DISCUSSION PAPER SERIES

No. 7079

AGGLOMERATION, BACKWARD AND FORWARD LINKAGES: EVIDENCE FROM SOUTH KOREAN INVESTMENT IN CHINA

Peter Debaere, Joonhyung Lee and Myungho Paik

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7079.asp**

ISSN 0265-8003

AGGLOMERATION, BACKWARD AND FORWARD LINKAGES: EVIDENCE FROM SOUTH KOREAN INVESTMENT IN CHINA

Peter Debaere, Darden Business School, University of Virginia and CEPR
Joonhyung Lee, University of Texas, Austin
Myungho Paik, University of Texas, Austin

Discussion Paper No. 7079
December 2008

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7079

December 2008

ABSTRACT

Agglomeration, Backward and Forward Linkages: Evidence from South Korean Investment in China*

With a firm-level dataset, we study the location decision of all South Korean multinationals across China's regions between 1992 and 2004, taking into account spatial aspects. Our conditional logit estimates confirm previous studies that stress the agglomeration effects along industry and along national lines in firms' location choice. In particular, South Korean investors target the place where there are more firms irrespective of their nationality and, at the same time, more affiliates from South Korean multinationals. More importantly, we decompose these agglomeration effects into a pure agglomeration effect and an upstream and downstream (backward and forward) linkage effect. We find that the presence of upstream and downstream South Korean affiliates significantly increases the likelihood that a South Korean multinational invests there. At the same time, however, backward and forward linkages with the companies irrespective of their nationality do not seem to matter. As such, our analysis of investors' location choice brings together two perspectives: (backward and forward) linkages and agglomeration along national lines.

JEL Classification: F1 and F2
Keywords: agglomeration and multinationals

Peter Debaere
Darden School of Business
University of Virginia
P.O. Box 6550
Charlottesville, VA 22906-6550
USA
Email:
DebaereP@darden.virginia.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157944

Joonhyung Lee
Department of Economics
University of Texas at Austin
1 University Station
Austin, TX 78712
USA
Email: joonlee@eco.utexas.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=164939

Myungho Paik
Center for Law, Business &
Economics
University of Texas at Austin727 E.
Dean Keeton St.
Austin, TX 3299
USA
Email: mpaik@law.utexas.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=169153

* An earlier version of the paper benefited from presentations at the Univ. of Nottingham, the Texas Econometric Camp, and the Darden Business School of University of Virginia. The usual disclaimers apply.

Submitted 26 November 2008

1 Introduction

The distribution of affiliates from South Korean multinationals across Chinese provinces is very uneven. As Figures 1 and 2 indicate, four regions out of 26 roughly cover about 75 percent of all South Korean Foreign Direct Investment (FDI) into China. As one can see, the distribution of South Korean affiliates across regions does not simply mimic that of other countries' affiliates. There is a raw correlation of merely 24 percent between South Korea's regional distribution of FDI into China and that of the rest of the world. This clustering of affiliates of multinationals along national lines is a well-known phenomenon that has also been observed for other countries. It has triggered empirical work on agglomeration, among others, by Head, Ries and Swenson (1995) that has been particularly influential. This type of work is suggestive of the impact that agglomeration externalities along national lines may have on a firm's decision to enter a particular market.

While the evidence of agglomeration is abundant, it is often not so clear what exactly is driving the agglomeration decision. Important in this respect is the work of Amiti and Javorcik (2008) and Javorcik (2004). While ignoring national linkages, both studies emphasize the role of backward/upstream and forward/downstream linkages at the industry level. To operationalize the upstream and downstream links, the authors use input-output tables that were advocated long ago by Hirschman (1958). After taking into account spatial aspects, Amiti and Javorcik (2008) finds the significant effect of forward and backward linkages in a multinational firm's location choice across Chinese regions. Firms choose to locate in a region where they can easily supply their intermediate goods to others or purchase intermediate goods from other firms. Javorcik (2004) studies the impact of inward FDI on domestic firm productivity considering forward and backward linkages between industries.

In our study of South Korean investment decisions across Chinese regions, we combine the approach of Head et al. (1995) and Amiti and Javorcik (2008). Using

input-output tables, we decompose agglomeration into a pure agglomeration effect and an upstream and downstream component. We investigate the extent to which the presence of South Korean upstream or downstream firms in an industry nearby increases the probability that a South Korean multinational will locate an affiliate in that Chinese region. At the same time, we explicitly compare and contrast this effect with the attraction of the total number of upstream and downstream companies in an industry irrespective of their nationalities. In doing so, we insert a national dimension into the analysis of Amiti and Javorcik (2008) and a linkages dimension into the work of Head et al. (1995). Interestingly, we find that linkages along national lines matter for South Korean investors. However, the linkages at the industry level irrespective of nationality do not matter. Note that the linkages at the industry level largely amount to linkages with firms from other nationalities, because of the relatively small number of South Korean investors in China's industries.

FDI has been among the fastest growing international indicators and its growth toward China and other emerging economies has been at the heart of many discussions about globalization. Increasingly, the full scale of the reorganization of production across national borders has become apparent and so are the many different ways in which this takes shape. Since the 1980s when the first formal theories of multinational activity were developed, empirical support has been found for the traditional explanations for multinational activity.[1] There is evidence that multinationals indeed relocate production to save transportation costs and to gain direct access to large markets, and that they open up affiliates in order to jump tariffs or that they move parts of their production to where resources are cheap. The work on

[1]Helpman (1984) is a key theoretical paper on vertical integration that has labor-intensive parts of production relocate to low-wage countries; Brainard (1997) and Markusen (1984) emphasize horizontal integration and distance for which exports and affiliate production are substitutes. Carr, Markusen and Maskus (2001) provides empirical evidence of horizontal multinational activity. Yeaple (2003) and Hanson, Mataloni and Slaughter (2005) are two papers that document vertical links. More detailed surveys of the literature are found in Navaretti and Venables (2004) and Markusen (2002).

geographic agglomeration that emphasizes how a multinational's location decision depends on the decisions of other multinationals addresses the sense, however, that the traditional factors do not exhaustively explain the full range of multinational activities and the allocation decisions that firms make.[2]

Firms are known to cluster in one geographic area.[3] Since Alfred Marshall, it has been pointed out that there could be different types of external economies to rationalize the geographic clustering of firms. Clustering may bring about knowledge and technology spillovers, the increasing availability of specialized labor and a growing pool of specialized input providers. Agglomeration also takes place when firms invest abroad. Foreign investors are more likely to choose locations where there are many local firms, since their presence may suggest the mentioned external economies. Interestingly, however, it has been observed that in the case of foreign direct investment, investors often agglomerate around investors from the same country of origin. Barry, Gorg and Strobl (2003) argue that investors may exhibit a tendency to imitate each other's location choice due to uncertainty. Since foreign investors face greater uncertainty than local firms in the host country, they may interpret the presence of firms from their home country as a positive signal of the location's attractiveness. It is this characteristic of foreign investment, in particular, that we investigate.

In the empirical literature, agglomeration is typically interpreted as a positive relation between a measure of the number of companies in a particular location and the probability that investors choose that location, which is why the particular path of history matters and why there is persistence in location decisions. We hypothesize that the presence of more downstream or upstream establishments makes it more likely that investors choose a particular location, which is quite intuitive. It suggests less costly access to suppliers and buyers. Moreover, the hypothesis could

[2] A recent survey by Blonigen (2005) emphasizes the need to go beyond the traditional theories of the multinational.

