Impact of aggregation on measuring FDI spillovers: a Monte Carlo appraisal

Sun, Sizhong

James Cook University

May 2009

Impact of Aggregation on Measuring FDI Spillovers: A Monte Carlo Appraisal

Sizhong SUN
School of Business
James Cook University
Townsville, QLD, 4811
Australia

E-mail: sizhong.sun@jcu.edu.au
Telephone: 61-7 4781 4710
Fax: 61-7 4781 4019
Mailing Address:
Sizhong Sun
School of Business
Faculty of Law Business and Creative Arts
James Cook University
Townsville, QLD, 4811
Australia

[ABSTRACT]

Using a Monte Carlo experiment, this paper explores the impact of data aggregation on measuring the FDI spillovers. We find the aggregation significantly covers up the spillovers, which is further exacerbated by the correlation between the foreign presence and the explanatory variable at the disaggregate level. However, if the FDI at the level of aggregation is proportional to its domestic counterparts, then the aggregation does not affect the measurement of spillovers.

[KEY WORDS] Aggregation, Spillovers, FDI

[JEL CLASSIFICATION] C81, D20

Impact of Aggregation on Measuring FDI Spillovers: A Monte Carlo Appraisal

[ABSTRACT]

Using a Monte Carlo experiment, this paper explores the impact of data aggregation on measuring the FDI spillovers. We find the aggregation significantly covers up the spillovers, which is further exacerbated by the correlation between the foreign presence and the explanatory variable at the disaggregate level. However, if the FDI at the level of aggregation is proportional to its domestic counterparts, then the aggregation does not affect the measurement of spillovers.

[KEY WORDS] Aggregation, Spillovers, FDI

[JEL CLASSIFICATION] C81, D20

1. Introduction

Since the pioneering work of Caves (1974), testing the spillovers of foreign direct investment (FDI) has become an active area where researchers have tested the FDI spillovers from different aspects and in different countries[①]. Nevertheless the findings are far from consensus in that some find positive spillovers, for example Caves (1974), Barrios and Strobl (2002), Barrios et al. (2004), Dimelis and Louri (2004), Sinani and Meyer (2004), and Branstetter (2005), while others find negative or no spillovers, for example Aitken and Harrison (1999), Braconier et al. (2001), and Sadik and Bolbol

[①] See Blomstrom and Kokko (1998), Saggi (2002), Gorg and Greenaway (2004), and Smeets (2008) for a survey.

(2001). The mixed findings also motivate researchers to explore the reasons, which can be summarized from the following two aspects.

First, the nature of FDI spillovers may be indeed heterogeneous such that the empirical exercises in different contexts will inevitably obtain different results, which is summarized by Smeets (2008) from three aspects, namely the spillover channels, mediating factors, and FDI heterogeneity. Three main channels have been identified through which the spillovers can occur, that is, the forward and backward linkage between FDI invested firms and domestic firms, the movement of employees trained by FDI invested firms, and demonstration and competition effects (Blomstrom and Kokko, 1998). Mixed findings from the empirical exercises that treat the spillover mechanism as a black box may be due to their failure of separating different channels. The survey of Smeets (2008) finds that studies that separate spillover channels yield more consistent results than those that do not. Besides, even with a significant level of FDI presence, spillovers do not necessarily take place, unless certain necessary conditions (mediating factors) are satisfied. For example, domestic firms need to have a minimum level of absorptive capacity to benefit from the FDI (Cohen and Levinthal, 1989; Cohen and Levinthal, 1990; Glass and Saggi, 1998). In addition, the activities of FDI invested firms may be heterogeneous such that domestic firms are affected differently. For example, shared foreign and domestic ownership is found to present better spillovers than that of sole foreign ownership in Romania (Javorcik and Spatareanu, 2008).

