Is Agglomeration a Free Lunch for New Exporters?

Evidence from Chile[†]

Youngho Kang[*]

Abstract

Unlike policy makers' belief that geographical proximity to incumbent exporters encourages non-exporters to start exporting through positive information externalities, the empirical evidence in early studies is mixed. To make a contribution to fill the gap, this paper reexamines the relationship between the spatial and industrial agglomeration of exporting firms and the probability of being an exporter, with an emphasis on the spirit of "no free lunch". According to results, the spatial and industrial agglomeration of exports positively affects the decision to start exporting. Furthermore, the squared term of agglomeration variable calculated by skilled workers is negatively and significantly associated with the decision to start exporting, which can be explained by the cost of agglomeration, *i.e.* the congestion costs (or the complexity of search) in a local labor market. As a result, these findings suggest the inverted U-shaped relationship between the agglomeration and the probability of being an exporter.

JEL Classification: F16; R12; C25; C23

Key Words: Export decision; Labor poaching; Agglomeration; Chile

[†] I am grateful to Jin-Hyuk Kim and Robert McNown for helpful comments and to Luis Castro for kindly sharing the Chilean plant-level data which comes from the Chilean National Statistics Institute (INE). I would like to thank seminar participants at Midwest International Economics Group (April 2011) and Asia Pacific Trade Seminars (June 2011). The views expressed herein are those of the author and do not necessarily reflect the views of the Samsung Economic Research Institute.

[*]Samsung Economic Research Institute, Republic of Korea, *Email address*: youngho.nomad@gmail.com

I. Introduction

Unlike policy makers' belief that geographical proximity to incumbent exporters encourages non-exporters to start exporting through positive information externalities, the following empirical evidence is divergent; that is, evidence from Colombia (Clerides et al., 1998), United Kingdom (Kneller, 2006), and France (Koenig et al., 2010) supports this belief, while evidence from Mexico (Atiken et al., 1997), Spain (Barrios et al, 2003), and the U.S. (Bernard and Jensen, 2004) fails to support it. To make a contribution to fill the gap, this paper reexamines the relationship between the spatial and industrial agglomeration of exporting firms and the probability of being an exporter, with an emphasis on the spirit of "no free lunch."

The simple linear relationship in previous studies is based only on the benefit of agglomeration, the local export spillover effect, which reduces entry costs for exporting.[1] However, the agglomeration of exporting firms cannot be the exception to the rule of "no free lunch"; that is, the agglomeration may cause costs. The possibility of agglomeration costs was already hinted at by Clerides et al. (1998), who suggest that the presence of other exporters appears to raise unit costs due to the increased prices of specialized inputs in response to rising export demand. The existence of agglomeration costs makes us turn attention to the nonlinearities because it causes the probability of being an exporter to be

[1] The entry costs for exporting are involved with gathering information on foreign consumers' preference, establishing the distribution system in foreign countries, improving infrastructure necessary to distribute the products abroad, search costs of identifying local bankers, networking, adopting the product to new standards, and so on.

dependent on the level of agglomeration.

This paper considers the potential congestion cost (or the complexity of search) in a local labor market to be the substantial cost of agglomeration. Generally, congestion costs stem from limited resources. Therefore, by supposing that heterogeneous skilled workers are distributed in a local labor market, I attempt to show how a local labor market is limited to non-exporters with the intention to export. Two facts narrow the pooling of skilled workers where non-exporting firms select to hire them at the time they decide to begin exporting. First, an exporting firm should set up a higher cut-off for worker ability than a non-exporting firm in order to be more productive due to additional costs for exporting, which causes an exporting firm to be confronted with a narrow pooling of skilled workers, as suggested in Helpman, Itskhoki, and Redding (2009). In the same vein, Filho and Muendler (2007) show interesting evidence that tariff cuts and additional imports trigger worker displacements, but that neither comparative-advantage sectors nor exports absorb trade-displaced workers. Accordingly, when the presence of local export spillovers enables the most productive among the non-exporters to start exporting, these firms will set up a higher cut-off for worker ability than under their previous status; therefore, they will select skilled workers from a narrow pooling.

Second, due to a productivity hierarchy, an exporting firm can make better offers for inside and outside workers than a non-exporting firm.[2] In fact, skilled workers who meet the ability criterion of an exporting firm can be found in non-exporting firms as well as in exporting firms. However, when non-exporting firms decide to sell abroad and so attempt to expand their capacity, they can poach skilled workers only from the other non-exporting firms, not from exporting firms. This is because exporting firms will make better offers to prevent their workers from being poached.[3] In this situation, non-exporting firms are always exposed to the threat of labor poaching by exporting firms;[4] that is, the asymmetry in labor poaching leads high-skilled workers to flow from non-exporting towards exporting firms.[5] As a result, the pooling where these non-exporting firms can select skilled workers is dependent on what fraction of local firms are non-exporters.

Consequently, the essence of non-linearity between the agglomeration and the probability of being an exporter is that as exporting firms are agglomerated spatially and industrially, the asymmetry in the labor poaching makes it more difficult for non-exporting firms to obtain high-skilled workers, *i.e.* the congestion costs (or the complexity of search) in local labor market. More specifically, when the agglomeration of exporting firms is low, the

---

[2] One of well-established empirical regularities in international trade is that firms selling abroad tend to be larger and more productive and to pay higher wages, especially to high-skilled workers (Bernard and Jensen, 1999). In particular, Kandilov (2009) shows that increased export activity in Chile can lead to a higher relative high-skilled wage.