[3] See Porter (1998).

be consistent with earlier findings by Belderbos and Carree (2002) that smaller firms tend to follow larger firms. In particular, our hypothesis provides the dependence on inputs from other firms as the reason why this might be the case. Needless to say, building clusters has been an integral part of the strategy to attract FDI in countries such as Ireland or Costa Rica (Larrain, Lopez-Calva and Rodriguez-Clare (2001)). Note, however, that the policy implications for attracting FDI are very different when the presence of upstream or downstream establishments as such is important versus when these establishments have to be of the same nationality. If the latter holds, bilateral strategies that specifically target certain countries may have a higher payoff than multilateral approaches.

To study agglomeration and its upstream and downstream dimension, we focus on South Korean multinationals and their initial investments in China. Together with the United States, China is the highest receiver of the world's FDI. Already for this reason is it important to better understand the location decisions in China. In addition, South Korea currently sends more than half of its FDI to China, and FDI into China has become one of the premier topics of policy debates in the region. It has fueled anxieties of the "hollowing out of Korea's production base as a result of the rush into China",[4] as the South Korean investment promotion agency *KOTRA* puts it. In this respect, it is also particularly relevant for our decomposition that South Korean affiliates abroad are increasingly active in high linkage industries such as machinery, transportation, and electronics that use and produce many intermediate inputs especially after the Asian currency crisis and the increased liberalization of outward FDI in its wake.

We use a relatively new dataset of South Korean FDI in China for the empirical analysis. Different from other datasets, ours is not limited in timespan or scope, which is, of course, related to the fact that China only opened up to South Korean FDI fairly recently. The data covers all South Korean investment in China between

[4]See, Economist, August 25, 2001.

1988 and 2004. The advantage of studying the distribution of FDI within a country rather than across countries is that country-specific factors can be taken as given. In addition, we can study firm location at a less aggregate level, which is particularly relevant for agglomeration issues. The challenge of empirical research on agglomeration is then to properly control for alternative explanations such as comparative advantage or government incentives to attract foreign investors that may explain the presence of clusters of affiliates. In our preferred specification, we include region-time-specific effects. We also include a specification with region-specific effects in addition to wages, a region's market potential, a measure for regional skill quality as well as controls for China's policy initiatives to attract FDI such as the foreign trade zones that have been created with the explicit objective of attracting FDI.

Our study most closely relates to the work of Head et al. (1995) who, together with Wheeler and Mody (1992), were among the first to study agglomeration effects for FDI. In particular, Head et al. (1995) examined the location of Japanese manufacturing investment across the US states in the 1980s. They also use conditional logit estimates that are well suited for an investigation into how the variation in location (state) attributes affects the probability that a multinational will choose to set up an affiliate in a particular state. Moreover, Head et al. (1995)'s specific analysis of agglomeration externalities within vertical *keiretsu* groupings for Japanese investment in the US paves the way for our more general analysis of forward and backward linkages that probes whether the influence of linkages extends beyond national lines.[5] As a matter of fact, we also investigate if our results are sensitive to whether firms are part of larger *Chaebols*.

We first discuss the approach in the next section before we turn to the data that

[5]Head, Ries and Swenson (1999) extend their previous analysis as they explicitly control for more factors that characterize the different regions - a key concern in conditional logit analysis. Blonigen, Ellis and Fausten (2005) find agglomeration effects across both horizontal and vertical *keiretsu* groupings. However, their analysis of Japanese multinationals is across countries, rather than states.

we use in section 3. In the last two sections, we explain the results and state our conclusions.

2 Approach

Conditional logits with the particular place that is chosen by the investor as dependent variable offer a straightforward way to implement location choice models. They allow us to investigate how the characteristics of the various locations affect the likelihood of investors investing in a particular place at the time of the first investment. We follow Head et al. (1995) who builds on McFadden (1974)'s result that logit choice probabilities can be derived from individual firm maximization decisions. In particular, the place that offers the highest expected profitability is chosen as destination. When the production function of the affiliate in a particular place is assumed Cobb-Douglas, agglomeration externalities from other companies in the place, together with other production inputs, will affect a firm's output and profitability in a multiplicative way. In this case, the expected profitability of an affiliate j in place p, Π_{jp}, is a log-linear function of the agglomeration measures and other attributes of the places, which are all captured by the vector X_{jp}.(We drop the time-subscripts for simplicity.)

$$\Pi_{jp} = \beta' X_{jp} + \varepsilon_{jp} \tag{1}$$

If ϵ_{jp} is Type-I Extreme Value random error, following McFadden (1974), the probability that j invests in place p equals the following expression:

$$pr(1_j = p) = \frac{\exp(\beta' X_{jp})}{\sum_p \exp(\beta' X_{jp})} \tag{2}$$

The most common formulation of equation (1) is as follows.

$$\theta_p + \alpha \ln A_p^s + \beta \ln Z_p + \epsilon_p \tag{3}$$

where θ_p represents place-specific effects, A_p^s stands for agglomeration externalities in sector s and place p, and Z_p represents other attributes of the different places. It could be argued that the geographic borders of provinces are arbitrary. Therefore, we construct a distance-weighted agglomeration variable for each province that also includes the agglomeration variables of the other provinces weighted by their relative distance.

$$WA_p^s = A_p^s + \sum_{l \neq p}^{P} \left(A_l^s * \frac{Dist_{lp}^{-1}}{\sum_{l \neq p}^{P} Dist_{lp}^{-1}} \right) \tag{4}$$

where $Dist_{lp}$ is the distance between capital cities. Hence, WA_p^s will be higher where there are many firms nearby.

As indicated, to capture agglomeration, we break WA_p^s up in two. On the one hand, we take it to be the number of Korean affiliates within an industry. On the other hand, we measure the total number of companies in an industry irrespective of nationality (including local Chinese companies) that are already located at the time that an investment decision is made.[6]

Place-specific effects are captured by place-specific dummies that control for time-invariant factors. These factors capture the geography, the proximity to South Korea, the infrastructure, or the presence of a South Korean expatriate community that all make a place more or less attractive to investors and that may be hard to measure. In addition, we include economic and demographic variables such as a place's educa-

[6]Since the number of South Korean establishments is small compared to the total number of companies in an industry, the industry agglomeration variable corresponds roughly to the non-South Korean companies.

tion levels, and its average wage rates that vary with the time of investment. These variables are known determinants of multinational activity. Larger markets tend to attract more (horizontal) FDI. Lower wages may be attractive for (vertical) FDI that takes advantage of low production costs to relocate parts of the production process that used to take place in the South Korean parent. We control for the variation in efforts to attract foreign direct investment by including the number of economics zones in the place. Since there is an issue about whether one can appropriately capture all characteristics that vary over time and place, our preferred specification includes place-time dummies. Needless to say, if there exists no agglomeration externality and if all relevant factors that distinguish places are controlled for, the α coefficients should be zero.

As indicated, we go beyond this baseline specification for agglomeration. We decompose the agglomeration effects into pure agglomeration and respectively backward and forward linkages, in order to capture the impact of increasing numbers of upstream suppliers of intermediate goods and downstream buyers of such goods. To generate these measures of forward and backward linkages, we use industry input-output tables and combine them with the number of companies in a particular place/industry. For the linkages with South Korean companies, we use the South Korean input-output tables and combine them with the number of South Korean affiliates across the industries in the Chinese regions. For the linkages at the industry level irrespective of nationality, we use the Chinese input-output tables combined with the total number of companies in a particular region/industry that are mostly of non-South Korean nationality.[7] In each case do we capture the strength of forward linkages as follows: $FL_p^m = \sum_n \delta_{mn} A_p^n$, where δ_{mn} is the proportion of sector m output supplied to sector n and $\sum_n \delta_{mn} = 1$.[8] Again, to take into account spatial

[7]We assume that the link with Korean establishments are reflected in the Korean IO table while that with non-Korean ones in the Chinese IO table. As a robustness check, however, we use either the Korean or the Chinese IO table for both cases and the results do not depend on it.