Second, the methodology and data sets employed may also play a role in finding the spillovers. The commonly used method in testing FDI spillovers is to regress a measurement of the performance of domestic industries or firms, for example the labor productivity or total factor productivity, against a measurement for the presence of FDI, usually called foreign presence, and control for the other factors that have a direct effect on the domestic performance, such as the capital intensity and human capital. The significance and magnitude of the coefficient of foreign presence is the focus of the studies. Thus it is important to choose an appropriate measurement for FDI. Conventionally, there are three kinds of measurements for FDI, that is, the share of foreign owned firms' equity in the whole industry, the share of foreign owned firms' employment in the whole industry, and the share of foreign owned firms' production in the whole industry. The equity share may be distorted by host country ownership restrictions (Kohpaiboon, 2006), for example the host country may impose a restriction that the share of foreign capital in total capital can not be higher than a fixed amount in some industries. FDI tends to invest in more capital-intensive industries compared with their domestic counterparts, particularly in developing countries, and the share of foreign owned firms' employment in the whole industry will be thus lower, so that this measurement will under-represent the presence of FDI (Kohpaiboon, 2006). As for the share of output of foreign owned firms in the whole industry, it is argued that when the dependent variable is productivity, which usually is calculated from the output, it is more appropriate to measure the foreign presence by inputs (Caves, 1974). These three measurements may thus distort the true presence of FDI. These distortions may contribute to the mixed empirical results, and make the estimation results sensitive to the choice of FDI measurement (Gorg and Greenaway, 2004).

Regarding the data sets, either cross-sectional or panel data sets are used in empirical studies, and these data sets are either on a disaggregate firm level or on an aggregate industry level. Gorg and Stobl (2001) find that studies using cross-sectional data find more technology spillovers from FDI than studies using panel data, which is later disputed by Lipsey and Sjoholm (2005). Nevertheless, using the panel data set, it is easier to control for the endogeneity of FDI. In addition, the cross-sectional data sets may not be able to capture all relevant aspects of FDI spillovers due to the dynamic nature of the spillovers themselves. The level of aggregation will also play a role in measuring the FDI spillovers. As shown by Caballero and Lyons (1989), spillovers at a lower level of aggregation may be internalized at a higher level of aggregation. Thus, estimations that use firm-level (disaggregate) data will tend to present contrasting results with those of industry-level (aggregate) data. In the survey of Gorg and Greenaway (2004), 12 papers out of the 40 papers that examine the intraindustry productivity spillovers use the aggregate industry level data, of which 2 find no significant spillovers. In contrast, 29 papers use the disaggregate data, of which 10 find positive spillovers, 16 find no significant spillovers, 4 find negative spillovers, and 3 papers that deal with multiple countries find some countries have positive spillovers while other countries have negative or insignificant spillovers.

This paper intends to address the impact of aggregation on measuring the FDI spillovers, using a Monte Carlo experiment approach. We first generalize the condition on which the aggregation does not affect measuring the spillovers, and then show that, given that the spillovers occur at the disaggregate level, the aggregation significantly reduces the probability of detecting the spillovers at the aggregate level. We also investigate whether this probability depends on the correlation between the measurement of FDI and the explanatory variable and the way that the measurement of FDI is constructed. The findings of this paper not only provide an answer from the data aspect to the mixed empirical results, but also present an implication to future empirical exercises in detecting the FDI spillovers. The remainder of this paper is organized into three sections. Section two provides a conceptual illustration of the impact of aggregation. In section three, we deploy the Monte Carlo experiment and discuss the results. Section four concludes.

2. Conceptual Illustration

In studying the role of external economies in the U.S. manufacturing sector, Caballero and Lyons (1989) show that the external economies at a lower level of aggregation can be internalized at a higher level of aggregation. We reproduce their result here in the context of FDI for the sake of a conceptual illustration.