[3] According to Fosfuri, Motta, and Ronde (2001) and Combes and Duranton (2006), pecuniary spillovers arise when the foreign affiliate pays the trained worker a higher wage to prevent her from moving to a local competitor.

[4] Davidson, Matusz, and Shevchenko (2008) show that the workers with high ability in non-exporting firms have the bigger motivation to transfer toward exporting firms due to the increased gap of wage, as economy become more open to trade.

[5] Kamal (2010) considers the productivity hierarchy of foreign-owned enterprises, private-owned enterprises, state-owned enterprises in China. In this hierarchy, he expects to see asymmetry in labor poaching effect-poaching would lead skilled workers to flow from less productive towards more productive enterprises.

local export spillover effects are more likely to dominate the congestion costs in a local labor market. On the other hand, when this agglomeration is high, the congestion costs in a local labor market are more likely to dominate the local export spillover effects. Particularly, since non-exporting firms with the intention to start exporting due to local export spillovers are likely to be on the threshold between exporters and non-exporters in terms of productivity, their decisions to export will be more vulnerable to the congestion costs in a local labor market. Subsequently, this paper constructs the testable hypothesis as follows: *Due to the congestion costs in a local labor market, the effect of agglomeration on the probability of being an exporter is non-linear*.

For the empirical approach, this paper uses the Chilean Manufacturing Census data collected by the National Statistics Institute (INE). Specifically, I use the data surveyed from 1999 to 2003 because it provides detailed information on plants' locations at three administrative level: macro-region, province and county. Agglomeration variables are calculated as the ratio of exporters' characteristics to those of total firms in the same industry and region, such as shipment, employment, and skilled workers, respectively. This paper pays more attention to the agglomeration variable calculated by skilled workers because the congestion costs in a local labor market stem from the asymmetry in poaching skilled workers. In addition, in line with early studies, this paper uses TFP (Total Factor Productivity), plant's size, export subsidies, foreign ownership, and real wage per worker as control variables.

Robert and Tybout (1999) deliver a theoretical explanation for the hysteresis in export participation generated by sunk costs. This means that a causal behavioral effect exists in the sense that the decision to export in one period itself enhances the probability of exporting in subsequent periods. To reflect it, this paper employs the dynamic binary panel model. However, the initial conditions problem in non-linear dynamic models makes it difficult to integrate out the unobservable heterogeneity in order to obtain consistent estimates. Therefore, I employ the Heckman (1981) and Wooldridge (2005) estimators.

This paper finds that the spatial and industrial agglomeration of exporters can positively affect the probability of being an exporter. However, the squared term of agglomeration variable calculated by skilled workers is also negatively and significantly associated with the decision to start exporting at the 5% level. These results imply that the relationship between the agglomeration and the probability of being an exporter is an inverted U-shape. It can shed crucial light on reconciling the mixed evidence shown in early studies; that is, above the vertex of the inverted U-shape, incumbent export activity inhibits the entry of new exporters, in line with Bernard and Jensen (2004); otherwise, the agglomeration encourages non-exporters to start exporting, as shown in Clerides et al. (1998) and Greenaway and Kneller (2006).

In the next section, I present the estimation strategy and econometric issues in

Section II. Section III contains the main results. Section IV concludes.

II. Estimation Strategy

Model Specification

To construct the empirical equations convincingly, I modify Robert and Tybout (1998) model by specifying the cost functions, especially related to agglomeration. According to them, the presence of sunk costs makes the export decision rule dynamic. So the value function of this dynamic problem can be expressed as:

$$V_{it} = \max_{Y_{it}\in\{0,1\}}(p_{it}q_{it} - c_{it}(q_{it}) - s_{it}(\alpha) - m_{it}(\alpha) - N(1-y_{it-1}) + \delta E(V_{it+1})), \qquad (1)$$

where p_{it} is the export price, q_{it} is the exported quantity, $c_{it}(q_i)$ is the production costs, $s_{it}(\alpha)$ is the congestion costs (or the complexity of search) for obtaining high-skilled workers in a local labor market, N is sunk costs for exporting, $m_{it}(\alpha)$ is entry costs which are incurred every time the firm decides to enter or reenter foreign markets, α is the ratio of exporting firms to total firms in the same region and industry, δ is a discount factor, and Y_{it} is a binary variable indicating whether a firm exports or not in period t. Thus, the solution to this problem is the decision rule as followings:

$$Y_{it} = \begin{cases} 1 & \text{if } p_{it}q_{it} + \delta[E(V_{it+1} \mid Y_{it}=1) - E(V_{it+1} \mid Y_{it}=0)] > c_{it} + s_{it}(\alpha) + m_{it}(\alpha) + N\cdot(1-Y_{it-1}) \\ 0 & \textit{otherwise} \end{cases} \qquad (2)$$

To explore what part of pooling of skilled workers is available to non-exporters, I introduce worker heterogeneity into the production function, $q_{it} = \theta\left(\sum_{i=c}^{h+c} h_i\right)^{\gamma}$. θ is the firm's productivity. A heterogeneous skilled worker (h_i) is assumed to be drawn from Pareto distribution with cumulative distribution function, $G_a(h) = 1-(h_{\min}/h)^k$ for $h \geq h_{\min} > 0$ and $k > 2$; that is, the higher the value of subscript is, the higher the ability of a worker is, and the scarcer this worker is. This distribution of skilled workers is drawn in <Figure 1>. Particularly, as suggested Helpman, Itskhoki, and Redding (2008), a local labor market has two cut-offs of skilled worker: $h^*_{non-\text{exporters}}$ and $h^*_{\text{exporters}}$, $h^*_{\text{exporters}} > h^*_{non-\text{exporters}}$. Skilled workers below $h^*_{non-\text{exporters}}$ would be unemployed. To be an exporter, non-exporters will attempt to raise the cut off up to the exporter's cut-off ($h^*_{\text{exporters}}$). This is the first factor to narrow the local pooling where the non-exporters with the intention to export select skilled workers to hire; that is, the area of A in <Figure 1> is excluded in the available pooling to potential new exporters.