[8]Consumer demand can be a factor in considering forward linkages. It is not straightforward

aspects, we construct distance-weighted forward linkages variables in the following way.

$$WFL_p^m = FL_p^m + \sum_n \delta_{mn} \sum_{l \neq p}^{P} \left(A_l^s * \frac{Dist_{lp}^{-1}}{\sum_{l \neq p}^{P} Dist_{lp}^{-1}} \right) \tag{5}$$

Hence, WFL will be higher in any place where many downstream firms are already located nearby. The variable for backward linkages is analogously represented by $BL_p^n = \sum_m \gamma_{mn} A_p^m$, where γ_{mn} is the proportion of sector m output supplied to sector n and $\sum_m \gamma_{mn} = 1$ and distance-weighted backward linkages, WBL, will be higher where many upstream establishments are already located nearby. As indicated, we will construct these upstream and downstream linkage variables specifically for South Korean affiliates as well as for the total number of companies in an industry irrespective of their nationality. Note that the variables differ by place/sector, since the usage of intermediate goods varies by industry. Rewriting the profitability equation (3), we obtain equation (6).

$$\theta_p + \sum_{i=\{SK,I\}} \alpha_{ia} \ln WA_{ip}^s + \sum_{i=\{SK,I\}} \alpha_{if} \ln WFL_{ip}^s + \sum_{i=\{SK,I\}} \alpha_{ib} \ln WBL_{ip}^s + \beta \ln Z_p + \epsilon_p \tag{6}$$

The new coefficients α_f and α_b should be significantly positive if an investor chooses a particular place because it has more South Korean (SK) upstream (downstream) affiliates for an industry or, more generally, because the total number of upstream (downstream) companies in this industry is higher irrespective of nationality (I).

to construct such a variable, however, because consumer demand will be expressed in monetary terms while the current linkage variable is based on the number of establishment. Therefore, we will include market potential as a control variable in some of the specifications.

3 Data and FDI from South Korea

The data of South Korean foreign affiliates is collected by the Export-Import Bank of Korea. It includes the full list of South Korean affiliates established worldwide since 1968. Relevant for our analysis are the first-time investments of multinationals in Chinese regions, which started in 1988. Figures 3 and 4 illustrate the dominant trends of South Korean outward FDI in terms of the amount invested and the number of newly established affiliates. Before the mid 1990s FDI gradually increased. Since then, there has been a massive outflow. The late 1990s were the only exception. At the time, South Korea was caught in the Asian financial crisis. As for outward FDI going into China, the data shows a significant increase in the number of affiliates established as well as in the amount of FDI since 1988. In particular, there was a dramatic increase in FDI moving into China around 1992 when Korea and China entered into diplomatic relations. Note that there may be some concern that the investments prior to 1992 were not merely a function of economic interests. Finally, as is clearly shown in Figures 3 and 4, a large percentage of foreign affiliates from South Korea are located in China. As of 2004, more than 50 percent of its new affiliates are established in China.

Figures 5 and 6 present the industry characteristics of FDI into China in terms of the number of affiliates. According to Figure 5, more than 80 percent of the affiliates over the entire period are active in manufacturing in China, which is significantly higher than the worldwide share (a bit above 60 percent). This is in line with the perception of China as the world's factory. We also categorized affiliates into high versus low linkage industries. The industries with high linkages include machinery, transportation, and electronics which use (provide) many intermediate goods from (to) other sectors.[9] The share of South Korean affiliates in China that belong to a

[9] Hanson et al. (2005) find strong vertical FDI activity in industries such as machinery, transportation, and electronics.

high or low linkage industry is presented in Figure 6. The shares add up to one. The trends are relatively stable before 1999 and show a fairly dramatic change after that. The share of high linkage industries has increased significantly while that of low linkage industries (by construction) has decreased.

China has 22 provinces (excluding Taiwan), 5 autonomous regions (Tibet, Xinjiang, Inner Mongolia, Ningxia, and Guangxi), and 4 municipalities (Beijing, Tianjin, Shanghai, and Chongqing). Figure 1 describes the regional distribution of South Korean FDI in China. From Figure 2, we can infer that the distribution of FDI/affiliates across the various regions differs significantly by country of origin. Overall, we obtain a raw correlation of only 24 percent between the regional distribution of FDI from South Korea in China and that from the rest of the world. Note that a correlation across region-industries is bound to be even lower.[10] The investments in Guangdong and Jiangsu, for example, amount to more than 40 percent of worldwide investment, while those investments account for less than 12 percent for Korea. Shandong is the premier destination for Korean investment. Interestingly, Liaoning, Jilin, and Heilongjiang seem to be attractive locations for Korean firms, but not for the other countries. This should perhaps not be so much of a surprise since many Korean-Chinese who speak Korean live in these provinces, which are adjacent to North Korea. As we will not use worldwide FDI in our conditional logits, but rather the number of companies from other nationalities in a particular region/industry irrespective of their nationality, we want to note that the correlation between the regional distribution of worldwide FDI and that of overall manufacturing output is, as one would expect, fairly high at 74 percent.

In the econometric analysis, we end up working with 19 regions. A constraint imposed by conditional logit analysis is that the different regions of China have to be chosen at least once, which is why we drop all the regions that do not report any investments. Five autonomous regions fall in this category. At the same time, we

[10] The source for worldwide multinational firms is the Chinese Statistical Yearbook 2006.

merge the data for the municipality of Chongqing with Sichuan Province, since it was separated in 1997. Finally, since our preferred specification involves region and time dummies, we exclude 6 provinces (Hainan, Shaanxi, Qinghai, Gansu, Yunnan, and Guizhou) where there are less than 10 data points.[11] With the assumption of the independence of irrelevant alternatives that is built into any conditional logit analysis, shrinking the number of choices the way we do should not affect the parameter estimates.

We measure agglomeration using the number of Korean affiliates already active in the region in the same manufacturing industry in the year before an investment takes place. The Export-Import Bank of Korea records South Korean outward FDI according to a total of 69 manufacturing industries, which is similar to Korea Standard Industry Classification (KSIC) 4- digit level. Accordingly, the agglomeration variable for South Korean affiliates is specified for 69 industries. For the total number of companies at the industry level that are overwhelmingly of non-South Korean nationality, we use the data from the China Industry Statistics Yearbook series. The data contains the total number of firms and output according to 18 manufacturing industries. In one specification, we will use both (different) classifications. In another, we will match as closely as we can the Chinese classification with the South Korean data by aggregating Korean data up to 18 (roughly) comparable industries.[12]

When we do the estimation, we will first focus on the period since 1999, before we turn to the entire period after 1992. There are two reasons for doing so. First, the period since the Asian financial crisis has seen the strongest increase in highly linked affiliates that are relevant for our analysis. Second, a data issue complicates the analysis for the entire period. The China statistics department changed the data

[11] It turns out that our conditional logit does not converge when these 6 provinces are included.

