View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

CENTRE D'ÉTUDES
ET DE RECHERCHES
SUR LE DÉVELOPPEMENT
INTERNATIONAL

Document de travail de la série

Etudes et Documents

E 2009. 09

OECD Imports: Diversification and quality search*

Revised version
February 2009

Olivier Cadot*
Céline Carrère+
Madina Kukenova±
Vanessa Strauss-Kahn§

* Research on this paper was supported by a grant from the World Bank. Without implicating them, the authors would like to thank Daniel Lederman, William Maloney, Marcelo Olarreaga, and an anonymous referee for useful comments.
* HEC Lausanne, CERDI, CEPR and CEPREMAP.
+ CERDI-CNRS, Université d'Auvergne.
± HEC Lausanne.
§ESCP-EAP and CEPR.

Abstract

This paper explores the evolution of OECD imports over time and as a function of income levels, measuring the concentration of those imports across origin countries at the product level. We find evidence of diversification followed, in the very last years of the sample period (post-2000), by a slight reconcentration. This reconcentration is entirely explained by the growing importance of Chinese products in OECD imports. We also find evidence of relatively more volatile concentration levels for differentiated goods, consistent with a simple model of adverse selection and screening of suppliers by OECD buyers. Finally, we find that "accession" to OECD markets occurs directly (rather than after acquiring prior export experience on other markets) for more than half of the (extra-OECD) exporter/product pairs, but that one to eight years of experience enhances subsequent survival on OECD markets. Exports that reach OECD markets after more than eight years of experience elsewhere tend to survive less.

Keywords: Import diversification, International trade, OECD,

JEL classification codes: F1, O11

1. Introduction

In spite of the rapid growth of emerging markets, OECD markets are still, today, the world's largest, providing key outlets for goods exported from developing countries. How much access there is for developing countries on OECD markets has been the subject of considerable attention (see e.g. Kee, Nicita and Olarreaga 2006 and references therein). By contrast, to our knowledge, not much has been written on the overall evolution of OECD imports. Yet, it matters whether they are opening up (in the sense of letting more extra-OECD exporters in) or concentrating on a few "preferred" suppliers. Contestable OECD markets would make it easier for entrants to get a foothold; on the contrary, if they exhibited strong incumbency advantages, they could create a two-track world among extra-OECD exporters (between countries that make it and countries that don't).

So far, a rapidly expanding literature has looked at the other side of the story, namely how export diversification (geographical and product-wise) interacts with economic development. Renewed interest in export diversification *per se* is motivated by the observation that a country with diversified exports is less vulnerable to terms-of-trade shocks (Ghosh and Ostry 1994). Terms-of-trade volatility has been shown to reduce long-term growth by Lutz and Singer (1994) and Easterly and Kray (2000). For that and other reasons, diversification is correlated with growth, as discussed in the recent book by Lederman and Maloney (2007).

Most of the literature has looked at product-wise diversification. Klinger and Lederman (2004) showed that the rate at which new products (defined at the HS4 or HS6 level) appear in a country's export portfolio, and found that it varies with economic development and peaks at middle income levels. In a subsequent paper (Klinger and Lederman 2005) they found that regulatory barriers to entry encourage new-product introduction. Their evidence is consistent with the

hypothesis of Hausmann and Rodrik (2003) whereby the private return to new-product introduction is reduced by informational externalities. Hummels and Klenow (2005, henceforth HK) introduced a decomposition of cross-country export variation into intensive and extensive margins that has been widely used since.[1] They showed that about 60% of the larger export volumes associated with country size is "explained" by the extensive margin. Cadot, Carrère and Strauss-Kahn (2007) showed that product diversification (measured by Herfindahl, Theil and Gini indices) evolves with income levels in a non-monotone way, with diversification followed by reconcentration beyond income levels around $20'000 at PPP, a pattern similar to what Imbs and Wacziarg (2003) found for production. Dutt, Mihov and van Zandt (2008) found no such non-monotonicity in Herfindahl indices calculated at a higher level of aggregation. Consistent with Hausmann, Hwang and Rodrik (2005), they found that export diversification (instrumented) correlates with future income levels, and, moreover, that the similarity of a country's export portfolio with that of the U.S. reinforces the effect on income levels.

A second strand of the literature has looked at the extensive margin defined geographically instead of product-wise. The first paper in that vein was Evenett and Venables (2002), who showed, on the basis of evidence for a limited set of developing countries, that about one third of the export growth observed during their sample period came from the expansion of existing exports to new markets. They found that the product-wise extensive margin accounted for only a small fraction of within-country export growth.[2] On the basis of a larger

[1] In words, a country's intensive margin is its market share in what it exports, whereas its extensive margin is the share of world trade accounted for by the goods in its export portfolio. Compared with simply counting the number of products a country exports, HK's extensive margin takes into account the weight of those products in world trade. That is, by HK's definition, a country that exports cars and computers will have a larger extensive margin than a country that exports carrots and potatoes, although both export just two goods.
[2] The seemingly inconsistent result of Evenett and Venables vs. HK come from the fact that the former considered within-country export growth whereas the latter looked at static cross-

sample, Brenton and Newfarmer (2007) found that the extensive margin accounted for only 19.6% of export growth; of that, 18% came from the export of existing products to new markets.

Another, time-honored strand of the trade literature, going back to the work of Hanson (1996), has emphasized the formation of regional production networks by multinational firms. According to this literature, a country's exports may be determined by the outsourcing decisions of multinationals based in other countries. Anecdotal evidence also suggests that it is retailers who decide which foreign suppliers (and hence countries) are included in cross-border supply chains. Thus, for producers located in developing countries, export opportunities are, at least partly, driven by the policies of large buyers in OECD countries. If those buyers decide to concentrate on a few suppliers in order, say, to simplify logistics or quality-control processes, opportunities will be fewer for entrants at every level of productivity and trade costs. Put differently, given the continued importance of OECD markets for developing-country exports, it seems difficult to understand how developing-country exports evolve without looking at how OECD imports evolve. This is what we set out to do in this paper.

Using a very large database of OECD imports at the SITC4 level since 1963, we find that, overall, OECD markets have been diversifying their sources of supplies (geographically) at the product level. This is reflected in decreasing concentration indices and a rising number of export sources. However, the trend in concentration has reversed itself in recent years. We show that this trend reversal is entirely explained by the rising share of Chinese products in OECD imports, as concentration indices decrease monotonically when China is excluded. We also find that the pattern of import diversification at the product level is broadly consistent with a simple model where buyers screen suppliers

---

country variation. Why results differ so much between the two approaches has not been explained so far in the literature.

for quality and toss them out when they under-perform. We test indirectly this conjecture by looking at how concentration indices vary across types of products. We find that they are more volatile, over time, for Rauch's differentiated products (where quality can be expected to be more heterogeneous across suppliers than for homogenous or reference-priced products). We also find, as implied by the model, that re-concentration, when it happens, is associated with a rise in unit values. That is, when buyers re-concentrate, they do so on higher-priced (and hence presumably higher-quality) suppliers. Finally, we show that a substantial chunk (more than half) of the new-product exports from developing-country exports are shipped to OECD markets without prior experience in other markets. The evidence is again (loosely) consistent with a story in which these North-South relationships are set up as part of vertically integrated supply chains. However, we also find that one to eight years of prior export experience on non-OECD markets enhances subsequent survival on OECD markets.

The paper is organized as follows. Section 2 sets up a simple model of supplier screening in the presence of adverse selection. Section 3 analyses the overall trend in OECD import concentration. Section 4 explores more specifically the model's implications for patterns of concentration and diversification. Section 5 deals with export-expansion paths and "waiting times" before exports are shipped to OECD markets. Section 6 concludes.

2. Supplier concentration and selection: Theory

2.1 Baseline model

In this section we explore how supplier concentration is affected by informational considerations in the presence of a selection problem. Consider a three-period setting where, in each period, a buyer needs to procure two units of a product from either one or two suppliers called X and Y. Each supplier has the

capacity to provide either one or two units, as the buyer wishes, at a constant price. Suppliers are of unknown quality, with a per-period probability of providing a non-defective product equal to λ^G for a good type and $\lambda^B < \lambda^G$ for a bad type (that is, the arrival of defective products follows an independent Bernoulli process for each supplier). The buyer knows λ^G and λ^B but not the type of each supplier, and assigns a prior probability p_1 on a good type. Let π^1 be the buyer's profit on a non-defective product and $\pi^0 < \pi^1$ on a defective one, payoffs being additive,[3] and let $\zeta^i = 1$ designate the event that the product is non-defective. Let

$$\bar{\pi}^G = \lambda^G \pi^1 + \left(1 - \lambda^G\right)\pi^0 \tag{1}$$

and similarly for $\bar{\pi}^B$. In periods 2 and 3, the buyer revises his beliefs about the quality of each supplier on the basis of information (defective product or not) he obtained by dealing with them (if he did) in the previous period. Let

$$p_t^i = \frac{\lambda_G p_{t-1}^i}{\lambda_G p_{t-1}^i + \lambda_B \left(1 - p_{t-1}^i\right)}, \quad t = 2,3 \tag{2}$$

be the revised probability that supplier *i* provides a non-defective product in period *t*, based on information from period *t-1*.

The buyer faces two sequential-sampling (or stopping-time) problems on two independent stochastic processes, but the decisions are not independent because sampling on one has consequences for the optimal stopping time on the other. The problem is thus potentially very complicated, but the limitation to two suppliers and three periods keeps it tractable.[4] Consider the third-period

---

[3] At this stage we consider only informational considerations. We introduce a taste for diversity and competition considerations in section 2.2.

[4] The problem of selecting the stochastic process that delivers the highest expected reward among a set of independent processes is known in the statistical-decision literature as a "multi-armed bandit" problem. One strategy, called "epsilon-first", consists of a sampling (exploratory) phase during which several "levers" are tried, after which the experimenter sticks to the lever for which he has the most optimistic belief based on information gathered during the sampling phase.

problem, and let V_3 be the buyer's expected profit. Suppose that he dealt with both suppliers in period 2. Then in period 3 he buys both units from the best, so

$$V_3(2) = p_3^*\left(2\overline{\pi}^G\right) + \left(1 - p_3^*\right)\left(2\overline{\pi}^B\right) \tag{3}$$

where

$$p_3^* = \max\left\{p_3^x, p_3^y\right\} \tag{4}$$

is the highest of the two posteriors. If he used just one of them in period 2, *i*, then he just keeps that one and

$$V_3(1) = p_3^i\left(2\overline{\pi}^G\right) + \left(1 - p_3^i\right)\left(2\overline{\pi}^B\right) \tag{5}$$

where p_3^i is the revised belief on supplier *i* used in period 2. Clearly, by definition of the max, $V_3(2) \geq V_3(1)$ and the expected difference, $\Delta V_3 = E\left[V_3(2)\right] - E\left[V_3(1)\right]$, is the value of information generated by keeping both suppliers in period 2.

