The productivity advantages of large markets: Distinguishing agglomeration from firm selection

Pierre-Philippe Combes*†
University of Aix-Marseille

Gilles Duranton*‡
University of Toronto

Laurent Gobillon*§
Institut National d'Etudes Démographiques

Diego Puga*¶
Universitat Pompeu Fabra

Sébastien Roux*‖
Institut National de la Statistique et des Etudes Economiques

Preliminary and incomplete draft, printed 27 June 2007

ABSTRACT: Firms and workers are more productive on average in larger markets. Two explanations have been offered: agglomeration economies (larger markets promote interactions that increase productivity) and firm selection (larger markets toughen competition allowing only the most productive to survive). To distinguish between them, we develop a model that nests a generalised version of a seminal firm selection model and a standard model of agglomeration. Stronger selection in larger markets left-truncates the productivity distribution whereas stronger agglomeration right-shifts the distribution. We use this prediction to assess the relative importance of agglomeration and firm selection using French establishment-level data.

Key words: agglomeration, firm selection, productivity, market size

JEL classification: C52, R12, D24

*We are grateful to Marc Melitz, Gianmarco Ottaviano, and John Sutton for comments and discussions.

†GREQAM, 2 Rue de la Charité, 13236 Marseille cedex 02, France (e-mail: ppcombes@univmed.fr; website: http://www.vcharite.univ-mrs.fr/PP/combes/). Also affiliated with the Centre for Economic Policy Research.

‡Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: gilles.duranton@utoronto.ca; website: http://individual.utoronto.ca/gilles/default.html). Also affiliated with the Centre for Economic Policy Research.

§Institut National d'Etudes Démographiques, 133 Boulevard Davout, 75980 Paris cedex 20, France (e-mail: laurent.gobillon@ined.fr; website: http://laurent.gobillon.free.fr/).

¶Department of Economics and CREI, Universitat Pompeu Fabra, Ramon Trias Fargas 25–27, 08005 Barcelona, Spain (e-mail: diego.puga@imdea.org; website: http://diegopuga.org). Also affiliated with the Instituto Madrileño de Estudios Avanzados (IMDEA) Ciencias Sociales and the Centre for Economic Policy Research (CEPR).

‖Centre de Recherche en Économie et Statistique (CREST), 15 Boulevard Gabriel Péri, 92245 Malakoff cedex, France (e-mail: sebastien.roux@insee.fr).

1. Introduction

Firms and workers are, on average, more productive in larger markets. For cities, this fact — first suggested by Adam Smith (1776) and Alfred Marshall (1890) — is now firmly established empirically (see Rosenthal and Strange, 2004, for a review and Henderson, 2003, and Combes, Duranton, and Gobillon, 2008, for recent contributions). Estimates of the magnitude of this effect range between a two and eight percent productivity increase from a doubling of city size, depending on the sector and details of the estimation procedure (Rosenthal and Strange, 2004, Combes, Duranton, Gobillon, and Roux, 2007). The evidence is also mounting when considering spatial markets at higher levels of aggregation, such as regions and even countries (Head and Mayer, 2004, Redding and Venables, 2004, Amiti and Cameron, 2007).

For a long time, the higher average productivity of firms and workers in larger markets has been attributed to 'agglomeration economies'. These agglomeration economies are thought to arise from a variety of mechanisms such as the possibility of similar firms sharing suppliers, the existence of thick labour markets ironing out firm-level shocks or facilitating matching, or the possibility to learn from the experiences and innovations of others (see Duranton and Puga, 2004, for a review).

More recently, an alternative explanation has been offered based on 'firm selection'. The argument builds on work by Melitz (2003), who introduces product differentiation and international or inter-regional trade in the framework of industry dynamics of Hopenhayn (1992). Melitz and Ottaviano (2005) incorporate endogenous price-cost mark-ups in this framework and show that larger markets attract more firms, which makes competition tougher.[1] In turn, this leads less productive firms to exit. This suggests that the higher average productivity of firms and workers in larger markets could instead result from a stronger Darwinian selection of firms.

