The Impact of Chinese Enterprises' Foreign Direct Investment on Their Total Factor Productivity—An Systematic Analysis Based on PSM-DID Model

Minghan Fu, Yiming Chen

School of Economics, Nankai University, Tianjin, 300270

Abstract: Although many evidences show that Transnational Company (TNCs) and Chinese companies have higher total factor productivity in foreign investment, no evidence to prove that foreign direct investment can promote the increase in total factor productivity of enterprises if they cannot eliminate their own selection effects. In order to eliminate the influence of the enterprise's self-selection effect, this paper uses the PSM-DID method to analyze the data of China's foreign direct investment enterprises since 2005. PSM-DID model would be used in this study to find the control group for the enterprises that started FDI in 2005, and then calculates the total factor productivity. Finally, it is concluded that Chinese enterprises' foreign direct investment (FDI) can significantly promote the increase of their total factor productivity. After getting the conclusion, the placebo test was carried out, and the test results showed that the model was relatively robust, so the conclusion was persuasive.

1. Introduction

At of the end of 2019, China's foreign direct investment stocks ranked third in the world, second only to the United States and the Netherlands. At the same time, since the implementation of the "Belt and Road" strategy, China's foreign direct investment in countries along the "Belt and Road" has also increased significantly, and has so far exceeded US$100 billion. Thus, "Bringing in and going out" has always been an important national policy of China.

2. Literature review

Most scholars believe that Chinese companies' outward direct investment has a reverse technology spillover effect, which can promote China's total factor productivity. Liu Mingxia (2009) and others used China's 2003-2007 inter-provincial panel data to examine the situation in China and found that foreign direct investment can significantly promote the increase of enterprises' total factor productivity. Shulin Liu (2011) used K-S test to study the distribution of the productivity of Chinese listed companies and found that the productivity of foreign direct investment companies is relatively high. But the above method does not really separate the self-selection effect from the two effects of becoming a multinational company.

Therefore, this article will focus on studying the effect of Chinese enterprises' outward direct investment on their total factor productivity, using the PSM-DID to analyze China's enterprises that began foreign direct investment in 2005, and measuring their total factor production, so as to further go out for Chinese enterprises to provide a reasonable reference for expanding foreign direct investment.

3. DID model

Dunning (1977) proposed the famous OIL theory in his representative work Trade Location of Economic Activities and the MNE: A Search for an Eclectic Approach. That is to say, the enterprise makes foreign direct investment because of its own ownership advantages, internalization advantages and location advantages. Therefore, in order to study the impact of FDI on TFP, it is necessary to separate its own advantages and consider it as a multinational company's role in

improving TFP, that is, the so-called "productivity" effect. The best way to study this problem is to use a double difference model.

In this paper, enterprises are divided into experimental group and control group according to whether the enterprise makes foreign direct investment. Ofdi=1 is used to represent the enterprises in the experimental group, and Ofdi=0 is used to represent the enterprises in the control group. At the same time, a variable $post_t$ is introduced in order to recognize the change of time. If $post_t$=1 indicates that the company has made foreign direct investment, $post_t$=0 indicates the year before the company made the foreign direct investment. TFP_{it} represents the total factor productivity of enterprise i in year t, meanwhile $\Delta TFP_i{}^1$ and $\Delta TFP_i{}^0$ represent the changes in total factor productivity of the enterprise before and after making foreign direct investment. Therefore, the actual effect of foreign direct investment on the enterprise's total factor productivity, after excluding its own advantages, is:

$$\gamma = \mathrm{E}(\gamma_i|post_t = 1) = E\left(\Delta TFP_i{}^1 \middle| post_t = 1\right) - E(\Delta TFP_i{}^0|post_t = 1)$$

However, since $E(\Delta TFP_i{}^0|post_t = 1)$ is unobservable, $E(\Delta TFP_i{}^0|post_t = 0)$ is used instead, and the above formula is converted into the following result:

$$\gamma = \mathrm{E}(\gamma_i|post_t = 1) = E\left(\Delta TFP_i{}^1 \middle| post_t = 1\right) - E(\Delta TFP_i{}^0|post_t = 0)$$

According to the model setting of double difference, this article uses the following equation to perform regression analysis:

$$\mathrm{TFP}_{it} = \beta_0 + \beta_1 * ofdi_i + \beta_2 * post_t + \beta_3 * owner + \beta_4 * size + \beta_5 * age + \beta_6 * fanet + \beta_7 * totalwage + \gamma * ofdi_i * post_t + \varepsilon_{it}$$

The coefficient that needs to be focused on is the coefficient γ of the model interaction term $ofdi_i * post_t$. Because the coefficient is the result of the coefficient after the double difference. If the coefficient is positive in the regression, it can indicate that the enterprise's outward direct investment behavior can promote the increase of its total factor productivity. If it is negative, it indicates that the foreign direct investment has a hindering effect on the enterprise's total factor productivity.

