100.0	11,323	100.0

Section 3

The challenge that this dataset poses for the identification of causal effects of FDI, is that the treatment (FDI) is not assigned randomly across firms but the result of a selection process. Investors will carefully decide which firms to invest in. Thus, treatment allocation will likely depend on observable firm characteristics in 2015, such that firms in the treatment group will structurally differ from firms in the untreated group. Table 2 shows that in the available dataset this imbalance in pre-treatment characteristics between treated and non-treated firms can be found across all available performance indicators from 2015. In addition, Table 4 presents the result of a logit regression of FDI on pre-treatment characteristics, which

Figure 1: Pre- and post-treatment characteristics by treatment group



corroborates the statistical significance of correlations with FDI assignment for all available pre-treatment variables except log wages. The set of pre-treatment firm characteristics is, consequently, correlated with post-treatment performance as well as with treatment allocation, which violates the conditional independence assumption necessary for causal identification. As this violation is found across multiple variables, the data suffers from what is called *the curse of dimensionality*. This can be circumvented by propensity score matching such that causal identification of FDI effects on firm performance can still be achieved.

Table 4: Logit estimation results of FDI on pre-treatment observables

	FDI in 2016	
	b	se
log wages in 2015	-0.0008	(0.008)
Total factor productivity in 2015	-0.0917***	(0.016)
log of employment in 2015	0.1905***	(0.011)
EXPORT INTENSITY in 2015	47.2347***	(0.765)
log of DEBTS in 2015	-0.3426***	(0.109)
R&D dummy in 2015=1	0.2892***	(0.097)
Medium low-tech industries	-2.1258***	(0.104)
Medium high-tech industries	-5.0274***	(0.112)
High-tech industries	-8.5163***	(0.164)
Subsidiaries	1.8301***	(0.148)
Independent	3.4707***	(0.150)
State	2.1772***	(0.151)
Ports within 500km	-3.0085***	(0.087)
Constant	-7.0707***	(0.212)
Observations	11323	

In order for propensity score matching to successfully alleviate any concerns of selection bias, a scoring function needs to be identified that achieves balancedness in terms of relevant observables, on one hand, and a good common support, i.e. overlap on the other. This is not trivial. Imagine a situation in which treatment is assigned according to one specific set of variables, z. Hence, all subject receiving treatment will exhibit high values in z. If scoring is performed based on this variable set then propensity scores will be high for firms with high values along this set of variables and matching of firms with similar propensity scores will produce very little differences in values of these variables. Yet, this will produce a most certainly bad overlap. Opposed to that using a variable that is not related to treatment assignment for the propensity score matching will produce a high overlap as propensity scores calculated based on an unrelated variable will be similarly distributed within treatment and non-treatment groups, but it will be unlikely that this matching procedure will be able to balance the previously unbalanced variables. Consequently, a scoring function specification has to be found that balances all variables relevant for treatment assignment but in a way that produces high overlap.

In a simple logit regression, almost all pre-treatment performance indicators showed relevance for selection. Figure 2 shows the overlap and Table 5 the balancedness check results

Table 5: Balance check of the propensity scoring using a logit model (linear variable inclusion) and a 1 on 1 matching procedure

	Diff:Raw	Diff:Match	Ratio:Raw	Ratio:Match
log wages in 2015	-.1300321	-.0515036	.9769191	.8845742
Total factor productivity in 2015	-.178877	.2139231	.9473458	.4003003
log of employment in 2015	.5654306	-.156559	.803081	.6229867
EXPORT INTENSITY in 2015	1.014184	.5934387	1.228659	1.336483
log of DEBTS in 2015	-.0529435	-.2515877	1.051101	1.021671
R&D dummy in 2015=1	.0356507	-.0076941	1.085768	.9795245
Medium low-tech industries	.1206088	-.0152956	1.263082	.9634731
Medium high-tech industries	-.2329159	-.1939862	.8156583	.7717778
High-tech industries	-.5425507	.4301909	.2855456	1.693899
Subsidiaries	-.018354	-.088095	.9769702	.858366
Independent	.0616272	-.0365731	1.02321	.9742405
State	.1016402	-.4001546	1.100951	.7088027
Ports within 500km	.4092869	-.3412843	1.253595	.8230184