Our main objective in this paper is to distinguish between agglomeration and firm selection in explaining why average productivity is higher in larger markets. To do so, our first step is to generalise the framework of Melitz and Ottaviano (2005) to many markets and free it from distributional assumptions. We then combine this model with a fairly general model of agglomeration in the spirit of Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002). This nested model allows us to parameterise the relative importance of agglomeration and selection. The main prediction of our model is that stronger selection effects in larger markets should lead to an increased left-truncation of the distribution of firm log productivities whereas stronger agglomeration effects should lead to a rightwards-shift of the distribution of firm log productivities. Put differently, we exploit the idea that although agglomeration and selection share the same prediction about average log productivity, they differ regarding higher moments of the log productivity distribution across markets.

We then use this prediction to assess the relative importance of agglomeration and firm selection for different sectors using data for all French firms. Our structural estimation is in two steps. We first estimate total factor productivity at the plant level. Next, we develop a new quantile approach to compare the distribution of plant log productivities for each sector across metropolitan areas of

[1] Bernard, Eaton, Jensen, and Kortum (2003) also develop a model with heterogenous firm productivity levels and endogenous mark-ups but, unlike in Melitz and Ottaviano (2005), this mark-ups are not affected by market size.

different size. As stipulated by the model, we estimate the extent to which the log productivity distribution in large markets is left truncated and right shifted compared to the log productivity distribution in small markets.

Our main finding is that productivity differences between French metropolitan areas are explained mostly by agglomeration.

[[Further details on results to be added here.]]

Our paper is related to the large agglomeration literature surveyed in Duranton and Puga (2004), Rosenthal and Strange (2004) and Head and Mayer (2004), which it extends by considering an entirely different reason for the higher average productivity in larger markets. It is also related to the pioneering work of Syverson (2004) who examines the effect of market size on firm selection in the ready-made concrete sector and the emerging literature that follows (Del Gatto, Mion, and Ottaviano, 2006, Del Gatto, Ottaviano, and Pagnini, 2008). A first difference with Syverson's work is that we jointly estimate agglomeration and selection effects rather than considering only selection. Our results below show that, unless we consider both simultaneously, we risk confounding the effects of selection with those of agglomeration. The second difference is that our procedure uses information from the entire distribution of productivities instead of its second moment only. This is important because predictions regarding the second moment are not general and depend crucially on distributional assumptions.[2] A final difference is that we consider firms in different sectors.

The rest of this paper is organised as follows. The next section proposes a generalisation of Melitz and Ottaviano (2005) and combines it with an agglomeration model. Section 3 describes our econometric approach. Section 4 discusses the data and the details of our empirical implementation. The main results are then presented in section 5. In section **??** we explore the relationship between the importance of selection effects and the tradability of goods. Finally, section 6 concludes.

2. A nested model of selection and agglomeration

To motivate our empirical approach, we nests a generalised version of the firm selection model of Melitz and Ottaviano (2005) and a model of agglomeration economies in the spirit of Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002).

An individual consumer's utility is given by

$$U = q_0 + \alpha \int_{i \in \Omega} q_i \mathrm{d}i - \frac{1}{2}\gamma \int_{i \in \Omega} (q_i)^2 \mathrm{d}i - \frac{1}{2}\eta \left(\int_{i \in \Omega} q_i \mathrm{d}i \right)^2 , \tag{1}$$

where q_0 denotes the individual's consumption of a homogenous numéraire good, and q_i her consumption of variety i of a set Ω of differentiated goods. The three positive demand parameters

[2]Like in our model, tougher competition in Syverson (2004) leads to a left-truncation of the distribution of productivities in larger markets. In turn, in his model this implies a decrease in the variance of productivities, a result that depends crucially on distributional assumptions. It holds in Syverson's model and will hold more generally for productivity distributions with log-concave density. However, the result would be reversed if one considered instead a productivity distribution with log-convex density, such as the popular Pareto distribution (on the relationship between the variance of a left-truncated distribution and log-concavity and log-convexity, see Heckman and Honore, 1990).

