Division of Labor, Market for Technology and Investment

Ashish Arora
Heinz School, Carnegie Mellon University, Pittsburgh

Andrea Fosfuri
Universidad Carlos III, Madrid, Spain

Alfonso Gambardella
University of Urbino, Italy

First Version: July 1996
This Version: February 1998

We thank the TEG Program at CEPR, Stanford University, for its hospitality during the early stages of the preparation of this paper. We are grateful to Dan Black, Zvi Griliches, Shane Greenstein, Patrick Hiron, Adam Jaffe, Sam Kortum, Ralph Landau, Jim Markusen, Nathan Rosenberg, Manuel Trajtenberg, Bill Vogt and participants at the applied economics seminars at the University of Pittsburgh, CMU, NBER, Univesidad Carlos III, Universitat Pompeu Fabra, University of Leuven and University of Urbino for helpful comments. Zhang Jianyu provided competent research assistance. Fosfuri and Gambardella also thank the Italian Ministry of University and Scientific Research (MURST), and CNR (Committees 10 and 11) for financial support. Financial support from the European Union, DG XII, TSER Project N.SOE1-CT97-1059, is also gratefully acknowledged. The customary disclaimers apply.

Abstract

Division of labor has benefits in addition to the direct productivity increases from specialization: In this paper we study how the development of specialized upstream technology suppliers improves access and lowers investment costs for latecomer downstream firms. Using a simple model, we test this idea using comprehensive data on investments in chemical plants in less developed countries (LDCs) during the 1980s. We find that investments in chemical plants in the LDCs are greater, the greater is the number of technology suppliers that operate in the first world. A major contribution of this paper is that we identify an important and understudied mechanism through which technology is made available, with beneficial effects for investment.

1. Introduction

That specialization and division of labor among firms promote economic welfare is almost an article of faith for economists. The legacy from Smith (1776) and Stigler (1951) is that when the optimal scale of production of an input is larger than the optimal size of the user firms, a specialized supplier can produce the input more efficiently because it can spread the fixed cost over a larger volume of output. Hence, division of labor is more extensive in larger markets. In turn, a more efficient production of the input induces the downstream industry to grow more rapidly (Young, 1929).

But while the literature has typically focused on the vertical division of labor within an industry, this paper argues that an important and almost unexplored part of the story lies in the benefits that a division of labor in one market can generate for other markets. Specifically, we focus on how division of labor in one country has beneficial effects for other countries. We argue that these benefits are likely to be important when the upstream input requires a large fixed cost but is highly 'portable' and can be transported cheaply across space. Technology is such an input *par excellence*. Thus, if technology suppliers arise in one country first, then, once the technology is developed, these technology suppliers can sell it to other countries at a small incremental cost (especially if compared with the cost of developing the technology in the first place). In this way, follower countries benefit from investments in leading countries.

To examine this issue we study investment in chemical plants in less developed countries (henceforth LDCs) during the 1980s. The chemical industry provides an ideal test-bed for the question that we want to examine here. Beginning in the 1930s and continuing into the 1960s, the modern chemical industry in the developed countries (henceforth 'first world') grew rapidly . This stimulated the growth of firms that specialized in the design and engineering of the chemical processes, the so-called 'specialized engineering firms' (SEFs), which are analogous to the software

engineering and computer consultancy firms that are more visible in the US economy today. SEFs have been important reservoirs of expertise in chemical plant technologies, which they provided in the form of engineering services to chemical firms. In some cases SEFs also developed radical process innovations but for the most part they sold improved versions of existing technologies packaged with engineering and design services. Other SEFs offered construction services in addition to engineering know-how. In the 1970s, and especially in the 1980s, as a modern chemical industry emerged in LDCs, it benefited from the presence of an upstream sector of technology suppliers in the first world. Simply put, the growth of the chemical industry in the first world created an upstream sector, which later spurred the growth of the chemical industry in the developing countries.[1]

As a guide to our empirical results, we develop a simple model. The model assumes that a larger number of technology suppliers increases the net surplus that the buyers get from investing in a chemical plant. This is a natural assumption since buyers should benefit from being able to choose from a larger pool of suppliers, and this is consistent with a large set of economic explanations that variously emphasize search costs, reduced bargaining power of sellers, and a better 'match'. The main result of the model is that if the existing SEFs in the first world are also potential suppliers of chemical firms in LDCs, then a larger number of first world SEFs in a given market for chemical process technology implies greater investment in that market in the LDCs. From the point of view of the LDCs, the number of potential suppliers (first world SEFs) is determined by the extent of division of labor in the first world. Thus, the organization of the industry in the first world, or to be precise, the extent of division of labor in the first world, enhances the growth of the market (chemical industry) in the LDCs.

[1] The story of SEFs and their role in the growth of the chemical industry is told in greater detail in Arora and Gambardella (1998). See also Freeman (1968).

The model also predicts that the larger the number of first world SEFs, the greater the number of plants in LDCs whose engineering services are 'bought' from SEFs, and the smaller the number of plants whose engineering services are 'made' in-house by the chemical firms. Moreover, the impact of an increase in the supply of SEFs are more pronounced for companies that have higher cost of 'making' the technology in-house. This suggests that SEFs are more beneficial for local third world companies than for the multinational enterprises that locate in these markets.

We test these propositions using data on chemical plant investments in 139 leading chemical technologies and 38 LDCs. These are drawn from a comprehensive data set of more than 20000 chemical plants announced and constructed during the 1980s worldwide.

The next section provides the conceptual underpinnings of our approach and links it to the literature on technology and international trade. Section 3 presents our model. Section 4 discusses the empirical specification and presents our empirical results. Section 5 summarizes our findings and concludes the paper. The appendix describes our data set.

2. Division of labor, markets for technology, and international trade

This paper is related to several strands in the literature. Our emphasis on the lower cost of using technologies compared with the cost of developing them is similar in spirit to the literature on endogenous growth (e.g. Romer, 1990 and 1996). Following Griliches (1979) and Jaffe (1986), there is a well established literature on technology spillovers, and Coe and Helpman (1995) have provided evidence of international technological spillovers (see also Eaton and Kortum, 1996, and Keller, 1997).

This paper does not examine the effects of spillovers on total factor productivity, as many of the papers in this literature do, but on the flow of investments, because we do not have measures of the former at this disaggregate level. However, our paper moves beyond the simple observation that

technological spillovers exist, and identifies a mechanism through which they take place – namely, the intermediation of an upstream sector which sells technology inputs through a full-fledged market for technology.[2] Note that both real and pecuniary effects may be involved. SEFs may imply lower technology transfer costs for LDCs. In addition, as already discussed, a larger number of SEFs results in LDC firms paying a lower price for technology and engineering services, a purely pecuniary effect.[3] Rich as our data are, we still cannot empirically distinguish between the two mechanisms because the net effect in both is to increase investment in LDCs. Thus, although our formalism focuses on the pecuniary externality, real externalities are likely to be involved as well.

The nature of the mechanism we study is similar to the one highlighted by the recent literature on 'general-purpose technologies', or GPTs (Bresnahan and Trajtenberg, 1995; Helpman and Trajtenberg, 1996). GPTs have high fixed costs, but they can be applied repeatedly at low incremental cost. This leads to spillovers across industries in which the growth of one user sector benefits other user sectors through the intermediation of the GPT industry. Our paper shows that this argument can be extended to the transmission of growth across countries, and it provides empirical evidence of these mechanisms.

It is natural to ask at this point whether our story is not simply a story about international trade. It is, but with one difference. While the standard Heckscher-Ohlin trade model locates comparative advantage in natural resources or factor endowments, we locate it in the fact that chemical engineering services are based on cumulative learning and experience, and that the (fixed) costs of acquiring this expertise are already sunk when the LDC markets arise. Put differently, the first world has a comparative advantage in engineering services simply because first world

[2] In other words, some of what appear to be "spillovers" may in fact be market mediated knowledge flows. In a recent paper, Zucker, Darby and Armstrong (1998) find that in biotechnology market mechanisms involving individual scientists may account for what appeared to be knowledge spillovers from universities to local firms.

engineering firms were founded more than 50 (and in some cases, more than a 100) years ago in response to the growth of the oil and chemical sectors in their own countries.

By stressing the historical sequence in the rise of new markets, our perspective also differs somewhat from Rivera-Batiz and Romer (1991), who argue that international economic integration increases growth because, with integration, the fixed cost of producing 'ideas' can be spread over a larger market. In our model, integration is beneficial to the follower countries even though the number of SEFs in the first world does not increase when the first world and LDC markets are integrated. Although analytically convenient, our primary reason for assuming that SEFs do not increase in response to LDC growth is that it is more faithful to history. As a historical fact, most SEFs arose to serve the first world market and their investments were not motivated by the hope of serving LDC markets that did not as yet exist.

Finally, our paper is related to the literature on the product life cycle (e.g. Vernon, 1979), whereby as technologies or products mature, they are transferred from the first world to LDCs. This literature has focused on one mechanism through which this transfer occurs – multinational enterprises (MNE) operating in final product markets. This has often been justified by the assumption that while downstream products are tradable, upstream inputs are not (e.g., Nadiri, 1993, Rodriguez-Clare, 1996). In fact, in the chemical industry as in many other high-tech industries, a key upstream input—intangible knowledge and expertise—is easier to move across locations, while the final products (chemicals such as ammonia and ethylene) are costly to transport. MNE are undoubtedly an important vehicle for technology transfer and for the growth of the host countries. However, when technologies are based on systematic body of knowledge like chemical engineering, MNE are not the only, or even the most efficient, way of transferring technology.

---

[3] Formally, if first world SEFs rationally anticipated the rise of the LDC market, and could practice dynamically optimal pricing, these pecuniary externalities would disappear. Such pricing strategies would require not merely extraordinary foresight but also extensive coordination between SEFs.

Instead, as in the chemical industry, specialized technology suppliers competing amongst themselves can be the predominant means of technology transfer.

3. The Theoretical framework

3.1 Description of the model

Figure 1 summarizes the two effects that we highlight in our model.