4. PSM model

The PSM model can help to find the appropriate control group.PSM is to select the appropriate variable first, and then use logit or Probit model to calculate the propensity score of the model. Then the corresponding control group can be obtained by matching the samples with similar scores with the corresponding matching method. In this paper, in order to ensure the number of control groups, the matching method of 1:3 K nearest neighbor is adopted.

5. Data

The source of the data in this article is divided into two parts. The specific data and various indicators of the enterprise come from the database of Chinese industrial enterprises, and the data of foreign investment come from the directory of foreign direct investment enterprises published by the Ministry of Commerce of China. First of all, this paper selects the enterprises with the approval date of 2005 from the directory of China's foreign direct investment enterprises from 1980 to 2013, and a total of 881 enterprises excluding duplicate values are obtained. However, considering that these companies may have had foreign direct investment from 1980 to 2004, the 882 companies were matched with the approved company names from 1980 to 2004, and 9 of the 2005 companies were found. Since 1980-2014, companies have already carried out foreign direct investment, so these 9 companies were eliminated, and 872 names of companies that made foreign direct investment for the first time in 2005 were obtained.

Next, in order to ensure the availability of enterprise data in each year, the 872 enterprise names

were intersected by the merge function in stata and the data from 2002-2007 in the industrial enterprise database. Enterprise names with incomplete data were removed, thus, 92 enterprises that made foreign direct investment for the first time in 2005 were obtained as the experimental group of this paper.

Next, the PSM method was used to match these 92 enterprises with all enterprises in the industrial enterprise database in 2005 according to the 1:3 k-nearest neighbor matching method, and a total of 276 enterprises were obtained. Then the 276 companies were intersected with the 2002-2007 database, and the companies with missing values in their total factor productivity were eliminated, and finally 94 companies were obtained as the control group for this study.

6. LP model

Total Factor Productivity Measurement Based on LP Method Total factor productivity is an important explained variable in the model, so the method of calculating total factor productivity is very important. Jiang Guanhong and Jiang Dianchun (2013) pointed out that the traditional method of calculating total factor productivity will have serious endogenous problems because it uses factor input as an explanatory variable, and there is a two-way relationship between factor input and total factor productivity. The Levinsohn and Pertin (2003) LP method uses the input of intermediate products as explanatory variables, which can avoid endogenous problems. The regression equation proposed by it is:

$$v_t = \alpha_l l_t + \varphi_t(k_t, m_t) + \eta_t$$

TFP is the residual value of the regression equation, that is, after removing the growth rate of capital and labor, the remaining part is the total factor productivity. Where v_t represents the industrial added value of the enterprise and l_t represents the labor employed by the enterprise, which is replaced by the annual average number of employees in the database of industrial enterprises. k_t represents the amount of capital used by the enterprise, which is expressed in the annual average balance of net fixed assets in the industrial enterprise database. m_t represents the enterprise's intermediate product input. This value can be found directly in the 2002-2007 industrial enterprise database. The subscript t indicates time. It should be pointed out that, because the data of industrial value added is missing in the industrial enterprise database of 2004, the method proposed by Liu Xiaoxuan and Li Shuangjie (2008) was used instead.

The data of industrial added value in 2004 is calculated according to this method. In the calculation, the industrial added value and intermediate input were deflated according to the industrial ex-factory price index with the base period of 2002, and the net value of fixed assets was deflated according to the fixed asset price index with the base period of 2002. The specific stata code refers to the Stata code provided in Pertin (2004) article. The regression results obtained after TFP calculation of the enterprises in the experimental group and the control group are as follows:

Table 1 LP regression coefficient

	Experiment	Control
employment	0.174*** (0.002)	0.178*** (0.009)
fanet	0.27*** (0.000)	0.21*** (0.005)

Employment represents the average annual number of employees, and fanet represents the average annual balance of net fixed assets. According to this regression result, we find that the coefficients of capital and labor are both positive and significant within 1%, so our TFP measurement results are reasonable.

7. PSM-DID model

First, the PSM method was used to match the data of the 92 experimental groups previously found with the data of the industrial enterprise database in 2005. When matching, the classic control variables of the company's business years, enterprise size, enterprise ownership, capital density, and unit wage level were first selected for matching, but the results of the matching were poor after trying various matching methods.