α, γ, and η are such that a higher α and a lower η increase demand for differentiated goods relative to the numéraire, while a higher γ reflects more product differentiation between varieties.[3]

Utility maximisation subject to the budget constraint yields the following inverse demand for differentiated variety i by an individual consumer:

$$p_i = \alpha - \gamma q_i - \eta \int_{i \in \Omega} q_i \mathrm{d}i \ , \tag{2}$$

where p_i denotes the price of variety i. It follows from (2) that varieties with too high a price are not consumed. This is because, by (1), the marginal utility for any particular differentiated variety is bounded. Let $\bar{\Omega}$ denote the set of varieties with positive consumption levels in equilibrium, M the measure of $\bar{\Omega}$, and $P \equiv \frac{1}{M} \int_{i \in \bar{\Omega}} p_i \mathrm{d}i$ the average price of varieties with positive consumption. Integrating equation (2) over all varieties in $\bar{\Omega}$, solving for $\int_{i \in \Omega} q_i \mathrm{d}i$, and substituting this back into equation (2), we can solve for an individual consumer's demand for variety i as:

$$q_i = \begin{cases} \frac{1}{\gamma + \eta M}(\alpha + \frac{\eta}{\gamma} MP) - \frac{1}{\gamma} p_i & \text{if } p_i \leqslant \frac{1}{\gamma + \eta M}(\gamma\alpha + \eta MP) \ , \\ 0 & \text{if } p_i > \frac{1}{\gamma + \eta M}(\gamma\alpha + \eta MP) \ . \end{cases} \tag{3}$$

The price threshold in equation (3) follows immediately from the restriction $q_i \geqslant 0$.

The numéraire good is produced under constant returns to scale using one unit of labour per unit of output and (when we consider more than one location) freely traded. This implies that the cost to firms of hiring one unit of labour is always unity.[4] Differentiated varieties are produced under monopolistic competition. By incurring a sunk entry cost s, a firm is able to develop a new variety that can be produced using h units of labour per unit of output. The unit labour requirement h differs across firms and for each of them it is randomly drawn from a distribution with known probability density function $g(h)$ and cumulative $G(h)$. Melitz and Ottaviano (2005) derive most of their results under the assumption that $g(h)$ is a Pareto distribution. For the results in our paper, no such assumption is needed, so we do not adopt any particular distribution for $g(h)$. Given the unit cost of each unit of labour, h is also the marginal cost. Firms with a marginal cost higher than the price at which consumer demand becomes zero are unable to cover their marginal cost and exit. The set of goods that end up being produced in equilibrium is therefore

$$\bar{\Omega} = \{i \in \Omega \mid h \leqslant \bar{h}\} \ , \ \text{ where } \bar{h} \equiv \frac{1}{\gamma + \eta M}(\gamma\alpha + \eta MP) \ . \tag{4}$$

Since all varieties enter symmetrically into utility, we can index firms by their unit labour requirement h instead of by the specific variety i they produce. Using (4) to re-write the individual consumer demand of (3) in terms of $\bar{h}$ and multiplying this by the mass of consumers C yields the following expression for the total sales of an individual firm:

$$Q(h) = Cq(h) = \begin{cases} \frac{C}{\gamma}[\bar{h} - p(h)] & \text{if } p(h) \leqslant \bar{h} \ , \\ 0 & \text{if } p(h) > \bar{h} \ . \end{cases} \tag{5}$$

[3] The specification in (1) is often referred to as the quadratic utility model of horizontal product differentiation. It has been used in industrial organization by, for instance, Dixit (1979) and Vives (1990) and has become popular in location modelling following Ottaviano, Tabuchi, and Thisse (2002).