In period 2, with two suppliers and a discount factor δ,

$$\begin{aligned} V_2(2) = {} & p_2^x p_2^y\left(2\overline{\pi}^G\right) + \left(1 - p_2^x\right)\left(1 - p_2^y\right)\left(2\overline{\pi}^B\right) \\ & + \left[p_2^x\left(1 - p_2^y\right) + p_2^y\left(1 - p_2^x\right)\right]\left(\overline{\pi}^G + \overline{\pi}^B\right) + \delta V_3(2). \end{aligned} \tag{6}$$

With one supplier,

$$V_2(1) = p_2^*\left(2\overline{\pi}^G\right) + \left(1 - p_2^*\right)\left(2\overline{\pi}^B\right) + \delta V_3(1) \tag{7}$$

where $p_2^* = \max\left\{p_2^x, p_2^y\right\}$.

In period 1, finally, the prior being the same on both suppliers, both are used, generating the information used to revise beliefs from p_1 to p_2^x and p_2^y respectively.

Clearly, the "interim" payoff collected in period 2 is higher, in expected value, with one supplier than with two, since in the former case the buyer buys only from the best whereas in the latter he carries both along. However, the expected

period-3 payoff is, as noted, higher when two suppliers are kept in period 2 because the information generated has a value. Thus, there is a trade off between concentrating on the most efficient supplier and keeping several in order to "test" them.

What does the value of the information depend on? Suppose that, at the end of period 1, the buyer kept only one supplier, the one with the highest probability of being good, and suppose (without loss of generality) that it was supplier *x*. Letting I_2 stand for the information available at the beginning of period 2, the conditional expectation of the period-3 gain is (see appendix):

$$E\left(V_3 \middle| I_2\right) = 2\left[p_2^x \bar{\pi}^G + \left(1 - p_2^x\right)\bar{\pi}^B\right]. \tag{8}$$

Let $\phi^y = \Pr\left(p_3^y > p_3^x \middle| p_2^y < p_2^x\right)$ be the probability that *y* would perform better than *x* if we could observe both in action in period 3. Using this, it can be shown that the value of the information is

$$\begin{aligned}\Omega &= \phi^y\left[E\left(V_3 \middle| p_3^y > p_3^x; I_2\right) - E\left(V_3^x \middle| I_2\right)\right] \\ &= 2\phi^y\left[E\left(p_2^y \middle| p_2^y > p_2^x\right) - p_2^x\right]\left(\bar{\pi}^G - \bar{\pi}^B\right).\end{aligned} \tag{9}$$

Thus, the value of the information depends on three multiplicative terms. The first is the probability that a good draw for the second-best supplier would reverse the ranking of beliefs. In our three-period model, the event that $p_2^y < p_2^x$ implies that *y* had a defect in period 1 while *x* did not. Then, if fortunes are reversed in period 2 (*x* has a defect while *y* has not), it is easily verified that posteriors at the beginning of period 3 will be just equal for *x* and *y*. So, at best, the buyer will be indifferent between *x* and *y* in period 3. In (9), we have thus $\phi^y = 0$ and, given the multiplicative form of Ω, the value of the information is nil: There is no reason to keep on sampling after period 1 and concentration *has* to take place. By contrast, in a 4-period framework, at the cost of tedious algebra it is (relatively) straightforward to show that a reversal of beliefs is possible with two successive lucky draws on *y* and two unlucky ones on *x*, and so, depending

on the parameters (λ and π) continued sampling (using both suppliers) can be optimal in period 2.

The second term is that in square brackets. Observe that it is decreasing in p_2^x; the better is the "front-runner" supplier (*x*) the less there is to gain from an eventual reversal of beliefs. In our 3-period setting, this doesn't say much, but in a multi-period setting it would have a potentially important consequence on which we will return. The third term, finally, is the difference in expected gains between a good and a bad supplier, which can be written as

$$\bar{\pi}^G - \bar{\pi}^B = \left(\lambda^G - \lambda^B\right)\left(\pi^1 - \pi^0\right). \qquad (10)$$

The first factor on the RHS of (10) is the difference between the prospects of a good and a bad supplier, a measure of their heterogeneity; the second is the effect of quality differences on profit, a measure of the industry's characteristics (quality-sensitivity). Thus, the value of information, which in our setting drives the search for quality, is increasing in their heterogeneity and in the sensitivity of buyers to product quality.

In order to get a better feel for what our simple model suggests empirically, we now turn to a few extensions of its baseline version.

2.2 Extensions

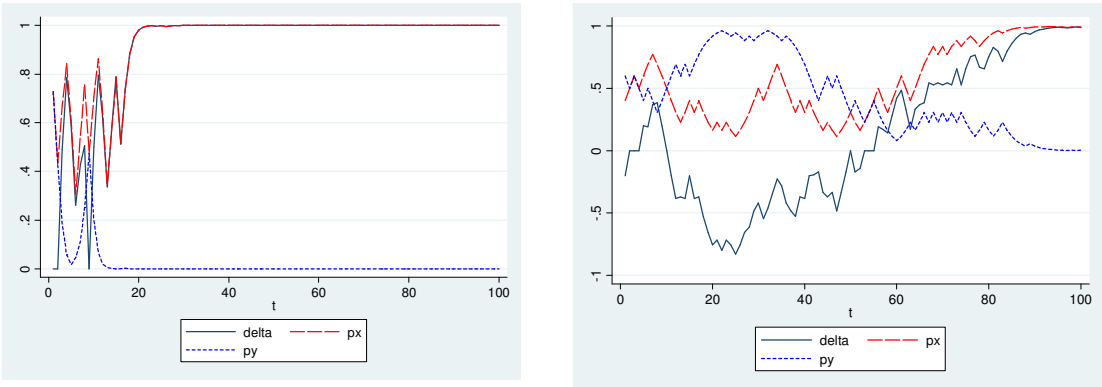
2.2.1 More than three periods

With more periods, the revision of beliefs (i.e. the difference between posterior and prior from one period to the next) becomes smaller over time as beliefs approach asymptotically zero or one, but how fast the process of revision converges depends, of course, on the parameters of the two processes. If the two distributions (good and bad) have similar parameters, it takes, in expectation, more time to tell apart the two types, which requires longer sampling. illustrates how the rate of convergence varies with the parameters. In the LHS panel the two distributions are characterized by sharply different parameters and beliefs

converge after twenty periods; in the RHS panel, the two distributions have similar parameters and the beliefs take almost a hundred periods to converge.

.Figure 1.
Random draws of Bernoulli processes in two cases

A pair of draws with $\lambda^G = 0.8, \quad \lambda^B = 0.3$ A pair of draws with $\lambda^G = 0.6, \quad \lambda^B = 0.4$



Notes: The dotted blue curve gives p_t^x, the revised probability that *x* is of the good type; the long-dashed red curve gives same thing for *y*, and the plain black curve gives the difference between the two. The same parameters are used to draw the observations and to update the beliefs (parameters are assumed common knowledge); *x* is of the good type and *y* of the bad type.

In the LHS case, positions tend to lock in fairly quickly. In the RHS case, longer sampling is needed to tell apart the two suppliers; however, note that the difference in expected returns ($\bar{\pi}^G - \bar{\pi}^B$), which is part of the value of the information, is also smaller, so the truth takes longer to appear but it matters less. Observe also that in the RHS panel, around iteration #20, supplier *x* has accumulated so many bad draws and supplier *y* so many good draws that the buyer is "almost certain" that *y* is of the good type, *even though this belief is false* (observe the dotted curve (*y*) approaching one between iterations #20 and #40). Going back to (9), we see that

$$\lim_{p_2^x \to 1} \Omega = \lim_{p_2^x \to 1} E\left(p_2^y \middle| p_2^y > p_2^x\right) - p_2^x = 0; \qquad (11)$$

that is, when the buyer becomes "almost certain" that his currently preferred supplier is of the good type, the value of information goes to zero and he stops sampling. If that were the case in the RHS panel of

.Figure 1 (where the preferred supplier around iteration 20 is y) the part of the red, long-dashed curve lying to the right of the stopping time would be censored. The remaining incumbent (here y) would then be the sole supplier until sufficient evidence accumulates to convince the buyer that he had bet on the wrong horse (in the figure, that becomes clear after about iteration 60 and the posterior on y finally converges to zero around iteration 90). The buyer would turn to an alternative supplier only when his revised opinion on the incumbent drops back below the evicted supplier's last posterior.[5]

2.2.2 New supplier entry

So far we did not consider the entry of new suppliers. The number of suppliers would enlarge if trade costs were coming down or if productivity was rising exogenously among producers in a pool of potential suppliers with heterogeneous productivity levels as in Helpman, Melitz and Rubinstein (2008). Suppliers would then appear progressively, creating scope for diversification of supplier sources at the extensive margin. As before, the repetition of transactions with incumbents would asymptotically reveal their quality, but strings of bad draws would always be possible even for good types, and their replacement would then set the clock back to zero for the new ones. With several entrants all characterized by similar priors, buyers would start by sampling all of them like at the beginning of our 3-period model, subsequently

[5] Note that in this setup there can be no "informational cascade". An informational cascade (Bikhshandani et al. 1992) can take place when a sequence of actors make binary decisions on a singe issue (say, buying or selling a stock) based on a noisy signal about the correct decision and on the observed behaviour of past players. Each player forms his own belief based on a weighted average of his signal and past players' actions, with weight on the latter that increases with the number of past players. Bikhshandani et al. show that there exists a critical number n such that, if n players observe the wrong signal and act accordingly, the n+1st will discard his own signal and follow the crowd. From then on, the herd behaviour cannot be reversed. Our setup is different for several reasons. First, the buyer is repeatedly getting information about his supplier, whereas in an informational cascade the individual experimenter gets only one signal that he compares with the actions of other (past) players. Second, the individual signal in an informational cascade is noisy, whereas it is not here.

concentrating on the best. Thus, episodes of diversification would be followed by episodes of concentration.

Thus, informational considerations in the multi-period setup suggest that, in sectors where quality matters and is not standard across suppliers, entrants will find it hard to unseat incumbents as long as those perform well. But, with stochastic quality draws, incumbents are bound to fail one day or another. When they fail sufficiently severely (i.e. with a string of bad draws in a row), a window of opportunity opens up for entrants, ushering in a new phase of diversification, quality search, and ultimate re-concentration on the best performers.[6] That is, diversification will happen by "bouts", as a result of repeated failures in established buyer-supplier relationships, rather than as a continuous phenomenon.