of a linear functional specification including all these variables in a logit scoring function. It can be seen that this functional specification is not suited to achieve sample balance in a 1 on 1 matching procedure and, most importantly, not providing sufficient overlap to identify an unbiased average treatment effect. Varying the functional specification in multiple ways (using probit instead of logit, interacting continuous performance indicators with categorical type determinants) does not produce any considerable improvements in overlap (compare Figure 4 and 5 in the Appendix). If we reduce the number of variables used in the scoring function to a subset of the pre-treatment observables, we find improvements in overlap as is expected based on the above presented argument (compare e.g. Figure 6). However, this comes at the cost of missing control over several variables that could be shown to significantly drive selection. In an attempt to reduce dimensionality of the information used while keeping control over all variables relevant for selection we opt for translating continuous performance indicators into less fine-grained categorical variables. In particular, we construct a variable "EXP2015_CAT" that assigns the values 0, 1 or 2 for firms that have an export intensity of 0, below 25% or larger 25% respectively. Using a scoring function in which the remaining continuous variables are interacted with the binary/ categorical variables including "EXP2015_CAT" achieves strong improvements in overlap compared to the other specifications based on the full set of variables (see Figure 3). In addition, we find acceptable balancedness levels in the 1 on 1 matching, with matched differences of close to or above 0.2 in only two cases (compare Table 2). Testing various alternative matching specifications such as matching with the two or four nearest neighbours or using a caliper (compare Tables 12 and 13 in the Appendix) shows that the best results are indeed achieved in a 1 on 1 matching.

Though we have managed to achieve a good balance for our treated and control groups and are able to improve overlap with this specification, it remains evident from Figure 3 that the achieved overlap remains partial. Given the vast space of possible functional specifications that could be implemented, it is difficult to say whether another specification might not perform better and to argue whether our propensity scoring function has indeed been correctly

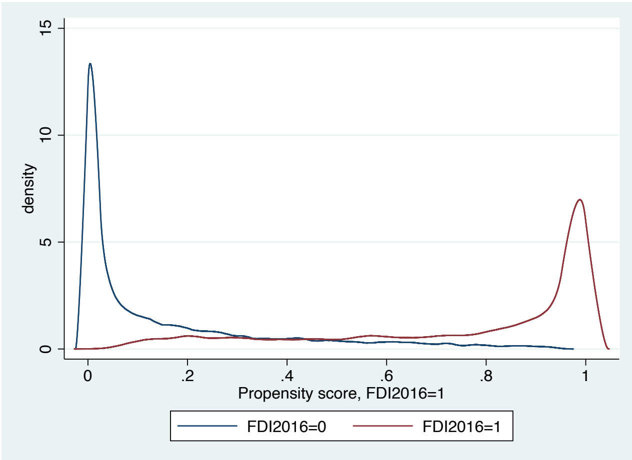


Figure 2: Overlap resulting from the propensity scoring using a logit model (linear variable inclusion).

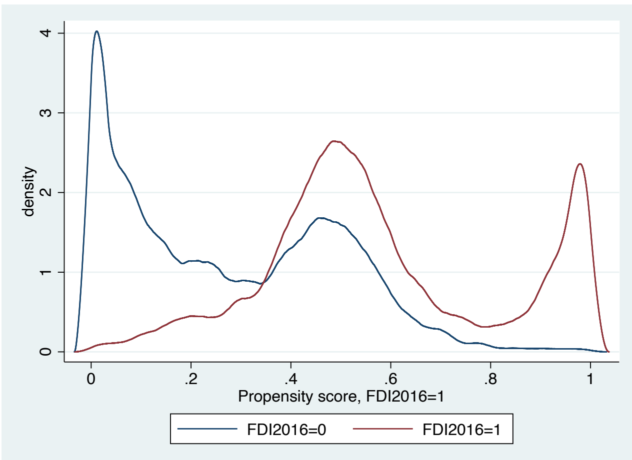


Figure 3: Overlap resulting from the propensity scoring using a logit model (interactions between the set of continuous variables excluding exports and the categorical/binary variables including exports)