Moreover, the asymmetry in labor poaching narrows this pooling more because non-exporters cannot poach the high-skilled workers from exporting firms. In deciding to export, the non-exporting firms will recognize the only area of C in <Figure 1> as the available pooling of skilled workers. This implies that as the ratio (α) is increasing, this pooling is narrower, and so the potential new exporters find it more difficult to obtain the skilled workers with ability above the cut-off (h^*_{exports}); that is, the congestion costs (or the complexity of search) is a increasing function of α, as defined by $\frac{\partial s_{it}^{nonexporter}(\alpha)}{\partial \alpha} > 0$ and $\frac{\partial^2 s_{it}^{nonexporter}(\alpha)}{\partial \alpha^2} > 0$, for $0 \le \alpha \le 1$. Additionally, due to the local export spillovers, the entry costs for exporting can be assumed to be a decreasing function of α, as defined by $\frac{\partial m_{it}^{nonexporter}(\alpha)}{\partial \alpha} < 0$ and $\frac{\partial^2 m_{it}^{nonexporter}(\alpha)}{\partial \alpha^2} > 0$, for $0 \le \alpha \le 1$. Consequently, we can know that the sum of the two costs is the U-shaped function of α. So, with other things the same, the relationship between the agglomeration of exporters (α) and the probability of being an exporter is an inverted U-shape over the value of α. This draws our attention to the equation (4) including the squared term of agglomeration index, not the equation (3) which is designed to test a simple linear relationship in early studies. More precisely, I expect the signs on β_1 and β_2 in the equation (4) to be positive and negative, respectively.

$$Y_{i,r,s,t} = 1(\theta Y_{i,t-1} + \beta_1 agg_{r,s,t-1} + \beta_3 X_{i,t} + \kappa_s + \rho_r + \delta_t + \omega_i + \eta_{i,s,t} > 0) \tag{3}$$

$$Y_{i,r,s,t} = 1(\theta Y_{i,t-1} + \beta_1 agg_{r,s,t-1} + \beta_2 agg^2_{r,s,t-1} + \beta_3 X_{i,t} + \kappa_s + \rho_r + \delta_t + \omega_i + \eta_{i,s,t} > 0) \tag{4}$$

where i is a plant index, s is an industry index, r is an region index, t is a time index, Y_{it} is the indicator variable for being the exporter, X_{it} is a vector of explanatory variables including plant-level TFP,[6] export subsidies, plant's size, real wage per worker, foreign ownership, $agg_{r,s,t-1}$ is the index of agglomeration, κ_s is industrial dummies, ρ_r is

[6] This paper employs Levinsohn and Petrin (2003) method which uses intermediate input like electricity as the instrument variable in order to address the endogeneity problem. Olley and Pakes (1996) also develop an estimator which uses investment as the instrument variable in order to avoid this issue. However, due to the substantial adjustment cost of investment and a number of zero investment firms in developing countries, this paper chooses Levinsohn and Petrin (2003) method instead of Olley and Pakes. In order to estimate the production function, I implement a STATA command (*i.e.*, levpet) introduced by Petrin, Poi, and Levinsohn (2004).

regional dummies to control for unobserved time-invariant regional effects such as access to ports, δ_t is time dummies to remove common macroeconomic shocks, ω_i is plant's level unobserved heterogeneity, and $\eta_{i,s,t}$ is the error term.

Data Description

To estimate the above empirical equation (3) and (4), I use the Chilean Manufacturing Census data collected by the National Statistics Institute (INE). This Census surveys plants with at least 10 employees and is updated annually by collecting information on plants' entry and exit during the year. Each plant has a unique identification number, which allows use of panel data analysis.[7] I use this data surveyed from 1999 to 2003 because the data in only this period have information on their locations at three administrative levels: macro-region, province and county. According to Saito and Gopinath (2009) that use the same data, the macro-region level is too aggregated and the county level has too many counties without a manufacturing plant. Consequently, among these levels, I also use 46 provinces as the geographical unit. Moreover, I exclude the cases that there is just one firm in the same region and industry because of being interested in the local export spillovers and the congestion costs, not firm's decision to locate.

The agglomeration of incumbent exporters, a variable of our interest, is calculated by

[7] However, there are some multiple plants which have the same identification number. So, I exclude these plants because the percentage of them in the survey is not significant.

plant's characteristics in the same province and SIC 3 industry such as shipment, employment, and skilled workers, respectively. Notably, to directly reflect the above discussion, I will pay more attention on the agglomeration variable calculated by skilled workers; that is, an inverted U-shaped relationship between the agglomeration and the decision to start exporting can be generated by the asymmetry in poaching high-skilled worker between exporters and non-exporters.

<Table 1> presents the descriptive statistics. The data set contains information for an unbalanced sample of 20,063 observations on a total of 4,907 plants, with 20% being exporters, and 80% non-exporters. Like earlier studies, this table shows that exporters have higher TFP, plant's (employment) size, and real wage per worker than non-exporters. Particularly, the higher real wage of exporters than that of non-exporters supports the possibility of asymmetry in labor poaching. Additionally, plants with export subsidies are 9.4% of all plants. Plants with at least 10% of foreign ownership are accounted for only 5.1% of all plants. Among these plants, the ratio of exporters is 63.2% and that of non-exporter is 36.8%. This implies that plants with high foreign ownership are more likely to be an exporter.

