Local Knowledge Spillovers in the Indonesian Manufacturing Industry

Bee Yan Aw
The Pennsylvania State University

and

Alfons Palangkaraya
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 18/04

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 0 7340 3160 2

August 2004

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

Theoretical models have long shown that knowledge spillovers are of great economic importance to sustained economic growth and innovation and that these spillovers may be facilitated by physical and technological proximity. However, local knowledge spillovers have not been identified using data from developing countries. In this paper, we examine the relationship between knowledge spillovers and both technological and geographical proximities using micro panel data of Indonesian manufacturing plants between 1990 and 1995. We find both physical and technological proximity are significant. Knowledge spillovers are stronger among plants in the same industrial sector and their magnitude decreases monotonically with geographical distance.

1. Introduction

As countries open their borders to foreign goods and services, information spillovers have been increasingly recognized as an important source of innovation and economic growth in developing countries. A number of theoretical arguments and some empirical findings indicate that such knowledge are concentrated in spatial proximity from their respective source, an idea that goes back to Marshall's (1920) hypothesis that geographical proximity increases the probability of knowledge spillovers.

The existence of spatial externalities embodied in Marshall's hypothesis relies on the assumption of human interaction as an important vehicle for knowledge transfers. In particular, geographical proximity is an important determinant of the diffusion of tacit knowledge (Audretsch and Feldman, 1996b). Similarly, Mills (1992) refers to this type of knowledge as ambiguous information that "requires an interactive and convergent set of exchanges before the final exchange can be consummated." In Mill's view, ambiguous information is information that requires negotiation to establish meaning. This includes informal exchanges that may have to occur among producers, managerial staff, local suppliers, marketing and legal experts and other groups (Beugelsdijk and Cornet, 2001). Since tacit or ambiguous knowledge is rarely published, and therefore cannot be accessed by studying books or other published materials, its diffusion requires face-to-face access and hence, human interaction. Human interactions, in turn, are facilitated by geographical concentration or clustering.

While there is some empirical evidence to support Marshall's hypothesis in developed countries such as the U.S., the same is not true for research on this issue in developing countries. The lack of any evidence of local knowledge spillovers in developing countries may be related to the fact that existing studies have concentrated on specific sources of knowledge spillovers such as research and development (R&D) expenditures, foreign direct investments or more recently, export experience. For instance, studies which focus on knowledge spillovers coming from say, foreign producers might miss the possibility of other domestic producers as the source of the spillovers. Finally, data-related problems including the paucity of firms formally involved in R&D or foreign direct investments, missing data, measurement error or the unreliability of the data are very likely contributors to this lack of spillover evidence in developing countries.

This paper follows Winston (2001) in taking a broad view of the sources of spillovers by characterizing a firm’s knowledge as anything that increases its total factor productivity (TFP). Differences in TFP among firms are assumed to reflect differences in product design, processing technologies, organizational technologies and/or managerial skills. If each of these differences can be interpreted as a part of a firm's collective knowledge, then it is also a potential source of knowledge spillovers enabling each firm in the vicinity to benefit from the spillovers of specialized knowledge and to have lower costs than if it operated in isolation. In this way, knowledge spillovers this period translate into lower TFP in the next period.

While the TFP measure does not allow us to identify the source of the knowledge spillover, it is able to capture the extent of knowledge accumulated and potentially available for other firms to learn from or imitate, whether or not it is in the interest of the agent with the knowledge to leak information about the new product or technology involved.

The empirical work in this paper is based on micro panel data from the Indonesian Census of Manufactures from 1990 to1995. We extend Winston’s paper in two important aspects. First, we include a physical distance measure between plants. Specifically, we follow Beugelsdijk and Cornet's (2001) "shell-model" idea in constructing our measure of external knowledge which varies by distance. Second, we analyze how the extent of knowledge spillover varies across different industries. Firms' ability to benefit from external knowledge depends on their absorptive capacity. This absorptive capacity is likely to be greater for firms within the same industrial sector than across different sectors. In addition, the industrial life cycle hypothesis asserts that tacit knowledge is more important in the early stage of the industry life cycle (Audretsch and Feldman, 1996b). Therefore, we expect spillover effects to be stronger in industries with relatively young firms.

The remainder of the paper is divided into five sections. In the next section, we summarize the theoretical and empirical frameworks. This is followed by a discussion on the measurements of geographical distance and firm internal and external knowledge in section three. The fourth section details the empirical model specifications. In section five we discuss and interpret the estimation results. In the final section we summarize the major findings and conclusions.

2. Theoretical and Empirical Framework

Our empirical framework is based on Hopenhayn's (1992) model of firm dynamics and Winston's (2001) extension to include knowledge spillovers from physical proximity. The model contains three basic elements. First, it specifies the knowledge creation process as one that generates internal and external knowledge. Second, it includes a measure of geographical proximity to capture how knowledge spills over. Third, since endogenous turnover determines which firms are observed in the data, the model incorporates an explanation for firm exit.

Given that the knowledge accumulation process and the spillovers associated with it is subject to uncertainty, new ideas are modeled as a random draw. In every period t, a representative plant i utilizes its current period stocks of internal (θ_{it}) and external knowledge in physical location L (Θ_{it}^{L}) to produce new ideas, κ_{it}.[1, 2] The quantity or quality of the new ideas and exogenous plant characteristics (x_{it}), such as plant age and size, determine the family of knowledge distributions, $K\left(\theta_{it}, \Theta_{it}^{L}, x_{it}\right)$ from which plant i can randomly draw its new knowledge. However, since the transformation of new ideas into new knowledge takes time, κ_{it} can only be incorporated into the next period stock of knowledge. "Old" knowledge may be rendered obsolete and replaced with new knowledge. Thus, the evolution of knowledge stock over time is specified as:

$$\theta_{it+1} = (1-\delta)\theta_{it} + \kappa_{it} \tag{1}$$

where δ denotes the proportion of the internal stock of knowledge destroyed in each period.

In every period t, given its own stock of production knowledge (θ_{it}), plant i produces total output (y_{it}) from labor input (w_{it}), capital input (k_{it}), and inputs of raw materials, fuel and electricity (m_{it}) according to a production technology represented by $Y_{it}\left(\theta_{it}, w_{it}, k_{it}, m_{it}\right)$. In addition, the plant forms a rational expectation of its future knowledge and, thus its future output, in order to decide whether or not to stay in the market. The exit-entry decision is based on an inter-temporal optimization of profits. In particular, in every period, plant i

---

[1] Henceforth, we will use 'plant' instead of 'firm' in order to be consistent with the level at which the data is collected in the Indonesian Census of Manufactures.

[2] The 'new-ness' of knowledge is with respect to plant i only. In other words, plant i's new knowledge may be another plant's old knowledge. Thus, there is no distinction between imitation or innovation in this framework.

compares the sum of all present discounted value future profits relative to its current scrap value. If the current scrap value exceeds the present discounted value of all future profits, then the firm will choose to exit. This implies a productivity threshold the plant uses to make exit decisions. This threshold level varies across plants and depends on plant's individual characteristics and its knowledge evolution process. The plant's decision rule is then specified as:

$$F_{it} = \begin{cases} 1 \text{ if } \theta_{it} \geq \underline{\theta}_{it}\left(z_{it}, \Theta_{it}^{L}\right) \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

where the plant's productivity threshold, $\underline{\theta}_{it}$, is a latent variable which depends on z_{it}, a vector of exogenous characteristics which affect the exit decision, such as age and the value of physical assets as well as its external knowledge stock.[3] The observed binary choice variable, F_{it}, takes on the value of 1 if the plant chooses to continue production and 0 otherwise.