	Diff:Raw	Diff:Match	Ratio:Raw	Ratio:Match
0b.RD2015xc.logwages2015	-.1195646	.0525046	.947075	1.062749
1.RD2015xc.logwages2015	.0055913	.0576109	.9912599	1.274914
1b.TECHxc.logwages2015	.4099088	.0288036	1.446385	1.012937
2.TECHxc.logwages2015	.0985765	.0376167	1.221177	1.078534
3.TECHxc.logwages2015	-.1947846	-.0195882	.7998561	1.077204
1b.OWNxc.logwages2015	-.2416651	-.0406993	.3646025	.8427318
2.OWNxc.logwages2015	-.0501523	-.078579	.8787442	.8758421
3.OWNxc.logwages2015	.0095374	.0352781	.9615021	1.109249
0b.PORTxc.logwages2015	-.3780525	.1273289	.9245549	1.141669
1b.EXP2015_CATxc.logwages2015	-.4723787	.0947225	1.179189	1.081208
0b.RD2015xc.TFP2015	-.1739665	-.0203347	.9100893	.9091296
1.RD2015xc.TFP2015	.0080044	.0443347	.9791256	1.135322
1b.TECHxc.TFP2015	.3830039	.0387406	1.517568	1.02464
2.TECHxc.TFP2015	.0592069	.0218559	1.09476	1.039168
3.TECHxc.TFP2015	-.2626395	-.0332445	.6142341	.9655526
1b.OWNxc.TFP2015	-.2670312	-.0571476	.2665297	.720536
2.OWNxc.TFP2015	-.064156	-.0762223	.8276227	.8391969
3.OWNxc.TFP2015	-.0408866	.0100326	.8831729	.9812749
0b.PORTxc.TFP2015	-.4416219	.0380993	.7259598	.88749
1b.EXP2015_CATxc.TFP2015	-.4566127	.0065822	.9833249	.8881018
0b.RD2015xc.logemp2015	.4513985	.1104609	1.015839	.7723211
1.RD2015xc.logemp2015	.1258157	.079598	1.551717	1.236144
1b.TECHxc.logemp2015	.4601552	.050231	1.534047	1.002554
2.TECHxc.logemp2015	.2274026	.0593328	1.926955	1.071029
3.TECHxc.logemp2015	.0899055	-.0276154	1.370232	.8064564
1b.OWNxc.logemp2015	-.0820011	.021356	.9009638	1.11227
2.OWNxc.logemp2015	.1399032	-.0275592	1.482211	.9386443
3.OWNxc.logemp2015	.2656301	.0867379	1.407778	1.052158
0b.PORTxc.logemp2015	.1339363	.2348676	1.28113	1.007255
1b.EXP2015_CATxc.logemp2015	.0840549	.190327	1.142787	.8433407
0b.RD2015xc.DEBTS2015	-.0687846	.0401788	1.018707	.9555409
1.RD2015xc.DEBTS2015	.0328123	.0331289	1.167688	1.141984
1b.TECHxc.DEBTS2015	.3620529	.0212764	1.493647	1.006909
2.TECHxc.DEBTS2015	.0875624	.0375397	1.216558	1.07361
3.TECHxc.DEBTS2015	-.1987245	-.0448235	.7404538	.9192316
1b.OWNxc.DEBTS2015	-.2451112	-.0710123	.3194455	.6600723
2.OWNxc.DEBTS2015	-.0444712	-.0653756	.8861299	.9017093
3.OWNxc.DEBTS2015	-.0148901	.0526032	.9654587	1.074223
0b.PORTxc.DEBTS2015	-.3147821	.125639	.9126556	1.089006
1b.EXP2015_CATxc.DEBTS2015	-.3607874	.0926031	1.078069	1.015868

Table 6: Balance check of the propensity scoring using a logit model (interactions between the set of continuous variables excluding exports and the set of categorical/binary variables including exports) and a 1 on 1 matching procedure

specified. One way to ease the uncertainty of which specification to use, is to compare results across a range of robust estimators. Hence, we calculate the regression adjustment estimator (RA), the inverse probability weighting estimator (IPW), the Augmented IPW (AIPW) and the regression adjusted IPW (IPWRA). We specify the outcome model in the same way as we specified the treatment model, where required. If we obtain similar results for all these models, we are sure that both models, the outcome and the treatment models have been correctly specified. If we discover that one of the models does not yield results consistent with the other results, we can rely on the double-robust estimator (AIPW) to estimate treatment effects as this estimator can give us confidence in our results, regardless of whether one of the models is incorrectly specified.

Table 7 reports the treatment effects of receiving FDI in 2016 on log employment in 2017 by estimator type. The average treatment effect on log wages is 0.79 in the propensity scoring estimation using *psmatch*. We find that the RA estimator, which estimates a linear outcome model and no treatment model yields an average treatment effect of 0.3 as well, similar to the IPW estimator which estimates a weighted mean for the outcome model, and a logit regression for the treatment model. Looking further, the augmented IPW which estimates a linear maximum likelihood outcome model, and a logit treatment model also yields the same result, 0.3, as well as the IPWRA estimator (which estimates a linear outcome model and a logit treatment model. A concerning observation with the four results is that for the IPW estimator, though the results are like those of the other 3 estimators, the standard errors are surprisingly high, rendering the estimate of the ATE insignificantly different from zero. This gives us a slight indication that the treatment model, estimated by logit might indeed be incorrectly specified. In the presence of this finding, we default to the double-robust estimator which is able to give us consistent estimators in the event that one of the models is incorrectly specified. This may be a telling clue as to why we have not been able to achieve a good overlap in the data. Taking a look at the balancing across variables (compare Table 14), we find that the there is slightly less balance in variables across groups as obtained with the *psmatch* approach. This reveals the balance overlap trade-off that we face when trying to consistently estimate ATEs using observational data. For the purposes of having confidence in our estimates, we proceed with the ATE estimated by the AIPW, which identifies an impact of FDI on log employment of 0.28. This translates into an increase in employment by 31.7% from FDI one year after the investment took place. We repeat this exercise to obtain the effects of FDI on wages. Results are displayed in Table 8. Using again the AIPW estimate we find average treatment effects for log wages to be 0.23. In light of the overall negative trend in wage changes from 2015 to 2017, thus, wages of firms that did receive FDIs fell by 25.6% less than wages of firms that did not receive any FDI.