Econometric Issues

This paper is interested in obtaining consistent estimates of the agglomeration variables in the equation (3) and (4). For this, two problems should be solved. The first is the endogenity problem due to the common shocks. To address this problem, Angrist and Pischke (2009) suggest using the ex ante value of plant characteristics.[8] Accordingly, I use the lagged value of agglomeration in estimating the equation (3) and (4).

The second problem is how to control for unobserved heterogeneity when it correlates with the covariates.[9] Since this correlation causes inconsistency in estimates, it should be controlled for. However, unlike the linear panel model, it is so complicated in that the dynamic binary panel model does not allow some data transformation to eliminate unobserved heterogeneity. Furthermore, if we treat the unobserved heterogeneity as parameters to be estimated, the incidental parameters problem would be raised as $N \to \infty$, especially more severe in dynamic model as shown in Heckman (1981). The inconsistency in estimates of unobserved heterogeneity has a 'knock-on effect' in the sense that the other estimates become inconsistent, too.

Alternatively, we can attempt to integrate the unobserved heterogeneity out by making a distributional assumption about the density of unobserved heterogeneity for given exogenous covariates. However, in this situation, the dynamic model would raise the issue of how we treat the initial observations, *i.e.* the initial conditions problem. To handle the initial

[8] Altough Angrist and Pischke (2009, p.196) deal with peer effects in labor economics, the way to compute peer effects is the same to that of agglomeration effect. Therefore, I use their suggestion, "The best shot a causal investigation of peer effects focuses on variation in ex ante peer characteristics, that is, some measure of peer quality that predates the outcome variable and is therefore unaffected by common shocks."

[9] A fixed effects (FE) model would be preferable because it assumes that the unobserved heterogeneity is random but leaves its distribution unspecified. For the dynamic FE logit model, Honore and Kyriazidou (2000) proposed a semi-parametric estimator, which is, however, extremely data demanding. Additionally, since the dynamic FE logit model does not assume the distribution of unobserved heterogeneity, we cannot compute the marginal effects to quantify our interest (Wooldridge, 2005). So I do not use the dynamic FE logit estimator.

conditions problem relevantly, this paper considers two different estimators: the Heckman (1981) and Wooldridge (2005) estimators.

Heckman (1981) suggests approximating the conditional density of the initial observations given a vector of exogenous covariates and the unobserved heterogeneity in order to address the initial conditions problem.[10] However, Wooldridge (2005) argues the computational difficulty of obtaining both parameter estimates and estimates of averaged effects in nonlinear models. So he proposes a conditional maximum likelihood estimator that considers the distribution conditional on the initial period observations and exogenous covariates.[11] This estimator is based on Mundlak's (1978) suggestion that the regression function of unobserved heterogeneity is linear in the means of all the time varying covariates. Therefore, instead of specifying a model for the initial conditions given observed covariates and the unobserved heterogeneity, this model is specified for the unobserved heterogeneity given observed covariates and the initial conditions. As shown in Stewart's (2006) Monte Carlo investigation, the Wooldridge estimator are found to perform as well as, and in some aspects better than, the Heckman estimator.

In addition, the estimations in this paper are based on an unbalanced sample. According to Akay (2009), the results from Monte Carlo experiment show that the

---

[10] I use Stewart's (2006) STATA code which suggests the short-cut implementation of Heckman's estimator, *i.e.* rebprob. Despite Stewart's shortcut implementation, the Heckman estimator still require considerably computer time.

[11] According to Stewart (2006), this implies that Wooldridge assumes normality of the conditional distribution of unobserved heterogeneity given the initial observation of dependent variable, while Heckman assumes bivariate normality for the unobserved heterogeneity and the initial observation of dependent variable.

Wooldridge estimator works very well for the panels with moderately long duration (longer than 5-8 periods), while the Heckman estimator is suggested in short panels (shorter than 5 periods). Consequently, since none of two estimators dominates the other in all cases, this paper will present the results from both the Heckman and Wooldridge estimators.

III. Estimation Result

<Table 2-4> show the empirical results of the standard (uncorrelated) random effect estimator, the Heckman estimator, and the Woodridge estimator, respectively. Before interpreting the results, this paper needs to check whether the model is correctly specified or not. The coefficients of lagged dependent variable in <Table 2-4> are significant at the 1% level. Therefore, the dynamic specification can be considered as being correct. In the dynamic specification, it is important to test exogeneity of the initial conditions. Exogeneity of the initial conditions can be tested by a simple significance test under the null of $\theta=0$ for the Heckman estimator and $Export(0)=0$ for the Wooldridge estimator, respectively. <Table 3-4> report that θ and $Export(0)$ are significant at the 1 % level, which implies that the initial conditions should be addressed in order to integrate the unobserved heterogeneity out. Otherwise, the estimates of lagged dependent variable would be inflated. Compared with the results of the standard (uncorrelated) random-effect model in <Table 2>, we can know that the coefficients of lagged dependent variable in <Table 3> and <Table 4>

are reduced by almost half. Accordingly, the Heckman and Wooldridge estimators are preferred in this paper.[12]

The Heckman estimates in <Table 3> are broadly in line with early studies; that is, plant's level TFP, plant's size, real wage per worker, export subsidies, and foreign ownership are positively and significantly associated with the probability of being an exporter at the 10% level. One can also expect the hysteresis of export decision because the coefficient of lagged export status is positive and significant at the 1% level. Moreover, the spatial and industrial agglomeration of exporting firms affects significantly and positively the probability of being an exporter at the 10% level; that is, the local export spillover effects exist.