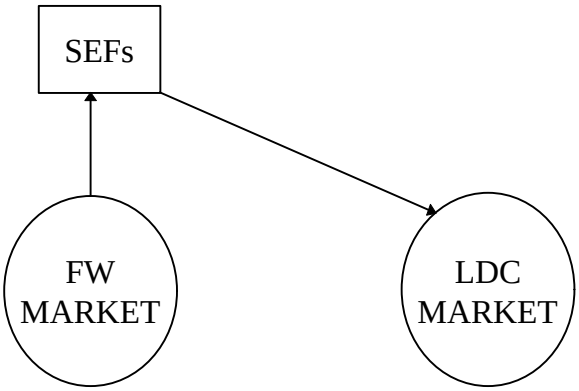


Figure Error! Unknown switch argument.**: The transmission of growth impulses**

First, the growth of the first world market for a given chemical process encourages the rise of engineering firms specialized in the design of chemical plants for that process. This result, which we do not formally prove here, is completely intuitive. It only requires that entry as an SEF have a fixed cost (corresponding to the cost of acquiring technical expertise), and that the price-cost margins (profits per unit of output) that SEFs earn, decline with the number of SEFs in that sector.

The second effect is from the SEFs in the first world to the size of the LDC market. To understand this effect suppose that first world SEFs could supply LDCs. Then, apart from relying on multinationals, either LDC firms would have to provide the services themselves or rely on domestic SEFs that may exist. In either case, LDC firms would face very high costs. Having fewer SEFs to choose from increases search costs, lowers the bargaining power of the buyer, reduces the likelihood of getting a more advanced or appropriate technology. As a result, the expected surplus

of setting up a plant would be lower, and this implies fewer investments in chemical plants. Given the high transportation costs for many chemical products, this would imply slower growth of chemicals, and industrial activity more generally in LDCS.

This simple story relies strongly on the assumption that the critical input is easily 'tradable' across countries. It is then important to understand why is this input tradable. Even though an ammonia plant in the US is a different object from an ammonia plant in India, what remains unchanged are the basic principles of how an ammonia plant should be designed and engineered. Clearly, applying what one has learnt in one place in another is not always easy, and the literature has shown that technology transfer is not costless (e.g. Teece, 1988). The important point, however, is that the transfer cost be substantially smaller than the cost of developing the technology, an assumption that fits especially well in the context of engineering services. It is in this sense that the fixed cost of developing or inventing the technology is paid by the industries or countries that emerge earlier, while the industries or countries that come later pay only the marginal cost.

This also points to the importance of independent suppliers that do not produce the downstream product.[4] Downstream producers (chemical firms) are less likely to sell technology or other key inputs to other producers. Unlike upstream specialists (SEFs), they have to offset the gains from selling technology against the loss in actual or potential revenues from selling the downstream product. Thus, in addition to the classical gains from productivity improvements, specialization and division of labor can have other benefits for industrial growth sometimes overlooked.

[4] Such specialized suppliers are also present in other industries, such as steel and power plants, but the division of labor there is not as extensive and widespread as in chemicals. As discussed elsewhere, the nature of the chemical technology, the development of chemical engineering as a discipline, and events in the history of the industry are important reasons why SEFs are more prominent in chemicals. (See Arora and Gambardella, 1998.) Note also how the story of SEFs parallels that of the extensive division of labor that we find today in many high-tech industries like software or biotechnology.

3.2 The model

We develop a simple stylized model to motivate our empirical analysis which focuses on the investment decision of chemical firms in LDCs. We posit that the first world market for chemical process technology has already emerged and a division of labor achieved. In other words, we take as given the number of SEFs, *k*, which have entered the market for engineering services in the first world.[5] We assume that first world SEFs can costlessly supply LDCs, with technology-specific costs already sunk.[6] Therefore, the number of SEFs which can potentially serve a LDC market is equal to *k*. Note that *k* is independent of the LDC market size.[7] We also assume that SEFs are ex-ante symmetric.

Let us now turn to the LDC market. We do not explicitly model downstream competition (i.e. competition in the chemical product market). Final demand generates a potential demand of inputs — chemical plants—that we take as given. We denote by N the exogenous potential demand for chemical processes and assume for simplicity that the investment decision of a given chemical company in a given plant is independent of the investment decision in any other plant, even by the same user. When considering whether to invest in a new plant, each downstream firm has three choices. It can either not invest (and hence earn zero), or it can 'buy' engineering services from SEFs, or it can 'make' (engineer) the plant by itself. We denote by $\Omega = \{B, M, f\}$ the set of possible alternatives where *B* stands for a plant bought from the upstream technological sector, *M*

[5] In an earlier working paper (Arora, Fosfuri and Gambardella, 1996) we treat *k*, the number of first world SEFs, as endogenous. By modeling the upstream sector as competing on price, with a fixed entry cost and stochastic variable cost of engineering a plant, we show that *k* increases with the size of the market for engineering services in the first world.

[6] This could be easily generalized where SEFs have to bear two types of fixed costs: product-specific, which are already sunk, and country-specific, which have to be incurred for each country in which SEFs want to penetrate. See our working paper for further details (Arora, Fosfuri and Gambardella, 1996).

[7] Strictly speaking, this is true only if the size of the LDC market is too small to induce further entry of SEFs. Also, we are assuming that first world SEFs are not fully forward-looking. Else they would anticipate that a given LDC market would arise in the future and they would adjust their optimal investment (entry) decision. As noted earlier, neither assumption is formally necessary for LDCs to benefit from first world SEFs.

for in-house engineering and f for no investment. Also, let $S = B \bigcup M$ where S denotes the observed size of the LDC market.

Let p_n be the net surplus to a given buyer (chemical firm) from a plant supplied by the nth SEF. We assume that p_n is ex-post idiosyncratic to the buyer-seller match. Therefore, p_n is an *i i d*. random variable with distribution function *G(.)*. Also, let $z = \underset{n}{Max}\, p_n$. Then, $\Pr(z \le t) = G^k(t)$, is the probability that the net surplus from running a plant bought from the upstream sector is less or equal than *t*. Finally, for any chemical firm let *a* be the net surplus from an in-house technology and let *a* be distributed according to the cumulative function *F(.)*.

We now analyze the investment decision of chemical firms in LDCs. By using the notation introduced so far, if $Max\{z, a\} \ge 0$, the firm will invest in the plant. Moreover, if $z \ge a$ the firm will buy the engineering services; and if $z < a$ it will supply them in-house. We show first that the greater the number of SEFs, the greater is the number of plants that are constructed. The expected number of plants, *E(size)*, in a given LDC market is:

$$E(size) = N\left[\Pr(\Omega = S)\right] \qquad (1)$$

where $\Pr(\Omega = S) = 1 - F(0)G^k(0)$ (2)

Result 1*: The total investment in a given LDC market, E(size), is greater the greater is the number of (potential) technology suppliers, k.*

<u>*Proof*</u>. By taking first differences with respect to *k* in expression (2) we get:

$$E(size|k+1) - E(size|k) = N\left\{F(0)G^k(0)\left[1 - G(0)\right]\right\} > 0 \qquad (3)$$

which suffices to show Result 1.■

It is also easy to see from (3) that E(*size*) is concave in *k*, so that the marginal increase in investment diminishes with *k*. We next show that SEFs increase the probability of 'buy' and

decrease the probability of 'make'. One can write the expected number of plants that are bought from the upstream technological sector, *E(buy)*, and the expected number of plants that are made in house by the chemical firms, *E(make)*:

$$E(buy) = N\left[\Pr(\Omega = B)\right] \tag{4}$$

$$E(make) = N\left[\Pr(\Omega = M)\right] \tag{5}$$

where $\Pr(\Omega = B) = 1 - F(0)G^k(0) - \int_0^\infty G^k(t)dF(t)$ (6)

and $\Pr(\Omega = M) = \int_0^\infty G^k(t)dF(t)$ (7)

Now, by writing $\frac{\Delta E(buy)}{\Delta k} \equiv E(buy|k+1) - E(buy|k)$ and similarly for $\frac{\Delta E(size)}{\Delta k}$ and $\frac{\Delta E(make)}{\Delta k}$ we can state our second result:

Result 2: $\frac{\Delta E(buy)}{\Delta k} > \frac{\Delta E(size)}{\Delta k} > 0 > \frac{\Delta E(make)}{\Delta k}$

<u>*Proof*</u>. First note that $\frac{\Delta E(buy)}{\Delta k} = \frac{\Delta E(size)}{\Delta k} + \frac{\Delta E(make)}{\Delta k}$. Then, by taking first differences in expression (4) with respect to *k* one obtains:

$$\frac{\Delta\left[E(buy)\right]}{\Delta k} = \frac{\Delta\left[E(size)\right]}{\Delta k} - \int_0^\infty \left[G(t) - 1\right]G^k(t)dF(t) > \frac{\Delta\left[E(size)\right]}{\Delta k} \tag{8}$$

which completes the proof.■

Our final result is useful because it develops an implication of the theory that is not obvious and so provides an additional test for the theory itself. Consider any variable *x* which increases the net surplus from running an in-house engineered plant (i.e. increases the profitability of the 'make' strategy). Formally, we assume that $x_1 \geq x_2$ implies $F(t;x_1) \leq F(t;x_2)$ for any *t*. For instance, *x* is

likely to be higher the greater is the level of in-house technological capability of the chemical firm. Thus, MNEs are likely to have a higher value of x than LDC chemical firms. Intuitively, SEFs ought to be more valuable for LDC firms than for MNEs. Our third result formalizes this intuition.

Result 3: $\left.\frac{\Delta\left[E(size)\right]}{\Delta k}\right|_{x=x_1} \leq \left.\frac{\Delta\left[E(size)\right]}{\Delta\ k}\right|_{x=x_2}$ *for all* $x_1 \geq x_2$.

<u>*Proof*</u>. One can rewrite $\left.\frac{\Delta\left[E(size)\right]}{\Delta k}\right|_{x=x_1} - -\left.\frac{\Delta\left[E(size)\right]}{\Delta\ k}\right|_{x=x_2}$ as $G^k(0)[1-G(0)][F(0;x_1)-F(0;x_2)]$

which is unambiguously non-positive. ■

4. Empirical analysis

4.1 Sample and variables

To test our model we used data on 38 LDCs and 139 chemical process technologies. This gave us 5282 'markets', where the unit of observation is a process-country pair. Plant-level data on the 139 chemical processes were obtained from a comprehensive data base, Chemical Age Project File (CAPF), which covers all new chemical plants (over 20,000 in all) announced all over the world during 1980-1989. The CAPF data base and the construction of the variables used in the empirical analysis are described in more detail in the Appendix. Our 38 countries (listed in the Appendix) are all the LDCs for which we could obtain complete data from two main sources: UN Statistical Yearbooks, and Barro-Lee (1994). These countries cover about 80% of all the chemical plants located in LDCs in the CAPF data base. In what follows we define 'first world' to be the Western European countries, the USA, Canada, Japan, Australia, and New Zealand. All the other countries are LDCs.