Table 7: Impact of FDI on Employment in 2017 (logs)

	(1) psmatch b/se	(2) RA b/se	(3) IPW b/se	(4) AIPW b/se	(5) IPWRA b/se
ATE	0.7867***	0.3074***	0.3275	0.2754***	0.2827***
	(0.194)	(0.008)	(0.410)	(0.020)	(0.012)
Observations	11323	11323	11323	11323	11323

* p<0.1, ** p<0.05, *** p<0.01

Table 8: Impact of FDI on Wages in 2017 (logs)

	(1) psmatch b/se	(2) RA b/se	(3) IPW b/se	(4) AIPW b/se	(5) IPWRA b/se
ATE	0.7456***	0.2260***	0.2997	0.2282***	0.2361***
	(0.181)	(0.008)	(0.400)	(0.016)	(0.009)
Observations	11323	11323	11323	11323	11323

* p<0.1, ** p<0.05, *** p<0.01

Section 4

When analyzing the impact of FDI on wages and employment at the firm level, the motive of the foreign investor might also play a role. For example, investors trying to sell their products in the destination country of their investment may be particularly concerned about their reputation in the host country, potentially leading to higher wage payments. Our dataset also provides information on the type of FDI and distinguishes across four categories: exports-oriented FDI (1), technology-intensive FDI (2), and domestic market seeking FDI (3). Of the firms receiving FDI, 44% receive domestic-market seeking FDI, follow by 35% that receive technology-intensive FDI. The share of firms receiving export-oriented FDI is 21%.

Figure 1 provides a graphical illustration of the main variables by type of FDI, displaying the means of wages, employment, TFP, export intesnity and R&D for 2015 and 2017, respectivly. In addition, the graph displays the means for firms not receiving FDI (0). The corresponding Tables can be found in the Appendix (Tables 9 and 10). While the differences in means across types of FDI are relativley small when it comes to wages, employment, export intensity and TFP both in 2015 and 2017, the differences across FDI types are large regarding R&D activities.

Section 5

In our research project, using firm-level data from 2015 to 2017, we find that FDI has positive effects on labor market outcomes. Thus, we find that receiving FDI increases both firm-level employment and wages by 31.7% and 25.6%, respectively. Interestingly, analyzing whether the motive for the foreign investment matters for these effects does not reveal any significant differences. More specifically, the impact of FDI on employment and wages, respectively, does not vary by type of FDI when we distinguish between domestic market seeking FDI, export-oriented FDI, and technology-intensive FDI.

While in principle the propensity score matching techniques we employ allow us to identify the causal effect of FDI on employment and wages, the reliability of our estimates rely on the balancedness and common support assumptions to be fulfilled. As presented in Sections 3 and 4, we tried optimizing our models to find the best balance between the two criteria. However, it might be possible to further optimize our models in this respect. For example, overlap is quite poor in some regions when performing the analysis by type of FDI. Trimming the sample could be one way to address this issue – but, of course, this comes at the cost of losing observations which is why we opted against it in our analysis.

In addition, our analysis is limited by data availability. There may be other factors driving FDI that, at the same time, also affect firm-level performance. For example, we do know the country of location of the firms in our dataset. Information on the location of the investor-origin and recipient country of FDI would enable us to control for both country-specific time-invariant characteristics by using country dummies, and time varying country variables such as macroeconomic conditions. Using growth variables relative to pre-treatment levels is another strategy to at least eliminate unobserved time-invariant confounders. We leave these optimizations to future ASP students.

References

Bajgar, M. and B. Javorcik (2020). Climbing the rungs of the quality ladder: Fdi and domestic exporters in romania. *The Economic Journal 130*(628), 937–955.

Borin, A. and M. Mancini (2016). Foreign direct investment and firm performance: an empirical analysis of italian firms. *Review of World Economics 152*(4), 705–732.

Chen, W. (2011). The effect of investor origin on firm performance: Domestic and foreign direct investment in the united states. *Journal of international economics 83*(2), 219–228.