The aim of this paper then is to test for non-linearity between the agglomeration and the decision to start exporting. According to the results of Eq. (4) in <Table 3>, the squared terms of agglomeration variables are negatively related to the probability of being an exporter. Particularly, the only coefficient of squared term of agglomeration variable computed by skilled workers is significant at the 1% level. In other words, this agglomeration variable appears to reflect well the asymmetry in labor poaching between exporters and non-exporters. Therefore, the results imply that the impact of agglomeration on the probability of being an exporter takes an inverted U-shape.

[12] Heckman's and Wooldridge's method need the set of covariates in order to approximate the conditional density of initial observations given a vector of the unobserved heterogeneity and use Mundlak suggestion, respectively. I construct the set of covariates by using log of TFP, log of plant's size, and log of real wage per worker. However, log of real export subsidies and foreign ownership are excluded because there are a number of zero values. Furthermore, the agglomeration index variables are region-industry level, not plant level. So, if these agglomeration variables include these sets, it could be overlapping with industry and region dummies. To avoid it, this paper excludes them.

This evidence can also help us figure out how the local export spillovers occur. In fact, the sign of this squared term is determined by how the local export spillovers occur. According to Fosfuri, Motta and Ronde (2001), the knowledge spillover effect between foreign firms and local firms may take mainly two forms: “demonstration effects” and the mobility of workers.[13] This can be applied to local export spillovers. If the latter is the main channel for this spillover, the coefficient of the squared term can be expected to be positive. This is because the high ratio of exporters means that there is a high possibility for high-skilled workers to move toward non-exporting firms. On the other hand, if we consider the demonstration effect as the main channel of local export spillovers due to the asymmetry in labor poaching, the coefficient of the squared term would be negative as argue in section II. Accordingly, the results of Eq. (4) in <Table 3> inform us that export local spillovers work through the demonstration effect.

The Woodridge estimates in <Table 4> confirm the Heckman estimates. Just as in Heckman estimates, the results of Eq. (3) show that plant’s size, real wage per worker, export subsidies, and foreign ownership are positively and significantly associated with the probability of being an exporter in the 10% significance, except plant’s TFP. Column Eq. (3) also shows that the agglomeration affects positively and significantly the probability of being an exporter at the 10% level. Furthermore, in column Eq. (4), the squared term of

[13] They explain that the former is that domestic competitors might successfully imitate technological innovations introduced by MNEs, while the latter arises when subsidiaries of foreign firms train local employees who later join local firms or set up their own companies, bringing with them all (or part of) the technological, marketing, and managerial knowledge that they have acquired.

agglomeration is negatively related to the probability of being an exporter.

This paper is also interested in interpreting the coefficients quantitatively. Therefore, <Table 5> reports the marginal effects calculated from the Wooldridge estimates. As shown in the third column, this still displays a clear inverted-U shape. To check what fraction of local exporters can change the direction in the impact of agglomeration, I attempt to calculate the value of export agglomeration where the vertex of this inverted-U shape is equal to zero; that is, $0.173 - 2 \times agg\_skilled_{i,t-1} = 0$ and so $agg\_skilled_{i,t-1}$=0.57. Roughly, this value is around the 80th percentile of export agglomeration measured by skilled workers. Therefore, it can reconcile the mixed evidence. If the exporting firms occupy above 57 percent of skilled workers in the same region and industry, incumbent export activity may inhibit entry into exporting, in line with Bernard and Jensen (2004). Otherwise, the agglomeration enables non-exporters to start exporting, as shown in Clerides et al. (1998) and Greenaway and Kneller (2006).

IV. Concluding remark

This paper attempts to reexamine the impact of agglomeration on the probability of being an exporter, by emphasizing on the spirit of “no free lunch”. The spirit of “no free lunch” implies that the agglomeration of exporting firms may cause costs. I consider the asymmetry in labor poaching between exporters and non-exporters to be the main source of agglomeration costs, i.e. the congestion costs (or the complexity of search) in a local labor market. The existence of congestion costs in a local labor market suggests that the relationship between the agglomeration and the probability of being an exporter is an inverted U-shape, not linear.

The findings indicate that the spatial and industrial agglomeration of exporters can positively affect the probability of being an exporter. However, the squared term of agglomeration variable calculated by skilled workers is negatively and significantly associated with the decision to start exporting at the 5% level. These results imply the inverted U-shaped relationship between the agglomeration and the probability of being an exporter. In particular, the evidence on non-linearity shed crucial light on reconciling the mixed evidence in early studies. Additionally, from a policy point of view, it is important to identify the coefficient of squared term of agglomeration variable. This is because the negative sign of squared term of agglomeration variable recommends a policy maker to pay more attention to a local labor market, broadly local input markets, in order to make the export promotion program effective.

Reference

Akay, Alpaslan (2009), “The Wooldridge method for the initial values problem is simple: what about performance?” *IZA Discussion Paper No. 3943*.

Angrist, Joshua D. and Jorn-Steffen Pischke (2009), *Mostly Harmless Econometrics*, Princeton University Press.

Atiken, B., Gordon H. Hanson, and Ann E. Harrison (1997), “Spillovers, foreign investment, and export behavior,” *Journal of International Economics*.