Therefore, all the data in 2005 were removed from the extreme values and data tailing. The specific processing method refers to the method of Xie Qianli (2008), which deletes the enterprise data that misses important financial indicators, and then deletes the enterprises with less than 10 people. At the same time, the above proportional indicators were replaced with overall indicators. Since capital density is the ratio of net fixed assets to the quantity of labor, it has been replaced with net fixed assets. The unit wage is the ratio of the total wage payment to the labor force, so it is replaced by the total wage expenditure. The size of the enterprise itself can measure the amount of labor employed by the enterprise, so I think this kind of replacement is more reasonable.

The 1:3 k-nearest neighbor matching method is used in the matching method, and the matching results are as follows:

Table 2. Result from PSM

variable	unmatched matched	mean treated	control	bias	reduct |bias|	t-test t	p>|t|
age	u	12.98	9.24	30.2		4.75	0.000
	m	12.98	13.25	-2.2	92.8	0.81	0.421
owner	u	8.3261	7.7471	23.1		9.4	0.000
	m	8.3261	8.2754	2.0	91.2	-0.51	0.61
size	u	2.3478	2.7864	-76.9		-3.35	0.001
	m	2.3478	2.3333	2.5	96.7	-0.06	0.954
fanet	u	2.8e+05	55527	18.0		5.46	0.000
	m	2.8e+05	2.9e+05	-0.8	95.8	0.58	0.562
totalwage	u	37071	7213.3	20.7		5.73	0.000
	m	37071	33213	2.7	87.1	0.62	0.533

It can be seen from the matching results that the deviation of each matching quantity is controlled within 5% after the matching. And it can be seen that, in the results of T-test, the t values before matching are significant, indicating the difference between the first two variables after matching is obvious. However, t value after matching is not significant, so it can be seen that the difference between the two variables after matching is not significant, which can be seen that the result of this match is better. Next, the k-nearest neighbor matching method of 1:3 is adopted to carry out the matching among enterprises. The number of enterprises in the experimental group was 92, and the number of enterprises in the control group after matching was 276.

8. Result from DID

The parallel trend test is the premise of DID. Only by ensuring that the trend of total factor productivity in the year before the foreign investment of the experimental group and the control group enterprises can make DID meaningful. In order to test whether the experimental group and the control group have a parallel trend, first draw a drawing analysis.
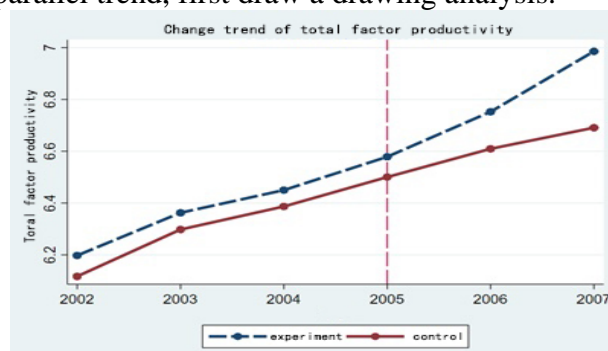
Figure 3 Parallel trend test

According to this figure, we can roughly see that the trend of total factor productivity of the experimental group and the control group before 2005 is roughly the same, while the experimental group can see its total factor productivity and control after direct foreign investment in 2005. Compared with the group, it has improved rapidly, so we can first judge from the figure that foreign direct investment has a certain role in promoting the total factor productivity of the enterprise. After drawing analysis, we use regression method to test the parallelism trend.

Table 4 Regression result

tfp	t	p>|t|
Pre_3	-4.55	0.000
Pre_2	-2.9	0.004
Pre_1	-2.01	0.045
current	-2.67	0.008

After regression, we found that the t values of the regression results of pre_3 (2002), pre_2 (2003, pre_1 (2004) and current are all within 5% level, so it can be shown from the regression results that the trend of experimental group and the control group is consistent from 2002 to 2005, that is, it meets the assumption of parallelism trend, so we can deal with the double difference.

Then we used two methods to do the DID, the first method is the basic regression model, and the second is the Individual time point bifixed effect model. The results are as follows.

Table 5 DID result

tfp	(1)	(2)
Ofdi*post	0.188*(1.74)	0.168*(1.81)
post	0.006(0.07)	0.44***(2.83)
ofdi	0.082(1.06)	-
owner	-0.057***(-5.12)	-0.0198(-0.89)
size	-0.99***(-21.95)	-0.34***(-4.39)
age	-0.007**(-2.35)	0.008(1.08)
fanet	1.64e-07***(9.21)	1.7e-07***(5.48)
totalwage	9.72e-08(0.67)	-1.24e-09(-0.01)

According to the results, we can find that the cross term ofdi*post are significant in both models, so we can think the OFDI can promote the enterprise's TFP considerably.