2.2.3 Taste for diversity

A taste for diversity can be introduced in the model by replacing the assumption of additive payoffs (see footnote 4 supra) by a utility function of the form

$$\Pi = \left[\sum_i (\pi_i)^\alpha \right]^{1/\alpha} \tag{12}$$

where $\pi_i \in \{\pi^0, \pi^1\}$ is the profit made on the purchase from supplier *i*. To see what happens to the model's basic predictions, consider period 3. The reasoning is similar for earlier periods. The period-3 payoff from using one supplier only (the preferred one), which was previously given by (3), is unchanged. That is,

$$V_3(2,1) = 2p_3^* \bar{\pi}^G + 2\left(1 - p_3^*\right)\bar{\pi}^B . \tag{13}$$

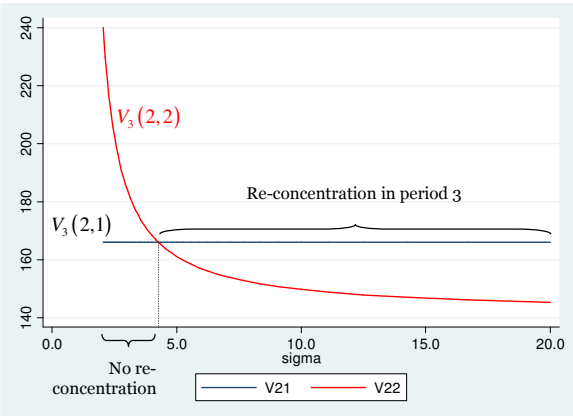
The corresponding payoff if the buyer uses both suppliers in period 3 is

[6] Failure may also be triggered endogenously by moral hazard if incumbents slacken the monitoring effort as time passes. For a reputational model with both selection and moral hazard, see e.g. Laeven and Perotti (2007).

$$V_3(2,2) = p_3^y \left\{ 2p_3^x \bar{\pi}^G + \left(1 - p_3^*\right)\left[\left(\bar{\pi}^G\right)^\alpha + \left(\bar{\pi}^B\right)^\alpha \right]^{1/\alpha} \right\}$$
$$+ \left(1 - p_3^y\right)\left\{ p_3^x \left[\left(\bar{\pi}^G\right)^\alpha + \left(\bar{\pi}^B\right)^\alpha \right]^{1/\alpha} + 2\left(1 - p_3^*\right)\bar{\pi}^B \right\} \tag{14}$$

Suppose, without loss of generality, that the preferred supplier is *x*. Replacing p_3^* by p_3^x in (3), it is easily verified that, for $\alpha = 1$, keeping one supplier is optimal (this is the benchmark case without a taste for diversity). However, as α goes down, the sign of the inequality is eventually reversed and the taste for diversity comes to dominate the selection effect. This is illustrated in Figure 2 where $V_3(2,1)$ and $V_3(2,2)$ are shown as functions of $\sigma = 1/(1-\alpha)$, the elasticity of substitution between the two suppliers, for assumed parameter values. For values of σ below 4.3, the taste for diversity dominates and keeping both suppliers is optimal; for values above 4.3, the selection effect dominates and keeping only one supplier is optimal.

Figure 2
Period-3 profit from one vs. two suppliers, as a function of the elasticity of substitution



Notes: Simulated parameter values are $\pi 1$=100, $\pi 0$=50, λG=0.8, λB=0.1, $p3x$=0.8, $p3y$=0.1.

What does this mean for our model? Essentially that the taste for diversity acts as a counterforce to the selection effect, generating situations where the Bayesian update of beliefs designates one supplier as preferable to others but the buyer nevertheless keeps several because he values diversity.

2.2.4 Competition among suppliers

Competition between suppliers affects the model in the same way the taste for diversity does. Formally, it can be introduced by assuming that when our buyer procures from two suppliers, they charge a duopoly price (presumably Bertrand since each of them is assumed to have the capacity to cover the buyer's entire needs) whereas if he buys from one supplier only, that supplier charges a monopoly price.[7] This can be easily accommodated by rewriting (12) as

$$\Pi = \begin{cases} \left[\sum_i \left(\pi_i^D\right)^\alpha\right]^{1/\alpha} & \text{with two suppliers} \\ 2\pi_i^M & \text{with only one} \end{cases} \tag{15}$$

and $\pi_i^D > \pi_i^M$. The algebra is the same as in the previous section but the $V_3(2,2)$ curve of Figure 2 shifts up, moving the crossing point to the right. That is, the range of the model's parameters where the buyer prefers keeping several suppliers in spite of their heterogeneity (as reflected in posterior beliefs) expands, because keeping several puts them in competition with each other.

All in all, our simple model suggests essentially this:

1. Diversification of import sources can be driven by three forces: (a) quality search in the presence of a selection problem; (b) an exogenous taste for diversity (extension 1), or (c) a desire to limit monopoly positions (extension 2).
2. When driven by quality search, diversification is only a temporary phenomenon, as the buyer will, at the end of each search phase, re-concentrate on the best supplier.
3. Incumbent suppliers' established positions will periodically be unseated by strings of bad quality draws, which will trigger the onset of new search phases.

---

[7] Closing the model would require, in addition, some ad-hoc limitation to contestability, like a switching cost, to make monopoly pricing sustainable.

Thus, whereas the competition and taste-for-diversity forces generate maximum diversification at all times (an essentially static prediction), quality search suggests alternating phases of diversification and re-concentration. The range of parameter values where these alternating phases take place depends on the counterforces, but their *existence* is implied only by the informational features of the model. Thus, volatility in concentration levels can be taken as a hallmark of informational phenomena.[8]

We now turn to an empirical exploration of three related questions: (i) how the concentration of OECD imports has evolved over time; (ii) whether it displays a volatility suggestive of the kind of informational considerations outlined in the simple model above; (iii) what (if any) implications this all has for access to OECD markets for extra-OECD producers.

3. Measuring geographical import concentration

3.1 Indices and data

As discussed in the introduction, we measure, product by product, the geographical concentration of imports across origin countries. Our measures are standard ones: Herfindahl and Theil.[9] The Herfindahl index for good k, normalized to range between zero and one, is

---

[8] However, supply shocks knocking out suppliers periodically could also create exogenous volatility at the extensive margin. This is to be kept in mind in the empirical exploration that follows, as baseline volatility is unlikely to be exactly zero.

[9] We decided not to use Gini coefficient because of the issues associated with this concentration index. The Gini coefficient is a numerical representation of the degree of concentration and represents the distance between the Lorentz curve and the 45° line (egalitarian distribution). There are two issues with Gini coefficients. First, they place more weight on changes in the middle part of the distribution. If a transfer occurs from a larger number of exporters to a smaller number of exporters, it has a greater effect on the Gini if these numbers of exporters are near the middle rather than at the extremes of the distribution. Second, if the Lorentz curves cross, it is impossible to summarize the distribution in a single statistic without introducing value judgements. While studying concentration of import across time these issue should be

$$H_k^* = \frac{\sum_i \left(s_k^i\right)^2 - 1/n_k}{1 - 1/n_k} \tag{16}$$

where $s_k^i = x_k^i / x_k$ is the share of origin country *i* in OECD imports of product *k* and n_k is the total number of countries exporting good *k* (we will discuss in more detail below alternative definitions of the set of exporting countries).

Theil's entropy index (Theil 1972) is given by

$$T_k = \frac{1}{n_k}\sum_{i=1}^{n_k}\frac{x_k^i}{\mu_k}\ln\left(\frac{x_k^i}{\mu_k}\right) \quad where \quad \mu_k = \frac{1}{n_k}\sum_{i=1}^{n_k} x_k^i \tag{17}$$

These indices are dependent on the definition of n_k, the number of "potential exporters". Our baseline definition of the set of potential exporters is the simplest one: it is the set of all countries having exported good *k* to *some* destination in the world (not necessarily OECD countries) at least two years in a row over the sample period. We impose the requirement of two consecutive years of exports instead of just one in order to ensure that the exporter is a successful one (Besedes and Prusa 2006a, 2006b show that two years is the median duration of export spells; only one year might signal failure rather than the capacity to export). This definition has the advantage of being time- and importer-invariant (the latter matters for the part of our analysis where we disaggregate OECD imports by importing country). In order to explore action at the extensive margin, we also consider the simple number of exporters of good *k* to OECD countries.

Our data is COMTRADE import data for OECD countries (either taken as a bloc or disaggregated by importer) at the product level. Our preferred product classification is SITC4. The alternative, HS6, is more disaggregated (with 4'990

---

relevant. Herfindahl and Theil indices are robust to these sensitivity issues [on this, see Sen (1997)].

to 5'016 lines depending on the year against 1'158 to 1'300 for SITC4), but the sample period is longer with SITC4, which also underwent fewer revisions. In terms of country coverage, SITC4 data covers 210 countries between 1962 and 2006 (44 years); HS6 coverage is nominally available starting 1988. Descriptive statistics for our sample are shown for our indices in Table 1 .

Table 1
Descriptive statistics

Variable	Obs	Mean	Std. dev.	Min	Max
n_k	54030	144	53	12	253
All exporters					
n_k^{OECD}	54030	57	35	1	223
Herfindahl	54030	0.19	0.13	0.03	1
Theil	54030	2.68	0.48	0.86	4.87
Extra-OECD exports only					
n_k^{OECD}	53769	35	31	1	194
Herfindahl	53769	0.35	0.22	0.03	1
Theil	53769	3.05	0.60	0.13	5.30

Note: All variables are defined at the product (SITC4) level. That is, the "number of countries with nonzero exports 2 consecutive years in the sample period" means the number of countries that exported a given product 2 consecutive years to somewhere in the world (i.e. the number of potential exporters of that product).