In an economy where there are *I* sectors, every sector has *J* industries, and every industry has *K* firms, the change of firm performance (*y*) is decomposed into change of the presence of FDI (λ) and change of all other factors (*x*), namely $dy_{ijk} = bdx_{ijk} + d\lambda_{ijk}$ where the subscripts *i*, *j*, and *k* denotes sector, industry, and firm respectively, and the presence of FDI is in turn determined by the performance of industry (y_{ij}), sector (y_i) and economy (y)[②], namely $d\lambda_{ijk} = \beta_1 dy_{ij} + \beta_2 dy_i + \beta_3 dy$ where the β 's are the 'externality' parameters. So at the firm level:

$$dy_{ijk} = bdx_{ijk} + \beta_1 dy_{ij} + \beta_2 dy_i + \beta_3 dy$$

② Empirical evidences (see the surveys by Blomstrom and Kokko, 1998; Saggi, 2002; Gorg and Greenaway, 2004; Smeets, 2008) have suggested that FDI tends to flow into high performance countries, sectors, and industries, which causes the endogeneity problem in these exercises.

Sum over *k*, we obtain the industry level equation:

$$dy_{ij} = \frac{b}{1-\beta_1} dx_{ij} + \frac{\beta_2}{1-\beta_1} dy_i + \frac{\beta_3}{1-\beta_1} dy$$

where *K* is normalized to 1. Higher level aggregations can be made similarly, which yields:

$$dy_i = \frac{b}{1-\beta_1-\beta_2} dx_i + \frac{\beta_3}{1-\beta_1-\beta_2} dy$$

$$dy = \frac{b}{1-\beta_1-\beta_2-\beta_3} dx$$

where *J* and *I* are normalized to 1. Hence, we see that as the level of aggregation increases the impact of FDI presence, namely the externality, decreases and eventually disappears.

In empirical exercises, researchers often regress a measurement of the performance of domestic firms against the foreign presence, controlling for other factors that have a direct effect on the performance of domestic firms. Considering this widely used empirical approach, we now suppose the foreign presence affects domestic firms in the following way:

$$y_{ijk} = a + bx_{ijk} + c\lambda_{ij} + \varepsilon_{ijk} \tag{1}$$

where the subscript *i* denotes the sector and $i = 1, \cdots, I$; the subscript *j* denotes the industry and $j = 1, \cdots, J$; the subscript *k* denotes the firm and $k = 1, \cdots, K$; *y* is the performance of the domestic firm *k* in the industry *j* of sector *i*, and in empirical exercises *y* is often the firm productivity; *x* represents the factors that directly affect *y*, for example in testing the productivity spillovers of FDI *x* often includes the labor and capital; λ_{ij} denotes the presence of FDI in the industry *j* of sector *i*, and

$\lambda_{ij} = \sum_{m \in M} z_{ijm} \Big/ \sum_{n \in N} z_{ijn}$, where z is the variable on which the foreign presence is constructed and often represents a firm's output, employees, or capital in empirical exercises, M and N denote the set of FDI invested firms and all firms in the industry j of sector i respectively and $M \subset N$; $\varepsilon_{ijk} \sim IIDN(0,1)$.

Aggregate at the industry level, we obtain:

$$y_{ij} = aK + bx_{ij} + cK\lambda_{ij} + \varepsilon_{ij} \tag{2}$$

where $y_{ij} = \sum_{k=1}^{K} y_{ijk}$, $x_{ij} = \sum_{k=1}^{K} x_{ijk}$, and $\varepsilon_{ij} = \sum_{k=1}^{K} \varepsilon_{ijk}$. Nevertheless due to the aggregation, the foreign presence at the industry level (λ_{ij}) becomes unobservable. Instead the presence of FDI is now constructed at the sector level, as follows:

$$\lambda_i = \frac{\sum_{j=1}^{J} \left(\frac{\lambda_{ij}}{1-\lambda_{ij}} z_{ij} \right)}{\sum_{j=1}^{J} z_{ij} + \sum_{j=1}^{J} \left(\frac{\lambda_{ij}}{1-\lambda_{ij}} z_{ij} \right)} = \lambda_{ij} + \varphi \tag{3}$$