[4] The unit cost for labour holds provided there is some production of the numéraire good everywhere. Given the quasi-linear preferences, this requires that income is high enough, which is easy to ensure.

Given that the entry cost is sunk when firms draw their value of h, active firms set prices to maximise operational profits given by

$$\pi(h) = [p(h) - h]Q(h) \; . \tag{6}$$

Maximising $\pi(h)$ in (6) subject to (5) yields the optimal pricing rule

$$p(h) = \frac{1}{2}(h + \bar{h}) \; . \tag{7}$$

Substituting (5) and (7) into (6) we obtain equilibrium operational profits:

$$\pi(h) = \frac{C}{4\gamma}(\bar{h} - h)^2 \; . \tag{8}$$

Entry into the monopolistically competitive industry takes place until *ex-ante* expected profits are driven to zero. The operational profits expected prior to entry must therefore be exactly offset by the sunk entry cost:

$$\frac{C}{4\gamma}\int_0^{\bar{h}} (\bar{h} - h)^2 g(h)\mathrm{d}h = s \; . \tag{9}$$

The free entry condition (9) implicitly defines the marginal cost cutoff $\bar{h}$ as a function of the distribution $g(h)$, the sunk entry cost s, and the degree of product differentiation parameter γ.

We now turn to the agglomeration components of the model. Workers are endowed with a single unit of working time each that they supply inelastically. Each worker is made more productive by interactions with other workers, so that the effective units of labour supplied by an individual worker in their unit working time is $a(I)$, where I is the population that a worker interacts with, $a' > 0$ and $a'' < 0$. We can think of such interactions as exchanges of ideas between workers, where being exposed to a greater diversity of ideas makes each worker more productive. This motivation for agglomeration economies based on interactions between workers can be found in, amongst others, Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002) — we introduce a discrete version of their spatial decay for interactions below. Duranton and Puga (2004) review micro-foundations of numerous alternative agglomeration mechanisms, which also result in a reduced form like $a(I)$. We assume that such interactions benefit workers across occupations, i.e., regardless of whether they produce any particular variety of the differentiated goods or the numéraire good. This, given the unit payment per effective unit of labour supplied, implies that the total income earnings of each worker in any occupation is $a(I)$. A firm with unit labour requirement h hires $l(h) = Q(h)h/a(I)$ workers at a total cost of $a(I)l(h) = Q(h)h$. The natural logarithm of the firm's productivity is then given by

$$\phi = \ln\left(\frac{Q}{l}\right) = \ln\left[a(I)\right] - \ln(h) \; . \tag{10}$$

The probability density function of firms' log productivities is then

$$f(\phi) = \begin{cases} 0 & \text{for } \phi < A - \ln(\bar{h}) \; , \\ \frac{e^{A-\phi} g\left(e^{A-\phi}\right)}{G(\bar{h})} & \text{for } \phi \geqslant A - \ln(\bar{h}) \; , \end{cases} \tag{11}$$

where

$$A \equiv \ln [a(I)] \ . \tag{12}$$

The numerator of $f(\phi)$, $e^{A-\phi} g\left(e^{A-\phi}\right)$ follows from using equation (10) and the change of variables theorem, while the denominator $G(\bar{h})$ takes care of the fact that firms with a unit labour requirement above $\bar{h}$ exit. The model can now be solved sequentially by first using the free entry condition of equation (9) to solve for the equilibrium cut-off unit labour requirement $\bar{h}$. We can then substitute $\bar{h}$ into (11) to obtain the equilibrium distribution of firm productivities.