3.2 Intensive and extensive margins: Prima-facie evidence

Figure 3 shows the evolution of simple averages over all products of our two concentration indices (Herfindahl and Theil) expressed as indices relative to the sample's initial year. That is, for Herfindahl, Figure 3 shows

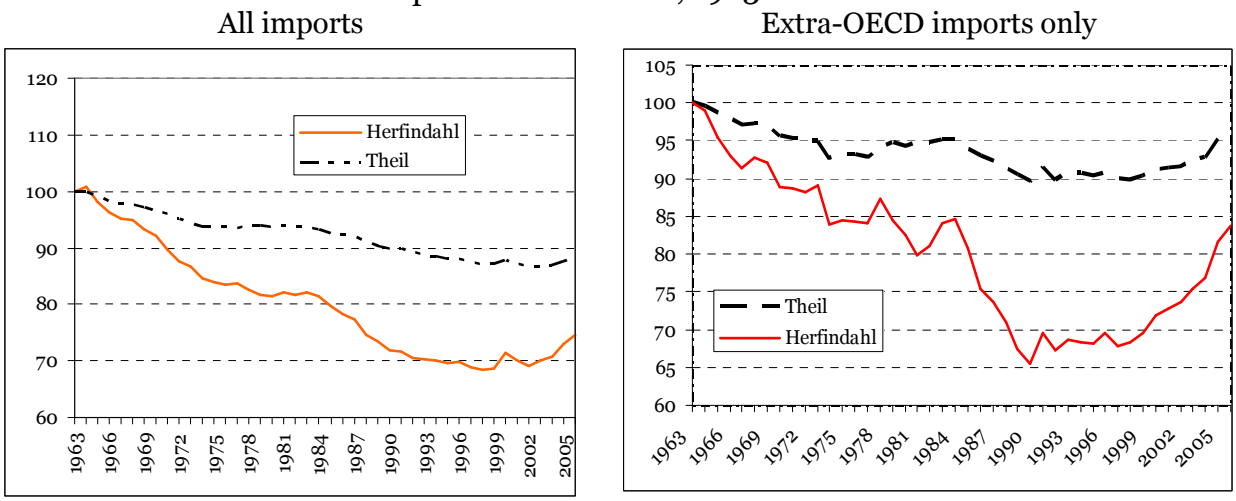
$$H_t = 100\bar{H}_t / \bar{H}_0 \tag{18}$$

where $\bar{H}_t = \sum_k H_{kt} / n_t$ is the simple average for year *t* of the Herfindahl indices calculated for all goods *k* imported at *t*. The calculation is the same for the Theil index.

Panel a) shows concentration indices calculated using all OECD imports (i.e. imports from all partners, including intra-OECD ones). A strong diversification trend is shown by both indices until 1999 (Herfindahl) and 2002 (Theil), after which both rise until 2006, the sample's last year (by 8.6% for Herfindahl and 1.5% for Theil). Panel b) shows concentration indices calculated using only extra-OECD partners (i.e. developing countries). Both Herfindahl and Theil indices decrease until 1990 (modestly for Theil, which goes down by about 10%

over the period) and then go up. Between 1999 and 2006, the Theil index rises by 7.4%, almost three times its coefficient of variation over the period 1963-99.

Figure 3
OECD import concentration, 1963-2006



Note: base 100, 1963; simple averages of indices over all products. Data from COMTRADE

The trend reversal is unmistakable as far as imports from non-OECD countries are concerned. However, it takes place quite late in the sample period. In order to verify whether it is statistically significant, and that it is not a pure composition effect between products (i.e. a sectoral shift away from widely-procured products toward narrowly-procured ones), we now turn to regressions of concentration indices on time and its square using fixed (product) effects. Results are shown in Table 2 . Columns (1)-(2) show results with concentration indices (the dependent variable) calculated over all imports (including intra-OECD) whereas columns (3)-(4) show results for extra-OECD imports only (a more interesting measure from a developmental perspective).

The within estimator confirms the convex time trend, as both time and its square are significant with opposite signs.

Table 2
Regression results, OECD import concentration on time trend

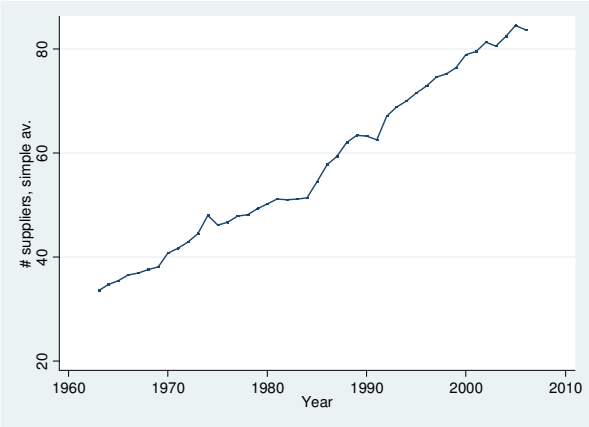
	All imports		Extra-OECD imports only	
	(1)	(2)	(3)	(4)
Regressors:	Herfindahl	Theil	Herfindahl	Theil
time	-0.002***	-0.013***	-0.007***	-0.023***
	(-13.19)	(-34.03)	(-31.80)	(-43.53)
timesq	0.000	0.000***	0.000***	0.000***
	(1.616)	(4.723)	(23.17)	(28.13)
Constant	0.218***	2.913***	0.425***	3.330***
	(202.6)	(861.4)	(221.1)	(686.9)
Observations	54030	54030	53769	53769
Number of index	1301	1301	1301	1301
R-squared	0.571	0.671	0.510	0.570
turning point	2001	2001	1993	1997
Product FE	yes	yes	yes	yes

Notes: t statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

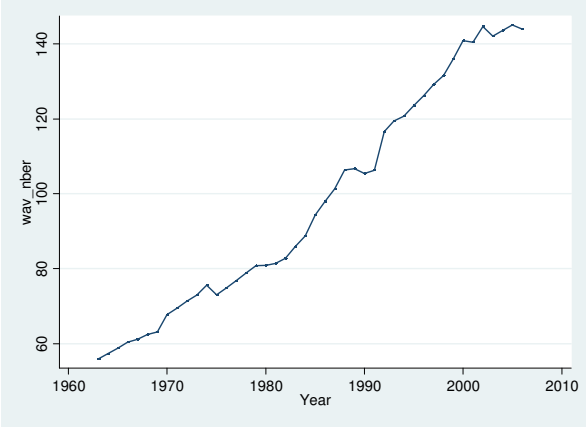
As for the extensive margin, Figure 4 shows the evolution of simple and import-weighted averages, across SITC4 lines, of the number of exporters to the OECD over the sample period.

Figure 4
Average number of exporters to OECD, 1963-2006
All suppliers

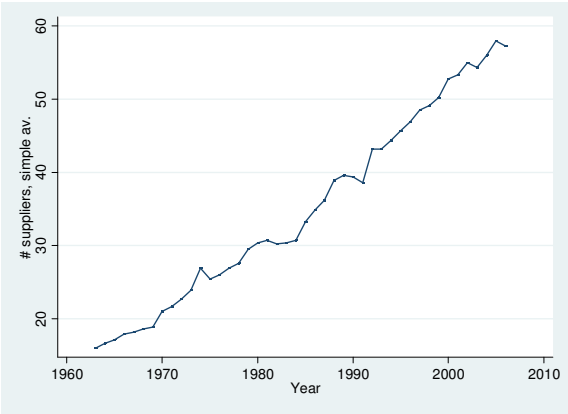
(a) Simple average

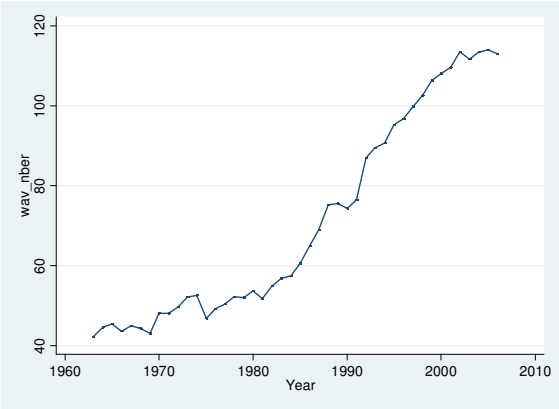


(b) Import-weighted average



Extra-OECD suppliers only





Note:
a/ Simple averages of number of exporters to OECD at the product (SITC4) level.
b/ Import-weighted averages (weights = shares of each SITC4 product in OECD imports in given year)

The extensive margin as measured by the average number of source countries does not seem to show the same kind of trend reversal that we observed in the concentration indices, which pick up action at both the extensive and intensive margins. Simple averages of the average number of OECD suppliers by SITC4 product category are rising monotonically over time. Import-weighted averages are leveling out after 2000, but this is not very surprising. The numbers on the vertical axis show that on a trade-weighted basis, the average number of suppliers per product was over 100. For many products, this is likely to exhaust the pool of potential exporters, so a leveling off is to be expected.

Table 3 reports the results of pooled and fixed-effects regressions of the number of exporters to the OECD on time, its square, and a specific time trend for the post-2000 period.

In the latter period, as expected from Figure 4, there is a decline in the rate of increase in the number of exporters to the OECD. This is reflected by the negative coefficient on Post 2000. This inflexion is however not strong enough to reverse the trend. The observed re-concentration of OECD imports thus seems to be entirely caused by action at the intensive margin.

Table 3
Regression results, Number of countries exporting to OECD

	(1)	*(2)*	*(3)*	*(4)*
Time	0.662	0.593	0.583	0.543
	(17.18)***	(12.71)***	(38.49)***	(29.66)***
Time, squared	0.008	0.010	0.008	0.009
	(9.19)***	(8.52)***	(24.76)***	(20.79)***
Post 2000		-0.391		-0.228
		(2.63)***		(3.91)***
Constant	15.103	15.481	16.560	16.777
	(40.80)***	(38.99)***	(113.19)***	(107.19)***
Observations	53'770	53'770	53'770	53'770
R-squared	0.17	0.17	0.54	0.55
Number of SITC4			1'301	1'301
Fixed (prod.) effects	no	no	yes	yes

Note: Dependent variable: Number of non-OECD exporters to OECD.
The time variable is an index starting as 1963 = 1. The post-2000 variable is another index starting at 2000 = 1. The panel is unbalanced.

In order to explore further what might be driving the apparent re-concentration of OECD imports, we now decompose OECD imports by importing country and construct a three-dimensional panel whose unit of observation (the basis for the calculation of our concentration indices) is a product imported by an OECD country in a year (a triplet importer × product × year). Looking at things this way allows us to look for another type of composition effects that would work as follows. Suppose that the OECD has two members, *A* and *B*, with *B* sourcing its imports of a given product more narrowly than *A*. A rise in B's share of OECD imports will raise the OECD-wide import concentration index for that product through a pure composition effect, although in our previous regressions this would be a within-product rise in the concentration index. Regression results are shown in Table 4 .