The empirical model is specified as the reduced form of equations (1) and (2) such that

$$\begin{aligned} \theta_{it+1} &= \alpha_0 + \alpha_1\theta_{it} + \Theta_{it}^{S,L}\beta + x_{it}\gamma + \varepsilon_{it+1} \\ F_{it}^{*} &= \lambda_0 + \lambda_1\theta_{it} + \Theta_{it}^{S,L}\varphi + z_{it}\psi + \mu_t \\ F_{it} &= \begin{cases} 1 \text{ if } F_{it}^{*} > 0 \\ 0 \text{ otherwise} \end{cases} \end{aligned} \tag{3}$$

where $\Theta_{it}^{S,L}$ is a vector of measures of external knowledge available for plant i in period t, x_{it} and z_{it} are vectors of plant characteristics. Furthermore, the set $\{\alpha_0, \alpha_1, \beta, \gamma, \lambda, \lambda_1, \varphi, \psi, \rho, \sigma\}$ contains the parameters to be estimated using Heckman's sample selection maximum likelihood estimation method.

An important observation from the derivation of equations (1) and (2) is that in every period t, the turnover of plants determines the stock of external knowledge available to plant i. As a result, it is possible for the unobservable terms, ε_{it+1} and μ_{it}, in equation (3) to be

[3] See Table 4 for the components of z_{it}.

correlated with each other. In the empirical estimation, we assume $(\varepsilon_{it+1}, \mu_{it})$ ~ bivariate normal $[0, 0, 1, \sigma, \rho]$.

To estimate the system of equations in (3), we need to construct the measures of external knowledge (Θ_{it}^{L}) bearing three properties. First, for any plant i, Θ_{it}^{L} must reflect the distribution of internal knowledge of all other plants located in each region. Second, the measures must capture within-sector spillover. Finally, the measures must reflect the level of possible human interactions between plant i and its neighbors in each location.

To fulfill the first property, we construct Θ_{it}^{L} for plant *i* as the median of the internal knowledge (θ_{jt}) of all j plants, $j \neq i$, producing in each location $L = 1,2,3$ with the closest ring represented by $L = 1$ and so on. Θ_{it}^{L} is constructed across all industrial sectors. In line with the second property, we construct $\Theta_{it}^{S,L}$ for each plant *i* as the median TFP of all j plants, $j \neq i$, in sector *S* and location *L*. Consistent with the final property, we construct $F_{it}^{S,L}$ and $W_{it}^{S,L}$ which represent the total number of plants and number of paid-workers, respectively, in each location *L* and sector *S*, excluding plant *i* from each measurement.

In the measures described above, the superscripts attached to the spillover measures, *S* and *L*, capture the degree of technological and geographical proximity, respectively. In addition to the number of neighboring plants (F_{it}^{L} and $F_{it}^{S,L}$) we also consider the total number of workers (W_{it}^{L} and $W_{it}^{S,L}$) in the neighboring plants as an alternative proxy for the level human interaction. We argue that having two neighboring plants with a combined total of 500 employees might provide greater opportunities for interaction than having five neighboring plants with a combined total of 100 employees.

We estimate four basic model specifications. Each specification uses different variations of the measure of external knowledge. The main parameters of interest are listed in Table 1. The + and the $\downarrow(L)$ symbols in the table represent the expectation, based on the theoretical model, about the coefficients in the equation for knowledge-evolution. In particular, the + signs indicate positive spillover effects and the $\downarrow(L)$ signs indicate that the strength of the spillover effects decreases as distance increases.

Knowledge spillovers exist in the model if Θ_{it}^{L}, $\Theta_{it}^{S,L}$, F_{it}^{L}, $F_{it}^{S,L}$, W_{it}^{L} and $W_{it}^{S,L}$ have positive effects on firm i's future knowledge stock, θ_{it+1}. Other things equal, a higher $\Theta_{it}^{S,L}$ implies a higher quantity or better quality of external knowledge from which new ideas can be drawn and converted into new knowledge, which translates into improved future internal knowledge and productivity. Similarly, a higher $F_{it}^{S,L}$ or $W_{it}^{S,L}$ implies greater opportunities of interacting with and learning from neighboring firms and, therefore, impacting on the quality of the firm's future stock of new ideas and thus, productivity.

Table 1: Parameters for the Knowledge-Evolution Equation

Measures of External Knowledge	Model 1A	Model 1B	Model 2A	Model 2B
Θ_{it}^{L}	+, $\downarrow$ (L)	+, $\downarrow$ (L)		
$\Theta_{it}^{S,L}$			+, $\downarrow$ (L)	+, $\downarrow$ (L)
F_{it}^{L}	+, $\downarrow$ (L)			
$F_{it}^{S,L}$			+, $\downarrow$ (L)	
W_{it}^{L}		+, $\downarrow$ (L)		
$W_{it}^{S,L}$				+,$\downarrow$ (L)

If both technological and geographical proximities matter, then we expect the positive spillover effects of external knowledge to vary across sectors, S and locations, L in a systematic way. In particular, Marshall's hypothesis can be translated in terms of the following inequality: $\Theta_{it}^{1} > \Theta_{it}^{2} > \Theta_{it}^{3}$. Furthermore, if technological proximity matters then we expect $\Theta_{it}^{S,L} > 0$ to hold. Table 2 lists the exogenous variables included in the vectors of firm characteristics x_{it} and z_{it}. The + and - are the expected signs for the respective coefficients. For example, if there is any vintage capital effect, then, ceteris paribus, we expect the coefficient on the age variable to be negative, implying that compared to younger plants, older plants have lower future productivity. Furthermore, if larger plants are more willing to accept losses in any economic downturn, then the coefficients of c_{it} (log value of total assets) will be positive. The set of vectors of dummy variables Yr_{it}, R_{it}, and Sd_{it} are included to capture other factors that may be important to account for in examining the correlation between a plant's future TFP and its current external knowledge. More specifically, Yr_{it} represent four year dummies which capture any contemporaneous shocks that affect all plants. R_{it} consists of four dummy variables based on the five major islands in

Indonesia and captures any common effects due to unobserved regional characteristics. Finally, Sd_{it} consists of eight dummy variables based on the 2-digit ISIC classification and captures specific shocks to common plants in each sector.

Table 2: Exogenous Variables

Exogenous variables	x_{it} (knowledge equation)	z_{it} (exit equation)
a_{it} (log of age in production)	- (vintage capital)	+ (experience)
N_{it} (1 if a new entrant)	- (inexperience)	- (inexperience)
c_{it} (log of the value of capital)		+ (size effects)
c_{it}^2 (the square of c_{it})		+ (size effects)
Yr_{it} (year dummy variables)	?	?
R_{it} (region dummy variables)	?	?
Sd_{it} (sector dummy variables)	?	?