Khandelwal, A. K., P. K. Schott, and S.-J. Wei (2013). Trade liberalization and embedded institutional reform: Evidence from chinese exporters. *American Economic Review 103*(6), 2169–95.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.

Appendix

Table 9: Summary statistics for pre-treatment variables by treatment categories

	0 No FDI Mean/SD	1 Exports-oriented Mean/SD	2 Techn. intensive Mean/SD	3 Dom. market Mean/SD
Wages in 2015 (logs)	7.529	6.833	7.006	7.146
	(3.849)	(3.741)	(3.814)	(3.824)
Total factor productivity in 2015	3.185	2.747	2.759	2.905
	(2.060)	(2.030)	(2.006)	(1.991)
Employment in 2015 (logs)	3.766	5.274	5.479	5.409
	(3.054)	(2.759)	(2.706)	(2.749)
Export intensity in 2015	0.131	0.201	0.199	0.209
	(0.068)	(0.073)	(0.076)	(0.077)
Debt in 2015 (logs)	0.511	0.488	0.505	0.485
	(0.349)	(0.353)	(0.367)	(0.352)
RD in 2015 (dummy)	0.117	0.061	0.089	0.191
	(0.321)	(0.239)	(0.285)	(0.393)
Observations	6863	940	1555	1965

Table 10: Summary statistics for post-treatment variables by treatment categories

	1	2	3	4
	Mean/SD	Mean/SD	Mean/SD	Mean/SD
Wages in 2017 (logs)	4.271	6.035	6.216	6.147
	(3.063)	(2.766)	(2.712)	(2.768)
Total factor productivity in 2017	3.683	3.548	3.543	3.703
	(2.073)	(2.053)	(2.033)	(2.014)
Emplyoment in 2017 (logs)	4.265	6.078	6.289	6.206
	(3.066)	(2.769)	(2.719)	(2.764)
Export intensity in 2017	0.247	0.303	0.296	0.313
	(0.086)	(0.125)	(0.126)	(0.130)
RD in 2017 (dummy)	0.377	0.417	0.430	0.491
	(0.485)	(0.493)	(0.495)	(0.500)
Observations	6863	940	1555	1965

Table 11: Characteristics of firms by FDI type received

	Treatment by FDI type									
Characteristics	*No FDI*		*Exports FDI*		*Tech. FDI*		*Domestic FDI*		*Total*	
	No.	%	No.	%	No.	%	No.	%	No.	%
Low-tech	1,869	27.2	530	56.4	805	51.8	990	50.4	4,194	37.0
Medium low	904	13.2	150	16.0	290	18.6	341	17.4	1,685	14.9
Medium high	2,432	35.4	220	23.4	381	24.5	506	25.8	3,539	31.3
High-tech	1,658	24.2	40	4.3	79	5.1	128	6.5	1,905	16.8
Total	6,863	100.0	940	100.0	1,555	100.0	1,965	100.0	11,323	100.0
No ports within 500km	4,988	72.7	479	51.0	871	56.0	1,028	52.3	7,366	65.1
Ports within 500km	1,875	27.3	461	49.0	684	44.0	937	47.7	3,957	34.9
Total	6,863	100.0	940	100.0	1,555	100.0	1,965	100.0	11,323	100.0
Listed companies	736	10.7	30	3.2	63	4.1	80	4.1	909	8.0
Subsidiaries	1,615	23.5	233	24.8	351	22.6	431	21.9	2,630	23.2
Independent	2,702	39.4	403	42.9	636	40.9	852	43.4	4,593	40.6
State	1,810	26.4	274	29.1	505	32.5	602	30.6	3,191	28.2
Total	6,863	100.0	940	100.0	1,555	100.0	1,965	100.0	11,323	100.0

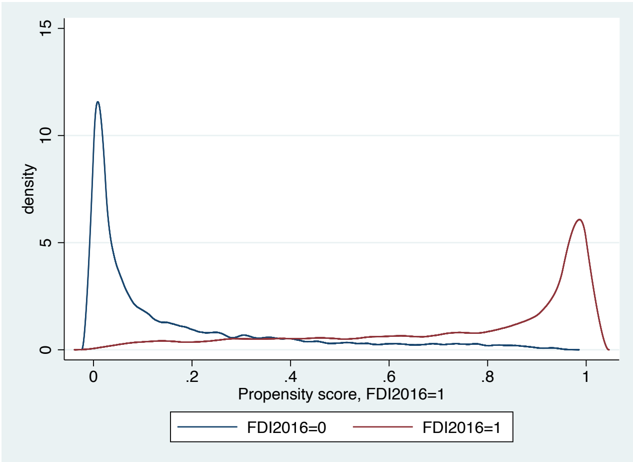


Figure 4: Overlap resulting from the propensity scoring using a logit model (interactions between the set of continuous variables and the set of categorical/binary variables)