[12] The 18 industries are general industry machinery, other machinery, non-metal mineral, textile, synthetic fiber, food, grain-mill products, beverages, instruments, automobile, electronic and electrical machinery, electronic and communication components, primary metal, fabricated metal products, printing and allied products, coke and petroleum, chemical and drug.

collection classification between 1997 and 1998. Due to that change, the total number of companies in 1998 dropped to 1/3 of the 1997 data across all the industries. To control for this anomaly, we will interact the aggregation variables with a dummy for the post 1998 years in our analysis when we focus on the entire period since 1992. Finally, to avoid missing values in a log transformation, we add one to this variable as previous studies have done.[13]

When we focus on the supply and demand of intermediate goods by South Korean affiliates, we use the Korean input-output table from the year 2000 to measure firm linkages.[14] To merge input-output tables and Korean FDI data, we have to concord the industry classifications as there is a slight variation between both sources.[15] For example, the input-output table specifies semiconductors (KSIC 3211) and other related devices (KSIC 3219) separately, while in the FDI data set, both industries are classified as semiconductors and related devices. In this case, we combine KSIC 3211 and KSIC 3219 to match the FDI classification. When some industries are more finely defined in the FDI data set than in the input-output table, we adjust the sectors accordingly. In the end, our adjusted input-output table consists of 53 industries. In order to be able to construct the second agglomeration variable at the industry level irrespective of nationality, we rely on the Chinese input-output table 1995, which comprises 15 industries.[16] Finally, in order to construct the linkage variables as described in section 2, we interacted Korean input-output table with the number of Korean affiliates and Chinese input-output table with the number of firms of all nationalities, and also add one to them for the same reason as with

[13]An easy way to rationalize this is to argue that the investing firm does take its own presence in the region into account as it decides whether to invest in a region. We follow Head et al. (1995).

[14]Since the industry shares do not change much during the sample period, the 2000 table is used for all years. Alternatively, one could use the IO tables that were published with 5-year intervals. A drawback of going that route is that the classification also changes over the time period.

[15]The input-output table is published by the Bank of Korea, while the FDI data come from the Export-Import Bank.

[16]We also used the table published by World Bank, which is based on GTAP 4 database. Input-output coefficients are fairly stable and analysis using either table produced the similar results.

the agglomeration variable. While it is not feasible to exactly match the Chinese and South Korean input output tables, we will also present estimates with a more aggregate South Korean input output table of 18 sectors.

We control for regional economic and demographic factors in the estimation in two ways. In our preferred specification, we include time-region dummies. We also show results for an alternative specification that includes some of the known determinants of FDI. It is well known, for example, that a larger regional economy attracts more FDI. At the same time, as Head and Mayer (2004) suggest, market potential is an important factor that may affect an investor's location decision. To control for market potential and a region's size, we include the distance-weighted real GDPs of all regions. We take the real GDP data from various issues of the China Statistical Yearbook. The distance between regions is measured as the distance between the provincial capitals, which is taken from yahoo.com. We control for the labor costs by including the average level of regional staff and worker wages from the China Statistical Yearbook. Lower wage rates could be more attractive to investors in search of cheap labor. Furthermore, to control for the quality of workers, we include the ratio of high school graduates to the total population. We compute this ratio from the China Statistical Yearbook. Finally, we also consider a variable that reflect the government's role in attracting FDI. We choose the number of the special economic zones (SEZ) in a region that were especially created for foreign companies.[17] There are many different types of economic zones such as Open coastal cities (OCCs), economic and technological development zones (ETDZs), open coastal areas (OCAs), technology industry development zones (TIDZs), bonded zones (BZs), border economic cooperation zones (BECZs), and export processing zones (EPZs). As of 2004, there was a lot of variation in the total number of zones in a region, with Guangdong having as many as 20 economics zones, for example.

Table 1 reports the summary statistics for the main variables of our analysis after

[17] Cheng and Kwan (2000) provide evidence of the significant role of special economic policy.

taking into account spatial aspects. Agglomeration measures the total number of firms in an industry, irrespective of their nationality. Agglomeration by SK affiliates only counts the South Korean affiliates in an industry. Similarly, we have forward and backward linkages involving all the firms in an industry as well as those only involving the South Korean affiliates.

4 Results

Table 2 reports the estimation results for the period after the Asian financial crisis. As mentioned before, this period saw an increase in the share of affiliates from high-linkage industries. At the same time, as far as the total number of firms goes, the data is the most consistent for this period. The results present the estimates of equation (6) that specifies the factors that determine an investor's decision to invest in a particular Chinese region. In all columns, except for the fourth one, the equation includes time-region-specific effects. The first two columns present familiar estimates of equation (6) that have been used in previous studies. The estimates in the first column suggest that firms agglomerate by industry regardless of nationalities. Those in the second column indicate that the decision to invest in a particular Chinese region is determined not only by the agglomeration of companies of any nationality in a given industry, but also in particular by the number of South Korean affiliates in that industry. The coefficients are significantly positive and in the range of previous studies. Note that the interpretation of the coefficient estimate as the average probability elasticity needs some care in a conditional logit model. It can be shown that the average probability of how any regressor impacts the location choice over all choosers and location choices should be calculated as follows: $(S-1)/S$ times the regressor's estimated coefficient, where S is the number of location choices.[18] Since there are 19 locations in our study, our estimates show

[18] see Head et al. (1995).

that a 10 percent increase in the distance-weighted number of Korean affiliates in one region will increase the probability that investors choose that region by around 10 percent (0.95 x 1.108 x 10).

As argued before, there are various ways to interpret these findings, and it is not clear why firms would agglomerate in a particular location. We therefore decompose the agglomeration effect into two parts: a pure agglomeration and an agglomeration effect related to the strength of the forward and backward linkages. Moreover, we consider both the agglomeration linkages for South Korean affiliates and the agglomeration linkages at the industry level for companies of any nationality (including local Chinese companies). In the third column of Table 2, we include the linkage variables that are described in section 2. The pure agglomeration externalities remain strongly positive. The result shows that both forward and backward linkage effects are significant for South Korean establishments. As expected, the magnitude of the pure agglomeration effect decreases significantly as forward and backward linkages are included. Interestingly enough, however, the linkage effects at the industry level across nationalities are not significant. This suggests that, while the presence of many companies in an industry matters, the nationality of the establishments is key for the specific upstream and downstream links that are directly aligned with the specific production process.

In the fourth column of Table 2, we drop the time-region effects and include region-specific dummies together with other more traditional determinants of industry location that vary over time. As noted, our measures of market potential, education, and the wage and economic policies meant to attract multinationals all enter with a positive sign. However, only the wage and the number of economic zones are statistically significant. The positive coefficient on the wage runs counter to our initial intuition that multinationals might seek low-wage regions. This may suggest that the wage also picks up the quality/education level of the labor force.

As mentioned, there is an issue about the different classification of the industries according to the Chinese versus the South Korean statistics. The fifth column reports estimates when we (imperfectly) map the South Korean industry classification into its Chinese counterpart and use a more aggregate South Korean input output table. For most part, the results hold up. Only the backward South Korean link loses significance, while maintaining the same sign.

Finally, we focus on the role of *Chaebol* for our results. It is well known that *Chaebol*, South Korea's conglomerates, play a prominent role in South Korea's industrial texture. We want to investigate whether our results are driven by these large corporations. It turns out that they are not. When we drop all the larger corporations from the sample, all variables of interest retain the same signs and significance.[19] These findings then add an interesting dimension to the existing literature. As Belderbos and Carree (2002) have noted, small firms are followers. What our results show is that smaller firms are followers especially because they go where suppliers of inputs and buyers of their intermediate goods are more plentiful. Moreover, Figure 7 shows that the fraction of big firms in total firms has been decreasing over time, suggesting in addition that initially larger multinationals went abroad.

In Table 3, we extend the sample. We present estimates for the post-1992 period. We estimate from 1992 onward rather than 1988, since there is some concern that the location choice before the diplomatic relations between China and South Korea were initiated in 1992 may not have been purely for economic reasons.[20] Note that we included early establishments going back to 1988 in the count of firms to construct the agglomeration variables when we estimated the location decision after 1992. The makeup of Table 3 mimics that of the previous table for the time since 1999, and the

[19] In the survey, the investing firms declare themselves large or not. Some 8 percent is large.