Barrios S., H. Gorg and E. Strobl (2003), “Explaining Firms’ Export Behavior: R&D, Spillovers and the Destination Market,” *Oxford Bulletin of Economics and Statistics*.

Bernard B., Andrew and J. Bradford Jensen (1999), “Exceptional exporter performance: cause, effect, or both?” *Journal of International Economics*.

Bernard, Andrew B. and J. Bradford Jensen (2004), “Why some firms export,” *The Review of Economics and Statistics*.

Chaney, Thomas. (2008), “Distorted Gravity: the Intensive and Extensive Margins of International Trade,” *American Economic Review*.

Clerides, Sofronis K., Saul Lach, and James R. Tybout (1998), “Is Learning by Exporting Important? Micro-dynamic Evidence from Colombia, Mexico, and Morocco,” *Quarterly Journal of Economics*.

Combes, P. P. and G. Duranton (2006), “Labour pooling, labour poaching, and spatial clustering,” *Regional Science and Urban Economics*.

Davidson, Carl, Steven J. Matusz, and Andrei Shevchenko (2008), “Globalization and firm level adjustment with imperfect labor market,” *Journal of International Economics*.

Filho, Narcoi and Marc Muendler (2007), “Labor Reallocation in Response to Trade Reform,” *CESifo Working Paper Series #1936.*

Fosfuri, A., M. Motta, and T. Ronde (2001), “Foreign direct investment and spillovers through workers’ mobility,” Journal of International Economics.

Gorg, H. and E. Strobl (2005), "Spillovers from Foreign Firms through Worker Mobility: An Empirical Investigation," Scandinavian Journal of Economics.

Greenaway, D. and R. Kneller, (2008), "Exporting, Productivity and Agglomeration," *European Economic Review*.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding (2008), "Wages, Unemployment and Inequality with Heterogeneous Firms and Workers," working paper.

Heckman, J.J. (1981), "The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process," in C.F. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, 114-78.

Honore, Bo and Ekaterini Kyriazidou (2000), "Panel data discrete choice models with lagged dependent variables," *Econometrica*.

Kandilov, Ivan T (2009), "Do Exporters pay more? Plant-level Evidence from an Export Refund Policy in Chile," *World Bank Economic Review*.

Kamal, Fariha (2010), "FDI, Labor Poaching, and Agglomeration Economies: Evidence from China," *mimeo, Syracuse University.*

Koenig, Pamina, Florian Mayneris and Sandra Poncet (2010), "Local export spillovers in France," *European Economic Review*.

Levinsohn, J. and A. Petrin (2003), "Estimating production functions using inputs to control for unobservables," *Review of Economic Studies*.

Mundlak, Y. (1978), "On the Pooling of Time Series and Cross Section Data," *Econometrica*.

Olley, S. and A. Pakes (1996), "The dynamics of productivity in the telecommunications equipment industry," *Econometrica*.

Petrin, A., B. Poi, and J. Levinsohn (2004), "Production Function Estimation in Stata using Inputs to Control for Unobservables," *the Stata Journal*.

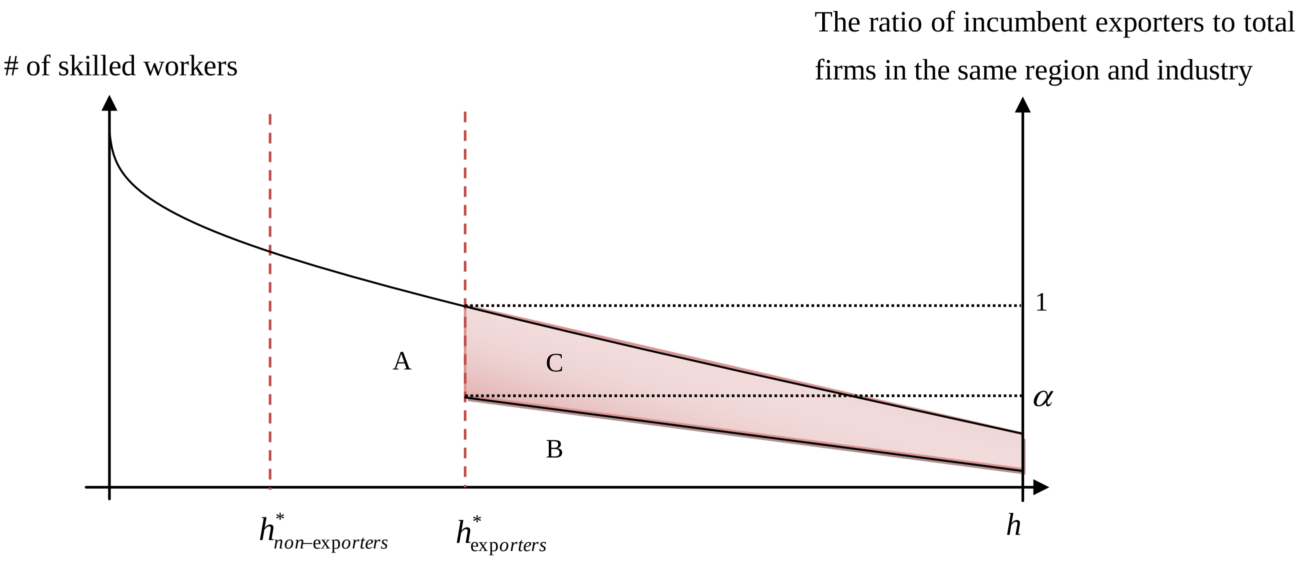
Robert, Mark and James Tybout (1997), "The Decision to Export in Colombia: An Empirical Model of Sunk Costs and the Decision to Export," *American Economic Review*.