We constructed the following variables. $SIZE_{ij}$ is the total investment in millions of US dollars in our 139 processes i and 38 countries j. This is obtained by multiplying the number of

plants in *ij* by the average investment cost of a plant in process *i* in all LDCs.[8] Similarly, we construct BUY_{ij} and $MAKE_{ij}$, which are the total dollar investments in plants whose engineering services are bought from an unaffiliated contractor, and the total investment in plants whose engineering services are made in-house (or by an affiliated SEF). DOM_{ij} and MNE_{ij} are the total dollar investments by LDC and multinational firms respectively.[9]

$SIZE_FW_i$ is the total value of investment in process *i* in the first world. This is obtained by multiplying the total number of plants in process *i* in the first world by $COST_FW_i$, the average investment cost of a plant in process *i* in the first world. SEF_FW_i is the number of firms (SEFs) that provide engineering services in process *i* in the first world.

We constructed two other process-specific variables, $NOVEL_i$ and $PROCPAT_i$. They are two measures of the nature of the technology. $PROCPAT_i$ is the total number of 1976-1997 US patents granted for the chemical process *i*. It covers only the patents relating to the process itself rather than to the use of the output produced by the process. $PROCPAT_i$ is a good measure of the complexity of the process technology, and of the potential for multiple inputs, pathways, and final product qualities. Three of our 139 chemical processes were very broad categories, with very large values for $PROCPAT_i$. For these three processes we used a dummy, $DPROCPAT_i$.

$NOVEL_i$ is the growth rate between the two periods 1986-1995 and 1976-1985 of all the US patents whose title contained the exact name of the process *i*. Unlike $PROCPAT_i$, $NOVEL_i$ does not distinguish between process and product patents. Thus, for instance, this variable also includes the development of new uses of the product. $NOVEL_i$ is then likely to be a measure of the rate of technological change. We use these two variables to control for the maturity and complexity of the

[8] Actual investment costs are reported for about 40% of the plants in our data base (see Appendix) and the average cost per plant in a given process was calculated from these observations.
[9] We adjusted for investments by subsidiaries of first world multinationals by including them in MNE_{ij}. See the Appendix.

technology. Our objectiv here is to rule out the possibility that the estimated coefficient of SEF_FW_i reflects the effect that LDCs are more likely to invest in older and more mature processes.

To address the potential endogeneity of SEF_FW_i, we have two instruments—$SLARGECO_i$ and $SGREEN_i$. The former is the share of plant counts by large non-oil companies in process *i* in the first world. These are the top 200 companies ranked in terms of total number of first world plants in the data base (in any process), and excluding the oil companies. $SGREEN_i$ is the share of 'greenfield' plant counts in the first world in process *i*. We discuss the choice of these two instruments in section 4.3 below.

Finally, CAPF classifies plants into 21 chemical sub-sectors, which we grouped in 9 sector dummies (listed in the Appendix). Our country-specific variables include measures like GDP, population, energy consumption, openness, and human capital, and are listed in table 1 along with their source. Table 2 presents descriptive statistics for the variables used in our analysis.

4.2 $SIZE_{ij}$, BUY_{ij}, and $MAKE_{ij}$: Specification and empirical results

We begin our empirical analysis by estimating three equations linking $SIZE_{ij}$, BUY_{ij}, and $MAKE_{ij}$ to the number of first world SEFs. We employed a logarithmic specification of the form

$$\log(X_{ij}) = const + aY_j + bZ_i + e_{ij} \qquad (9)$$

where X_{ij} is $SIZE_{ij}$, BUY_{ij}, or $MAKE_{ij}$, Y_j is a vector of country-specific characteristics, Z_i is a vector of process specific characteristics, and e_{ij} is an error term. We note that the results are robust to alternative specifications such as linear or exponential. A logarithmic specification is justified on several grounds. First, our theory in section 3 suggests that the marginal effect of SEFs on investments diminishes with the number of SEFs. Further, as table 1 shows, SEFs are very often the source of technology services for plants and the average number of SEFs per process is about 12. As Bresnahan and Reiss (1991) have shown, even four to five suppliers in a market may be

enough to make the market competitive. This suggests a functional form that accounts for the diminishing returns of an extra supplier.[10]

As country-specific characteristics we used measures of the economic size of the country ($\log(GDP_j)$, $\log(POP_j)$, $\log(ENERGY_j)$, $\log(AREA_j)$), and other variables -- human capital ($HKAP_j$), openness to imports of intermediate goods ($OPEN_j$), geographic area dummies. We also used dummy variables for the presence in the country of oil or gas reserves and interacted them with the sector dummies for oil refining plants and gas plants ($DOIL_{ij}$ and $DGAS_{ij}$). We also interacted $\log(AREA_j)$ with a dummy for the three sectors inorganic chemicals, agricultural chemicals, and minerals & metallurgy ($ICHEM_i + AGRI_i + MM_i$). This is because larger countries are more likely to possess natural resources which are the basis for production in inorganic chemicals and minerals. Similarly, it is more likely that bigger countries have large areas for extensive agricultural production. Apart from $\log(SEF_FW_i)$, we used the following process-specific characteristics as controls: sector dummies, $\log(SIZE_FW_i)$, $\log(COST_FW_i)$, $\log(PROCPAT_i)$, and $NOVEL_i$.[11]

Table 3 presents the results of our OLS estimations of equation (9) for $SIZE_{ij}$, BUY_{ij}, and $MAKE_{ij}$. The key parameters are the elasticities of the three dependent variables with respect to SEF_FW_i. The estimated elasticities of $SIZE_{ij}$ and BUY_{ij} are respectively 0.46 and 0.52, and they are statistically significant. By contrast, the elasticity of $MAKE_{ij}$ is very small, and insignificantly different from zero. The marginal effects of SEF_FW_i , computed across all observations in our sample, are $3.97 millions for SIZE, $ 4.32 millions for BUY, and a mere $6,600 dollars for

[10] Since our dependent variables can take zero values, we redefined $SIZE_{ij}$ as $1+SIZE_{ij}$, and to keep with the adding up restruction we redefined BUY_{ij} and $MAKE_{ij}$ as $0.5+BUY_{ij}$ and $0.5+MAKE_{ij}$. This is just to make the results of the three equations comparable. Using $1+BUY_{ij}$ and $1+MAKE_{ij}$ did not produce any significant difference. We also tried some alternative models, including a semi-log specification and threshold models with constant or stochastic thresholds. These produced similar qualitative results. The results of these more sophisticated specifications are available from the authors upon request.

[11] Two of the processes had $SEF_FW_i = 0$. For these we set $SEF_FW_i = 1$, and used $\log(SEF_FW_i)$.

MAKE. As predicted by Result 2 in our model, the marginal effect of SEFs is greater for the 'buys' than for the chemical plant investments as a whole, and the effect on the 'makes' is almost zero.

Thus, the important result of table 3 is that the effects of SEF_FW on SIZE and BUY are of sizable magnitude and they are statistically significant in spite of our extensive controls for the size of the chemical process and country markets. In this respect, note also that our measures of market size (e.g., GDP, ENERGY, AREA, SIZE_FW) are generally significant in all three equations. Finally, in all three equations, the cost of a 'typical' plant ($COST_FW_i$), and our measures of technological change and complexity ($NOVEL_i$ and $PROCPAT_i$) have the expected sign. As predicted by the product lifecycle theory, more costly and technological advanced or complex processes are associated with lower LDC investments.

4.3 Measurement error and unobserved heterogeneity: Instrumental variable estimation of $SIZE_{ij}$

Our empirical procedure in the previous section raises two important issues. First, our measure of potential suppliers, SEF_FW_i ignores the possibility that an SEF operating in a certain market could be a potential supplier for a related process. Our assumption is certainly plausible. For instance, our data base shows that three different sets of SEFs supply the markets for three different types of polyethylene—high-density, low-density, and linear low density polyethylene—with very little overlap. If our assumption does not hold, it implies that SEF_FW_i is measured with error. If so, our estimates are likely to be biased towards zero. In turn, this means that our estimates of the effects of SEFs are likely to understate the true impact. As a robustness check, we estimated a specification (not reported here) where we use SEFs operating in an entire sector (e.g., synthetic fibers) in the first world as being potential suppliers for all processes (e.g., polyester and nylon) belonging to that sector. The estimated impacts of SEF_FW_i on $SIZE_{ij}$, BUY_{ij}, and $MAKE_{ij}$

were qualitatively similar to those reported in table 3. Therefore our measure of the supply of SEFs is not likely to be key to our empirical results.

A more important problem is that we are likely to measure N_{ij}, the potential size of the market, with error. If this error is correlated with SEF_FW_i, our estimates will be biased. For instance, if unobserved increases in the size of the LDC process markets induce more SEFs to operate in those process markets in the first world, a typical endogeneity bias arises. In fact, there is both qualitative and quantitative evidence that our estimates are not seriously biased.

First, we already noted that SEF_FW_i has a substantially different effect in the $MAKE_{ij}$ equation vis-à-vis the other two equations. This suggests that the effect of SEF_FW_i on $SIZE_{ij}$ is a genuine one, rather than being created spuriously by unobserved correlations between SEF_FW_i and market size. If SEF_FW_i captured only unobserved market size effects one would expect a more sizable and statistically significant impact on $MAKE_{ij}$ as well. By contrast, the effect on $MAKE_{ij}$ is negligible, while in the same equation measures of market size (particularly $SIZE_FW_i$, GDP_j) are larger and statistically significant.

Evidence from the industry's history provides further support. Most of the SEFs, and virtually all the major ones in the US, Europe or Japan, were founded before World War II or right afterwards, and most chemical process technologies were developed 20 to 40 years ago when the LDC chemical markets were still quite small or non existent—and by that time SEFs had already accumulated considerable expertise in these fields (Arora and Gambardella, 1997). Moreover, first world SEFs maintain strong linkages with their country of origin, and from our database we confirmed that it is rare for SEFs to supply engineering services to LDCs in a process unless they also do so in their home markets.