where $\varphi = \frac{\sum_{m=1}^{J} \left(\frac{\lambda_{im} - \lambda_{ij}}{1-\lambda_{im}} z_{im} \right)}{\sum_{m=1}^{J} \left(\frac{1}{1-\lambda_{im}} z_{im} \right)}$. So given the true model of equation (2), the following equation is used to estimate the spillovers:

$$y_{ij} = aK + bx_{ij} + cK\lambda_i + \eta_{ij} \tag{4}$$

where $\eta_{ij} = \varepsilon_{ij} - cK\varphi$. Thus in general cases, the estimation of the coefficient cK will be biased due to the presence of φ . Nevertheless if the FDI at the industry level is proportional to its domestic counterparts in every sector, that is, $z_{ij,FDI} \big/ z_{ij} = \alpha_i$

where α_i is the proportion and can vary across sectors and $z_{ij,FDI}$ and z_{ij} denote the FDI and domestic firms in the industry *j* of sector *i* respectively, then the aggregation has no effect on measuring the spillovers, as follows:

$$\begin{aligned}\lambda_i &= \frac{\sum_{j=1}^{J} z_{ij,FDI}}{\sum_{j=1}^{J}\left(z_{ij} + z_{ij,FDI}\right)} \\ &= \frac{\alpha_i}{1+\alpha_i} \\ &= \lambda_{ij}\end{aligned}$$

3. The Monte Carlo Experiment

If the FDI is not proportional to its domestic counterparts at the level of aggregation, then whether the spillovers will be detected at the aggregate level depends on φ, which is a function of *z*. Hence a guess is that the impact of aggregation will possibly depend on the correlation between the presence of FDI at the disaggregate level and the explanatory variable *x*, and the way that the aggregate foreign presence is constructed, namely whether *z* is an exogenous variable or explanatory variable (*x*) or dependent variable (*y*). We investigate this hypothesis in the following Monte Carlo Experiment.

We look at an economy where there are 10 sectors ($I = 10$), every sector has 20 industries ($J = 20$), and every industry have 45 domestic firms ($K = 45$). The disaggregate foreign presence, λ_{ij}, is generated from a uniform distribution, namely $\lambda_{ij} \sim U(0,1)$. Note that every domestic firm in the same industry is faced with the same level of foreign presence, that is, $\lambda_{ijm} = \lambda_{ijn} \quad \forall\, m,n \in [1 \cdots K]$. The explanatory variable x_{ijk} is generated from the following equation:

$$x_{ijk} = \tilde{x} + \lambda_{ij} \frac{100\rho}{12\sigma(\lambda_{ij})\sqrt{1-\rho^2}}$$

where $\tilde{x} \sim U(0,10)$, $\sigma(\lambda_{ij})$ is the standard deviation of λ_{ij}, and ρ is the correlation between λ_{ij} and x_{ijk}, which is set to take three different values (0, 0.2, and 0.8) in the experiment. The domestic firm performance, y_{ijk}, is then generated following equation (1), where *a* =1, *b* = 0.5, and *c* is set to be 0.8 or -0.8 in different scenarios to examine the impact of aggregation when there are positive or negative spillovers at the disaggregate level respectively.

Given the data generated at the disaggregate firm level, we first regress the *y* against *x* and λ, test the significance, and check the sign of the estimated coefficient of λ. Then we aggregate them to the industry level. The aggregate foreign presence is constructed following equation (3), where *z* is equal to *x*, *y*, or an exogenous uniform distribution, depending on which scenario we examine. We again regress the *y* against *x* and λ, where the robust standard errors are used to accommodate the heteroskedasticity due to the aggregation, test the significance, and check the sign of the estimated coefficient of λ at the aggregate level. We then use two indicator variables to capture the regression outcomes. One indicator variable (*A*) takes a value of one if the regression at the disaggregate level obtains significant estimate with correct sign, and zero otherwise. The other indicator variable (*B*) takes a value of one if both regressions obtain significant estimate with correct sign, and zero otherwise. This exercise is repeated 100,000 times, which produces an empirical distribution for these two indicator variables. The conditional probability of detecting significant spillovers at the aggregate level given the significant spillovers are detected at the disaggregate level is then equal to the mean value of *B* divided by that of *A*.