Saito, Hisamitsu and Munisamy Gopinath (2009), "Plants' self-selection, agglomeration economies and regional productivity in Chile," *Journal of Economic Geography*.

Stewart, M.B. (2006), "-Redprob- a Stata program for the Heckman Estimator of the Random Effects Dynamic Probit Model," *mimeo, University of Warwick.*

Wooldridge, J. M. (2005), "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics*.

<Figure 1> Pooling of skilled workers



where h_i is a skilled worker with ability i; the higher the subscript (i) is, the higher the ability is; α is the ratio of exporting firms to total firms in the same region and industry, $0 \le \alpha \le 1$.

<Table 1> Statistical Summary

Variable	Description.	Average	Number	%
Number of observation			20,063	100.0
Exporter	Total number of exporters		16,057	80.03
Nonexporter	Total number of nonexporters		4,006	19.97
$\ln tfp_{i,t}$	Log of Total Factor Productivity			
Exporter		5.09		
Nonexporter		4.51		
$\ln size_{i,t}$	Log of total contract workers: plant's size			
Exporter		4.56		
Nonexporter		3.19		
$\ln rsubsidy_{i,t}$	Log of real export subsidies			
Total number of plant with subsidies			1,886	9.40
$\ln pwage_{i,t}$	Log of real wage per wage			
Exporter		8.48		
Nonexporter		7.94		
$foreign_{i,t}$	Percentage of foreign ownership			
Number of plants with above 10% foreign ownership			1,029	5.13
Exporter			650	63.2
Nonexporter			379	36.8
Agglomeration index: the ratio of exporters to total plants in the same region and industry				
$agg\_ship_{r,s,t-1}$	In terms of shipment	0.52		
$agg\_emp_{r,s,t-1}$	In terms of employment	0.43		
$agg\_skill_{r,s,t-1}$	In terms of skilled worker	0.39		

Notes: i is a plant index, s is an industry index, r is a region index, and t is a time index.

<Table 2> Standard Random Effects Model: dependent var.: 1=exporters, 0=non-exporters

	Eq. (3)	Eq. (4)	Eq. (3)	Eq. (4)	Eq. (3)	Eq. (4)
$Export_{i,t-1}$	2.271***	2.270***	2.262***	2.263***	2.273***	2.271***
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
$\ln tfp_{i,t}$	0.055**	0.055**	0.056**	0.056**	0.057**	0.057**
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
$\ln size_{i,t}$	0.279***	0.279***	0.278***	0.277***	0.278***	0.278***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
$\ln rsubsidy_{i,t}$	0.623***	0.624***	0.624***	0.626***	0.624***	0.625***
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
$\ln pwage_{i,t}$	0.261***	0.260***	0.261***	0.261***	0.261***	0.258***
	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)	(0.040)
$foreign_{i,t}$	0.280***	0.281***	0.282**	0.282***	0.278***	0.280***
	(0.094)	(0.094)	(0.095)	(0.095)	(0.094)	(0.094)
$agg\_ship_{r,s,t-1}$	0.176*	0.736**				
	(0.095)	(0.356)				
$agg\_ship^2_{r,s,t-1}$		-0.574				
		(0.350)				
$agg\_emp_{r,s,t-1}$			0.275**	0.528***		
			(0.109)	(0.202)		
$agg\_emp^2_{r,s,t-1}$				-0.260		
				(0.175)		
$agg\_skilled_{r,s,t-1}$					0.176*	0.970***
					(0.096)	(0.335)
$agg\_skilled^2_{r,s,t-1}$						-0.857**
						(0.345)
Log likelihood	-2198.67	-2197.31	-2197.17	-2196.06	-2198.70	-2195.60
Observations	14,408	14,408	14,408	14,408	14,408	14,408

Notes: Each column includes time dummies, industrial dummies, and regional dummies; Significant variables at 10%, 5%, and 1% significance level are marked with *, **, and ***, respectively; $agg\_ship_{i,t-1}$, $agg\_emp_{i,t-1}$, and $agg\_skilled_{i,t-1}$ are measured by shipment, employment, and skilled workers, respectively.

<Table 3> The Heckman Estimator: dependent variable: 1=exporters, 0 =non-exporters

	Eq. (3)	Eq. (4)	Eq. (3)	Eq. (4)	Eq. (3)	Eq. (4)
$Export_{i,t-1}$	1.595***	1.597***	1.581***	1.581***	1.592***	1.594***
	(0.082)	(0.082)	(0.084)	(0.083)	(0.083)	(0.083)
$\ln tfp_{i,t}$	0.068*	0.067*	0.068*	0.069*	0.068*	0.069**
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
$\ln size_{i,t}$	0.531***	0.529***	0.532***	0.532***	0.531***	0.530***
	(0.047)	(0.047)	(0.048)	(0.048)	(0.048)	(0.048)
$\ln rsubsidy_{i,t}$	0.707***	0.706***	0.709***	0.710***	0.709***	0.707***
	(0.047)	(0.047)	(0.048)	(0.048)	(0.048)	(0.048)
$\ln pwage_{i,t}$	0.427***	0.425***	0.428***	0.428***	0.427***	0.422***
	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)
$foreign_{i,t}$	0.430***	0.430***	0.431***	0.431***	0.429***	0.434***
	(0.123)	(0.123)	(0.123)	(0.123)	(0.123)	(0.123)
$agg\_ship_{r,s,t-1}$	0.254**	0.763				
	(0.124)	(0.465)				
$agg\_ship^2_{r,s,t-1}$		-0.514				
		(0.452)				
$agg\_emp_{r,s,t-1}$			0.376***	0.593**		
			(0.142)	(0.261)		
$agg\_emp^2_{r,s,t-1}$				-0.222		
				(0.221)		
$agg\_skilled_{r,s,t-1}$					0.267*	1.463***
					(0.141)	(0.457)
$agg\_skilled^2_{r,s,t-1}$						-1.34***
						(0.483)
ρ	0.373***	0.370***	0.377***	0.377***	0.376***	0.373***
	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)	(0.055)
θ	2.291***	2.304***	2.259***	2.254***	2.272***	2.262***
	(0.455)	(0.460)	(0.448)	(0.447)	(0.450)	(0.453)
Log likelihood	-3921.85	-3921.21	-3920.44	-3919.94	-3922.19	-3918.28
Observations	20,063	20,063	20,063	20,063	20,063	20,063