[20] It turns out that estimates for the entire sample, starting from 1988 are qualitatively the same, but somewhat weaker. Alternatively, one could argue that the early years of the data set are not of major interest to uncover agglomeration effects, since there were no South Korean affiliates in 1988.

results largely correspond to those of the more recent period. The main difference is that the estimates are somewhat weaker and less precisely estimated.

The last issue that we address is the robustness of the results. In particular, we investigate the independence of irrelevant alternatives (IIA) assumption that is implied in a conditional logit analysis. In a conditional logit, the relative probability of choosing between two alternatives should not depend on the availability of a third alternative. The IIA assumption hinges upon the identical and independent error terms. As argued by Head et al. (1995), the inclusion of alternative specific constants (in our case, regional dummies) allows for conditional logits in the presence of violations of IIA, as long as investors have uniform perceptions about the substitutability between states. At the same time, the regional dummies complicate formal testing of IIA since they yield different numbers of parameters across specifications. We therefore compare the estimates of the critical variables of interest as we exclude several regions with the baseline estimates for the full set of choices. When the coefficients and significance levels are relatively stable, we regard the IIA assumption as valid, as is done in the previous studies. We exclude Shandong, three Northeast provinces (Jilin, Liaoning, and Heilongjiang), and three municipalities (Beijing, Tianjin, and Shanghai) in turn in the second and third column of Table 4. Note that the first column for each time period is the standard estimate (4) from Tables 2 and (10) from Table 3 that includes regional variables with region-specific dummies. For reference, Shandong is the location of most Korean firms ; the Northeast provinces are attractive regions for Korean investors since there are many Korean Chinese; and the municipality itself has economic significance. For the period after 1999, the coefficients and significance levels are relatively stable. For the entire period after 1992 we have comparable results that, as before, tend to be somewhat smaller and somewhat less precisely estimated.

5 Conclusion

China, together with the United States, has topped the list of recipients of FDI for a number of years. This has heightened the interest to gain a better understanding of what drives FDI into China and what explains the local dispersion of FDI as China is rapidly being integrated in the world economy. With an unpublished data source for all of Korea's affiliates across China's regions, we investigate whether and how agglomeration affects the allocation decision along industry lines and along national lines. In particular, we extend the usual agglomeration analysis and decompose the agglomeration effect as we introduce backward and forward linkages. Since the work by Hirschman (1958), there has been an active interest in forward and backward linkages. Moreover, the idea of forming clusters of economic activity has been central in the attempts to attract FDI in countries such as Ireland and Costa Rica.

We find that forward and backward linkages interacted with the presence of other Korean affiliates in China play a significant role in determining the location of Korean FDI in China. At the same time, however, the general forward and backward linkages at the industry level, irrespective of nationality, do not seem to matter. These findings are fairly robust and not driven by the South Korean *Chaebol*. Moreover, our results imply that a bilateral policy targeted to a specific country in order to attract foreign FDI may have an additional payoff compared to a multilateral approach.

References

Amiti, Mary and Beata Smarzynska Javorcik (2008), 'Trade costs and location of foreign firms in China', *Journal of Development Economics* **85**(1-2), 129–149.

Barry, Frank, Holger Gorg and Eric Strobl (2003), 'Foreign direct investment, agglomeration and demonstration effects: an empirical investigation', *Weltwirtschaftliches Archiv* **139**, 583–600.

Belderbos, Rene and Martin Carree (2002), 'The location of Japanese investments in China: Agglomeration effects, *Keiretsu*, and firm heterogeneity', *Journal of the Japanese and International Economics* **16**, 194–211.

Blonigen, Bruce A. (2005), 'A Review of the Empirical Literature on FDI Determinants', *NBER working paper* (11299).

Blonigen, Bruce A., Christopher J. Ellis and Dietrich Fausten (2005), 'Industrial groupings and foreign direct investment', *Journal of International Economics* **65**(1), 75–91.

Brainard, S Lael (1997), 'An empirical assessment of the proximity-concentration trade-off between multinational sales and trade', *American Economic Review* **87**(4), 520–44.

Carr, David L., James R. Markusen and Keith E. Maskus (2001), 'Estimating the knowledge-capital model of the multinational enterprise', *American Economic Review* **91**(3), 693–708.

Cheng, Leonard K. and Yum K. Kwan (2000), 'What Are The Determinants of The Location of Foreign Direct Investment? The Chinese Experience', *Journal of International Economics* **51**(2), 379–400.

Hanson, Gordon H., Raymond J. Mataloni and Matthew J. Slaughter (2005), 'Vertical production networks in multinational firms', *The Review of Economics and Statistics* **87**, 664–678.

Head, Keith, John Ries and Deborah Swenson (1995), 'Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States', *Journal of International Economics* **38**(3-4), 223–247.

Head, Keith, John Ries and Deborah Swenson (1999), 'Attracting foreign manufacturing: Investment promotion and agglomeration', *Regional Science and Urban Economics* **29**(2), 197–218.

Head, Keith and Thierry Mayer (2004), 'Market Potential and the Location of Japanese Investment in the European Union', *The Review of Economics and Statistics* **86**(4), 959–972.

Helpman, Elhanan (1984), 'A simple theory of international trade with multinational corporations', *Journal of Political Economy* **92**(3), 451–71.

Hirschman, Albert O. (1958), *The Strategy of Economic Development*, Yale University Press.

Javorcik, Beata Smarzynska (2004), 'Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages', *American Economic Review* **94**(3), 605–627.

Larrain, Felipe B., Luis F. Lopez-Calva and Andres Rodriguez-Clare (2001), *Intel: A Case Study of Foreign Direct Investment in Central America*, In Larrain, Felipe (eds.), Growth and Internationalization, Harvard University Press.

Markusen, James R. (1984), 'Multinationals, multi-plant economies, and the gains from trade', *Journal of International Economics* **16**(3-4), 205–26.

Markusen, James R. (2002), *Multinational Firms and the Theory of International Trade*, MIT Press, Cambridge.

McFadden, Daniel S. (1974), *Conditional logit analysis of qualitative choice behavior*, In: Zarembka, P. (Ed), Frontiers in Econometrics, Academic Press, New York.