Nonetheless, it is still possible that the growth of LDC markets led to further accumulation of expertise by SEFs or that LDCs demand prevented some SEFs from exiting the industry. If so,

unobserved variations in LDC demand for chemical plants are correlated with our measure of SEFs. This means that we have to control for such factors.[12]

The standard approach in this case is to find instruments for SEF_FW_i. However, distinguishing among countries according to the extent to which they were open to the inflow of foreign technologies is difficult. Even countries like South Korea that have protected downstream markets have been open to imports of technology and engineering services. Aggregate measures of openness, such as the ratio of exports and imports to GDP, or even the more nuanced measure we use here, $OPEN_j$, cannot capture this subtlety. Further, the rise of West European and Japanese SEFs has made the market for chemical processes a truly global market. Even a country like Libya, which is not directly supplied by US SEFs, is served by a large number of non US SEFs and European subsidiaries of US SEFs. In short, there seems to be insufficient cross country variation in the effective supply of SEFs for this to be a useful way to identify their impact.

By contrast, differences across processes in the nature of the technology has been an important source of variations in the number of SEFs.[13] We have two instruments that exploit this variation—$SLARGECO_i$ and $SGREEN_i$. The first instrument – i.e. the share of first world plants in process *i* by the top 200 non-oil companies – is justified by the fact that the top chemical companies are more likely to have the capabilities to produce their engineering services in-house. The top 200 non-oil companies include virtually all the important chemical companies in the first world. We exclude oil companies because historically the oil companies, including the large ones, are those that have first made systematic use of SEFs. (See Arora, Landau, and Rosenberg, 1998.) Thus, the top oil companies would have a sort of ambiguous effect – on the one hand, they are large enough

[12] We also estimated a lower bound for the estimated elasticity based on the assumption that the unmeasured demand variation had a constant elasticity for different types of investors (LDC firms versus multinationals). Our estimates, available on request, indicate a lower bound to the elasticity of 0.38 of SIZE with respect to SEF_FW, which is very close to the OLS and GMM estimates reported here in table 5.

to produce in-house, but on the other hand they have been more active in outsourcing these services. The second instrument – i.e. the share of first world greenfield plants in process *i* vis-a-vis expansion or revamping of existing plants – has a similar justification. Greenfield plants are likely to be more complex projects than mere expansions or revamping of existing capacities. As a result, they are more likely to require the work of specialist contractors. In sume, we have two instruments that are should be uncorrelated with unmeasured variations in market size but are likely to affect the demand for SEF services in the first world.

Table 4 shows that SEF_FW_i s negatively correlated with $SLARGECO_i$, and positively correlated with and $SGREEN_i$.[14] As a check against the possibility that, since large firms are more likely to move abroad, $SLARGECO_i$ may have a direct impact on investments in LDCs we also estimate a specification where we control for the share of plants by multinational enterprises in process *i* in the first world.[15] The effect of this additional variable was small and insignificant, and it did not affect our estimates of the impact of SEF_FW_i.

Table 5 presents the results of a GMM estimation of the $SIZE_{ij}$ equation using $SLARGECO_i$ and $SGREEN_i$ as instruments.[16] Table 5 also reports the OLS regression for SIZE from table 3 for comparison, and the *chi-square* statistics of a Hausman specification test comparing the OLS and GMM estimates of the elasticity of $SIZE_{ij}$. (See Hausman, 1978, for details).

Table 5 shows that the potential endogeneity bias of SEF_FW_i is quite modest. Our GMM estimator of the elasticity with respect to SEF_FW_i is 0.44, which is very close to our original OLS

---

[13] For instance, some processes are based more than others on standard and codified technologies, which encourages specialization and increases the number of SEFs.

[14] In other specifications, we also used the share of the top 200 oil firms in first world plants as a separate instrument, with little change in the results. We also tried using the share of first world plants by the chemical (and/or oil) companies among the top 50 and top 100 companies in terms of first world plants. Our results are robust to such variations.

[15] Results available on request. Multinational enterprises were defined to be all the companies operating in six or more countries in the first world.

estimate of 0.46. Moreover, based on the Hausman test, one cannot reject the nul hypothesis that the potential endogeneity of SEF_FW_i had no effect on the OLS estimate of this elasticity. The computed *chi-square* statistics is 0.02, which is far smaller than the theoretical *chi-square* values of 3.84 and 6.64 at the 5% and 1% confidence levels and one degree of freedom.

4.4 Assessing the 'marginal' effect of SEFs

To get a sense of the order of magnitude of the impact of an upstream sector of technology suppliers on the total investment of the downstream market, we use our GMM estimate to compute the effect of one additional SEF in a typical process market on the expected total dollar value of investment in the LDCs in that process market. As figure 2 shows, the sample average of the effect of one extra SEF across all our processes and countries is $3.75 millions over 1981-90, with the increase being larger for larger countries like China and India, and smaller in smaller countries. For the 38 LDCs as a whole, the increase in investment in a typical process would then be $142.5 millions, or 4.75% of the average investment per process of about $3 billion in these 38 countries over the ten year period. The average investment cost of a plant in LDCs is about $121 million. Thus, our results imply that for a typical process, one additional SEF will result in slightly more than one extra plant over a ten year period for the 38 LDCs taken as a group. The impact of an additional SEF is statistically significant, but small.

[FIGURES 2, 3, AND 4 ABOUT HERE]

Recall however, that this effect is the average effect over the sample: It varies varies by the size and nature of the process. For instance, figure 3 shows that the effect diminishes with size. Processes where there is a large number of third world plants are likely to be less affected than others. There are two forces at work here. On the one hand, a given percentage increase in

[16] A non-linear GMM procedure, suggested by Mullahy (1997), of the form $SIZE_{ij} = Y^a u + e$, where Y is our vector of regressors, *a* is the set of parameters to be estimated, and *u* and *e* are stochastic error termss, produced an implausibly high estimate of the elasticity.

investment implies a larger dollar increase if the base level investment is high. On the other hand, an additional SEF in the market is more important when the number of SEFs is small than when the number is large. Our results suggests that the "diminishing returns to the number of SEFs" effect dominates when there is a large number of plants. The variation in $SIZE_{ij}$ due to an extra SEF for the typical process with total investment in the 38 countries equal to our average of $3 billions is 0.44*3/*k* where *k* is the number of existing first world SEFs in that process. Thus, if there were only one SEF in that process, the increase in investment due to an additional SEF would be of about $1.3 billions, which corresponds to more than 10 extra plants on average in the 38 countries over the ten year period. For $k = 12$ (i.e. the mean of SEF_FW_i in the sample, from table 1), the increase would be of $109 millions, i.e. less than one extra plant over the 10 year period in the 38 countries. Thus, the results indicate that division of labour is important, but that in many sectors it has proceed far enough that further entry of specialised suppliers would have only a smaller impact.

Finally, given that technological maturity increases both the expected number of plants, and the average investment per plant, our estimates imply that the impact of an additional SEF is likely to be greater in more mature processes. This is borne out by figure 4. The figure shows that processes where the technological frontier is moving rapidly are likely to be affected less than more mature processes.

4.5 The differential effect of SEFs on domestic and multinational firms

Another interesting issue is whether SEFs have different impacts on different types of investors. Specifically, we ask whether SEFs are more likely to benefit domestic LDC firms rather than first world multinational enterprises. Result 3 in our model predicts a chemical company better able to perform in-house engineering should be less affected by the presence of SEFs compared with a company with lower in-house engineering capabilities. It is natural to think that first world

chemical multinationals (MNE)have better in-house capabilities for process engineering than local LDC companies, and thus the impact of SEFs be greater for LDC firms than for MNEs.

We estimated two equations for the total dollar investment by domestic firms and multinationals—DOM_{ij} and MNE_{ij} – using the same specification and variables employed earlier. Table 6 presents the results of these estimations using OLS and GMM, where in the latter we use $SLARGECO_i$ and $SGREEN_i$ as instruments. The key result in table 6 is that the OLS and GMM estimates of the impact of SEF_FW_i are sizable and significant in the domestic firm equation, whereas they are small and even negative in the MNE equation. This is an interesting result for it suggests that SEFs are likely to have encouraged entry of LDC firms into new chemical markets by transferring them the chemical technologies. This result is consistent with other findings that the presence of SEFs lowers entry barriers and makes markets more competitive (e.g., Lieberman, 1987). Relatedly, the fact that SEFs do not have a similar impact on MNE suggests that first world SEFs are a sources of increased competition for the very same first world companies that operate into the LDC markets, and that gave rise to these SEFs in earlier periods.

Table 6 also shows that in the DOM equations the OLS and GMM estimates of the impact of SEF_FW_i are very similar. Moreover, as reported in table 6, the Hausman test statistics for the GMM vis-à-vis the OLS estimate is very small. The test statistics does not reject the null hypothesis that, for domestic firms, there is no endogeneity bias in our measure of SEFs. By contrast, the estimated elasticities of MNE with respect to SEF_FW_i differ between OLS and GMM. The Hausman test statistics of 7.51 is above the theoretical *chi-squared* of 3.84 and 6.64 at 5% and 1% confidence level and one degree of freedom.

Finally, to measure the effects of SEFs on LDC firms and multinationals, note from table 1 that the sample average of the total dollar investment by domestic firms in *ij* is $62.8 millions, which implies $2.4 billions for the 38 countries as a whole, over the ten year period. Thus, using the

GMM estimate of the elasticity of DOM_{ij}, the additional dollar investment by LDC firms due to an extra SEF for a typical process with *k* SEFs is 0.46*2.4/*k*. For *k* = 1, the additional investment would be of about $1 billion over 1981-90, or about 9 extra plants. For the average number of SEFs in our data, i.e. *k* = 12, the marginal effect on one extra SEF would be of $92 millions, i.e., about three fourths of an average plant.

Similar calculations for MNE indicate that, using our GMM estimate, the effect of an extra SEF on the average process is –0.25*578/*k* where $578 millions is the average across processes of the total dollar investment in the 38 countries by MNE. For *k* = 1, this expression is equal to $145 millions—that is if there were only one SEF, an extra SEF would reduce the investments by MNE in our 38 countries by slightly more than one plant. For *k* = 12, the reduced investment by MNE would be of only $12 millions.