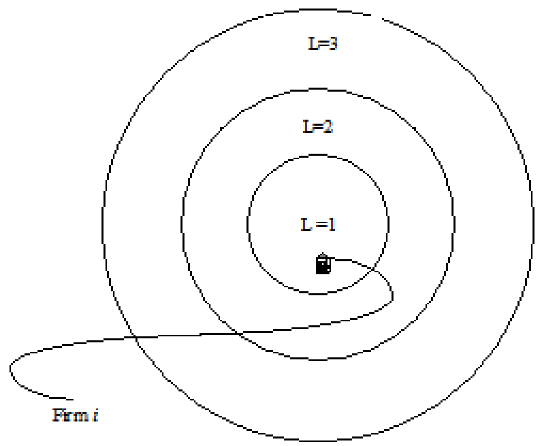
3. Measurement Methodology

3.1. Ring-model of geographical proximity

To test Marshall's hypothesis, we need to construct a measure of external knowledge that varies with a measure of physical distance. In this paper, we follow Beugelsdijk and Cornet (2001) by drawing consecutive rings around each plant i as shown in Figure 1. We then construct separate measures of external knowledge for each of the regions defined by these rings. As shown in Figure 1, the three rings define the three shell-locations ($L = 1, 2, 3$) used in our empirical estimation.

The alternatives on how rings are defined depend on the available location data. According to Beugelsdijk and Cornet, ideally the rings should be defined around plant i on an equal cost of distance base. This way, the regions are defined consistently in terms of how costly it is for plant i to interact with its neighbors in each different region. Alternatively, rings may be defined simply as circles with radius based on the actual physical distance from plant i. For

Figure 1: Rings around a representative firm i defining i's neighbors located in three different locations with increasing distances from i's location.



example, $L=1$ may be defined as the area within 0-10 miles from plant i's location, and so on. Thus, regions are defined consistently for every plant in a metric sense. Beugelsdijk and Cornet, for lack of better data, used four-digit Dutch zip code to define the first ring ($L=1$), the three-digit zip code to define the second ring ($L=2$) and so on. Unfortunately, the use of zip codes might result in inconsistencies in the way the locations are defined.

Our empirical estimation is based on micro panel data from the Indonesian Census of Manufactures from 1990 to1995. We define the rings to coincide with three levels of regional borders defined by the Indonesian government. At the most aggregated level, the country is divided into several provinces. Each province is divided into several *kabupaten*.[4] Each *kabupaten* is divided into several *kecamatan*. For practical purposes we can think of *kecamatan* as representing counties and *kabupaten* as districts. We then define $L=1$ to be the *kecamatan* in which plant i is located. Consequently, $L=2$ is the region outside plant

---

[4] The definition of a *kabupaten* is approximately equivalent to the combination of several neighboring counties in the U.S.

i 's *kecamatan* but still in the same *kabupaten* as plant *i*'s. Finally, $L = 3$ is defined as the region outside plant i 's *kabupaten* but still in the same province as plant *i*'s.[5]

The use of administrative region as a measure of physical distance is not perfect in the sense that there may be inconsistency across plants. However, it captures the effect of physical proximity in the sense that opportunities for interaction decrease or becomes more costly with each movement away from the inner shell. Furthermore, this measure is probably superior to the use of zip codes in that different administrative region is likely to imply different administrative rules and regulations and therefore different costs of doing businesses. Any information or knowledge that can be used to reduce certain costs can certainly be shared or spilled over to neighboring plants. On the other hand, two different regions defined by two different zip codes are not necessarily located in different administrative areas. Therefore, the implied extent of the locality of knowledge spillovers, at least in the sense of sharing information that may reduce business administrative costs, is smaller.

3.2. The measure of knowledge

We assume that that a plant's knowledge can be represented by an index of TFP. Basically, this implies that any variation in plant performance as measured by its TFP is a reflection of the variation in the plant's general knowledge: knowledge of production, business operation, marketing and/or management.

The TFP index is used as a single measure of the plant *i*'s relative efficiency in a year, a proxy for θ_{it} , in the theoretical model. A *TFP* index captures many factors that can lead to profit differences across plants, including differences in technology, age or quality of the capital stock, managerial ability, scale economies combined with differences in size, or differences in output quality. Our interest is in the relationship between this broad-based performance measure and the extent of knowledge spillovers as well as the exit decision.

As our index of TFP we adopt the multilateral index developed by Caves, Christensen and Diewert (1982), extended by Good, Nadiri, and Sickles (1997) and empirically used in Aw, Chung and Roberts (2000). The multilateral index relies on a single reference point that is constructed as a hypothetical plant that has the arithmetic mean values of log output, log

---

[5] In Sjoholm (1999), the smallest regions defined as 'districts' are actually the *kabupaten*.

input, and input cost shares over all plants in each year. Each plant's logarithmic output and input levels are measured relative to this reference point in each year and then the reference points are chain-linked over time.

The total factor productivity index for plant *i* in year *t* is defined as:

$$\begin{aligned}\ln TFP_{it} \quad = \quad & \left(\ln Y_{it} - \overline{\ln Y_t}\right) + \sum_{\tau=2}^{t}\left(\overline{\ln Y_\tau} - \overline{\ln Y_{t-1}}\right) - \\ & \sum_{f=1}^{n} \tfrac{1}{2}\left(a_{it}^{f} + \overline{a_{it}^{f}}\right)\left(\ln X_{it}^{f} - \overline{\ln X_{it}^{f}}\right) - \\ & \sum_{\tau=2}^{t}\ \sum_{f=1}^{n} \tfrac{1}{2}\left(\overline{a_{\tau}^{f}} + \overline{a_{\tau-1}^{f}}\right)\left(\ln X_{\tau}^{f} - \overline{\ln X_{\tau-1}^{f}}\right)\end{aligned} \tag{4}$$

where X^f and a_{it}^f denote plant i's value and cost share of the specific factor input f.

The first term in the first line of equation (4) measures plant i's output (Y_i) relative to a hypothetical plant output, constructed as the average (represented by the overbar) of total output over all plants. The second term in the same line measures how much the output of the hypothetical plant has changed between period t *-1* and period *t*. Similarly, the second and third line measure plant i's relative inputs, weighted by the appropriate cost or revenue shares, a_{it}^f. In short, for every plant i, equation (4) measures the plant's performance relative to the performance of a hypothetical plant in each period and over time.[6]

4. Empirical Results

Two sets of regressions are estimated to determine the importance of local knowledge spillovers using variations in the measures of external knowledge listed in Table 1. The first set investigates the effects of geographical and technological proximity on the extent of knowledge spillovers. The second set examines sectoral variation of those effects.

Table 3 reports the estimation results of the baseline model and two specifications that take geographical and technological spillovers into account. All the regressions in the table are based on data pooled over all industrial sectors and years. We do not report the results of the coefficients on the sector and year dummies.

---

[6] Appendix 1 provides a more detailed description of the data set and the measurement of firm inputs and output.