Altogether we examine 18 scenarios, which are categorized from three aspects, namely whether there exist positive or negative spillovers at the disaggregate level in the data generating process, the three levels of correlation (ρ) between the foreign presence and the explanatory variable *x* at the disaggregate level, and three ways of constructing the aggregate foreign presence. Table 1 presents the probability of detecting significant spillovers at the aggregate level, conditional on being able to detect significant spillovers at the disaggregate level when there exist positive spillovers at the disaggregate level in the data generating process. Table 2 presents the conditional probability estimate when the spillovers are negative in the data generating process.

Three observations can be found from the tables. First, it is evident from both the Tables 1 and 2 that the probability of detecting the FDI spillovers at the aggregate industry level, conditional on being able to detect the spillovers at the disaggregate firm level, is rather low, no matter which scenarios we use. Hence the aggregation indeed can significantly cover up the spillovers.

The second observation is that the correlation between the disaggregate foreign presence and the explanatory variable plays an important role in detecting spillovers at the aggregate level. For example, in the scenarios where positive spillovers exist at the disaggregate level and the aggregate foreign presence is constructed over the explanatory variable *x*, the conditional probability of detecting significant spillovers at the aggregate level decreases nearly five times, from 16.18 per cent when the correlation is 0 to 3.36 per cent when the correlation is 0.2. When we continue to increase the correlation to 0.8, the conditional probability continues to decrease but at a smaller magnitude. By comparing the conditional probabilities in the scenarios with zero correlation and non-zero correlation, we can find that the correlation is detrimental to detecting significant spillovers at the aggregate level. There exists a significant decrease in the conditional probability when the correlation is increased from zero to non-zero.

Third, by comparing the scenarios where the correlation is fixed, we can find that the way we construct the aggregate foreign presence appears not to significantly affect the conditional probability of finding significant spillovers over the aggregate data. For example, when the correlation is zero and there exist positive spillovers at the disaggregate level (see Table 1), the conditional probabilities are quite similar across the three ways of constructing the aggregate foreign presence (16.18 per cent, 15.59 per cent, and 16.08 per cent respectively).

<insert Table 1 here>

<insert Table 2 here>

4. Conclusion

The findings can be summarized as four points: (1) the aggregation significantly covers up the spillovers; (2) nevertheless, if the FDI at the level of aggregation is proportional to its domestic counterparts, then the aggregation does not affect the measurement of spillovers; (3) the correlation between the foreign presence and the explanatory variable at the disaggregate level exacerbates the covering-up impact of aggregation; (4) fortunately the way we construct the foreign presence appears not to significantly affect the impact of aggregation. The implication for future empirical exercises in testing the FDI spillovers, or more broadly testing the impact of a variable where the variable is constructed from the data set, is that the disaggregate data is preferred to the aggregate data. However if the disaggregate data is not available and the aggregate data is used instead, the distribution of FDI shall be paid attention to.

Table 1 Conditional Probability of Detecting Spillovers at the Aggregate Level When Positive Spillovers Exist at the Disaggregate Level

		$\rho = 0$	$\rho = 0.2$	$\rho = 0.8$
Aggregate foreign presence constructed over:				
	x	0.1618	0.0336	0.0279
	y	0.1559	0.0328	0.0272
	z	0.1608	0.0325	0.0281

Note: *z* denotes an exogenous variable.
Source: Monte Carlo experiment output.

Table 2 Conditional Probability of Detecting Spillovers at the Aggregate Level When Negative Spillovers Exist at the Disaggregate Level

		$\rho = 0$	$\rho = 0.2$	$\rho = 0.8$
Aggregate foreign presence constructed over:				
	x	0.1621	0.0319	0.0271
	y	0.1646	0.033	0.0287
	z	0.1618	0.0316	0.0281

Note: *z* denotes an exogenous variable.
Source: Monte Carlo experiment output.