6. Conclusions

Technological spillovers play an important role in the process of economic growth. (See Griliches, 1979, Jaffe, 1986, or Coe and Helpman, 1995.) But the typical description of the mechanism of these spillovers is, in Alfred Marshall's often used phrase, one where "the secrets ... are in the air". Important as this ethereal mechanism may be, there are other mechanisms through which technology and knowledge are transferred across sectors and countries that are more material and more amenable to economic analysis. In this paper, we argued that an important institutional form through which these apparent spillovers take place is division of labor. In addition to improving productivity through specialization, a division of labor can also serve as a mechanism whereby technology is made available to latecomers at low costs.

The economics of this mechanism is very simple. As Romer (1990) has emphasized, the development of technological capability is a fixed cost activity, while the productive application of

the technological capability is a (low) marginal cost activity. In our story, firms in the upstream sector invest in learning about the production process. If the upstream sector is competitive these costs are ultimately paid by customers downstream. The expertise and the technologies that they supply are process—and not location—specific, and thus, can be made available to downstream firms in other countries. Competition between these firms implies that the benefits of the acquired expertise will be made available to users in other countries, or in other sectors of the economy, at prices close to marginal cost because the development costs have already been sunk. However, as our results also show, the benefits for the downstream industry of an upstream sector of technology suppliers diminish as the number of upstream suppliers increases. As also noted by Bresnahan and Reiss (1991) in a different context, the advantages to consumers of an extra supplier decrease sharply as the number of suppliers rises.

In short, what our results imply is not that the observed rates of investment in chemical plants in LDCs are being fueled solely by SEFs and could never be achieved without SEFs. Rather, they imply that the investment is taking place earlier and more rapidly than if LDCs had to rely solely upon chemical producers in the first world to transfer the technology, or even worse, if they had to 're-invent the wheel' – i.e., develop process technologies and the broader engineering expertise required to design and construct chemical plants domestically. In short, the organization of industry in the first world 'matters' not just for the growth of the first world but also for the growth of the LDCs.

References

Amemiya, T., 1985, *Advanced Econometrics*, Harvard University Press, Cambridge MA

Arora, A., Fosfuri, A., and Gambardella, A., 1996, "Division of Labor and the Transmission of Growth", CEPR Working Paper, Stanford University, July.

Arora, A., and Gambardella, A., 1997, "Domestic Markets and International Competitiveness: Generic and Product-Specific Competencies in the Engineering Sector", *Strategic Management Journal*, Special Issue (Summer), 53-74.

Arora, A., and Gambardella, A., 1998, "Evolution of Industry Structure in the Chemical Industry", in Arora, A., Landau, R., and Rosenberg, N. (eds.) *Chemicals and Long-Term Economic Growth*, John Wiley & Sons, New York.

Barro, R., and Lee, J. H., 1994, "Data Set for a Panel of 138 Countries". Downloaded from ftp://www.nber.org/pub/barro.lee.

Bresnahan, T. and Reiss, P., 1991, "Entry and Competition in Concentrated Markets", *Journal of Political Economy*, Vol.99(5), pp.977-1009.

Bresnahan, T. and Trajtenberg, M., 1995, "General-Purpose Technologies – Engines of Growth?", *Journal of Econometrics*, Vol. 65(1), pp. 83-108.

Chemical Age Project File (CAPF), 1989, Pergamon Financial Data Services, London.

Coe, D.T., and Helpman, E., 1995, "International R&D Spillovers", *European Economic Review*, Vol. 39, pp. 859-887.

Eaton, J, and Kortum, S., 1996, "Trade in ideas: Patenting and productivity in the OECD", Journal of International Economics, Vol. 40, pp. 251-278.

Freeman C., 1968, "Chemical Process Plant: Innovation and the World Market", *National Institute Economic Review*, No. 45, pp. 29-51, August.

Griliches Z., 1979, "Issues in Assessing the Contribution of Research and Development to Productivity Growth", *Bell Journal of Economics*, Vol. 10(1), pp. 92-116.

Hausman, J., 1978, "Specification Tests in Econometrics", *Econometrica*, Vol.46, pp.1251-1272.

Helpman, E., and Trajtenberg, M., 1996, "Diffusion of General-Purpose Technologies", NBER Working Paper, No. 5773, September.

Keller, W., 1997, "Are international R&D spillovers trade-related? Analyzing spillovers among randomly matched trade partners", NBER Working Paper, No. 6065, forthcoming in *European Economic Review*.

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

Landau, R., 1966, *The Chemical Plant. From Process Selection to Commercial Operation*, Reinhold Publishing Co., New York.

Lenz, A. J., 1996, *The U.S. Chemical Industry: Performance in 1996 and Outlook*, Chemical Manufacturer's Association, Washington DC.

Lieberman, M., 1987, "Patents, Learning by Doing, and Market Structure in the Chemical Processing Industries", *International Journal of Industrial Organization*, Vol.5, pp.257-276.

Mullahy, J., 1997, "Instrumental Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior", *Review of Economics and Statistics*, pp.586-593.

Nadiri, M. I., 1993, "Innovation and Technological Spillovers", NBER Working Paper, No. 4423, August.

Predicast's Company Thesaurus, 1991, World Headquarters, Cleveland Ohio, Annual Edition.

Rivera-Batiz, L., and Romer, P., 1991, "Economic Integration and Endogenous Growth", *Quarterly Journal of Economics*, Vol. 106, pp. 531-555, May.

Rodriguez-Clare, A., 1996, "Multinationals, Linkages, and Economic Development", *American Economic Review*, Vol. 86(4), pp. 851-873.

Romer, P., 1990, "Endogenous Technological Change", *Journal of Political Economy*, Vol. 98(5), pp. S71-S102.

Romer, P., 1996, "Why, indeed, in America? Theory, history, and the origins of modern economic growth", *American Economic Review*, Vol. 86(2), pp. 202-206, May.

Rosenberg, N., 1976, *Perspectives on Technologies*, Cambridge University Press, Cambridge UK.

Smith, A., 1776, *The Wealth of Nations*, Penguin Books, Harmondsworth, UK, 1983 edition.

Stigler, G., 1951, "The division of labor is limited by the extent of the market", *Journal of Political Economy*, Vol. 59(3), June.
Teece, D., 1988, "Technological Change and the Nature of the Firm", in Dosi, G., et al. (eds), *Technological Change and Economic Theory*, Francis Printer, London.
Vernon, R., 1979, "The Product Cycle Hypothesis in a New International Environment", *Oxford Bulletin of Economics and Statistics*, Vol. 41(4), pp. 255-267, November.
UN Statistical Yearbook, United Nations, New York, 38th and 39th issues.
World Atlas, 1990, Istituto Geografico De Agostini, Novara, Italy.
Young, A., 1929, "Increasing Returns and Economic Progress", *Economic Journal*, Vol. 38, 539, December.
Zucker, L., Darby, M., and Armstrong, J., 1998, " Geographically localized knowledge: Spillovers or markets?", *Economic Inquiry*, Vol XXXVI, January, 65-86.

Appendix

Description of the data

Plant data. Our main source of plant level data is Chemical Age Project File (CAPF, 1989), a data base compiled by Pergamon Press, London. CAPF provides information on 20581 plants announced or constructed all over the world in the broadly defined chemical sector during 1980-1989. The data base is organized by plants. It reports the name of the company that ordered the plant, the name of the engineering company for that plant (or 'staff' for in-house engineering), the location of the plant (city and country), the name of the chemical process or of the product being produced, the date in which the investment was first reported in the specialized trade press, along with other information. For about 40% of the plants, CAPF also reports the total cost of investment in the plant in US million dollars. Finally, the data base reports the status of the plant along with the date in which the information was last updated. In most of the cases the information was updated in 1988-1989, which suggests that we can reasonably assume that this was the status of the plant at the end of our sample period. There are 14893 plants in the data base which are either 'completed' or 'under construction'. The others are 'planned', 'under study', 'abandoned', 'canceled', 'delayed', or other.

We focused our analysis on the plants that were either completed or under construction. Thus, $SIZE_{ij}$, BUY_{ij}, $MAKE_{ij}$, DOM_{ij}, MNE_{ij}, and $SIZE_FW_i$ were obtained by using only the 14893 plants that are completed or constructed. In counting SEF_FW_i, we used information from all the 20581 plants in the data base. This is because even plants that were planned, under study, abandoned, or other, provided useful information about whether a given SEF was a potential supplier for that technology. For similar reason, we used all the available information about plant costs to compute $COST_FW_i$ and the average plant investment cost in LDCs which was used to compute the dollar values of the variables mentioned above. The vast majority of the firms counted in SEF_FW_i are independent SEFs, or they engineering subsidiaries of larger chemical groups (especially for the European and Japanese companies). Since these normally act as independent companies on the open market, they can also be considered as full-fledged SEFs. There are a few chemical firms that provide engineering services to other (non-affiliated) chemical companies. Since they do so, we also included them as potential suppliers of engineering services for that process.

Chemical processes. CAPF plants cover 2081 different chemical processes. However, we focused our analysis on the 139 largest processes with 20 or more plants in the data base. Gathering information for all processes for our measures of technological novelty and complexity, is very difficult, and often, prohibitively so. In fact, our sample is a comprehensive set of all the important and widely diffused chemical technologies in the world. It covers 10145 plants that are completed or in construction, i.e. almost 70% of all the 14893 plants completed or in construction in the data base. In addition, in an earlier working paper we report qualitatively similar empirical results using a larger sample of processes, albeit without the extensive controls that we use here (Arora, Fosfuri and Gambardella, 1996). We also performed similar analyses using a cross-section of all the 2081 processes and aggregate variables for the LDCs as a whole. Again, the qualitative results did not change.

Sectors. CAPF classifies its plants in the following 21 sectors (in parenthesis the number of plants that are completed or under construction in each sector): Agricultural Chemicals (116), Air Separation (596), Coal Refining (32), Desalination (40), Engineering Materials (110), Environmental Technologies (75), Fertilizers (1000), Food Products (308), Gas Handling (1014), Inorganic Chemicals (1249), Industrial Gases (613), Minerals and Metallurgy (532), Miscellaneous (505), Organic Chemicals (1114), Oil Refining (2246), Petrochemicals (2155), Pharmaceuticals (745), Plastics and Rubber (1474), Pulp and Paper (396), Synthetic Fuels (135), Textiles and Fibers (438). The sector dummies that we actually used in all our regressions were obtained, however, after aggregating these 21 sectors in 9 classes of relatively homogeneous sectors. The 9 aggregate sectors are: OIL REFINING, PETROCHEMICALS, MINERALS & METALLURGY, PLASTICS & RUBBER, INORGANIC CHEMICALS, AGRICULTURE (Agricultural Chemicals and Fertilizers), GAS (Gas Handling, Air Separation, and Industrial Gases), ORGANIC CHEMICALS (Organic Chemicals, Explosives, Textile and Fibers, Food Products, and Pharmaceuticals), and MISCELLANEOUS (all the rest).