Table 3: Estimates for the knowledge-evolution equation based on all sectors data

Regressors	Baseline	Model 1A	Model 1B	Model 2A	Model 2B
N_{it}	-0.012** (0.006)	-0.013** (0.006)	-0.009 (0.006)	-0.003 (0.009)	0.002 (0.009)
a_{it}	-0.015*** (0.002)	-0.013*** (0.002)	-0.012*** (0.002)	-0.010*** (0.003)	-0.009*** (0.003)
θ_{it}	0.658*** (0.004)	0.653*** (0.004)	0.652*** (0.004)	0.635*** (0.006)	0.636*** (0.006)
Θ_{it}^{1}		0.079*** (0.008)	0.076*** (0.008)	0.101*** (0.015)	0.100*** (0.015)
$\Theta_{it}^{S,1}$				0.021** (0.010)	0.026*** (0.010)
Θ_{it}^{2}		0.056*** (0.013)	0.040*** (0.012)	-0.057** (0.021)	-0.070*** (0.021)
$\Theta_{it}^{S,2}$				0.054*** (0.010)	0.064*** (0.010)
Θ_{it}^{3}		-0.068** (0.027)	-0.052** (0.025)	-0.093** (0.038)	-0.084** (0.034)
$\Theta_{it}^{S,3}$				0.043*** (0.015)	0.051*** (0.015)
F_{it}^{1} or W_{it}^{1}		-0.002 (0.002)	0.003*** (0.001)	0.016*** (0.003)	0.007*** (0.002)
$F_{it}^{S,1}$ or $W_{it}^{S,1}$				-0.021*** (0.003)	-0.004** (0.002)
F_{it}^{2} or W_{it}^{2}		0.005*** (0.002)	0.008*** (0.0021	-0.002 (0.004)	0.012*** (0.003)
$F_{it}^{S,2}$ or $W_{it}^{S,2}$				-0.003 (0.003)	-0.004* (0.002)
F_{it}^{3} or W_{it}^{3}		-0.009*** (0.003)	-0.007*** (0.002)	-0.004 (0.005)	-0.004 (0.015)
$F_{it}^{S,3}$ or $W_{it}^{S,3}$				-0.006* (0.003)	-0.001 (0.002)
Const	0.030*** (0.014)	0.073*** (0.026)	0.005 (0.034)	0.024 (0.044)	-0.054 (0.054)
Observations	37158	34347	34347	18017	18017

*, **, *** = significant at 10%, 5%, and 1% level.

Figures in parenthesis are standard errors.

All regressions include year-, region-, and sector-dummy variables

The estimated coefficients of the baseline model in column 2 indicate that current level of TFP is a strong and statistically significant determinant of future TFP. The estimated coefficient of internal knowledge ($\widehat{\theta}_{it}$) implies that a 10% difference in the mean of current period TFP distribution is positively associated with 6.58 % difference in the mean of future TFP distribution from which the plant can draw for the next period. The negative sign of the estimated coefficient of age (a_{it}) indicates a statistically significant effect similar to that predicted by the vintage capital argument. However, the magnitude of the effect is small. Finally, we find evidence that newer plants are drawing from a TFP distribution with a mean less than those for the incumbent plants. This is consistent with the fact that, in contrast to incumbent plants, recent entrants include many that are low productivity plants and likely to fail in the future.

Table 4 reports the estimated parameters of the selection equation. In addition to the parameters of the evolution equation, the selection equation includes the natural log of capital and its square. The results are consistent with the predictions of the theoretical model regarding plants' endogenous exit decisions in that older, larger and more productive plants are more likely to survive while recent entrants are more likely to exit. These patterns are consistent with the predictions of Hopenhayn's theoretical model. For example, his model predicts that new entrants are more likely to exit, especially when sunk costs are small. In addition, it is also consistent with the empirical finding based on U.S. manufacturing firms (Bernard and Jensen, 2002). However, given that the capital squared term is never statistically significant, there is no evidence that the plant size effect on survival is nonlinear.

4.1. Physical Proximity matters

Going back to Table 3, the third column of the table reports the coefficient estimates of the external knowledge variables in the knowledge-evolution equation: $\Theta_{it}^{L}, \Theta_{it}^{S,L}, F_{it}^{L}$, and $F_{it}^{S,L}$. In particular, model 1 shows how much geographical proximity affects the extent knowledge spillovers and model 2, the effect of technological proximity.

There is strong evidence of localized knowledge spillovers when we consider median TFP of the neighboring plants as a measure of external knowledge. The coefficient estimates for median TFP of the immediate neighbors located in the first shell-location ($\widehat{\Theta}_{it}^{1}$) are positive and statistically significant. A 10% higher median TFP of the closest neighbors is associated with 0.8% higher future TFP. The relatively small values of $\widehat{\Theta}_{it}^{1}$ compared to the coefficient of the internal knowledge ($\hat{\theta}_{it}$) indicate that the latter is a much more important factor in the process of knowledge-evolution.[7]

Table 4: Estimated coefficients for the exit-rule equation using all sectors data

Regressors	Baseline	Model 1A	Model 1B	Model 2A	Model 2B
N_{it}	-0.615*** (0.032)	-0.626*** (0.034)	-0.636*** (0.034)	-0.617*** (0.052)	-0.625*** (0.052)
a_{it}	0.142*** (0.011)	0.134*** (0.012)	0.133*** (0.012)	0.111*** (0.016)	0.108*** (0.016)
c_{it}	0.053 (0.040)	0.095** (0.043)	0.100** (0.043)	0.125* (0.068)	0.115* (0.069)
c_{it}^2	0.002 (0.002)	0.001 (0.002)	0.001 (0.002)	0.000 (0.003)	0.000 (0.003)
θ_{it}	0.270*** (0.025)	0.325*** (0.028)	0.324*** (0.028)	0.265*** (0.043)	0.251*** (0.043)
Θ_{it}^1		-0.072 (0.048)	-0.098** (0.048)	0.039 (0.098)	-0.017 (0.099)
$\Theta_{it}^{S,1}$				0.032 (0.063)	0.015 (0.140)
Θ_{it}^2		-0.037 (0.072)	0.004 (0.073)	0.506*** (0.137)	0.557*** (0.140)
$\Theta_{it}^{S,2}$				-0.004 (0.060)	-0.030 (0.060)
Θ_{it}^3		-2.368*** (0.153)	-2.244*** (0.143)	-3.482*** (0.266)	-2.916*** (0.221)
$\Theta_{it}^{S,3}$				-0.058 (0.101)	-0.013 (0.101)
F_{it}^1 or W_{it}^1		0.075*** (0.009)	0.038*** (0.007)	0.044** (0.020)	0.033** (0.014)
$F_{it}^{S,1}$ or $W_{it}^{S,1}$				0.076*** (0.020)	0.032** (0.020)
F_{it}^2 or W_{it}^2		-0.039*** (0.012)	-0.039*** (0.009)	-0.028 (0.024)	-0.053 (0.018)
$F_{it}^{S,2}$ or $W_{it}^{S,2}$				0.011 (0.018)	0.017 (0.013)
F_{it}^3 or W_{it}^3		-0.029 (0.019)	0.019 (0.016)	-0.136*** (0.037)	-0.064** (0.030)
$F_{it}^{S,3}$ or $W_{it}^{S,3}$				-0.023 (0.019)	-0.041 (0.015)
Const	1.186*** (0.261)	0.986*** (0.273)	1.121*** (0.339)	2.255*** (0.583)	2.500*** (0.624)
ρ	-0.238*** (0.023)	-0.214*** (0.025)	-0.217*** (0.004)	-0.233*** (0.035)	-0.246*** (0.034)

*, **, *** = significant at 10%, 5%, and 1% level.
Figures in parenthesis are standard errors.
All regressions include year-, region-, and sector-dummy variables

[7] This finding is in contrast to that of Winston (2001) who found the effect of external knowledge in the Taiwanese electronics industry was almost three times as large as the internal knowledge effect. He attributed his finding to the possibility of omitting the effects of local public goods available to all firms in a given location. Also, instead of annual date, his data is based on census collected every 5 years.

By comparing the magnitude of $\widehat{\Theta}_{it}^{L}$ across *L*, we will be able to determine the extent of knowledge spillovers across the three shells representing increasing measures of physical distance. It is clear from examining the magnitudes of $\widehat{\Theta}_{it}^{L}$ in column 3 that physical proximity to knowledge source matters. In fact, $\widehat{\Theta}_{it}^{1} > \widehat{\Theta}_{it}^{2} > \widehat{\Theta}_{it}^{3}$ implies that knowledge spillover weakens monotonically with increasing distance from plant *i*. The null hypothesis that all these three coefficients equal each other is rejected at the 1% significance level.