Table 4
Regression results, OECD import concentration on time trend
Herfindahl

	(1)	*(2)*	*(3)*	*(4)*
Time	-0.010	-0.009	-0.009	-0.007
	(115.74)***	(86.33)***	(140.45)***	(92.51)***
Time, squared	0.000	0.000	0.000	0.000
	(51.13)***	(27.11)***	(41.94)***	(3.42)***
Post 2000		0.005		0.008
		(14.76)***		(39.35)***
Constant	0.807	0.802	0.798	0.789
	(932.13)***	(859.28)***	(1320.17)**	(1216.18)***
Observations	1'154'420	1'154'420	1'154'420	1'154'420
R-squared	0.07	0.07	0.13	0.13
Fixed effects a/	no	no	yes	yes

Theil

	(1)	*(2)*	*(3)*	*(4)*
Time	-0.012	-0.009	-0.019	-0.015
	(63.78)***	(41.27)***	(155.47)***	(100.10)***
Time, squared	0.000	-0.000	0.000	-0.000
	(7.94)***	(7.32)***	(35.37)***	(7.05)***
Post 2000		0.012		0.020
		(19.11)***		(47.64)***
Constant	4.236	4.222	4.358	4.336
	(2336.35)***	(2160.58)***	(3626.10)***	(3363.56)***
Observations	1'154'420	1'154'420	1'154'420	1'154'420
R-squared	0.05	0.05	0.18	0.18
Fixed effects a/	no	no	yes	yes

Number of partners

	(1)	*(2)*	*(3)*	*(4)*
Time	0.093	0.153	0.062	0.091
	(31.30)***	(42.34)***	(40.07)***	(48.33)***
Time, squared	0.003	0.001	0.004	0.003
	(46.90)***	(14.21)***	(129.70)***	(75.63)***
Post 2000		0.305		0.147
		(29.10)***		(27.01)***
Constant	3.259	2.917	3.033	2.871
	(111.72)***	(92.81)***	(196.73)***	(173.47)***
Observations	1'154'420	1'154'420	1'154'420	1'154'420
R-squared	0.09	0.09	0.29	0.29
Fixed effects a/	no	no	yes	yes

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%
a/ fixed effects by importer × product pair

Several observations come out of Table 4 . First, the re-concentration apparent in the Herfindahl and Theil indices seems robust to the introduction of fixed effects by importer × product pair. The news comes from the extensive margin, where not only the square term on time preserves the mononicity of diversification but even the post-2000 time trend no longer indicates a trend inflexion in the very last years. The disappearance of the trend inflexion (apparent in Table 3 which included fixed effects by products but not by importing country because the unit of observation was all-OECD imports) suggests that the inflexion resulted from a composition effect between importers as described above.

Thus, the observed reconcentration of Table 2 is robust to the decomposition of OECD imports by importing country. However, as Tables 3 and 4 show, it does not occur at the extensive margin, all of the action being at the intensive margin.

3.3 Import concentration and income

We now explore the relationship between import concentration and income levels. "Within" importers, income levels are of course highly correlated with time. However, the correlation is not perfect, and looking at income levels also adds information in the between-country dimension. Accordingly, Table 5 reports both pooled and fixed-effect regression results for the relationship between the level of income of OECD importers and the concentration of their imports.

Income squared has a positive coefficient in all concentration regressions and a negative one in the number-of-partners regression, and this is robust to the introduction of fixed effects by country × product pair. Thus, the convexity (concavity for number of products) appears not just in the between-importer dimension but also in the within. Estimated turning points are shown in Table 6 and illustrated in Figure 5.

Table 5
Import concentration and income levels

	Herfindahl				Theil				# non-OECD partners			
	(1)		*(2)*		*(3)*		*(4)*		*(5)*		*(6)*	
Income	-1E-05	***	-2.10E-05	***	-2.75E-05	***	-5.08E-05	***	3.24E-04	***	8.44E-04	***
	(-54.31)		(-91.81)		(-77.26)		(-111.8)		(45.46)		(152.4)	
Income, squared	1E-10	***	2.46E-10	***	3.42E-10	***	5.65E-10	***	-8E-10	***	-5.31E-09	***
	(32.50)		(57.98)		(45.57)		(67.05)		(-5.321)		(-51.64)	
Constant	7.42E-01	***	9.26E-01	***	4.33E+00	***	4.74E+00	***	2.24E+00	***	-7.14E+00	***
	(353.9)		(319.1)		(1068)		(822.7)		(27.58)		(-101.6)	
Observations	735'000		735'000		735'000		735'000		735'000		735'000	
R-squared	0.061		0.096		0.046		0.134		0.086		0.227	
# groups	0.015		0.036		0.030		0.059		0.040		0.213	
Fixed effects	no		yes		no		yes		no		yes	

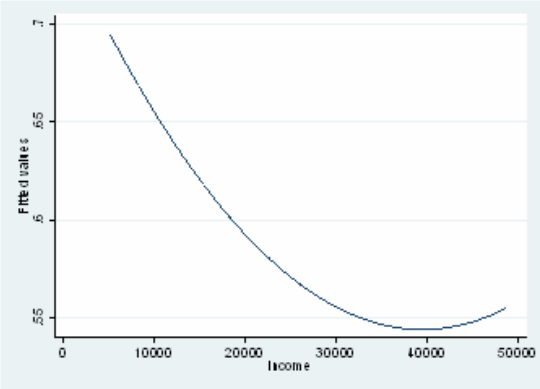
Notes
a/ Fixed effects by importer × product pair.

Table 6
Estimated turning points, 2005 PPP dollars

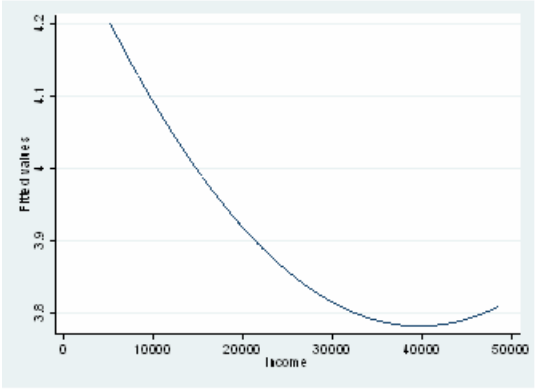
	Pooled	Within
Herfindahl index	39'679	42'703
Theil index	40'250	44'958
# suppliers	n.a.	n.a.

Figure 5
Predicted concentration indices (pooled)

Herfindahl



Theil



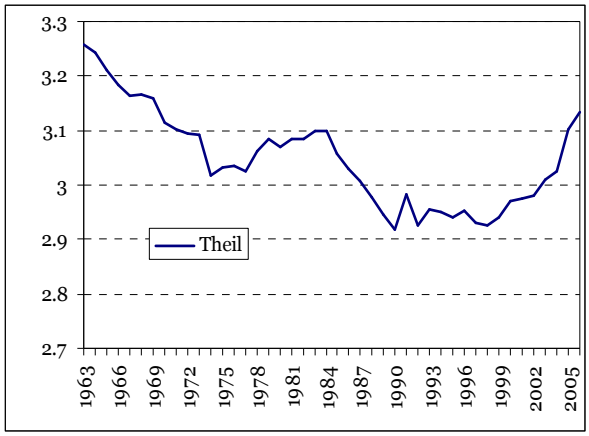
Obviously the turning point is very high in terms of income, leaving only Norway since 1997 and the US since 2004 to the right of it.

3.4 The China effect

Considering the rising importance of OECD trade with China over the last decade, we must control for the role that China may play in that reconcentration. Figure 6 shows the evolution of the Theil index for extra-OECD imports, both with and without China. The figures show that China is indeed driving the observed reconcentration. Further evidence is provided in Table 7, which shows that the coefficient on time squared loses its significance when China is excluded from the sample.

Figure 6
Theil index for OECD imports excluding China, 1963-2006

Extra-OECD imports



Extra-OECD imports excluding China

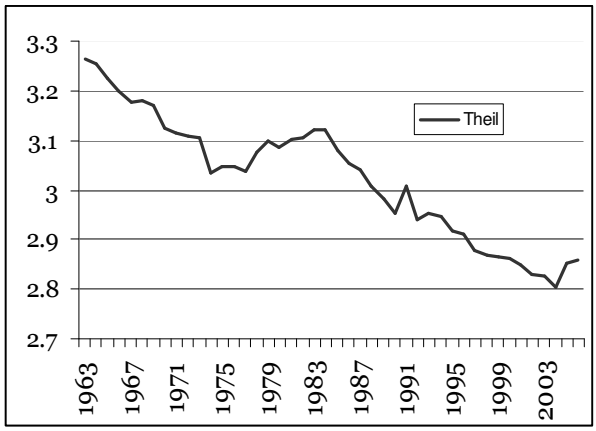


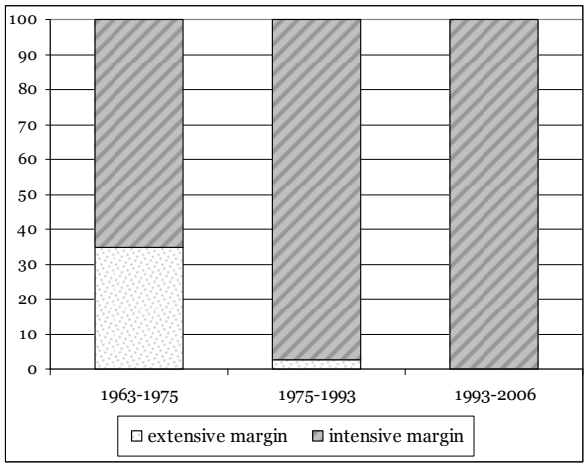
Table 7
Regression results, OECD import concentration on time trend, excluding China

Regressors:	Herfindahl	Theil
time	-0.004***	-0.013***
	(-17.08)	(-24.61)
timesq	0.000*	0.000
	(1.925)	(0.0720)
Constant	0.419***	3.302***
	(218.4)	(677.4)
Observations	53763	53763
R-squared	0.531	0.581
Product FE	yes	yes

Notes: t statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

Figure 7 confirms that the increased concentration of OECD imports on China occurs at the intensive margin: In recent years, no new product line opened between China and its OECD trade partners.

Figure 7
Contribution of the intensive and extensive margin to China-OECD imports growth, 1963-2006



Thus, although OECD imports are increasingly concentrating on Chinese products, evidence from section 3 shows that OECD countries continue to open new imports lines with extra-OECD countries. Put differently, China's expanding exports to the OECD do not seem (yet) to crowd out the entry of new exporter/product pairs on OECD markets.

The evidence in this section shows diversification taking place over time and as a function of income levels, subject to two caveats: (i) China's growing share of OECD imports, which caused a recent re-concentration, and (ii) an ultimate reversal (i.e. a re-concentration) at the intensive margin at very high levels of income. However, the evidence so far does not say what drives this progressive diversification, nor what counter-forces, if any, may be at play, preventing it from being instantaneous. That is, prima-facie evidence does not tell us when the doors of OECD markets open and when they close. In order to deal with this question, we now explore empirically the implications of the quality-search model of section 2.