Nationality of companies and subsidiaries. We used Predicast's (1991) and other company thesauruses to group all the companies that were subsidiaries of other companies in the data base under the names of their mothers companies, and assign nationalities to companies. We treated the SEFs that were subsidiaries of larger chemical groups as independent firms. However, when an SEFs provided services for its parent company, this was counted as a 'make'. We do not find any SEFs from LDCs operating in the first world.

Limitation of the data base. CAPF is a commercial data base, and is constructed from various sources such as questionnaires and reports in the trade press. Its vast coverage suggests that biases are unlikely. Its most serious limitation is that for about 17% of the plants in our sample the name of the engineering company is not given. Conversations with data providers in the industry suggested that these blanks could arise for a number of reasons. Companies may still be looking for suitable engineers (including possibly doing the engineering in-house), or they do

not want to disclose the name, or the information is simply missing. However, we also used information about these plants from another data base (Hydrocarbon Processing Unit – HPI, compiled by Gulf Publishing, Texas) to see whether we could classify the blanks as 'buys' or 'makes'. The details of this diagnostic check are available from the authors upon request. The check was inconclusive because a large number of the identified plants in the other data base were still blanks. However, since most of the identified plants which did report the name of an engineering company in the other data base were 'buys', our check ruled out the possibility that the blanks are predominately 'makes'.

We performed all our empirical analyses under different assumptions about the blanks – i. e. all the blanks are 'buys', all the blanks are 'makes', the blanks are 50% 'buys' and 50% 'makes', the blanks are distributed between 'buys' and 'makes' in the same proportion as in the case in which the name of the engineering company (or 'staff') is observed. The results presented here are those where we assume that all the blanks are 'buys'. It turns out that all other assumptions about the blanks had even more favorable results for our theory.

Construction of $PROCPAT_i$. $PROCPAT_i$ was constructed as follows. We selected all relevant patents using a keyword search with the process as the keyword. From these, we selected and read the full abstracts of patents that exactly fit our criterion. The patent classes (and sub-classes) into which these patents were classified were examined to ensure that the invention was in fact a process invention. These subclasses of the US patent classification system were used along with the process name as the basis for Boolean queries of the US patent database to generate the final set of patents, one set for each process. (The details of the Boolean queries are available upon request.) The titles (and some abstracts selected at random) of the patents in the final sample for each process were examined to ensure that the final sample did not contain irrelevant patents.

Countries. In all our analyses, and for all the variables that we constructed, we defined first world to be all the OECD countries except Mexico, the Czech Republic, Hungary, Poland, South Korea and Turkey. These countries joined the OECD only very recently and for our purposes, it was more appropriate to include them in the LDC category. Therefore our first world countries are all the Western European countries, the USA and Canada, Japan, Australia and New Zealand. The 38 countries in our sample are: Algeria, Argentina, Bangladesh, Brazil, Burma, Chile, China, Colombia, Ecuador, Egypt, Hong Kong, Hungary, India, Indonesia, Iran, Iraq, Kuwait, Malaysia, Mexico, Morocco, Nigeria, Pakistan, Peru, Philippines, Poland, Saudi Arabia, Singapore, South Africa, South Korea, Sri Lanka, Sudan, Syria, Taiwan, Thailand, Tunisia, Turkey, Venezuela, Yugoslavia.

Table 1: List of country characteristics.

GDP_j	Real GDP of country in 1985 in billions of US dollars. Obtained from per capita GDP of country (Barro-Lee) times population.
POP_j	Population of country in 1985 in millions. (Barro-Lee)
$ENERGY_j$	Total energy consumption of country (1985-1987 average) in thousand metric tons of coal equivalent. (UN Statistical Yearbook)
$AREA_j$	Size of land in thousands square Km. (Barro-Lee) In regressions we also used $AREA_j$ interacted with sector dummies for inorganic chemicals, agricultural chemicals, minerals & metallurgy ($ICHEM_i$+$AGRI_i$+MM_i)
$HKAP_j$	Human Capital. Average schooling years of population over 25 in the country. (Barro-Lee) Equal to zero if data is missing in Barro-Lee (missing data for: China, Egypt, Morocco, Nigeria, Saudi Arabia).
$DHKAP_j$	Dummy equal to 1 for countries for which $HKAP_j$ is missing in Barro-Lee.
$OPEN_j$	Own-import weighted tariff rates of the country on intermediate inputs and capital goods. (Barro-Lee) This is Barro-Lee's variable OWTI. Equal to zero if data is missing in Barro-Lee (missing data for: Burma, Hungary, Poland, South Africa).
$DOPEN_j$	Dummy equal to 1 for countries for which $OPEN_j$ is missing in Barro-Lee.
$DOIL_{ij}$	Dummy for countries with oil reserves interacted with dummy for processes in the oil refining sector. Countries with oil reserves: Algeria, Argentina, Brazil, China, Colombia, Ecuador, Egypt, India, Indonesia, Iran, Iraq, Kuwait, Malaysia, Mexico, Nigeria, Saudi Arabia, Syria, Venezuela. (Main countries with oil reserves listed in World Atlas, 1990.)
$DGAS_{ij}$	Dummy for countries with natural gas reserves interacted with dummy for processes in the gas processing sector. Countries with natural gas reserves: Algeria, Argentina, Indonesia, Mexico, Venezuela (Main countries with natural gas reserves listed in World Atlas, 1990.)
Geographical Area Dummies	Africa, Eastern Europe, Middle East, Central and South America, Far East.

Table 2: Descriptive statistics.

Variable	No. of Obs.	Mean	Std Dev	Min	Max
$SIZE_{ij}$[1]	5282	78.1	431.4	0	17751.7
BUY_{ij}[1]	5282	75.3	422.9	0	17751.7
$MAKE_{ij}$[1]	5282	2.8	33.9	0	1321.1
DOM_{ij}[1]	5282	62.8	364.6	0	13313.7
MNE_{ij}[1]	5282	15.2	129.7	0	4437.9
$SIZE_FW_i$[1]	139	2761.4	6277.5	2.0	45555.6
$COST_FW_i$[1]	139	76.5	145.5	0.8	1190
SEF_FW_I	139	11.94	11.77	0	60
$NOVEL_I$	139	0.17	0.65	-0.78	3.600
$PROCPAT_i$ [2]	136	61.19	60.65	1	345
$SLARGECO_i$	139	0.50	0.22	0	1
$SGREEN_I$	139	0.63	0.22	0.08	1
GDP_j	38	171.28	323.40	19.63	1918.79
POP_j	38	84.90	204.40	1.70	1059.50
$ENERGY_j$	38	61.26	126.70	1.48	765.18
$AREA_j$	38	1311.4	2039.4	1.0	9537.0
$OPEN_j$ [3]	34	0.237	0.23	0.00	1.32
$HKAP_j$ [3]	33	4.85	2.18	0.91	10.75

Notes:
[1] In millions of US dollars.
[2] Missing values for ‘specialty chemicals’, ‘resins’, and ‘refinery’.
[3] Missing values. See Table 1.

Table 3: OLS Regressions – $SIZE_{ij}$, BUY_{ij}, $MAKE_{ij}$.

$$\log(X_{ij}) = const + aY_j + bZ_i + e_{ij}$$

	$SIZE_{ij}$	**BUY_{ij}**	**$MAKE_{ij}$**
Constant	-9.95	-10.95	-4.21
	(1.44)	(1.62)	(0.55)
$DOIL_I$	0.51	0.51	0.14
	(0.14)	(0.16)	(0.05)
$DGAS_I$	0.34	0.37	-0.01
	(0.21)	(0.24)	(0.08)
$DHKAP_j$	-0.12	-0.11	-0.09
	(0.12)	(0.14)	(0.05)
$(1- DHKAP_j) \times HKAP_j$	-0.04	-0.05	0.01
	(0.02)	(0.03)	(0.01)
$Log(GDP_j)$	0.44	0.42	0.26
	(0.16)	(0.18)	(0.06)
$Log(POP_j)$	-0.29	-0.31	-0.07
	(0.11)	(0.12)	(0.04)
$Log(ENERGY_j)$	0.39	0.47	-0.05
	(0.07)	(0.08)	(0.03)
$Log(AREA_j)$	0.07	0.08	-0.02
	(0.03)	(0.04)	(0.01)
$(ICHEM_i+AGRI_i+ MM_I) \times$	0.13	0.15	0.01
$Log(AREA_j)$	(0.03)	(0.04)	(0.01)
$DOPEN_j$	-0.18	-0.22	0.11
	(0.15)	(0.17)	(0.06)
$(1-DOPEN_j)\times OPEN_j$	0.62	0.72	0.07
	(0.18)	(0.20)	(0.07)
$Log(SIZE_FW_i)$	0.38	0.41	0.07
	(0.05)	(0.06)	(0.02)
$Log(COST_FW_i)$	-0.31	-0.35	-0.07
	(0.06)	(0.06)	(0.02)
$NOVEL_I$	-0.16	-0.17	-0.03
	(0.05)	(0.05)	(0.02)
$DPROCPAT_I$	-0.35	-0.49	-0.01
	(0.20)	(0.23)	(0.08)
$(1-DPROCPAT_i) \times$	-0.13	-0.15	0.00
$Log(PROCPAT_i)$	(0.03)	(0.03)	(0.01)
$Log(SEF_FW_I)$	0.46	0.52	0.02
	(0.05)	(0.06)	(0.02)
Adj. R^2	0.25	0.24	0.04
No. Of observations	5282	5282	5282

Note: Standard errors in parenthesis. All regressions include sector dummies and dummies for geographical areas of country.