F_{it}^{L} is defined as the natural log of the number of plants in the location *L* regardless of sector classification and acts as a proxy for the extent of human interaction or number of opportunities a plant has to combine knowledge and produce new ideas. While $\hat{F}_{it}^{1}$ is not statistically different than zero, the coefficient estimate of F_{it}^{2} is positive, statistically significant and greater in magnitude than $\hat{F}_{it}^{3}$. Thus, in this case the distance effect on the extent of knowledge spillover is less clear than that measured by external knowledge.[8] Perhaps this is because F_{it}^{L} also captures a stronger negative local competition effect from the closest neighbors. Our subsequent specifications attempt to control for such effects by using the number of workers in neighboring firms, separating firms according to industrial sectors, and differentiating firms according to their level of knowledge.

The inclusion of Θ_{it}^{L} and F_{it}^{L} has predictable effects on the remaining parameters of the model. In particular, there is negligible change in the effect of a plant's own productivity on its future productivity, $\widehat{\theta}_{it}$, suggesting little correlation between current TFP and measures of external knowledge stocks and therefore alleviating the concern that the $\widehat{\Theta}_{it}^{L}$ may be picking up the effects of local public goods available to all plants in a given location.

[8] When we collapse the model into two locations by combining locations 1 and 2 and treating them as a single location, our results (not reported in the table) indicate that $F_{it}^{1+2} > F_{it}^{3}$. That is, stronger effects that are also statistically significant (0.004) come from a greater number of neighbours in the inner shells. The magnitude of the effects, however are tiny compared to the effect from the external knowledge spillovers generated by high productivity neighbours.

The corresponding results based on the natural log of the number of workers in the neighboring plants (W_{it}^{L} and $W_{it}^{S,L}$) are reported in column 4 of Table 3. If we compare these results to those reported in the third column of the same table, we can see little difference in terms of $\widehat{\Theta}_{it}^{L}$ for all L. This supports our observation about the robustness of $\widehat{\Theta}_{it}^{L}$ with respect to the presence of other independent variables. The main difference of the two results is in the coefficient estimates for F_{it}^{1} and W_{it}^{1}. While the coefficient estimate of F_{it}^{1} is negative and insignificant, $\hat{W}_{it}^{1}$ is positive and statistically significant. This is probably because number of workers captures less of the negative local competitive effects and/or more of the positive spillover externalities.

To summarize, the estimation results provide evidence of the importance of physical proximity to sources of new knowledge in future productivity. Furthermore, there is some evidence that having more plants to interact with in a closer location has a positive, although very small effect on future productivity.

4.2. Technological Proximity Matters More as Physical Distance Increases

Our next model specification incorporates a test for technological proximity. In order to test the hypothesis that local knowledge spillovers are stronger among plants that share the same 4-digit ISIC industrial sector, we construct the natural log of median TFP of all neighboring plants, $\Theta_{it}^{S,L}$, and the natural log of the number of plants, $F_{it}^{S,L}$, for each location/sector combination and added them to the previous specifications.

The results are displayed in the fifth[9] column of Table 3. Three key features stand out from among the figures. First, $\widehat{\Theta}_{it}^{1}$ is even higher than before (0.101) relative to $\widehat{\Theta}_{it}^{2}$ and $\widehat{\Theta}_{it}^{3}$, reinforcing the importance of physical proximity to neighboring plants with high TFP. Second, while we find evidence of positive intra-sector knowledge spillover in all three locations, those in the outer two shells ($\Theta_{it}^{S,2}$ and $\Theta_{it}^{S,3}$) are larger in magnitude compared to intra-sector spillover in the closest shell ($\Theta_{it}^{S,1}$) as well as the their respective knowledge spillover from being in a specific location (Θ_{it}^{2} and Θ_{it}^{3}). These results indicate that as

[9] The last column of Table 5 summarizes the estimates of the corresponding model based on the number of workers. Since the coefficient estimates of both models are quite similar, we discuss only the results based on the number of firms.

physical distance increases, technological proximity is a more important determinant of knowledge spillovers relative to geographical proximity. Finally, the coefficients on both, F_{it}^{L} and $F_{it}^{S,L}$, the number of plants in each location and each sector/location combination, respectively, are not statistically significant. The exceptions are the number of plants in the closest location bearing a positive coefficient and the number of plants in the same sector in that location bearing a negative coefficient.[10] The magnitudes of both coefficients are very small relative to the coefficients representing median TFP of any given location.

In our final specifications, we take a closer look at the effect of increased opportunities to interact with other plants by breaking down the number of plants in each location, F_{it}^{L} productivity quartiles. If high productivity neighbors are better sources of local knowledge spillovers than low productivity neighbors, then the spillovers associated with the number of high productivity neighbors should be greater than those associated with the number of low productivity neighbors. To test this hypothesis, we construct TFP quartiles, $Q_k F_{it}^{L}$ for *k=1,2,3,4*. For instance if *k=1*, then $Q_1 F_{it}^{L}$ is the number of plants in the same location that fall in the top quartile of the TFP distribution of all plants in that location L[11].

The results are reported in the third column of Table 5. The estimated coefficients of the number of plants in the top one and/or two TFP quartiles are generally positive and significant while those of the bottom-half are either negative and significant or not statistically significant. This is strong evidence that a plant's future productivity is affected by those plants located in the top half of the productivity distributions in each location.

---

[10] The negative coefficient is more consistent with the importance of local competition rather than local knowledge spillovers. For example, it is possible that local plants operating in the same industry are competing for similar factor inputs. This competition would result in higher input costs, lowering TFP, so that if the negative competitive effect from having more neighbours dominates its positive knowledge spillovers effects, accounting for negative estimated coefficient.

[11] We also constructed the corresponding measures based on the number of workers and ran separate regressions. These results are available upon request.

Table 5: Estimates for knowledge-evolution equations - all sector quartiles

Regressors	**Model 1A**	**Model 1A** All-Quartiles
N_{it}	-0.013** (0.006)	-0.006 (0.009)
a_{it}	-0.013*** (0.002)	-0.012*** (0.003)
θ_{it}	0.653*** (0.004)	0.689*** (0.006)
Θ_{it}^{1}	0.079*** (0.008)	0.078** (0.033)
Θ_{it}^{2}	0.056*** (0.0125)	-0.230*** (0.053)
Θ_{it}^{3}	-0.068** (0.027)	-0.221 (0.150)
F_{it}^{1}	-0.002 (0.002)	
F_{it}^{2}	0.005*** (0.002)	
F_{it}^{3}	-0.009*** (0.003)	
$Q_1F_{it}^{1}$		0.011** (0.004)
$Q_2F_{it}^{1}$		-0.002 (0.004)
$Q_3F_{it}^{1}$		-0.004 (0.004)
$Q_4F_{it}^{1}$		0.004 (0.004)
$Q_1F_{it}^{2}$		0.040*** (0.007)
$Q_2F_{it}^{2}$		0.021*** (0.007)
$Q_3F_{it}^{2}$		-0.040*** (0.007)
$Q_4F_{it}^{2}$		-0.008 (0.007)
$Q_1F_{it}^{3}$		-0.024 (0.024)
$Q_2F_{it}^{3}$		0.094*** (0.024)
$Q_3F_{it}^{3}$		-0.099*** (0.024)
$Q_4F_{it}^{3}$		0.007 (0.017)
Const	0.073*** (0.026)	0.090** (0.035)
Observations	34347	18203

*, **, *** = significant at 10%, 5%, and 1% level.
Figures in parenthesis are standard errors.
All regressions include year-, region-, and sector-dummy variables

4.3. Sectoral Variation

Tables 6 and 7 repeat the specifications in models 1 and 2 in Table 3 separately for seven key two-digit ISIC to see if the importance of local knowledge spillovers differs across industrial sectors. The only difference in the regressions in these tables is that we reduce the number of locations from three to two because the number of observations in each location falls significantly in going to the individual sector analysis.