To understand how selection and agglomeration forces contribute to determining the distribution of firms' log productivities it is instructive to think of what is the relevant mass C of consumers that each firm sells to and what is the relevant mass I of people that each worker interacts with. For clarity, consider two polar possibilities for each (to be generalised to intermediate cases later) in an economy with two markets (to be generalised to multiple markets below). In terms of demand, at one extreme we can think of firms selling only to consumers in their location and thus competing with other local firms only (*local product-market competition*). At the other extreme, firms can sell with equal ease to consumers anywhere and thus compete with firms everywhere (*global product-market competition*). In terms of interactions, at one extreme we can think of workers interacting exclusively with other workers living in the same location (*local interactions*). At the other extreme, workers can interact with equal ease with workers living anywhere (*global interactions*). The combination of these possibilities gives us four cases. We now compare in each of the four cases the distribution of firms' log productivities across two locations of different population size (a large location 1 where $C = C_1$ and $I = I_1$ and a small location 2 where $C = C_2$ and $I = I_2$).

Case 1 (*global product-market competition and global interactions*)**.** When every firm competes with the same intensity with firms from everywhere and every worker enjoys interactions with the same intensity with workers from everywhere (i.e., $C_1 = C_2$ and $I_1 = I_2$), the distribution of firms' log productivities in a location with a large population is exactly the same as in a location with a small population. Panel (a) in figure 1 plots the probability density function of the distribution of firms' log productivities, $f(\phi)$, in this benchmark case where it is independent of location size. Note that the fact that the distribution of firms' log productivities does not depend on location size in this case does not imply that there are no selection or agglomeration effects. It simply implies that selection and agglomeration effects are equally strong everywhere. If there were no selection or agglomeration effects, the distribution of firm productivities would be given by $f(\phi) = e^{-\phi} g\left(e^{-\phi}\right)$, plotted as a dotted line in panel (a). Relative to this underlying distribution, the actual distribution of firm productivities is both left-truncated and right-shifted for reasons that will be clear once we consider case 4.[5]

Case 2 (*local product-market competition and global interactions*)**.** Panel (b) in figure 1 plots the distribution of firms' log productivities in a location with a small population (continuous line) and in a location with a large population (dashed line) in the case where firms only sell in their local

[5]These graphs are drawn using a Fisher distribution, but recall that our analytical results are distribution-independent. We use a Fisher distribution for the graphs because it matches well the empirically-observed distributions presented later in the paper.