Table 4: OLS Estimation: Dependent Variable – Log(SEF_FW_i)

Variable	Coefficient	Std. Error
Constant	-0.94	0.37
AGRI	0.24	0.26
GAS	0.43	0.24
REFINING	0.62	0.22
ORG_CHEM	-0.01	0.24
PETRO_CHEM	0.34	0.22
MM	-0.25	0.26
PLAS&RUBBER	0.13	0.25
ICHEM	0.08	0.27
NOVEL	-0.15	0.08
DPROC_PAT	-0.15	0.36
LOG(PROC_PAT)	-0.03	0.05
SIZE_FW	0.79	0.06
SLARGECO	-0.64	0.27
SGREEN	0.41	0.25
LOG(COST_FW)	-0.69	0.08

N = 139

Adj R-squared = 74

Table 5: OLS and GMM Estimation—$SIZE_{ij}$

$$\log(X_{ij}) = const + aY_j + bZ_i + e_{ij}$$

	OLS	**GMM**
Constant	-9.95	-10.1
	(1.44)	(1.36)
$DOIL_I$	0.51	0.49
	(0.14)	(0.15)
$DGAS_I$	0.34	0.39
	(0.21)	(0.28)
$DHKAP_j$	-0.12	-0.11
	(0.12)	(0.12)
$(1- DHKAP_j) \times HKAP_j$	-0.04	-0.04
	(0.02)	(0.02)
$Log(GDP_j)$	0.44	0.46
	(0.16)	(0.15)
$Log(POP_j)$	-0.29	-0.32
	(0.11)	(0.11)
$Log(ENERGY_j)$	0.39	0.40
	(0.07)	(0.06)
$Log(AREA_j)$	0.07	0.07
	(0.03)	(0.03)
$(ICHEM_i+AGRI_i+MM_i) \times$	0.13	0.13
$Log(AREA_j)$	(0.03)	(0.03)
$DOPEN_j$	-0.18	-0.14
	(0.15)	(0.15)
$(1-DOPEN_j)\times OPEN_j$	0.62	0.67
	(0.18)	(0.21)
$Log(SIZE_FW_i)$	0.38	0.39
	(0.05)	(0.14)
$Log(COST_FW_i)$	-0.31	-0.32
	(0.06)	(0.13)
$NOVEL_I$	-0.16	-0.17
	(0.05)	(0.05)
$DPROCPAT_I$	-0.35	-0.53
	(0.20)	(0.21)
$(1-DPROCPAT_i) \times$	-0.13	-0.13
$Log(PROCPAT_i)$	(0.03)	(0.03)
$Log(SEF_FW_I)$	0.46	0.44
	(0.05)	(0.17)
Hausman test statistic for $Log(SEF_FW_I)$	--	0.02
No. Of observations	5282	5282

Note: Standard errors in parenthesis. All regressions include sector dummies and dummies for geographical areas of country. All variables and $SLARGECO_i$ and $SGREEN_i$ used as instruments for $Log(SEF_FW_i)$ in GMM: Hausman test statistic computed as the square of the difference between the OLS and GMM estimates, divided by the difference between their variances.

Table 6: OLS and GMM Estimation—DOM_{ij}, MNE_{ij}

$$\log(X_{ij}) = const + aY_j + bZ_i + e_{ij}$$

	DOM_{ij}		**MNE_{ij}**	
	OLS	**GMM**	**OLS**	**GMM**
Constant	-7.81	-8.01	-7.40	-7.66
	(1.54)	(1.44)	(1.04)	(1.12)
$DOIL_I$	0.63	0.60	-0.06	-0.06
	(0.15)	(0.17)	(0.10)	(0.11)
$DGAS_I$	0.16	0.23	0.35	0.35
	(0.23)	(0.29)	(0.15)	(0.22)
$DHKAP_j$	0.02	0.03	-0.25	-0.25
	(0.13)	(0.13)	(0.09)	(0.09)
$(1- DHKAP_j) \times HKAP_j$	-0.01	-0.01	-0.05	-0.05
	(0.03)	(0.02)	(0.02)	(0.02)
$Log(GDP_j)$	0.17	0.19	0.52	0.51
	(0.17)	(0.16)	(0.12)	(0.12)
$Log(POP_j)$	-0.18	-0.21	-0.32	-0.32
	(0.12)	(0.12)	(0.08)	(0.08)
$Log(ENERGY_j)$	0.48	0.48	0.02	0.02
	(0.08)	(0.06)	(0.05)	(0.04)
$Log(AREA_j)$	0.10	0.11	0.00	0.00
	(0.03)	(0.03)	(0.02)	(0.02)
$(ICHEM_i + AGRI_i + MM_I) \times$	0.14	0.14	0.03	0.04
$Log(AREA_j)$	(0.04)	(0.03)	(0.02)	(0.02)
$DOPEN_j$	-0.51	-0.46	0.40	0.40
	(0.16)	(0.16)	(0.11)	(0.11)
$(1-DOPEN_j) \times OPEN_j$	0.85	0.90	0.28	0.28
	(0.19)	(0.23)	(0.13)	(0.16)
$Log(SIZE_FW_i)$	0.38	0.35	0.23	0.47
	(0.06)	(0.15)	(0.04)	(0.10)
$Log(COST_FW_i)$	-0.31	-0.29	-0.22	-0.43
	(0.06)	(0.14)	(0.04)	(0.09)
$NOVEL_I$	-0.17	-0.18	-0.09	-0.14
	(0.05)	(0.06)	(0.03)	(0.04)
$DPROCPAT_I$	-0.48	-0.62	-0.02	-0.05
	(0.22)	(0.22)	(0.15)	(0.16)
$(1-DPROCPAT_i) \times$	-0.15	-0.14	-0.02	-0.02
$Log(PROCPAT_i)$	(0.03)	(0.03)	(0.02)	(0.02)
$Log(SEF_FW_i)$	0.44	0.46	0.06	-0.25
	(0.05)	(0.18)	(0.04)	(0.12)
Adj R^2	0.23		0.07	
Hausman test statistic	--	0.01	--	7.51
No. Of observations	5282	5282	5282	5282

Note: Standard errors in parenthesis. All regressions include sector dummies and dummies for geographical areas of country.

Figure 2: Impact of an additional SEF
Additional investment per process, by country, in millions of US dollars.

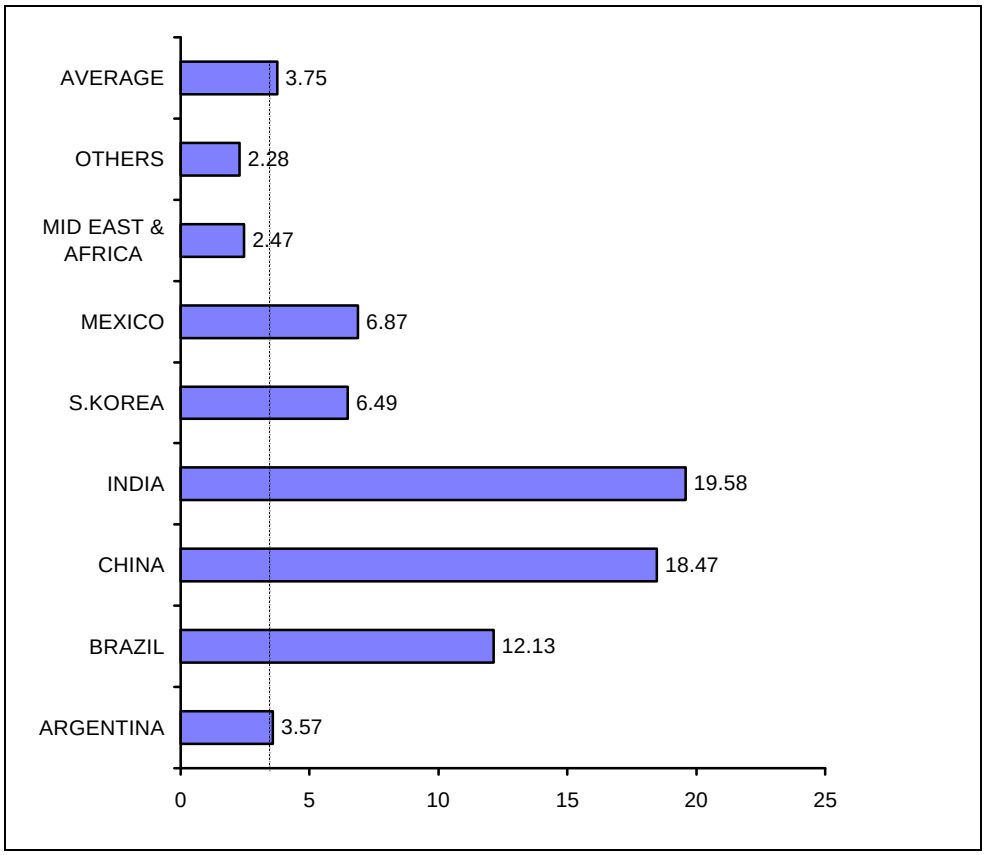


Figure 3: The impact of an additional SEFs on investment in LDC
By size of process market, in millions of US dollars.