In examining the coefficient estimates for the external TFP variables Θ_{it}^{L} in Table 6, our results indicate that the degree and significance of physical proximity matters in the majority of the sectors under consideration, with $\Theta_{it}^{1} > \Theta_{it}^{2}$. $\hat{\Theta}_{it}^{1}$ is positive and statistically significant in 4 out of the 7 sectors, ranging from .05 in textile and wood products industries to 0.147 in food. Physical proximity does not appear to matter at all in the three industries of paper/printing/publishing sector, the nonmetallic mineral products sector, and fabricated metal products, electronics, machinery and transport equipment sector.

As at the aggregate level, the results of the estimated coefficients on F_{it}^{L} at the level of the individual sectors are either not statistically significant or a fraction of the magnitude of the effect captured by Θ_{it}^{L}. For instance, in the wood products sector, the coefficient on number of plants in the closer physical location (0.013) is larger than its counterpart in the farther location (0.004) but significantly smaller than the coefficient on $\Theta_{it}^{S,1}$ (0.048).

Overall, the sectoral estimation results reinforce the pattern found at the aggregate level although it is clear that the extent of knowledge spillovers and how it is affected by geographical distance does vary across sectors. This finding is consistent with the findings of other studies on knowledge spillovers, such as Anselin, Varga, and Acs (2000) and Audretsch and Feldman (1996b), where the importance of sectoral variations are also documented.

5. Summary and Conclusions

In this paper, we develop empirical measures of knowledge spillovers that are broader then existing measures in order to test the dynamic productivity effects of knowledge stocks based on physical and technological proximity. Instead of focusing on specific sources of knowledge, we follow Winston’s (2002) approach and use TFP to measure a plant’s knowledge.

In our model, knowledge spillovers is determined by the level of knowledge accessed as well as the extent of interactions the plant has with other plant in each location and/or industrial sector. The evolution of a firm’s knowledge is specified as a reduced form equation that estimates a plant’s future TFP as a function of its current TFP, location-specific knowledge stocks and other plant characteristics. The estimation accounts for endogenous plant exit by employing the Heckman selection model.

Table 6: Estimated knowledge-evolution (Model 1A) based on 2-digit sector data, 2-location shells

	Sector							
	All	**31**	**32**	**33**	**34**	**35**	**36**	**38**
N_{it}	-0.013**	-0.010	-0.006	-0.026	0.036	-0.020	-0.031	0.000
	(0.006)	(0.010)	(0.013)	(0.022)	(0.036)	(0.025)	(0.020)	(0.022)
a_{it}	-0.013***	-0.013***	-0.026***	0.009	-0.016	-0.002	-0.022***	-0.002
	(0.002)	(0.003)	(0.004)	(0.008)	(0.010)	(0.007)	(0.006)	(0.007)
θ_{it}	0.646***	0.665***	0.635***	0.573***	0.758***	0.711***	0.494***	0.695***
	(0.004)	(0.007)	(0.010)	(0.016)	(0.022)	(0.088)	(0.014)	(0.014)
Θ_{it}^{1}	0.132***	0.147***	0.050**	0.048**	-0.004	0.112***	-0.008	0.041
	(0.012)	(0.081)	(0.024)	(0.023)	(0.047)	(0.025)	(0.027)	(0.030)
Θ_{it}^{2}	-0.061**	-0.098**	0.038	-0.086*	0.084	-0.046	-0.016	-0.016
	(0.025)	(0.049)	(0.045)	(0.045)	(0.081)	(0.059)	(0.051)	(0.059)
F_{it}^{1}	0.004**	-0.018***	0.001	0.013**	-0.005	0.007	-0.017***	0.020***
	(0.002)	(0.003)	(0.003)	(0.006)	(0.010)	(0.005)	(0.004)	(0.005)
F_{it}^{2}	-0.009***	-0.007*	0.003	0.004	-0.017	-0.002	0.015*	0.017*
	(0.003)	(0.004)	(0.006)	(0.008)	(0.018)	(0.008)	(0.008)	(0.010)
Const	0.071***	0.130***	0.037	-0.093*	0.135	0.002	0.011	-0.136**
	(0.025)	(0.027)	(0.034)	(0.049)	(0.084)	(0.049)	(0.048)	(0.061)
Obs.	36926	12002	6955	4264	1258	3790	3810	2798

*, **, *** = significant at 10%, 5%, and 1% level.
Figures in parenthesis are standard errors
All regressions include year-, and region-dummy variables

Our results indicate that the extent of knowledge spillovers among Indonesian manufacturing plants is monotonically decreasing with geographical distance. Furthermore, we also find that as distance increases, the extent of knowledge spillovers becomes more dependent on the technological proximity. Finally, we find sectoral variations in the relationship between knowledge spillovers and both measures of proximity. Together these findings are consistent with policies that promote clustering for certain manufacturing sectors.

Our paper can be extended in a couple of ways that may sharpen the role of physical distance, technological similarity and human interactions in determining the magnitude of knowledge spillovers. Using more detailed location information or finer disaggregation of sectors may enable us to obtain an improved measure of physical and technological distance between any two plants, which can then be used to weight each plant's knowledge contribution to the other.

Since knowledge is measured as an index of TFP, the methodology used here does not require data of specific sources of knowledge, such as R&D expenditures or direct foreign investments, information that is often not available or unreliable for many developing countries. Consequently, the model can be applied to a wide range of micro-level data sets from countries where knowledge spillovers may not only exist, but have important economic effects despite the absence of information about the extent of formal foreign or R&D investments.

Table 7: Estimated knowledge-evolution (Model 2A) based on 2-digit sector data, 2-location shells