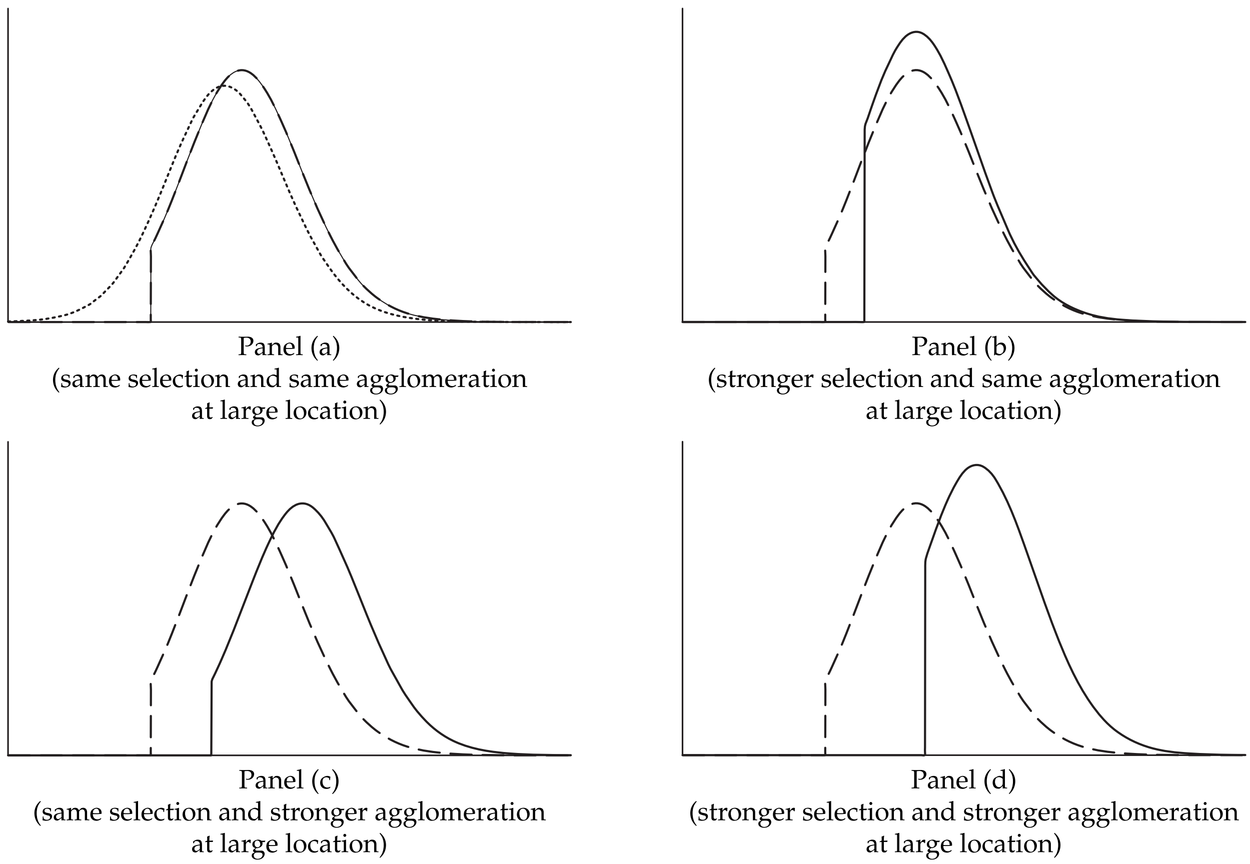

Figure 1. Log-productivity distributions in a large (continuous line) and small location (dashed)

market and workers enjoy interactions with the same intensity with workers from everywhere (i.e., $C_1 > C_2$ and $I_1 = I_2$).[6] Compared with the distribution in the small location, the large-location distribution is left-truncated as a consequence of firm selection (this left-truncation implies that the peak of the large location distribution is higher than that of the small location distribution and that the two peaks occur at the same level of productivity). To understand this greater truncation in the large location note that if the number of active firms in the large location was the same as in the small location, every large-location firm would sell proportionately more. Formally, using equations (5) and (7), total sales for an individual firm can be expressed as $\frac{C}{2\gamma}(\bar{h} - h)$, which, for a given number of firms M and hence for a given $\bar{h}$, increase proportionately with the population of consumers C. However, the larger individual firm sales associated with a larger C make further entry profitable and, by equation (9), they must be offset by a lower $\bar{h}$ to restore zero *ex-ante* expected profits. To understand how firms in different ranges of the productivity distribution are affected by location size, note that from (5) and (7), the price elasticity of demand faced at equilibrium by a firm with unit labour requirement h can be written as follows:

$$\epsilon(h) \equiv -\frac{p(h)}{Q(h)} \frac{\mathrm{d}Q(h)}{\mathrm{d}p(h)} = \frac{\bar{h} + h}{\bar{h} - h} \,. \qquad (13)$$

[6]To facilitate visual comparisons, we re-scale the combined size of the large and small locations from panel to panel to keep the sets C and I for the small location constant across cases, thus making the distribution of firms' log productivities in the small location identical in all four panels. This is done for the purpose of plotting the graph only, and does not change the qualitative comparison between the small-location and large-location distributions.

Demand becomes more price-elastic as h increases or as $\bar{h}$ decreases. Thus, each firm in the large location (where $\bar{h}$ is lower) faces a more elastic demand, and hence charges a lower markup, $[p(h) - h]/h$, than a firm with the same h in the small location. The combination of more consumers, entry, and the ensuing lower markups in the large location affects firms' sales differently depending on their h.[7] Firms with high productivity, and hence high markups, enjoy smaller profit margins but larger sales than their small-location counterparts. Low productivity firms, however, have both smaller profit margins and smaller sales in the large location than in the small location. Some low-productivity firms that would have been able to survive in a small market cannot lower their prices any further and must exit in the large market. It is this exit at the low-productivity end that leads to the large-location log productivity distribution being a left-truncated version of the small-location distribution (this can be seen from equation 11, given the lower $\bar{h}$ in the large location).