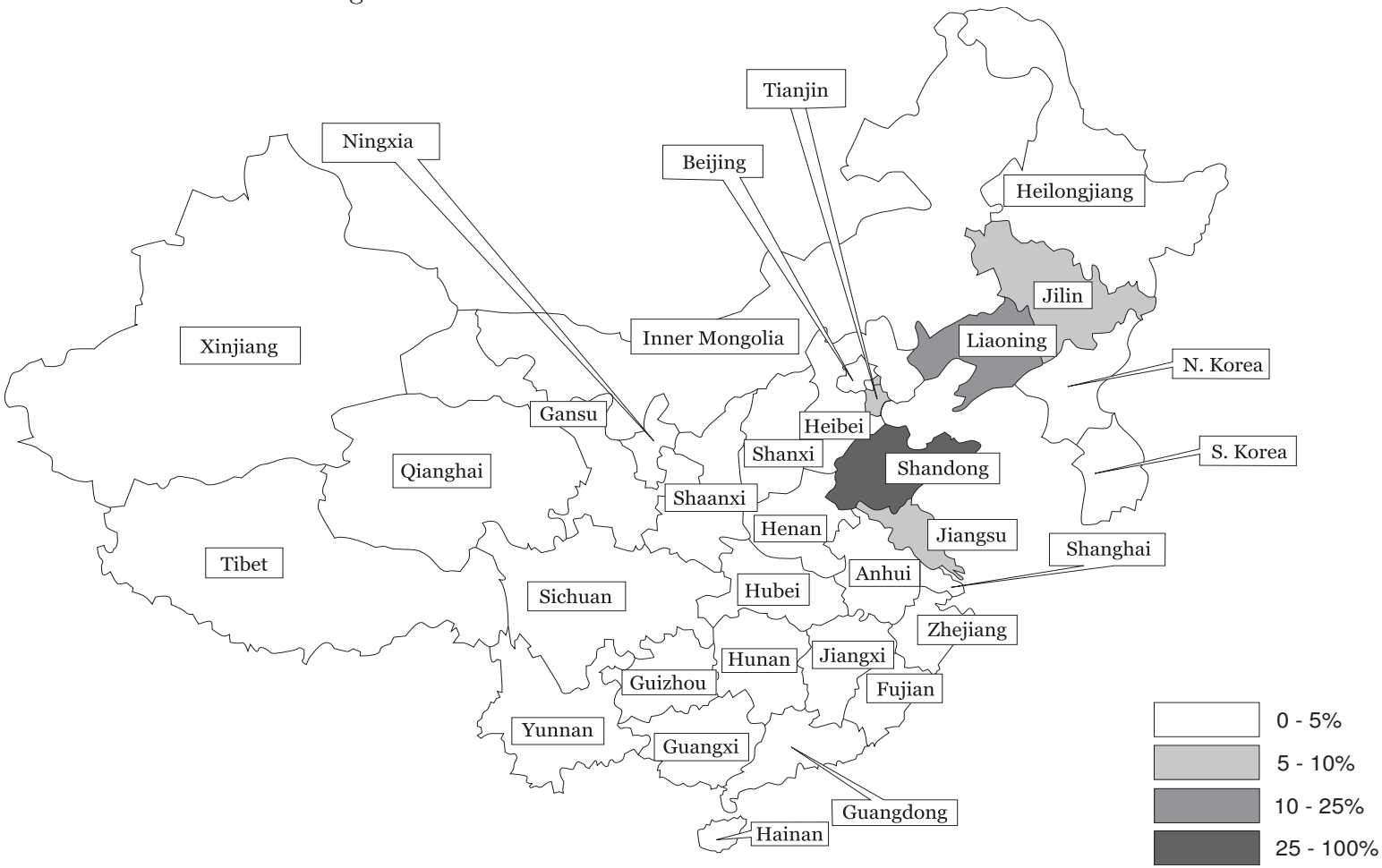
Navaretti, Giorgio Barba and Anthony J. Venables (2004), *Multinational Firms in the World Economy*, Princeton University Press.

Porter, Micheal (1998), 'Clusters and the new economics of competition', *Harvard Business Review* (November-December).

Wheeler, David and Ashoka Mody (1992), 'International investment location decisions : The case of U.S. firms', *Journal of International Economics* **33**(1-2), 57–76.

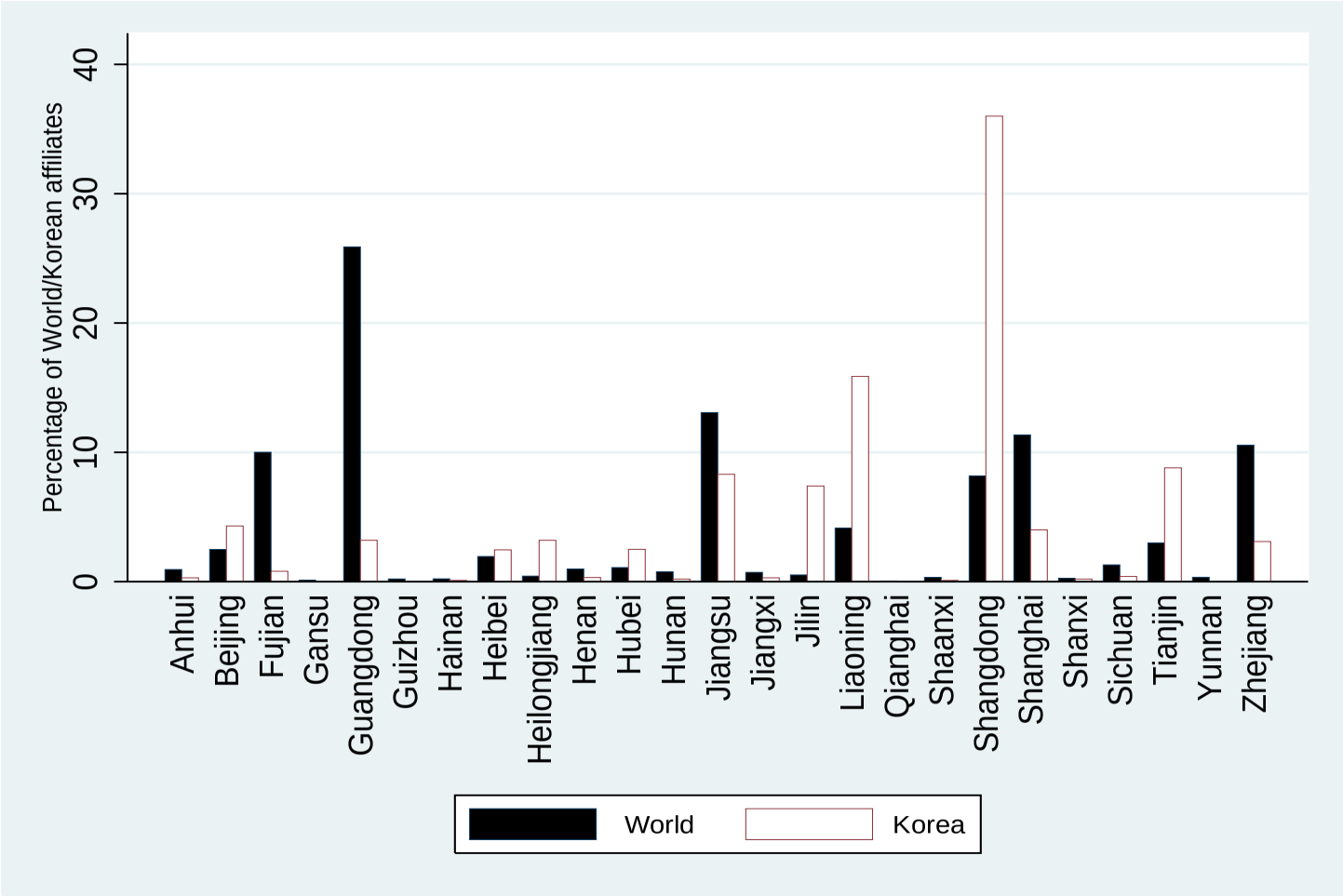
Yeaple, Stephen Ross (2003), 'The Role of Skill Endowments in the Structure of U.S. Outward Foreign Direct Investment', *The Review of Economics and Statistics* **85**(3), 726–734.