4. Concentration and quality search: The evidence

Quality screening of suppliers can be expected to be more important for products whose quality matters and where it is not standardized across suppliers. We attempt here to identify this type of effect by exploring how product type correlates with the concentration of supply sources using Rauch's classification of goods into homogenous, reference-priced, and differentiated. Rauch's classification has the advantage of being, in principle, orthogonal to economies of scale, as it characterizes goods rather than their production process (although the two may be related). Regression results are shown in Table 8.

Table 8
Regression results, concentration and product differentiation

	Herfindahl (1)	Theil (2)	# partners (3)
Time	-0.006	-0.011	-0.002
	(19.17)***	(16.36)***	(0.15)
Time^2	0.000	0.000	0.004
	(5.52)***	(1.30)	(17.22)***
Reference price	-0.023	0.013	0.188
	(10.58)***	(2.85)***	(2.39)**
Differentiated	-0.138	-0.204	5.151
	(64.69)***	(45.50)***	(66.65)***
Importer GDP	0.000	0.000	0.000
	(5.12)***	(2.60)***	(13.49)***
Constant	0.898	4.742	-1.357
	(97.80)***	(245.33)***	(4.07)***
Observations	99'638	99'638	99'638
R-squared	0.15	0.19	0.21
Importer FE	yes	yes	yes

Compared to homogenous goods (the omitted category), differentiated goods are characterized by significantly lower geographical concentration, although the effect is quantitatively small.

The model of section 2 also suggested that, for products subject to supplier screening, diversification should take place by "bouts", followed by re-

concentration as buyers and suppliers establish relationships based on accumulated quality records. Thus, there should be more variability in the degree of concentration for those goods. Accordingly, we report in Table 9 regression results of the standard deviation of concentration indices "within products" (i.e. calculated over the entire sample period for each product) on Rauch's index of production differentiation. Note that, in so doing, we reduce the sample's dimensionality from three (importer × product × time) to two (importer × product). We estimate the regressions with importer fixed effects.

Table 9
Regression results, volatility of concentration on product differentiation

	σ_herf (1)	σ_theil (2)	σ_nber (3)
Reference price	0.010	0.034	0.741
	(2.89)***	(4.90)***	(6.56)***
Differentiated	0.025	0.100	3.264
	(7.57)***	(15.39)***	(20.02)***
Constant	0.160	0.290	0.741
	(20.58)***	(20.80)***	(4.21)***
Observations	3'122	3'122	3'122
R-squared	0.07	0.13	0.24
Importer FE	yes	yes	yes

Notes: Robust t-statistics in parentheses. Dependent variable: standard deviation of concentration index (Herfindahl, Theil or number of suppliers); sample is a cross section of SITC4 products (not all SITC4 goods are coded in Rauch's database, hence the relatively small number of observations). Results reported use Rauch's "liberal" classification; using his "conservative" classification yields almost identical results.

Results are as suggested by the model. The volatility of concentration indices is higher for differentiated products, and the difference between categories is very large. For instance, the standard deviation of the number of suppliers is, on average, 2.35 for homogenous goods. For differentiated goods, ceteris paribus, it rises to 5.61 (2.35 + 3.26), a 138% increase. This indeed suggests alternating periods of diversification and concentration.

The model also suggests that at the end of quality-search phases, concentration should take place on the best-performing suppliers. This is a hypothesis that is

difficult to verify, as quality is largely unobserved, but the following exercise provides a rough indication about it. For each good and year, we calculate an import-weighted average of unit values. If re-concentration, when it takes place, is on the best performers, year-on-year changes in the Theil index should be correlated with changes in the average unit value (that is, when concentration rises, the average unit value rises as well, through a composition effect, because high-quality suppliers are given a higher share). Results of a regression of first differences in Theil indices on first differences in weighted-average unit values are shown in Table 10.

Table 10
Regression results, change in Theil on change in unit values

	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Dep. Var.	Δ Theil	Δ Theil	Δ Theil	Δ Theil	Δ Theil
Regressors					
Δ UV	4.50e-07*	4.88e-07**	4.79e-07**	4.44e-07*	4.80e-07**
	(2.47e-07)	(2.41e-07)	(2.41e-07)	(2.46e-07)	(2.41e-07)
Constant	-0.0107***	-0.0107***	-0.0138***	-0.00623***	-0.0133***
	(0.000268)	(0.000265)	(0.00262)	(0.00185)	(0.00187)
Observations	*1'059'984*	*1'059'984*	*1'059'984*	*1'059'984*	*1'059'984*
Number of index	*36'016*	*36'016*		*36'016*	*36'016*
R-squared	*0.000*	.	*0.003*	*0.002*	.

Notes: Robust t-statistics in parentheses. Column 1 uses fixed effects (by product × importer); column 2 uses random effects; column 3 uses dummies by year, product and importer, and columns 4 and 5 are like 1 and 2 but with time dummies.

It can be seen that positive year-on-year changes in the Theil index are, by and large, associated with positive changes in unit values, providing further support to the model's basic prediction. A similar regression performed using the exporter's GDP per capita instead of unit values gives a qualitatively similar result, suggesting that when re-concentration takes place, it is on suppliers located in higher-income countries. The evidence in Table 10 is suggestive of a quality-search process rather than a price-search one (where the price-search phase would settle on the lowest-price supplier).

5. Getting to OECD markets

For most if not all export products, OECD markets are the largest. Even if, on aggregate, they are not necessarily the fastest growing, at the level of the firm they typically offer the largest expansion potential. However they are also often difficult in the sense of requiring product and service quality standards that not all firms and countries can offer. Whether export promotion strategies should target those markets directly, or whether they should proceed in steps, using regional LDC markets as stepping stones for exporters, is a recurrent debate among policy makers in developing countries. We explore in this section what the aggregate evidence has to say, i.e. whether the norm among non-OECD exporters is to try OECD markets first or whether those are reached only after exporting experience has been accumulated on non-OECD markets.

5.1 Prima-facie evidence

We use again bilateral trade data at the SITC4 level. Our unit of observation is the first year of a bilateral export spell of an SITC4 product, i.e. a vector $\mathbf{x}_{kt}^{ij}$ = [exporting country *i*, importing country *j*, product *k*, initial year *t*]. For each non-OECD exporter and product, using mirrored data[10] we isolate (i) the first year of export to any destination in the world, and (ii) the first year of export to any OECD country. A "new export" is a product that is exported for the first time in the sample period (*not* an export spell starting after a temporary trade interruption). Thus, a spell starting in 1964 after zero values in 1962-63 is a new export; but a spell starting in 1966 after a 1962-63 spell interrupted in 1964-65

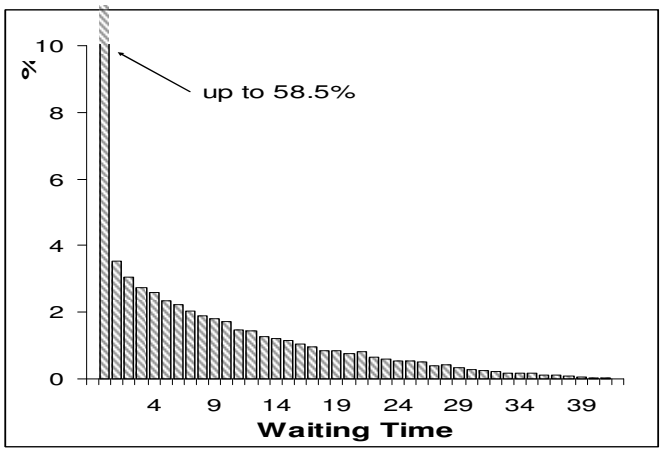
[10] "Mirroring" consists of using the importing country is the reporter and the exporting one as the partner. This improves the reliability of trade data as national customs administrations (the ultimate source of COMTRADE data) usually monitor imports more closely than exports. Mirroring can however lead to distortions in the presence of smuggling or under-reporting. In order to circumscribe the problem, CEPII reconciles import and export data on a case-by-case basis in the BACI database. We stick to COMTRADE data here because BACI does not cover enough years; in addition, for many exercises in this paper our reporter countries are OECD countries whose data is usually reliable.

is not. We treat left-censoring at the sample period's initial year (1962) by excluding from new exports the spells starting in 1962 or 1963.[11] The number of years between the first year of export to the world and to OECD countries is the "waiting time" during which a country exports the product in question before shipping it to OECD markets. If it is zero, exports go directly to the OECD. In assessing the frequency of instances of "OECD-first", "OECD after one year" and so on, the unit of observation is a country-year pair. However multiple observations will appear in the database when a country starts exporting a product to several non-OECD destinations simultaneously, or when it starts to export it to several OECD destinations. In order to avoid double counting, we collapse those multiple-destination occurrences into single observations. That is, if in 1990 Ghana exports mil for the first time and does so simultaneously to Burkina Faso and Mali, we count [Ghana, . , mil, 1990] as a single observation where the dot stands for the export destination (Burkina Faso and Mali in this case).

Overall, 76% of non-OECD new exports reach the OECD markets at some point in time. Among these exports that are eventually shipped to the OECD, more than half (58.5%) are first-export instances (i.e., the first year of export to the world and to OECD countries is the same). Figure 8 depicts the frequency of the "waiting time" for the sub-sample of non-OECD countries engaged in trade with the OECD. When first-export instances are to non-OECD countries, the proportion of observations with one year before the first export to an OECD market is 3.5% after which it decreases exponentially as a function of the waiting time.

[11] We tested for alternative assumptions (i.e., 3 and 4 years of no trade prior to first appearance in the database). Results are similar to the ones presented here and are available upon request.

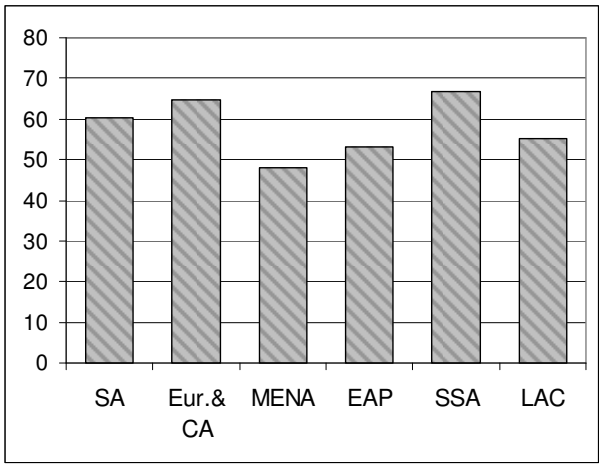
Figure 8
Frequency of the waiting time



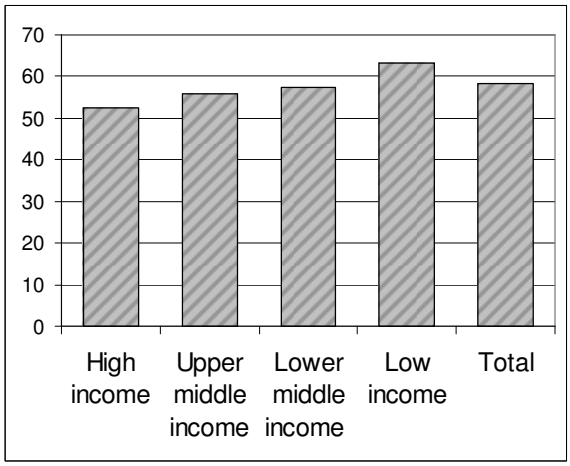
As shown in Figure 9, the percentage of new exports that go directly to the OECD is remarkably similar across regions and levels of income.