Notes: Each column includes time dummies, industrial dummies, and regional dummies; Significant variables at 10%, 5%, and 1% significance level are marked with *, **, and ***, respectively; Standard errors in parentheses; $agg\_ship_{i,t-1}$, $agg\_emp_{i,t-1}$, and $agg\_skilled_{i,t-1}$ are measured by shipment, employment, and skilled workers, respectively.

<Table 4> The Wooldridge Estimator: dependent variable: 1=exporters, 0 =non-exporters

	Eq. (3)	Eq. (4)	Eq. (3)	Eq. (4)	Eq. (3)	Eq. (4)
$Export_{i,t-1}$	1.222***	1.224***	1.210***	1.211***	1.216***	1.224***
	(0.097)	(0.097)	(0.097)	(0.0972)	(0.097)	(0.097)
$\ln tfp_{i,t}$	0.000005	-0.002	0.0004	-0.0005	0.0002	0.0005
	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)	(0.054)
$\ln size_{i,t}$	0.286**	0.285**	0.293**	0.297**	0.290**	0.288**
	(0.123)	(0.124)	(0.124)	(0.124)	(0.124)	(0.124)
$\ln rsubsidy_{i,t}$	0.761***	0.761***	0.762***	0.764***	0.763***	0.761***
	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
$\ln pwage_{i,t}$	0.053	0.052	0.054	0.054	0.052	0.048
	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)
$foreign_{i,t}$	0.297**	0.299**	0.299**	0.300**	0.297**	0.302**
	(0.139)	(0.139)	(0.139)	(0.139)	(0.139)	(0.139)
$agg\_ship_{r,s,t-1}$	0.230*	0.632				
	(0.129)	(0.476)				
$agg\_ship^2_{r,s,t-1}$		-0.407				
		(0.464)				
$agg\_emp_{r,s,t-1}$			0.369**	0.629**		
			(0.149)	(0.274)		
$agg\_emp^2_{r,s,t-1}$				-0.263		
				(0.232)		
$agg\_skilled_{r,s,t-1}$					0.298**	1.443***
					(0.148)	(0.479)
$agg\_skilled^2_{i,s,t-1}$						-1.28***
						(0.504)
$Export(0)$	1.803***	1.797***	1.803***	1.802***	1.808***	1.791***
	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)
Log likelihood	-2065.855	-2065.47	-2064.33	-2063.69	-2065.40	-2062.16
Observations	14,408	14,408	14,408	14,408	14,408	14,408
# of nui	4,907	4,907	4.907	4,907	4,907	4,907

Notes: Each column includes time dummies, industrial dummies, and regional dummies; Significant variables at 10%, 5%, and 1% significance level are marked with *, **, and ***, respectively; Standard errors in parentheses; $agg\_ship_{i,t-1}$, $agg\_emp_{i,t-1}$, and $agg\_skilled_{i,t-1}$ are measured by shipment, employment, and skilled workers, respectively.

<Table 5> The Marginal Effect calculated from Eq. (4) in <Table 4>

	Eq. (4)		
$Export_{i,t-1}$	0.247***	0.243***	0.245***
	(0.043)	(0.043)	(0.043)
$\ln tfp_{i,t}$	-0.0002	-0.00006	0.00007
	(0.007)	(0.007)	(0.006)
$\ln size_{i,t}$	0.034**	0.036**	0.034**
	(0.015)	(0.015)	(0.015)
$\ln rsubsidy_{i,t}$	0.092***	0.092***	0.091***
	(0.012)	(0.013)	(0.012)
$\ln pwage_{i,t}$	0.0063	0.0065	0.006
	(0.012)	(0.012)	(0.012)
$foreign_{i,t}$	0.036**	0.036**	0.036**
	(0.017)	(0.017)	(0.017)
$agg\_ship_{r,s,t-1}$	0.077		
	(0.058)		
$agg\_ship^2_{r,s,t-1}$	-0.049		
	(0.056)		
$agg\_emp_{r,s,t-1}$		0.076**	
		(0.033)	
$agg\_emp^2_{r,s,t-1}$		-0.032	
		(0.028)	
$agg\_skilled_{r,s,t-1}$			0.173***
			(0.058)
$agg\_skilled^2_{r,s,t-1}$			-0.153**
			(0.061)

Notes: Significant variables at 10%, 5%, and 1% significance level are marked with *, **, and ***, respectively; Standard errors in parentheses; $agg\_ship_{i,t-1}$, $agg\_emp_{i,t-1}$, and $agg\_skilled_{i,t-1}$ are measured by shipment, employment, and skilled workers, respectively.