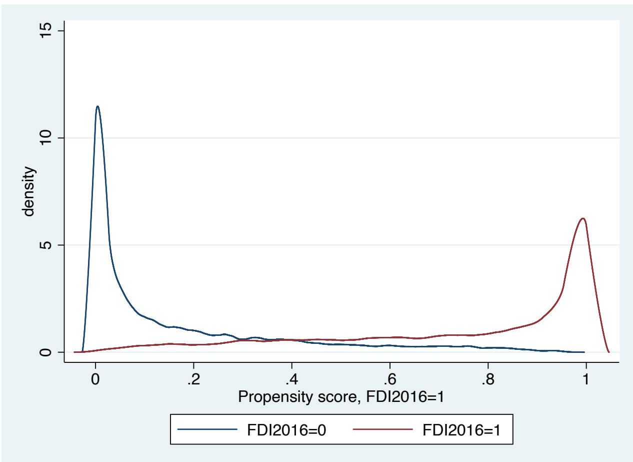


Figure 5: Overlap resulting from the propensity scoring using a probit model (interactions between the set of continuous variables and the set of categorical/binary variables)

	Diff:Raw	Diff:Match	Ratio:Raw	Ratio:Match
0b.RD2015xc.logwages2015	-.1195646	.0533576	.947075	1.083427
1.RD2015xc.logwages2015	.0055913	.0622468	.9912599	1.316477
1b.TECHxc.logwages2015	.4099088	.0368431	1.446385	1.028314
2.TECHxc.logwages2015	.0985765	.0250057	1.221177	1.058754
3.TECHxc.logwages2015	-.1947846	-.0206775	.7998561	1.052487
1b.OWNxc.logwages2015	-.2416651	-.0436965	.3646025	.827438
2.OWNxc.logwages2015	-.0501523	-.073387	.8787442	.8821358
3.OWNxc.logwages2015	.0095374	.0913069	.9615021	1.220857
0b.PORTxc.logwages2015	-.3780525	.1312182	.9245549	1.171242
1b.EXP2015_CATxc.logwages2015	-.4723787	.1005806	1.179189	1.120247
0b.RD2015xc.TFP2015	-.1739665	.0116561	.9100893	1.012206
1.RD2015xc.TFP2015	.0080044	.0449328	.9791256	1.146995
1b.TECHxc.TFP2015	.3830039	.0380806	1.517568	1.029373
2.TECHxc.TFP2015	.0592069	.0059181	1.09476	.9937183
3.TECHxc.TFP2015	-.2626395	-.0239952	.6142341	1.017338
1b.OWNxc.TFP2015	-.2670312	-.0545362	.2665297	.7563862
2.OWNxc.TFP2015	-.064156	-.0591638	.8276227	.8711978
3.OWNxc.TFP2015	-.0408866	.0506196	.8831729	1.080465
0b.PORTxc.TFP2015	-.4416219	.0664575	.7259598	.989018
1b.EXP2015_CATxc.TFP2015	-.4566127	.0384141	.9833249	.9922028
0b.RD2015xc.logemp2015	.4513985	.1458985	1.015839	.8251611
1.RD2015xc.logemp2015	.1258157	.0770499	1.551717	1.230311
1b.TECHxc.logemp2015	.4601552	.0557388	1.534047	1.003494
2.TECHxc.logemp2015	.2274026	.0438903	1.926955	1.024059
3.TECHxc.logemp2015	.0899055	-.0193275	1.370232	.8302818
1b.OWNxc.logemp2015	-.0820011	.0187964	.9009638	1.111048
2.OWNxc.logemp2015	.1399032	-.0080772	1.482211	1.009055
3.OWNxc.logemp2015	.2656301	.1355491	1.407778	1.151253
0b.PORTxc.logemp2015	.1339363	.2636523	1.28113	1.094793
1b.EXP2015_CATxc.logemp2015	.0840549	.2223448	1.142787	.9004376
0b.RD2015xc.DEBTS2015	-.0687846	.0529298	1.018707	.9896502
1.RD2015xc.DEBTS2015	.0328123	.0247653	1.167688	1.089759
1b.TECHxc.DEBTS2015	.3620529	.0286521	1.493647	1.017304
2.TECHxc.DEBTS2015	.0875624	.0298535	1.216558	1.069462
3.TECHxc.DEBTS2015	-.1987245	-.0287849	.7404538	.9846953
1b.OWNxc.DEBTS2015	-.2451112	-.0661205	.3194455	.6951376
2.OWNxc.DEBTS2015	-.0444712	-.0452431	.8861299	.9559771
3.OWNxc.DEBTS2015	-.0148901	.0841772	.9654587	1.139031
0b.PORTxc.DEBTS2015	-.3147821	.1120791	.9126556	1.079805
1b.EXP2015_CATxc.DEBTS2015	-.3607874	.0860133	1.078069	1.028403