Figure 9
Percentage of OECD-first instances

(a) By region



(b) By level of income



Notes: SA: South Asia; Eur. & CA: Europe and Central Asia; MENA: Middle East and North Africa; EAP: East Asia and Pacific; SSA: Sub-Saharan Africa; LAC: Latin America and Carribeans. Vertical axis measures the percentage of (exporter, year, product) observations that are exported first to one or several OECD markets.

Interestingly, the percentage is slightly higher for low-income countries (63.1%) although policy-makers and business leaders in these countries are often

intimidated by the suggestion of trying OECD markets directly. For sub-Saharan African exporters, the percentage (66.7%) is higher than any other region. This may reflect a scenario where European buyers directly approach and train African producers to fill specific niches in the value chain under the benefit of preferential trade regimes.[12]

In order to verify if tariff preferences have anything to do with the decision to shoot directly for OECD markets, we restricted the sample to SSA exporters and ran a probit of a binary variable equal to one when the first export is to E.U. markets on the E.U.'s MFN tariffs, which approximate the tariff preference margins available to SSA producers under the EPA/EBA regimes. Results are shown in Table 11.

Table 11
Incidence of "EU-first" exports of SSA products and tariff preference margins: probit marginal effects

	(1)	(2)	(3)	(4)
EU MFN tariff	0.004	0.004	0.003	0.003
	(5.80)***	(5.13)***	(4.27)***	(4.08)***
Upper middle income	-0.242	-0.251		
	(8.08)***	(8.41)***		
Lower middle income	-0.025	-0.014	-0.501	0.039
	(0.97)	(0.55)	(5.81)***	(0.64)
Low income	-0.048	-0.036	-0.219	0.002
	(1.99)**	(1.57)	(4.30)***	(0.03)
Sector controls	no	yes	no	yes
Exporter controls	no	no	yes	yes
Observations	19'328	19'328	19'200	19'192
Pseudo R2	0.025	0.1332	0.2206	0.3186

Notes: Dependent variable: probability of a direct export to OECD markets; probit marginal effects ; for dummy variables (exporter income levels) give effect of changing the level from zero to one on the dependent variable; z-statistics in parentheses.

[12] Preferential regimes available to SSA exporters on EU markets include the EBA (Everything But Arms) initiative, which gives tariff-free access for exports from LDCs except in "special-regime" sectors like sugar, beef, dairy, or bananas, and the EPAs (Economic Partnership Agreements) which replace the Cotonou Convention regime for those countries in SSA that have signed these agreements (for others, the legal regime of exports is presently uncertain).

The first column reports pooled probit regression results, the second with broad sector controls at the SITC1 level (9 sectors), the third with exporter controls, and the fourth with sector and exporter controls. In all cases, the partial correlation of E.U. MFN tariffs (and hence preferential margins) with the probability of a direct export to the E.U. is positive, giving support to the conjecture above, although the effect is very small (the marginal effects reported in Table 11 show that a one percentage point rise in the preference margin raises the probability of a first export to E.U. markets by only 0.004). The probit, by construction, does not allow for negative values. The negative signs on dummies' coefficients are curves shifter and should not be interpreted literally. We may however compare the coefficient across income groups. As we are considering the sub-sample of sub-Saharan African countries, it does not include high-income countries. The three included dummies exhaust the sample and the constant is omitted. Compared to upper middle income countries, lower income countries have higher probability of assessing the EU market at first incidence. This may reflect the fact that the least developed countries are the one having greater access to preferential trade regimes.

5.2 Export-expansion paths and contracts

The regressions reported above included only coarse sectoral dummies that do not tell us much on the relationship between the good exported and the decision to export to OECD markets first. Antras (2003) showed that vertical integration between final-good producers and their suppliers is more efficient than outsourcing as a device to overcome moral hazard in a context of incomplete contracts in capital-intensive industries. The reasoning, roughly, goes as follows. Consider a North-South relationship between the Northern buyer and the Southern supplier of an intermediate good. Assume that there is moral hazard in the provision of the intermediate good (quality is not contractible) and that cost-sharing is feasible for capital investments but not for labor investments, because buyers cannot effectively meddle with labor management in a foreign, developing country. With incomplete contracts, moral hazard will

lead to under-investment in both labor and capital. Compared to outsourcing, vertical integration will alleviate the moral-hazard problem (although not entirely), but will also weaken the bargaining position of the Southern subsidiary, making it unwilling to hire as much as would be jointly optimal. In Antras' model, headquarters in the North cannot "command" the optimal level of hiring and so must live with this sub-optimal outcome. If the industry is capital intensive, Antras shows that the ensuing distortion is less costly than if it is labor intensive. Econometric evidence motivating the theory shows that intra-firm trade is more prevalent in capital-intensive industries. The relevance for our investigation of trade patterns is this: when Northern firms set up subsidiaries in Southern countries, those subsidiaries are likely to engage in direct trade relationships with the Northern parent company. This will show in our data as a higher incidence of direct exports to OECD. If this reasoning is correct, we should observe a correlation, at the product level, between capital-intensiveness and the probability of a direct export.

Antras uses a direct measure of capital intensiveness (the stock of capital per employee), but such measures are typically available only at the level of broad sectors in UNIDO and other databases. We take a different approach here and use indices of revealed factor intensity calculated at the SITC4 level in Cadot, Shihotori and Tumurchudur (2008). Briefly, the method used to calculate those indices goes as follows. Step one consisted in constructing a systematic database of country endowments (capital per head, human capital per head, and land per head) updating Easterly and Levine (2000) for capital and Barro and Lee (1993) for human capital. Step two consisted in calculating revealed comparative advantage indices for each product and country at the SITC4 level. Step three consisted in calculating revealed factor intensity indices for each SITC4 good as follows: Index goods by *k* and countries by *i*, let $\omega_{ik} = (X_{ik} / X_i)/(X_k / X)$ be country *i*'s revealed comparative advantage (RCA) index for good *k*, and let κ_i be country i's capital endowment (stock of capital per head). Good *k*'s revealed

capital intensity is calculated as a weighted average of the capital endowment of exporters of good *k*, each of them weighted by its RCA in good *k*:

$$\kappa_k = \sum_i \omega_{ik} \kappa^i \ . \tag{19}$$

Good *k*'s revealed intensity in human capital is similarly calculated as

$$h_k = \sum_i \omega_{ik} h^i \tag{20}$$

The relationship between the probability of a direct export and the exported product's factor intensity is shown in Table 12. The dependent variable is, as in Table 11, a dummy variable marking first-export instances where the product (defined at the SITC4 level) is shipped to an OECD country, and we run a probit of that variable on characteristics of the good (revealed factor intensity indices) and the exporting country (level of income).

Table 12
Incidence of « OECD-first » exports and factor intensities: Probit marginal effects

	(1)	(2)	(3)
HK intensity	0.036	0.040	0.037
	(10.10)***	(11.26)***	(12.50)***
Capital intensity	0.000	0.000	0.000
	(9.51)***	(9.17)***	(8.28)***
Land intensity	-0.246	-0.244	-0.171
	(14.97)***	(14.86)***	(12.63)***
Upper middle income		-0.136	
		(5.77)***	
Lower middle income		0.067	
		(3.49)***	
Low income		0.056	
		(2.74)***	
Exporter controls	no	no	yes
Observations	19'758	19'758	19'758
Pseudo-R2	0.1046	0.1315	0.3313

Notes: Dependent variable: probability of a direct export to OECD markets; probit marginal effects.

The results are consistent with the vertical-integration conjecture. OECD-first instances are positively correlated with the capital intensity of the product exported, and the correlation is significant at 1%. The positive correlation with the human-capital intensity tells a similar story: in a vertically integrated relationship where the Northern parent company can share training costs

(which is plausible) but not labor management, moral hazard will distort, as before, the Southern subsidiary's hiring decision but not so much the Northern parent's training/skilled hiring decision. This will be less of a problem for skill-intensive goods. Thus, Southern skill-intensive exports are more likely to be intra-firm trade in intermediates and therefore to go directly to OECD markets for further processing.[13] [14]

5.3 Export-expansion paths and survival

As a last exercise, we explore whether getting to OECD markets in steps (first gaining experience on extra-OECD markets) improves subsequent survival on OECD markets. Figure 11 plots the average number of years of active export on OECD markets (distinct from the average length of export spells, which is much shorter) against the waiting time (between the first export to any destination and the first export to the OECD), by exporter × product. Primary evidences show two important results. First, short periods of "waiting time" correlate with longer survival. Second, further "waiting" reduces drastically the likelihood of survival in OECD markets. As can be seen in Figure 10, gaining experience by exporting to non-OECD countries for one year prior to accessing OECD markets increases the survival of exporting to OECD countries. Longer "waiting times", up to 8 years for middle income countries, might still be beneficial. Waiting "too" much in non-OECD markets reduces however the chance of success in OECD markets (this is true for waiting time above one year for the less developed countries).

[13] Intermediates may also be shipped to other, non-OECD countries for further processing; thus, some of the extra-OECD exports to other extra-OECD markets can also be intra-firm trade controlled by multinationals headquartered in OECD.
[14] The interpretation of the coefficients on countries income group dummies follows the same vein as for Table 10.