Regressor	Sector							
	All	**31**	**32**	**33**	**34**	**35**	**36**	**38**
N_{it}	-0.013**	-0.022**	-0.003	-0.018	0.013	-0.008	-0.018	-0.003
	(0.006)	(0.011)	(0.013)	(0.018)	(0.037)	(0.031)	(0.020)	(0.026)
a_{it}	-0.013***	-0.011***	-0.027***	0.010	-0.022**	-0.005	-0.013*	-0.004
	(0.002)	(0.003)	(0.004)	(0.007)	(0.010)	(0.009)	(0.007)	(0.008)
θ_{it}	0.646***	0.627***	0.635***	0.571***	0.732***	0.699***	0.497***	0.702***
	(0.004)	(0.009)	(0.011)	(0.014)	(0.022)	(0.014)	(0.016)	(0.017)
Θ_{it}^{1}	0.132***	0.055**	0.106***	0.091**	-0.029	0.092**	-0.023	0.048
	(0.012)	(0.023)	(0.037)	(0.043)	(0.077)	(0.039)	(0.040)	(0.055)
$\Theta_{it}^{S,1}$	0.132***	0.084***	-0.026	-0.021	0.047	0.023	0.042	0.001
	(0.012)	(0.013)	(0.025)	(0.035)	(0.058)	(0.024)	(0.028)	(0.027)
Θ_{it}^{2}	-0.061**	-0.298***	0.016	-0.143**	0.236**	-0.127*	0.055	-0.054
	(0.025)	(0.061)	(0.057)	(0.064)	(0.120)	(0.077)	(0.084)	(0.081)
$\Theta_{it}^{S,2}$	-0.061**	0.132**	0.027	-0.012	-0.082	0.056*	-0.112***	-0.005
	(0.025)	(0.019)	(0.036)	(0.042)	(0.094)	(0.031)	(0.039)	(0.034)
F_{it}^{1}	0.004**	-0.018	0.010*	-0.011	-0.007	0.009	0.030***	0.023**
	(0.002)	(0.005)	(0.006)	(0.012)	(0.017)	(0.009)	(0.011)	(0.009)
$F_{it}^{S,1}$	0.004**	-0.015***	-0.013**	0.016	0.017	-0.007	-0.038***	-0.013
	(0.002)	(0.003)	(0.005)	(0.011)	(0.017)	(0.011)	(0.009)	(0.010)
F_{it}^{2}	-0.009***	-0.003	0.003	-0.000	0.003	-0.001	0.035***	0.027
	(0.003)	(0.004)	(0.007)	(0.012)	(0.021)	(0.015)	(0.008)	(0.017)
$F_{it}^{S,2}$	-0.009***	-0.013***	0.004	-0.003	-0.004	0.001	-0.019***	-0.022**
	(0.003)	(0.004)	(0.005)	(0.009)	(0.016)	(0.010)	(0.008)	(0.010)
Const	0.071***	0.116***	0.035	-0.028	0.044	0.006	-0.098	-0.117**
	(0.025)	(0.033)	(0.034)	(0.056)	(0.080)	(0.073)	(0.057)	(0.079)
Observations	36926	9846	6162	3652	1095	2942	3284	1996

*, **, *** = significant at 10%, 5%, and 1% level.
Figures in parenthesis are standard errors.

All regressions include year, region, and sector dummy variables.

Appendix 1 Data Description

The empirical estimation of this paper is based on annual plant-level census data of all medium and large plants in the Indonesian manufacturing sector.[12] The focus of our analysis is from 1990 to1995, a period during which the country embarked on an export-oriented industrialization strategy. This annual data set contains detailed information on plant characteristics, expenditures on various inputs, output and other revenues. More importantly, the data set allows us to track each individual plant's performance from year to year and thus, allows us to estimate the exit-decision equation.

Overall, our raw data set contains annual observations of plants for the six-year period. For instance, there are 16,030 and 21,714 records of manufacturing plants in 1990 and 1995, respectively. Several data cleaning steps were performed to ensure that our empirical analysis is consistent with fundamental theories of production. First, all records with zero or negative value of output were eliminated. Similarly, all records with zero or negative factor input values were dropped. Finally, records of plants with sales of more than 75% of their output in industrial services to other plants were also eliminated. The last step avoids counting of factor inputs usage twice, since those plants are primarily subcontractors whose factor inputs are provided by plants using their services.

As a result of the cleaning, we are left generally with about fifty percent of the number of observations present in the raw data. For instance, in 1990, the cleaned data set contains 7240 records, forty five percent of the total records in the raw data set.[13] Appendix Tables 1 and 2 provide the breakdown of the 1990-1994 average number of plants and values of plant characteristics by province and by industrial sector.[14] Appendix Figure 1 provides a map of the provincial distribution of the manufacturing plants in the study in 1992.

---

[12] Defined as plants with total employment of at least 20 workers.

[13] Data for 1995 is not included in Appendix Table 1 since in order to construct the *t+1* measures for that year, we need to use data for 1996 which we do not have.

[14] During the sample period, Indonesia consisted of 27 provinces, including East Timor. However, as of May 2002, East Timor has become an independent country. In addition, three of the remainder provinces (Maluku, West Java, and Riau,) have, since 1999, 2000, and 2002, respectively, been split into several provinces resulting in a total of 30 provinces.

Appendix Table 1: Plant characteristics by location

Provinces	**Plants**	**Age** (years)	**Labor** (workers)	**Capital** (log)	**TFP** (log)
Aceh	36	13	193	12.9	0.044
North Sumatra	446	16	160	12.4	0.002
West Sumatra	45	15	112	11.7	0.122
Riau	72	9	416	14.0	0.057
Jambi	47	10	340	12.9	0.162
South Sumatra	82	6	155	12.4	-0.011
Bengkulu	2	14	437	12.5	0.108
Lampung	66	14	189	12.6	-0.037
Jakarta	604	14	196	13.1	0.090
West Java	2077	14	216	12.8	-0.022
Central Java	1414	20	107	11.2	-0.110
Yogyakarta	135	20	82	11.8	-0.097
East Java	1965	16	139	11.7	-0.098
Bali	159	11	75	11.7	-0.033
West Nusa Tenggara	63	12	55	10.8	0.019
East Nusa Tenggara	6	18	32	12.1	-0.221
East Timor	6	22	27	10.9	0.017
West Kalimantan	57	13	476	13.9	0.117
Central Kalimantan	30	10	237	13.3	0.160
South Kalimantan	83	11	280	12.9	0.096
East Kalimantan	51	13	523	14.1	0.128
North Sulawesi	21	9	131	11.9	0.056
Central Sulawesi	15	13	83	11.7	0.188
Southeast Sulawesi	109	14	95	11.7	-0.129
South Sulawesi	25	9	34	10.4	0.192
Maluku	8	10	678	14.3	0.108
Irian Jaya	12	9	367	14.1	-0.123
Total	7637	16	168	12.2	-0.042

All figures are rounded average across 1990-1994.

To construct the index of total factor productivity $\ln TFP_t^i$ defined in the text we need to construct output, input, and cost-share variables for each plant-year observation. The value of plant output is measured as the sum of total revenues from sales, repairing and fixing services, the revenue from performing subcontracted work, and the change in inventory of final goods between the beginning and end of the year. The value of output is deflated by a producer price index defined at the 2-digit industry level.

Appendix Table 2: Plant characteristics by industrial sector

Sector	Plants	Age (years)	Labor (workers)	Capital (log)	TFP (log)
31 Food, beverages, and tobacco	2491	18	102	11.5	-0.097
32 Textile, garments, and leathers	1463	15	236	12.1	0.043
33 Woods and wood products	951	10	246	12.6	0.033
34 Paper, printing and publishing	293	18	161	13.0	-0.010
35 Chemical, rubber and petroleum products	825	16	232	13.4	0.069
36 Nonmetallic mineral products	826	17	81	11.4	-0.156
37 Metals (iron, steel, nonferrous)	36	10	159	13.9	0.083
38 Fabricated metal products	617	14	179	13.0	0.032
39 Others	135	13	202	12.2	0.020
Total	7637	16	167	12.2	-0.042