9. Heterogeneity analysis

Because the size of the firm and the net fixed assets value of the firm are significant within 1% in both models, this paper will use these two factors to conduct heterogeneity analysis. When conducting heterogeneity analysis, this paper did not classify enterprises according to size and net fixed assets in accordance with other articles, but introduced interactive items for analysis. Therefore, we introduce the cross-term interaction terms with enterprise size and fixed net asset value respectively. And the new function is:

$$\text{TFP}_{it} = \beta_0 + \beta_1 * ofdi_i + \beta_2 * post_t + \gamma_0 * ofdi_i * post_t + \beta_3 * owner + \beta_4 * size + \beta_5 * age + \beta_6 * fanet + \beta_7 * totalwage + \beta_8 * ofdi * post * size + \beta_9 * ofdi * post * fanet + \varepsilon_{it}$$

The regression result of the model is as follows:

Table 6 Heterogeneity model

tfp	coef	t	p>|t|
Ofdi*post	0.553**	2.5	0.013
post	0.377**	2.42	0.017
ofdi	-	-	-
owner	-0.335	-1.54	0.125
size	-0.292***	-3.45	0.001
age	0.007	1.07	0.286
fanet	1.61e-07***	4.97	0.000
totalwage	-7.91e-09	-0.06	0.956
Ofdi*post*size	-0.175*	-1.85	0.067
Ofdi*post*fanet	5.92e-08***	4.56	0.000

As can be seen from the regression results, the coefficient of ofdi* post *size is negative and significant within the level of 10%.Since the larger the value in the database of industrial enterprises, the smaller the enterprise scale is, the negative coefficient of this interaction term indicates that with the increase of enterprise scale, the more obvious the effect of OFDI on the improvement of total factor productivity of enterprises will be. The coefficient of the interaction term ofdi * post * fanet is positive and significant within 1%, which indicates that with the increase of the net fixed assets of enterprises, the more obvious the effect of OFDI on the improvement of total factor productivity of enterprises will be.

10. Robustness test

In order to test this model, we used the placebo test. We assume the beginning year is 2004 and 2003 to test its robustness, the results are as follows.

Table 7 Placebo test

tfp	(1)post=2004	(2)post=2003
Ofdi*post	0.148(1.29)	0.12(0.83)
post	0.098(1.15)	0.181(1.71)
ofdi	0.077(0.82)	0.076(0.57)
owner	-0.057(-5.62)	-0.06(-6.19)
size	-1.0(-22.05)	-1.0(-22.05)
age	-0.007(-2.38)	-0.007(-2.42)
fanet	1.65e-07(9.26)	1.65e-07(9.28)
totalwage	9.12e-08(0.63)	8.73e-08(0.61)

As can be seen from the regression results, when enterprises started ofDI after 2003 and 2004, the coefficient of the interaction term is not significant, which indicates that the improvement of total factor productivity of enterprises is the result of their ofDI rather than due to some random factors. Therefore, the model has strong robustness.

11. Conclusion

Through the empirical analysis of the ABOVE PSM-DID model and the placebo test, it can be found that Chinese enterprises' OFDI can significantly improve their total factor productivity. And the extent of improvement is positively correlated with the level of net fixed assets, scale. At the same time, through the subsequent heterogeneity analysis, it can be found that the higher the size of enterprises or the stock of net fixed assets, the more obvious the effect of their OFDI on the

improvement of total factor productivity.

References

[1] Liu Mingxia, Wang Xuejun. Research on the reverse technology spillover effect of China's foreign direct investment [J]. World Economic Research, 2009 (9): 57-62.

[2] Liu Shulin, Huang Jingbo. Foreign Direct Investment and Enterprise Productivity——An Empirical Analysis Based on Chinese Listed Companies [J]. International Economic and Trade Exploration, 2011, 27(2): 64-68.

[3] Dunning J H. Trade Location of Economic Activities and the MNE: A Search for an Eclectic Approach,[in:] B. Ohlin, PO Hesselborn, PM Wijkman (eds.) The International Allocation of Economic Activities[C]//Proceedings of a Nobel Symposium held in Stockholm. 1977.

[4] Petrin A, Poi B P, Levinsohn J. Production function estimation in Stata using inputs to control for unobservables[J]. The Stata Journal, 2004, 4(2): 113-123.

[5] Levinsohn J, Petrin A. Estimating production functions using inputs to control for unobservables[J]. The review of economic studies, 2003, 70(2): 317-341.