Figure 1: Distribution of South Korean investment in China



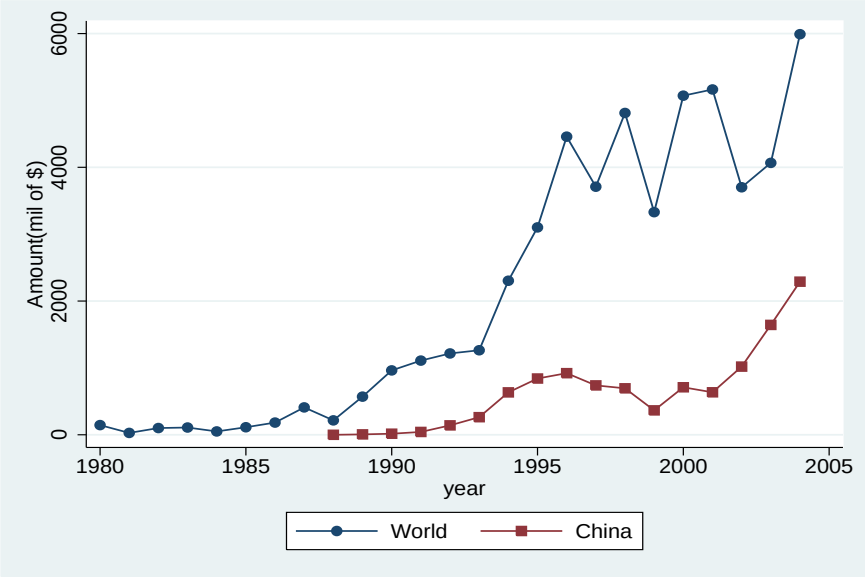
Source: Export-Import Bank of Korea

Figure 2: Distribution of World/Korea FDI in China



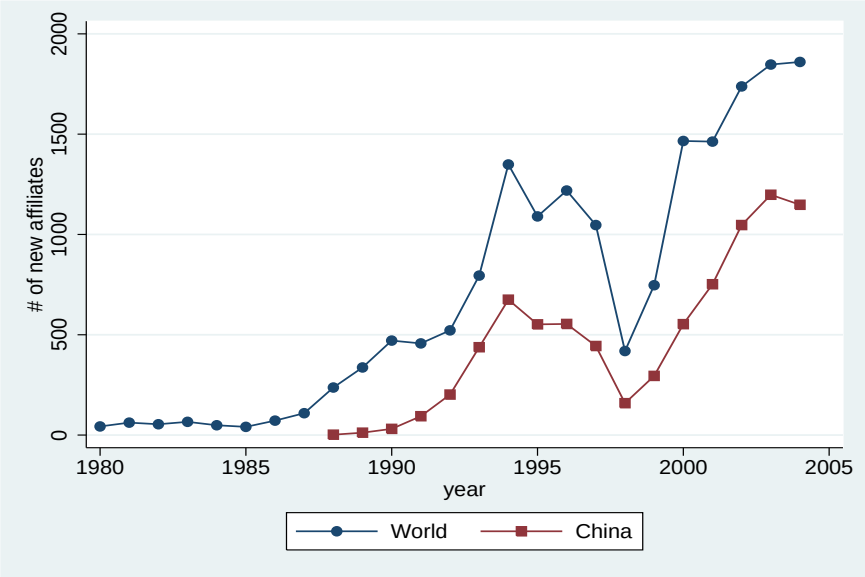
The source for worldwide multinational firms is China Statistics Yearbook 2006. The data for Korea is from Export-Import Bank of Korea.

Figure 3: Amount of South Korean investment to the world and China



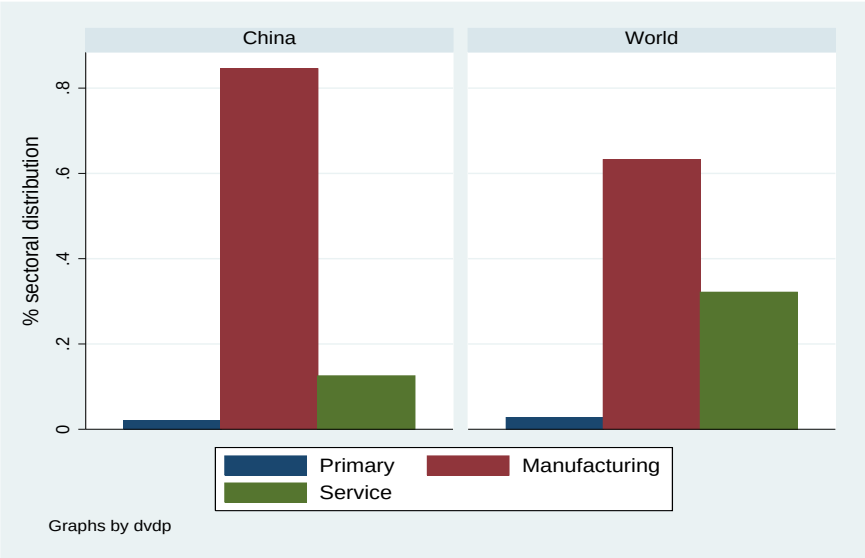
Source: Export-Import Bank of Korea

Figure 4: Number of new South Korean affiliates in the world and China



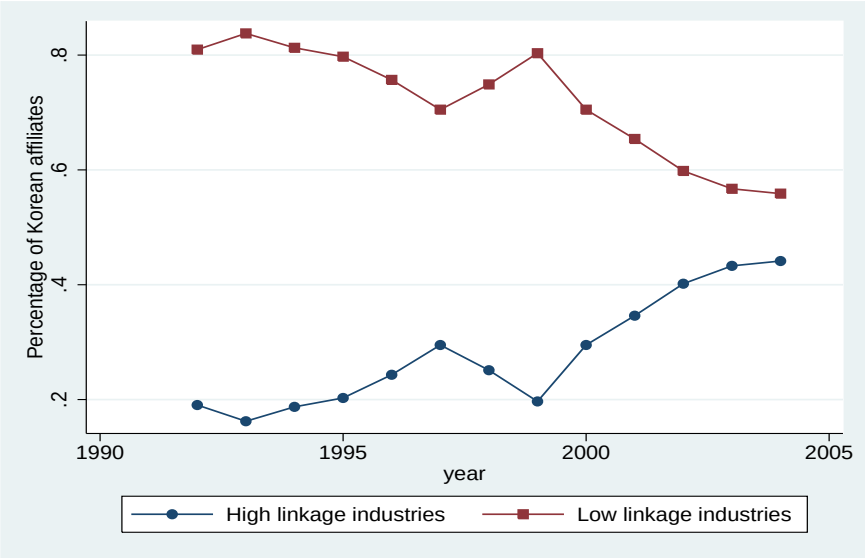
Source: Export-Import Bank of Korea

Figure 5: Sectoral Shares of South Korea's new affiliates in China vs. for the world as a whole, in percent



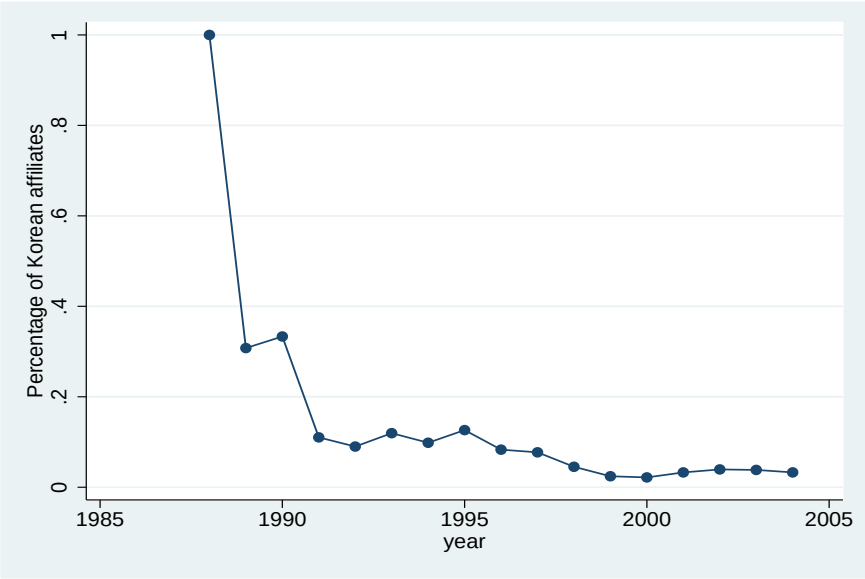
Source: Export-Import Bank of Korea

Figure 6: The share of high/low linkage sectors in manufacturing



Source: Export-Import Bank of Korea. High linkage sectors include machinery, transportation, and electronics. Low linkage sectors comprise the remaining industries.