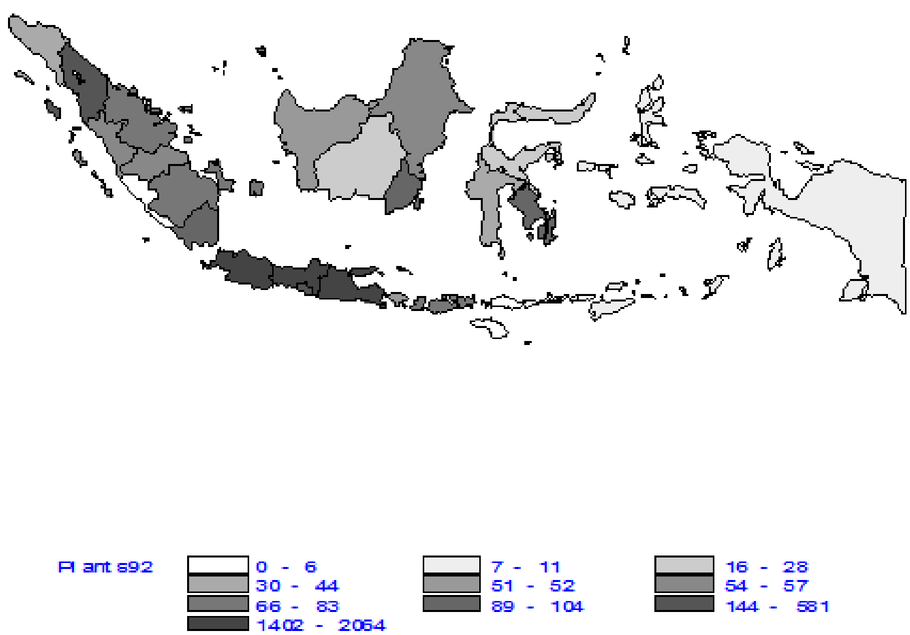
All figures are rounded average across 1990-1994

In our analysis, each producer uses four inputs in production: labor, capital, intermediate materials, and subcontracting services. The labor input is measured as the number of production and non-production workers. Total payments to labor are measured as total salaries to both groups (unfortunately, these do not include fringe benefits and pensions). The cost share of labor is the ratio of total payments to labor to the value of plant output.

The capital input is estimated as the book values of tangible assets, including building, machinery, tools, and transport equipment at the beginning of the year. To control for price level changes in new capital goods, using the 1988 book values (1986 in Taiwan) as the basis, we deflate the changes in each plant's book values between the censuses by the producer price indices for capital goods. By adjusting these deflated changes to the 1988 book values, we scale the 1983 and 1993 (1981 and 1991 in Taiwan) book values of capital goods to the 1988 basis. The cost share of capital is measured as the residual after subtracting the shares of labor, material, and subcontracting services.

The material input includes raw materials, fuel, and electricity used by the plant. Expenditures on raw materials are deflated by the producer price index for manufacturing raw materials. Fuel expenditures are deflated by an energy producer price index, and electricity expenditures are deflated by an electricity producer price index. The cost share of materials is the ratio of total expenditures on intermediate materials to the value of plant output. The cost incurred for the work that is subcontracted out to other plants is included as an input expenditure since it comprises the principle's payments to subcontractors for the labor, capital services, and expenditures on fuel and electricity by the latter. These costs are deflated by the producer price index of the industry to construct a subcontracting input. The cost share of subcontracting services is the ratio of the principle's payments to the plant's output value.

Appendix Figure 1: The Distribution of Indonesian Manufacturing Plants in 1992.



References

Acs, Zoltan J.. David B. Audretsch and Maryann P. Feldman (1992), “The Real Effects of Academic Research: Comment,” *American Economic Review*, 82, pp. 363-367.

Adams, J. D. and A. B. Jaffe (1996), "Bounding the Effects of R&D: An Investigation Using Matched Establishment-Firm Data," *RAND Journal of Economics*, 27 (4), 700-721.

Aghion, P. and P. Howitt (1998), *Endogenous Growth Theory*, Cambridge, Massachusetts: The MIT Press.

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

Anselin, L., A. Varga, and Z. Acs (2000), "Geographical Spillovers and University Research: A Spatial Econometric Perspective," *Growth and Change*, 31 (4), 501-515.

Audretsch, D. B. and M. P. Feldman (1996a), "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86, 630-640.

Audretsch, D. B. and M. P. Feldman (1996b), "Innovative Clusters and the Industry Life Cycle," *The Review of Industrial Organization*, 11, 253-273.

Aw, B. Y., S. Chung, and M. J. Roberts (2000), "Productivity and Turnover in the Export: Micro-level Evidence from the Republic of Korea and Taiwan," *World Bank Economic Review*, 14(1), pp. 65-89.

Aw, B. Y., A. Palangkaraya, and M. J. Roberts (2001), "Export Market Turnover and Selection: Evidence from Indonesian Panel Data," Department of Economics Working Paper, The Pennsylvania State University.

Barro, R. J. and X. Sala-i-Martin (1995), *Economic Growth*, New York: McGraw-Hill, Inc.

Bernard, A. B., and J. B. Jensen (2002), “The Deaths of Manufacturing Plants,” Department of Economics Working Paper, Dartmouth University.

Beugelsdijk, S. and M. Cornet (2001), "How Far Do They Teach? The Localization of Industrial and Academic Knowledge Spillovers in the Netherlands," Center Discussion Paper No. 2001-47.

Caves, D. W., L. Christensen and E. Diewert (1982), "Output, Input, and Productivity Using Superlative Index Numbers," *Economic Journal*, 92, 73-96.

Clerides, S., S. Lach and J. R. Tybout (1998), "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico and Morocco," *Quarterly Journal of Economics*, 113 (3), 903-947.

Caniels, M. C. J., and B. Verspagen (2001), "Barriers to Knowledge Spillovers and Regional Convergence in an Evolutionary Model," *Journal of Evolutionary Economics*, 11 (3), 307-29.

Feenstra, Robert (1996), "Trade and Uneven Growth," *Journal of Development Economics*, Vol. 49 (1), pp. 229-256.

Good, D. H., M. I. Nadiri and R. Sickles (1996), "Index Number and Factor Demand Approaches to the Estimation of Productivity," in H. Pesaran and P. Schmidt (eds.) *Handbook of Applied Econometrics: Microeconometrics, Vol.II.* Blackwell, Oxford.

Griliches, Z. (1992), "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, 94, 29-48.

Hopenhayn, H. (1992), "Entry, Exit and Firm Dynamics in Long-Run Equilibrium," *Econometrica*, 60, 1127-50.

Krugman, P. (1991), *Geography and Trade*, Cambridge, Massachusetts: MIT Press.

Marshall, A. (1920), *Principles of Economics*, 8th Ed, London: Macmillan.

Mills, E. S. (1992), "Sectoral Clustering and Metropolitan Development," in Edwin S. Mills and John F. McDonald (eds.), *Sources of Metropolitan Growth*. New Brunswick: Rutgers University Center for Urban Policy Research.

Jaffe, A. B. (1986), "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value," *American Economic Review*, 76 (5), 984-1001.

Jaffe, A., M. Trajtenberg and R. Henderson (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 557-598.

Sjoholm, F. (1999), "Productivity Growth in Indonesia: The Role of Regional Characteristics and Direct Foreign Investment," *Economic Development and Cultural Change*, Vol. 47, pp. 559-584.

Winston, T. (2001), "Local Knowledge Spillovers in the Taiwanese Electronic Industry," Department of Economics Working Paper, The Pennsylvania State University.

Zucker, L. G. and M. R. Darby (2001), "Capturing Technological Opportunity via Japan's Star Scientists: Evidence from Japanese Firms' Biotech Patents and Products," *Journal of Technology Transfer*, 26 (1-2), 37-58.