References:

Aitken, B. J. and A. E. Harrison, "Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela," *The American Economic Review* 89:3 (1999), 605-618.

Barrios, S., S. Dimelis, et al., "Efficiency Spillovers from Foreign Direct Investment in the EU Periphery: A Comparative Study of Greece, Ireland, and Spain," *Review of World Economics/Weltwirtschaftliches Archiv* 140:4 (2004), 688-705.

Barrios, S. and E. Strobl, "Foreign Direct Investment and Productivity Spillovers: Evidence from the Spanish Experience," *Review of World Economics/Weltwirtschaftliches Archiv* 138:3 (2002), 459-481.

Blomstrom, M. and A. Kokko, "Multinational corporations and spillovers," *Journal of Economic Survey* 12:2 (1998), 1-31.

Braconier, H., K. Ekholm, et al., "In Search of FDI-Transmitted R&D Spillovers: A Study Based on Swedish Data," *Review of World Economics/Weltwirtschaftliches Archiv* 137:4 (2001), 644-665.

Branstetter, L., "Is Foreign Direct Investment a Channel of Knowledge Spillovers? Evidence from Japan's FDI in the United States," *Journal of International Economics* 68:2 (2005), 325-344.

Caballero, R. J. and R. K. Lyons, "The role of external economies in U.S. manufacturing," *NBER working paper*:no.3033 (1989).

Caves, R. E., "Multinational firms, competition, and productivity in host-country markets," *Economica* 41:162 (1974), 176-193.

Cohen, W. M. and D. A. Levinthal, "Innovation and learning: The two faces of R&D," *The Economic Journal* 99:397 (1989), 1309-1329.

Cohen, W. M. and D. A. Levinthal, "Absorptive capacity: A new perspective on learning and innovation," *Administrative Science Quarterly* 35:1 (1990), 128-152.

Dimelis, S. and H. Louri, "Foreign Direct Investment and Technology Spillovers: Which Firms Really Benefit?," *Review of World Economics/Weltwirtschaftliches Archiv* 140:2 (2004), 230-253.

Glass, A. J. and K. Saggi, "International technology transfer and the technology gap," *Journal of Development Economics* 55:2 (1998), 369-398.

Gorg, H. and D. Greenaway, "Much ado about nothing? do domestic firms really benefit from foreign direct investment?," *World Bank Research Observer* 9:2 (2004), 171-197.

Gorg, H. and E. Strobl, "Multinational companies and productivity spillovers: a meta-analysis," *The Economic Journal* 111 (2001), F723-F739.

Javorcik, B. S. and M. Spatareanu, "To share or not to share: Does local participation matter for spillovers from foreign direct investment?," *Journal of Development Economics* 85:1-2 (2008), 194-217.

Kohpaiboon, A., "Foreign direct investment and technology spillover: A cross-industry analysis of Thai manufacturing," *World Development* 34:3 (2006), 541-556.

Lipsey, R. E. and F. Sjoholm, "The impact of inward FDI on host countries: Why such different answers?" in T. H. Moran, E. Graham and M. Blomstrom (Eds.), *Does Foreign Direct Investment Promote Development?* (Washington D.C.: Institute for International Economics and Center for Global Development, 2005).

Sadik, A. T. and A. A. Bolbol, "Capital flows, FDI, and Technology Spillovers: Evidence from Arab Countries," *World Development* 29:12 (2001), 2111-2125.

Saggi, K., "Trade, Foreign Direct Investment, and international technology transfer: a survey," *World Bank Research Observer* 17:2 (2002), 191-235.

Sinani, E. and K. E. Meyer, "Spillovers of Technology Transfer from FDI: the Case of Estonia," *Journal of Comparative Economics* 32 (2004), 445-466.

Smeets, R., "Collecting the pieces of the FDI knowledge spillovers puzzle," *The World Bank Research Observer* 23:2 (2008), 107-138.