Table 12: Balance check of the propensity scoring using a logit model (interactions between the set of continuous variables excluding exports and the set of categorical/binary variables including exports) and a 2 on 1 matching procedure

	Diff:Raw	Diff:Match	Ratio:Raw	Ratio:Match
0b.RD2015xc.logwages2015	-.1195646	.0512606	.947075	1.061216
1.RD2015xc.logwages2015	.0055913	.0554608	.9912599	1.266693
1b.TECHxc.logwages2015	.4099088	.0453662	1.446385	1.037856
2.TECHxc.logwages2015	.0985765	.011763	1.221177	1.008307
3.TECHxc.logwages2015	-.1947846	-.0112131	.7998561	1.071675
1b.OWNxc.logwages2015	-.2416651	-.00517	.3646025	.9595119
2.OWNxc.logwages2015	-.0501523	-.0700267	.8787442	.8936324
3.OWNxc.logwages2015	.0095374	.0653864	.9615021	1.159856
0b.PORTxc.logwages2015	-.3780525	.120801	.9245549	1.140006
1b.EXP2015_CATxc.logwages2015	-.4723787	.0957885	1.179189	1.085749
0b.RD2015xc.TFP2015	-.1739665	-.0176556	.9100893	.9616481
1.RD2015xc.TFP2015	.0080044	.0449186	.9791256	1.144561
1b.TECHxc.TFP2015	.3830039	.0394225	1.517568	1.019847
2.TECHxc.TFP2015	.0592069	.0021288	1.09476	.9735207
3.TECHxc.TFP2015	-.2626395	-.0295988	.6142341	.9941437
1b.OWNxc.TFP2015	-.2670312	-.0223185	.2665297	.8203311
2.OWNxc.TFP2015	-.064156	-.0677646	.8276227	.8347698
3.OWNxc.TFP2015	-.0408866	.0193216	.8831729	1.021586
0b.PORTxc.TFP2015	-.4416219	.041847	.7259598	.9348855
1b.EXP2015_CATxc.TFP2015	-.4566127	.0159171	.9833249	.9485514
0b.RD2015xc.logemp2015	.4513985	.1416871	1.015839	.8061887
1.RD2015xc.logemp2015	.1258157	.078879	1.551717	1.246292
1b.TECHxc.logemp2015	.4601552	.0592596	1.534047	1.00785
2.TECHxc.logemp2015	.2274026	.0425132	1.926955	1.032892
3.TECHxc.logemp2015	.0899055	-.0237793	1.370232	.8161867
1b.OWNxc.logemp2015	-.0820011	.0683805	.9009638	1.402103
2.OWNxc.logemp2015	.1399032	-.0186314	1.482211	.9648858
3.OWNxc.logemp2015	.2656301	.1108782	1.407778	1.093449
0b.PORTxc.logemp2015	.1339363	.2622673	1.28113	1.085431
1b.EXP2015_CATxc.logemp2015	.0840549	.2246056	1.142787	.8873473
0b.RD2015xc.DEBTS2015	-.0687846	.0601309	1.018707	1.00674
1.RD2015xc.DEBTS2015	.0328123	.0212035	1.167688	1.081422
1b.TECHxc.DEBTS2015	.3620529	.0308845	1.493647	1.013492
2.TECHxc.DEBTS2015	.0875624	.0220798	1.216558	1.041233
3.TECHxc.DEBTS2015	-.1987245	-.0172314	.7404538	1.008782
1b.OWNxc.DEBTS2015	-.2451112	-.0274972	.3194455	.8041278
2.OWNxc.DEBTS2015	-.0444712	-.047437	.8861299	.9560576
3.OWNxc.DEBTS2015	-.0148901	.0678248	.9654587	1.114978
0b.PORTxc.DEBTS2015	-.3147821	.1155958	.9126556	1.103188
1b.EXP2015_CATxc.DEBTS2015	-.3607874	.0898824	1.078069	1.043331

Table 13: Balance check of the propensity scoring using a logit model (interactions between the set of continuous variables excluding exports and the set of categorical/binary variables including exports) and a 4 on 1 matching procedure