Figure 7: Fraction of affiliates from large multinationals over time



Source: Export-Import Bank of Korea

Table 1: Summary statistics of regressors

Variable	Mean	St. Dev	Min	Max
Agglomeration	6.835	0.792	3.514	8.902
Agg. by SK affiliates	2.075	1.194	0	5.777
Forward Linkages	7.232	0.808	5.399	10.496
Backward Linkages	7.290	0.722	5.542	10.229
For. Link. by SK affiliates	1.896	0.97	0	5.288
Back. Link. by SK affiliates	1.759	0.903	0.005	4.93

Source: Korean data is from Export-Import Bank of Korea and Bank of Korea. Non-Korea data is from Chinese Statistical Yearbook and China Industry Statistical Yearbook. All variables are in log term.

Table 2: Estimation after 1999

	(1)	(2)	(3)	(4)	(5) Aggregation	(6) Non-*Chaebol*
Agglomeration	1.108	0.754	0.494	0.453	0.552	0.466
	[0.075]***	[0.077]***	[0.128]***	[0.125]***	[0.137]***	[0.135]***
Agg. by SK affiliates		0.800	0.374	0.347	0.567	0.379
		[0.051]***	[0.092]***	[0.090]***	[0.154]***	[0.095]***
Forward Linkages			0.040	-0.036	0.243	0.000
			[0.251]	[0.242]	[0.258]	[0.259]
Backward Linkages			0.153	0.243	-0.314	0.240
			[0.312]	[0.298]	[0.319]	[0.321]
For. Link. by SK affiliates			0.479	0.554	0.547	0.490
			[0.148]***	[0.144]***	[0.260]**	[0.154]***
Back. Link. by SK affiliates			0.531	0.543	0.244	0.486
			[0.140]***	[0.136]***	[0.258]	[0.146]***
Market Potential				2.208		
				[3.780]		
High School Graduates				0.045		
				[0.620]		
Wage rate				7.145		
				[1.246]***		
Economic Zones				1.916		
				[0.660]***		
Region dummy	no	no	no	yes	no	no
Region & Year dummy	yes	yes	yes	no	yes	yes
No. of choices	19	19	19	19	19	19
No. of investors	4245	4245	4245	4245	4245	4012
Pseudo-R2	0.32	0.33	0.33	0.32	0.32	0.34
Log-likelihood	-8552.79	-8428.90	-8410.13	-8501.17	-8467.30	-7851.72

Source: Korean data is from Export-Import Bank of Korea and Bank of Korea. Non-Korean data is from Chinese Statistical Yearbook and China Industry Statistical Yearbook. Distance is from Yahoo.com. All variables except for dummies are in log term. Standard error is in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3: Estimation after 1992

	(7)	(8)	(9)	(10)	(11) Aggregation	(12) Non-*Chaebol*
Agglomeration	0.919	0.686	0.439	0.338	0.444	0.482
	[0.102]***	[0.101]***	[0.124]***	[0.113]***	[0.126]***	[0.134]***
Agg. *D(year>= 1998)	0.187	0.078	0.022	0.518	0.090	-0.021
	[0.126]	[0.125]	[0.126]	[0.085]***	[0.126]	[0.135]
Agg. by SK affiliates		0.772	0.446	0.497	0.472	0.433
		[0.039]***	[0.070]***	[0.067]***	[0.112]***	[0.073]***
Forward Linkages			0.039	-0.126	0.009	-0.043
			[0.212]	[0.196]	[0.214]	[0.221]
Backward Linkages			0.265	0.194	0.114	0.356
			[0.263]	[0.241]	[0.261]	[0.273]
For. Link. by SK affiliates			0.319	0.265	0.339	0.335
			[0.111]***	[0.104]**	[0.190]*	[0.117]***
Back. Link. by SK affiliates			0.452	0.347	0.311	0.441
			[0.105]***	[0.098]***	[0.200]	[0.111]***
Market Potential				1.177		
				[0.622]*		
High School Graduates				0.102		
				[0.230]		
Wage rate				0.207		
				[0.335]		
Economic Zones				0.409		
				[0.149]***		
Region dummy	no	no	no	yes	no	no
Region & Year dummy	yes	yes	yes	no	yes	yes
No. of choices	19	19	19	19	19	19
No. of investors	6836	6836	6836	6836	6836	6285
Pseudo-R2	0.30	0.31	0.31	0.30	0.31	0.32
Log-likelihood	-14072.41	-13876.75	-13854.52	-14099.51	-12563.45	-13960.18

Source: Korean data is from Export-Import Bank of Korea and Bank of Korea. Non-Korean data is from Chinese Statistical Yearbook and China Industry Statistical Yearbook. Distance is from Yahoo.com. All variables except for dummies are in log term. Standard error is in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4: Robustness

	1999-2004				1992-2004			
	(4)	(13)	(14)	(15)	(10)	(16)	(17)	(18)
Excluding		Shandong	Northeast	Municipalities		Shandong	Northeast	Municipalities
Agglomeration	0.453	0.560	0.446	0.471	0.338	0.375	-0.047	0.472
	[0.125]***	[0.139]***	[0.141]***	[0.149]***	[0.113]***	[0.127]***	[0.151]	[0.138]***
Agg. by SK affiliates	0.347	0.448	0.337	0.262	0.497	0.494	0.535	0.483
	[0.090]***	[0.112]***	[0.107]***	[0.102]**	[0.067]***	[0.084]***	[0.081]***	[0.075]***
Forward Linkages	-0.036	-0.257	0.646	-0.259	-0.126	-0.323	0.703	-0.384
	[0.242]	[0.294]	[0.307]**	[0.285]	[0.196]	[0.239]	[0.261]***	[0.227]*
Backward Linkages	0.243	0.413	-0.503	0.499	0.194	0.453	-0.551	0.503
	[0.298]	[0.349]	[0.369]	[0.341]	[0.241]	[0.280]	[0.309]*	[0.272]*
For. Link. by SK affiliates	0.554	0.365	0.531	0.770	0.265	0.245	0.275	0.305
	[0.144]***	[0.174]**	[0.169]***	[0.167]***	[0.104]**	[0.128]*	[0.127]**	[0.118]***
Back. Link. by SK affiliates	0.543	0.632	0.385	0.443	0.347	0.374	0.138	0.326
	[0.136]***	[0.179]***	[0.164]**	[0.154]***	[0.098]***	[0.125]***	[0.120]	[0.109]***
Region dummy	yes	yes	yes	yes	yes	yes	yes	yes
No. of choices	19	18	16	16	19	18	16	16
No. of investors	4245	2552	3564	3445	6836	4373	5285	5543
Pseudo-R2	0.32	0.24	0.38	0.41	0.30	0.24	0.38	0.39
Log-likelihood	-8501.17	-5730.74	-6176.58	-5687.48	-14099.51	-9787.18	-9319.34	-9524.14

Source: Korean data is from Export-Import Bank of Korea and Bank of Korea. Non-Korean data is from Chinese Statistical Yearbook and China Industry Statistical Yearbook. Distance is from Yahoo.com. All variables except for dummies are in log term. Standard error is in parentheses. All specifications include market potential, wage rate, the ratio of high school graduates, and the number of economics zones. * significant at 10%, ** significant at 5%, *** significant at 1%.