Figure 10
Years of active exports with OECD as a function of waiting time

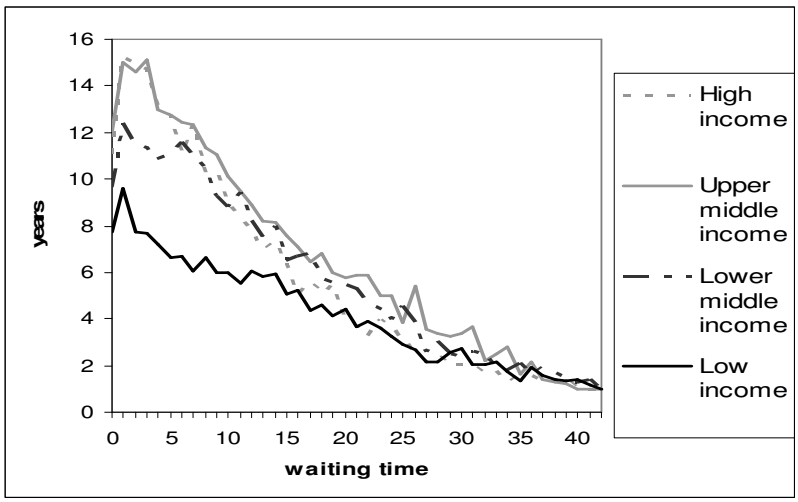
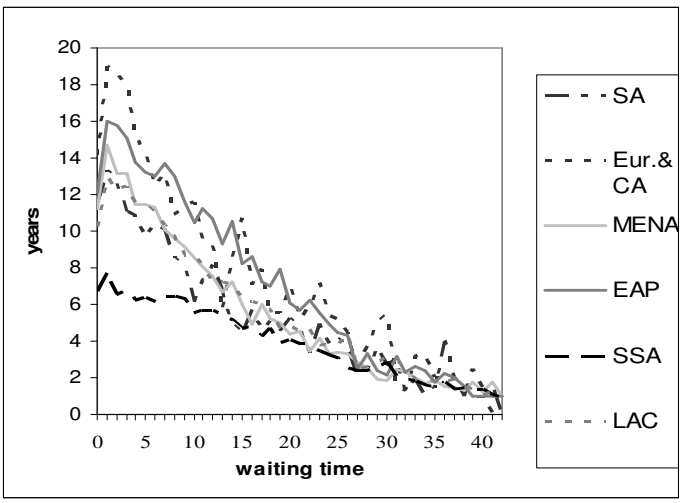


Figure 11 shows the same information arranged by region. Results are similar. For all regions (except SSA), a short experience correlates with longer survival. The positive effect of waiting before entering the OECD market picks at one year of experience and fades away after five to nine years depending on the regions.[15] Longer waiting times are correlated with lower chance of surviving on OECD markets. This latter result might reflect a selection bias with the less-efficient/ worth-quality producers taking their chance on OECD markets only after long periods.

Thus, prima facie evidence is suggestive of two important findings which would merit further analysis: a benefit of acquiring export experience on extra-OECD markets for short periods (one year seems to be the optimum), and a risk of waiting too long thereby reducing success of survival on OECD markets.

[15] SSA increases its survival on OECD market only if it gathered experience for one year.

Figure 11
Years of active exports with OECD as a function of waiting time



Notes: SA: South Asia; Eur. & CA: Europe and Central Asia; MENA: Middle East and North Africa; EAP: East Asia and Pacific; SSA: Sub-Saharan Africa; LAC: Latin America and Carribeans.

6. Concluding remarks

Looking at the evolution of OECD imports, at a high degree of disaggregation (over a thousand product lines) on a forty-year period where data is available, we found striking evidence of geographical diversification at the product level. That is, OECD countries have been sourcing each good from increasingly large pools of suppliers. We also found evidence of re-concentration of imports in the last five years or so, but this trend reversal is attributable to the growing share of China in OECD imports. Put together with Besedes and Prusa's (2006a, 2006b) findings of high churning rates among exporters, our findings suggest that OECD markets are increasingly contestable for developing-country exporters.

As for the drivers of diversification vs. re-concentration, we find that geographical concentration is higher but also more volatile for differentiated goods where quality presumably matters more and is more heterogeneous across suppliers. Moreover, when re-concentration takes place, it tends to be on

higher-priced national varieties. Put together, these observations lend support to a model of quality search by OECD buyers, discussed in section 2 of this paper, which generates alternating periods of concentration and diversification. This means that the contestability of OECD markets varies across time and products, with closed-door periods (characterized by strong incumbency advantages) alternating with open-door periods (characterized by contestability). In terms or policy implications, our results highlight the importance of raising exporter quality-management capacities in developing countries, as periods of open doors appear to be essentially periods of quality search.

Finally, we find that about 60% of extra-OECD suppliers ship their goods for the first time directly to OECD markets. We find evidence supportive of the view that OECD-first exports are contractual relationships driven by Northern buyers. While a short period (one to eight years) of prior export experience gathered on non-OECD markets enhances subsequent survival on OECD markets, further waiting is associated with a lower likelihood of survival on these markets. Products designed for non-OECD markets are, probably, less suited to OECD buyers than those that were designed according to specifications by OECD buyers, reinforcing our argument about their role in developing-country export expansion.

References

Antras, Pol (2003), "Firms, Contracts, and Trade Structure"; Quarterly Journal of Economics 118, 1375-1418.

Barro, Robert J. and J.W. Lee (1993), International Comparisons of Educational Attainment, Journal of Monetary Economics, 32, pages 5-23.

Besedes, Tibor, and T. Prusa (2006a), "Surviving the U.S. Import Market: The Role of Product Differentiation", Journal of International Economics 70, 339-358.

– and – (2006b), "Ins, Outs, and the Duration of Trade", Canadian Journal of Economics 39, 266-95.

Bikhchandani, Sushil; D. Hirshleifer, and I. Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades"; *Journal of Political Economy* 100, 992-1026.

Brenton, Paul, and R. Newfarmer (2007), "Watching More than the Discovery Channel: Export Cycles and Diversification in Development"; mimeo, The World Bank.

---, ---, and P. Walkenhorst (2007), "Export Diversification: A Policy Portfolio Approach"; paper prepared for the Growth Commission Conference on Development.

Broda, Christian, and D. Weinstein (2006), "Globalization and the Gains from Variety"; Quarterly Journal of Economics 121, 541-585.

Cadot, O, M. Shihotori and B. Tumurchudur (2008), "Revealed Factor Intensity Indices" ; mimeo, UNCTAD.

Cadot, O., Carrère, C and V. Strauss-Kahn (2007), "Export Diversification: What's Behind the Hump?"; mimeo, University of Lausanne.

Dutt, Pushan; I. Mihov and T. van Zandt (2008), "Trade Diversification and Economic Development"; mimeo, INSEAD.

Easterly, William, and A. Kray (2000), "Small States, Small Problems? Income, Growth, and Volatility in Small States"; World Development 28, 2013-2027.

--- and R. Levine (2000), "It's not Factor Accumulation: Stylized Facts and Growth Models",

Ethier, Wilfred (1982), "National and International Returns to Scale in the Modern Theory of International Trade"; American Economic Review 72, 389-405.

Evenett, Simon, and A. Venables (2002), "Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows"; mimeo, 2002.

de Ferranti, David; G. Perry, D. Lederman and W. Maloney (2002); From Natural Resources to the Knowledge Economy; The World Bank.

Hanson, Gordon (1996), "Localization Economies, Vertical Organization and Trade", American Economic Review 86, 1266-1278.

Hausmann, Ricardo; J. Hwang and D. Rodrik (2005), "What You Export Matters"; NBER working paper 11905.

-- and D. Rodrik (2003), "Economic Development as Self-Discovery"; Journal of Development Economics 72, 603-633.

Helpman, Elhanan; M. Melitz and Y. Rubinstein (2008), "Trading Partners and Trading Volumes"; Quarterly Journal of Economics, forthcoming.

Hummels, David, and P. J. Klenow (2005), "The Variety and Quality of a Nation's Exports"; American Economic Review 95, 2005, 704-723.

Imbs, Jean, and R. Wacziarg (2003), "Stages of Diversification"; American Economic Review 93, 63-86.

Jansen, Marion (2004), "Income Volatility in Small and Developing Countries: Export Concentration Matters"; Discussion Paper #3, Research Dpt, World Trade Organization.

Kee, Hiau Looi; A. Nicita and M. Olarreaga (2006) "Estimating trade restrictiveness indices", Policy Research Working Paper Series 3840, The World Bank.

Klinger, Bailey, and D. Lederman (2004), "Discovery and Development: An Empirical Exploration of 'New' Products"; mimeo, The World Bank.

-- and – (2005), "Diversification, Innovation, and Imitation off the Global Technology Frontier"; mimeo, The World Bank.

Laeven, Luc, and R. Perotti (2008), "Confidence Building in Emerging Stock Markets"; forthcoming, World Bank Economic Review.

Lederman, Daniel, and W. F. Maloney (2007); Natural resources: Neither curse nor destiny; Stanford: Stanford University Press.

Lutz, Mathias, and H. W. Singer (1994), "The Link Between Trade Openness and the Terms of Trade: An Empirical Investigation"; World Development 22: 1697-1709.

Shepherd, Ben (2008), "Geographical Diversification of Developing Country Exports"; MPRA paper 11267.

Appendix

The expression for the expected period-3 gain, as of the beginning of period 2, given that the buyer kept only one supplier, *x*, is

$$E\left(V_3 \middle| I_2\right) = q_2^x E\left(V_3 \middle| \zeta_2^x = 1\right) + \left(1 - q_2^x\right) E\left(V_3 \middle| \zeta_2^x = 0\right) \tag{21}$$

where the probability of no defect in period 3 given information in period 2, q_2^x, is

$$q_2^x \equiv \Pr\left(\zeta_3^x = 1 \middle| I_2\right) = p_2^x \lambda^G + \left(1 - p_2^x\right) \lambda^B, \tag{22}$$

and the expected gain in period 3 is

$$E\left(V_3 \middle| \zeta_2^x = 1\right) = p_3^x\left(\zeta_2^x = 1\right)\left(2\bar{\pi}^G\right) + \left[1 - p_3^x\left(\zeta_2^x = 1\right)\right]\left(2\bar{\pi}^B\right) \tag{23}$$

given no defect in period 2 and

$$E\left(V_3 \middle| \zeta_2^x = 0\right) = p_3^x\left(\zeta_2^x = 0\right)\left(2\bar{\pi}^G\right) + \left[1 - p_3^x\left(\zeta_2^x = 0\right)\right]\left(2\bar{\pi}^B\right) \tag{24}$$

given a defect in period 2. Finally, the probability of supplier *x* being of the good type is, by Bayes' rule,

$$p_3^x\left(\zeta_2^x = 1\right) = \Pr\left(G \middle| \zeta_2^x = 1\right) = \frac{\lambda^G p_2^x}{\lambda^G p_2^x + \lambda^B\left(1 - p_2^x\right)} \tag{25}$$

given no defect in period 2 and

$$p_3^x\left(\zeta_2^x = 0\right) = \Pr\left(G \middle| \zeta_2^x = 0\right) = \frac{\left(1 - \lambda^G\right) p_2^x}{\left(1 - \lambda^G\right) p_2^x + \left(1 - \lambda^B\right)\left(1 - p_2^x\right)} \tag{26}$$

given a defect. Substituting these expressions into (21) and simplifying gives expression (8) in the text.