	Diff:Raw	Diff:Match	Ratio:Raw	Ratio:Match
0b.RD2015xc.logwages2015	-.1195646	.0075964	.947075	1.111949
1.RD2015xc.logwages2015	.0055913	.0678727	.9912599	1.29589
1b.TECHxc.logwages2015	.4099088	.0882317	1.446385	1.06342
2.TECHxc.logwages2015	.0985765	.0323807	1.221177	1.063771
3.TECHxc.logwages2015	-.1947846	-.1157132	.7998561	.979123
1b.OWNxc.logwages2015	-.2416651	-.0474861	.3646025	.81565
2.OWNxc.logwages2015	-.0501523	-.0520133	.8787442	.9046346
3.OWNxc.logwages2015	.0095374	.0802055	.9615021	1.167214
0b.PORTxc.logwages2015	-.3780525	.1182721	.9245549	1.138085
1b.EXP2015_CATxc.logwages2015	-.4723787	.1161452	1.179189	1.043137
0b.RD2015xc.TFP2015	-.1739665	.0042368	.9100893	.9940214
1.RD2015xc.TFP2015	.0080044	.0585258	.9791256	1.20869
1b.TECHxc.TFP2015	.3830039	.0836885	1.517568	1.071151
2.TECHxc.TFP2015	.0592069	.0318663	1.09476	1.067882
3.TECHxc.TFP2015	-.2626395	-.0953681	.6142341	.9447969
1b.OWNxc.TFP2015	-.2670312	-.0664198	.2665297	.6787129
2.OWNxc.TFP2015	-.064156	-.0446857	.8276227	.8718027
3.OWNxc.TFP2015	-.0408866	.0486999	.8831729	1.057604
0b.PORTxc.TFP2015	-.4416219	.0625049	.7259598	.9615733
1b.EXP2015_CATxc.TFP2015	-.4566127	.0591635	.9833249	.9540943
0b.RD2015xc.logemp2015	.4513985	-.0328685	1.015839	.5844724
1.RD2015xc.logemp2015	.1258157	.0947427	1.551717	1.318069
1b.TECHxc.logemp2015	.4601552	.0884816	1.534047	.9960945
2.TECHxc.logemp2015	.2274026	.0507278	1.926955	1.019688
3.TECHxc.logemp2015	.0899055	-.2195556	1.370232	.4494659
1b.OWNxc.logemp2015	-.0820011	.030118	.9009638	1.240997
2.OWNxc.logemp2015	.1399032	-.003491	1.482211	.9428245
3.OWNxc.logemp2015	.2656301	.1296749	1.407778	1.105473
0b.PORTxc.logemp2015	.1339363	.2443691	1.28113	1.037298
1b.EXP2015_CATxc.logemp2015	.0840549	.2410722	1.142787	.8458924
0b.RD2015xc.DEBTS2015	-.0687846	-.0252322	1.018707	.9472902
1.RD2015xc.DEBTS2015	.0328123	.0394126	1.167688	1.141701
1b.TECHxc.DEBTS2015	.3620529	.0882979	1.493647	1.077583
2.TECHxc.DEBTS2015	.0875624	.0364184	1.216558	1.08119
3.TECHxc.DEBTS2015	-.1987245	-.1889552	.7404538	.7389029
1b.OWNxc.DEBTS2015	-.2451112	-.0669317	.3194455	.6746464
2.OWNxc.DEBTS2015	-.0444712	-.0128759	.8861299	1.01678
3.OWNxc.DEBTS2015	-.0148901	.071986	.9654587	1.111722
0b.PORTxc.DEBTS2015	-.3147821	.12112	.9126556	1.099798
1b.EXP2015_CATxc.DEBTS2015	-.3607874	.1209194	1.078069	1.019338

Table 14: Balance check of AIPW estimation (using interactions between the set of continuous variables excluding exports and the set of categorical/binary variables including exports)

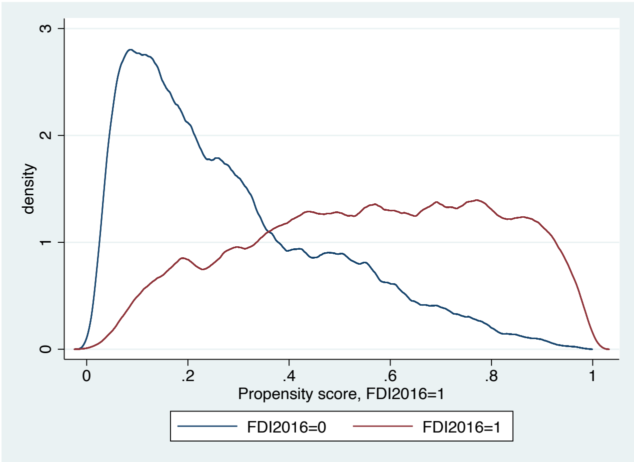


Figure 6: Overlap resulting from the propensity scoring using a logit model (linear inclusion of continuous variables only)