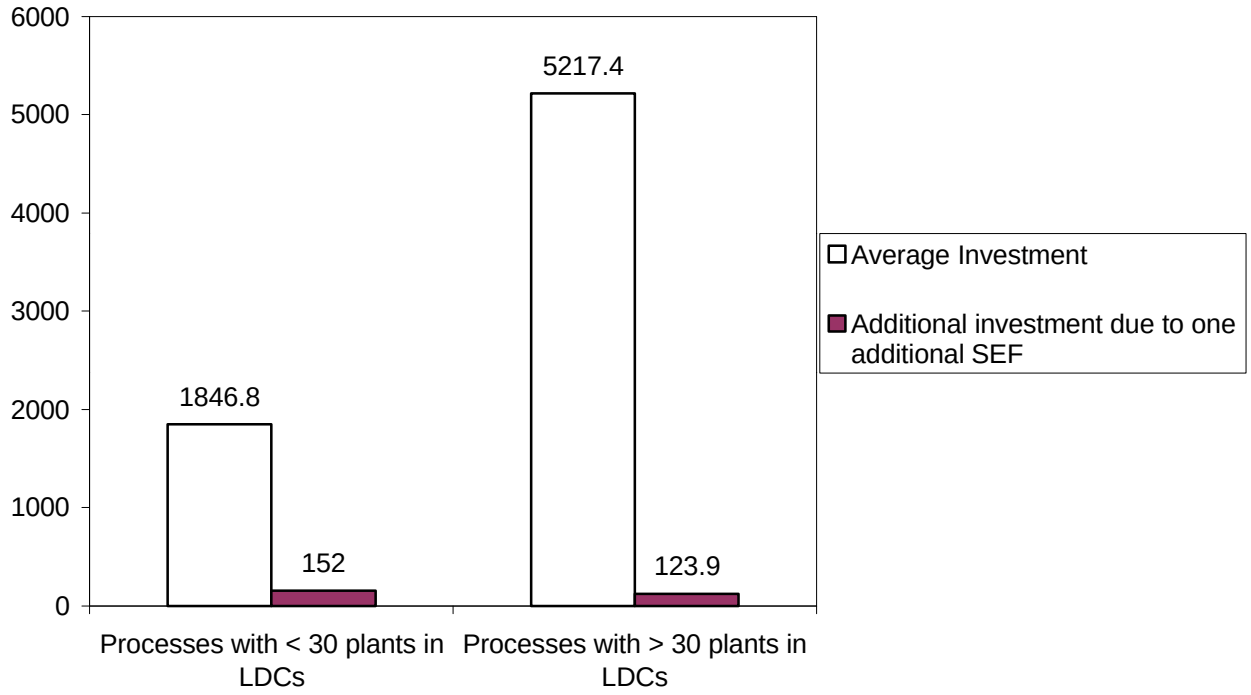
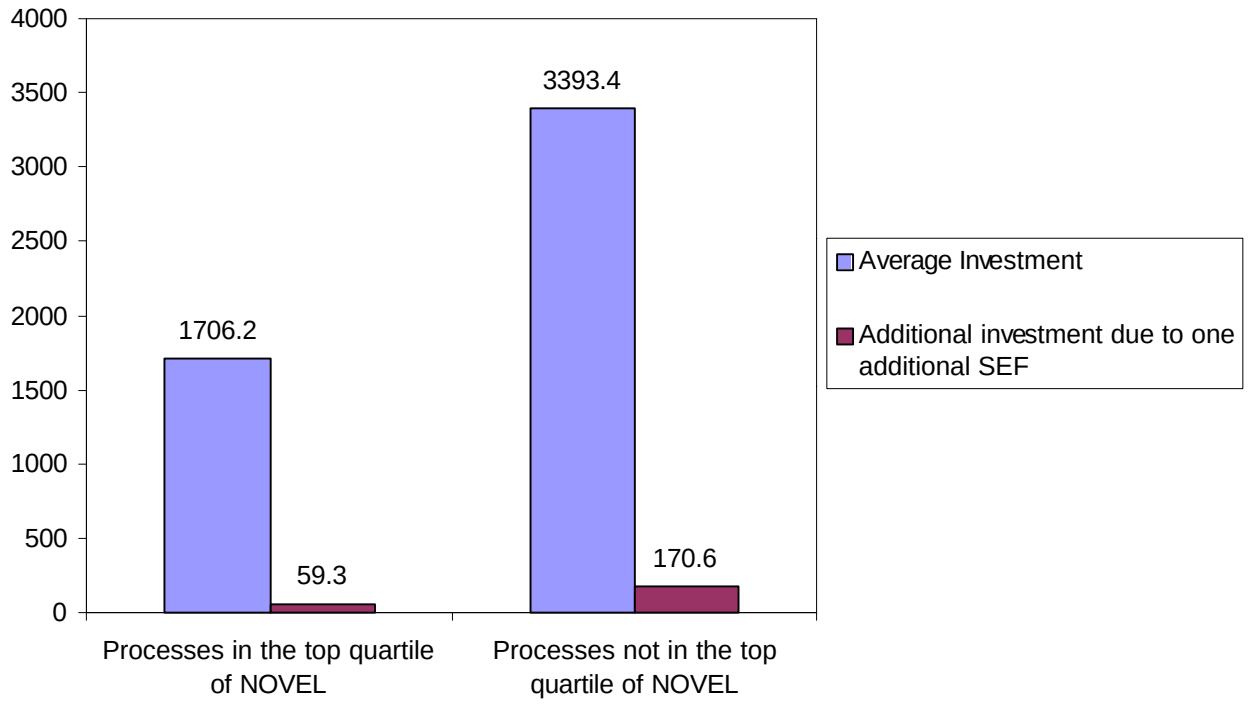


Figure 4: The impact of SEFs on investment, by type of process, in millions of US dollars.



APPENDIX II (Not for publication)

II.1 Bounding the effect of SEF_FW$_i$ when potential market size is measured with error.

This appendix provides a way of assessing the question of unobserved variation in market size that does not rely on the use of instruments (the method followed in the paper.) Consider investments by domestic LDC firms and MNEs. Let

$$Y_1 = X'\beta_1 + (Z + \zeta)\alpha_1 + \varepsilon_1$$

$$Y_2 = X'\beta_2 + (Z + \zeta)\alpha_2 + \varepsilon_2$$

where Y_1 = Domestic plants, Y_2 = MNE plants, X = matrix of explanatory variables, $(Z+\zeta)$ = potential market size with Z = measured component of potential market size, ζ = unobserved component of potential market size. All relevant variables are measured in logs so that the coefficients are to be interpreted as elasticities. As usual, we assume that ε_1 and ε_2 are *i.i.d.* random errors. Moreover, since the unobserved component of market size varies across processes but not countries, it is natural to specify the equations at the level of the process. Accordingly, in this Appendix we show OLS estimates obtained after aggregating across countries, so that the unit of observation is the process, not the process-country pair.

Letting $\{\mathbf{b_i}, \mathbf{a_i}\}$ represent the OLS estimate of $\{\beta_i, \alpha_i\}$, and E{} representing the expectation operator, we can write the normal equations as (e.g., see Johnston, 1984, pp 111-113)

$$E\,\{(X'X)(\mathbf{b_i} - \beta_i) + (X'Z)(\mathbf{a_i} - \alpha_i)\} = E\{(X'\zeta)\alpha_i\} \qquad \text{(II.1)}$$

$$E\,\{(Z'X)(\mathbf{b_i} - \beta_i) + (Z'Z)(\mathbf{a_i} - \alpha_i)\} = E\{(Z'\zeta)\alpha_i\},\ i = 1, 2 \qquad \text{(II.2)}$$

Note that if the number of SEFs are correlated with unobserved variations in market size, we have $E\{(X'\zeta)\alpha_i\} \neq 0$, which leads to biased estimates of β and α. By manipulating (II.1), we get

$$E\{(\mathbf{b_1} - \beta_1) - (\mathbf{b_2} - \beta_2)\} = -E\{(X'X)^{-1}(X'Z)[(\mathbf{a_1} - \mathbf{a_2}) - (\alpha_1 - \alpha_2)]\} \qquad \text{(II.3)}$$

or

$$\beta_1 - \beta_2 = E\{(\mathbf{b_1} - \mathbf{b_2}) + (X'X)^{-1}(X'Z)[(\mathbf{a_1} - \mathbf{a_2}) - (\alpha_1 - \alpha_2)]\} \qquad \text{(II.4)}$$

Assuming that $\alpha_1 = \alpha_2$, one can consistently estimate the extent of the bias because $(X'X)^{-1}(X'Z)$ is simply the OLS coefficient obtained by regressing SIZE_FW$_i$ on the remaining explanatory variables. One can then use (II.4) to form lower bounds on β_1, and hence, on the average impact of SEF_FW$_i$ on SIZE. An alternative is to directly regress $Y_1 - Y_2$ on X. This has virtually no effect on the computed lower bounds. From the DOM- MNE estimates (OLS, table 5), we find that $\mathbf{b_1} - \mathbf{b_2} =$ 0.38 and $\mathbf{a_1} - \mathbf{a_2}$ = 0.15. The OLS coefficient of SEF_FW$_i$ in the SIZE_FW$_i$ equation (not shown here) is 0.7. This implies $\beta_1 - \beta_2 = 0.38 + 0.7\times0.15 = 0.48$. Thus, as long as β_2 is positive, β_1 is no smaller than 0.48. We can also derive a lower bound for β, the elasticity of SIZE with respect to SEF_FW$_i$ by noting that $\frac{\P \log(SIZE)}{\P \log(KFW)} = \theta_1 \frac{\P \log(DOM)}{\P \log(KFW)} + (1-\theta_1)\frac{\P \log(MNE)}{\P \log(KFW)}$, where θ_1 is the share of DOM plants in the total. To determine a lower bound we assume that $\frac{\P \log(MNE)}{\P \log(KFW)} = 0$. Then using $\frac{\P \log(DOM)}{\P \log(KFW)} = 0.48$, as computed above, and by evaluating the elasticity at the average value $\theta_1 = 0.78$, a lower bound of β, the elasticity of SIZE with respect to SEF_FW$_i$, is 0.36. If $\alpha_1 \neq \alpha_2$, then, the lower bound. However, it is easy to see that even for large values of the difference between α_1 and α_2, the true coefficient is bounded away from zero.

II.2 GMM ESTIMATES: $SIZE_{ij}$, CONTROLLING FOR SHARE OF MNE INVESTMENT IN THE FIRST WORLD

VARIABLE	ESTIMATE	STD. ERROR
CONSTANT	-10.14	1.37
E. EUROPE	-0.06	0.19
FAR EAST	0.36	0.12
MID EAST	-0.26	0.10
S. ASIA	-0.13	0.11
DOIL	0.49	0.15
DGAS	0.40	0.28
LOG(AREA)	0.08	0.03
LOG(AREA)* (ICHEM + AGRI + MM)	0.13	0.03
DHKAP	-0.11	0.12
HKAP	-0.04	0.02
LOG(GDP)	0.46	0.15
LOG(POP)	-0.32	0.11
LOG(ENERGY)	0.40	0.06
DOPEN	-0.14	0.15
OPEN	0.67	0.20
AGRI	-0.01	0.23
GAS	-0.64	0.20
REFINING	-0.38	0.23
ORG_CHM	0.15	0.13
ICHEM	-0.86	0.22
PETRO_CHM	-.056794	0.14
MM	-0.70	0.23
PLAS&RUBBER	-0.01	0.13
NOVEL	-0.17	0.06
LOG(PROCPAT)	-0.13	0.03
DPROCPAT	-0.53	0.21
LOG(SIZE_FW)	0.34	0.18
LOG(SHARE_MNE)	0.21	0.21
LOG(SEF_FW)	0.50	0.23
COST_FW	-0.27	0.17

N = 5282
Standard Errors computed from heteroscedastic-consistent matrix