Case 3 (*global product-market competition and local interactions*)**.** Panel (c) in figure 1 plots the distribution of firms' log productivities in a location with a small population (continuous line) and in a location with a large population (dashed line) in the case where every firm competes with the same intensity with firms from everywhere and workers only interact with workers in their location (i.e., $C_1 = C_2$ and $I_1 > I_2$). Compared with the distribution in the small location, the large-location distribution is right-shifted. Since interactions are local, workers in the larger location benefit from being exposed to a wider range of ideas than workers in the small location and this makes them more productive. As a result, all large-location firms achieve higher log productivity than their small-location counterparts (i.e., log productivity ϕ is higher in the large location for every h). Since product-market competition is global, all firms can sell to consumers everywhere and this eliminates the firm selection mechanism of the previous case. Hence, the log productivity cut-off $\ln[a(I)] - \ln(\bar{h})$ simply moves rightwards to the same extent as every point of the log productivity distribution. Consequently the large-location log productivity distribution is simply a right-shifted version of the small-location distribution (this can be seen from equation 11, given the higher I in the large location and the fact that $a' > 0$). Thus, agglomeration acts like the tide that lifts all boats.

Case 4 (*local product-market competition and local interactions*)**.** Finally, panel (d) in figure 1 plots the distribution of firms' log productivities in a location with a small population (continuous line) and in a location with a large population (dashed line) in the case where firms only sell in their local market and workers only interact with workers in their location (i.e., $C_1 > C_2$ and $I_1 > I_2$). Compared with the distribution in the small location, the large-location distribution is both left-truncated and right-shifted. With local product-market competition, large-location markups are lower and this left-truncates the distribution of firms' log productivities to exactly the same extent as under case 2. With local interactions, large-location workers are more productive and this right-

[7] This is best seen by considering the effect on a firm's sales $Q(h)$ of a small increase in C. From (5) and (7), $\frac{\mathrm{d}Q(h)}{\mathrm{d}C} = \frac{1}{2\gamma}\left[\bar{h} - h + C\frac{\mathrm{d}\bar{h}}{\mathrm{d}C}\right]$. From the free entry condition of (9), $\frac{\mathrm{d}\bar{h}}{\mathrm{d}C} = -2\gamma s/[C^2 \int_0^{\bar{h}}(\bar{h} - h)g(h)\mathrm{d}h]$. It follows that $\frac{\mathrm{d}Q(h)}{\mathrm{d}C} > 0$ if and only if $(\bar{h} - h)\frac{C}{2\gamma}\int_0^{\bar{h}}(\bar{h} - h)g(h)\mathrm{d}h > s$. The expression on the left-hand side of this inequality is twice the firm's markup times *ex-ante* expected sales. Since there are zero expected *ex-ante* profits, for firms near the top of the productivity distribution (those with the lowest values of h and thus highest markups) this inequality necessarily holds, so their sales increase as market size increases. For firms near the bottom of the productivity distribution (those with a value of h close to the cutoff $\bar{h}$) the inequality necessarily fails to hold, so their sales fall as market size increases.

shifts the distribution of firms' log productivities (truncated by firm selection) to exactly the same extent as under case 3.[8]

Parameterising the strength of selection and agglomeration

For expositional clarity, we have so far focused on the polar possibilities of either local or global product-market competition and either local or global interactions. We now generalise our analysis to also consider intermediate cases for both. We do so by parameterising both the spatial decay of product market competition, creating differences in firm selection across locations, and the spatial decay of interactions, creating differences in agglomeration economies across locations. As before, we compare the distribution of productivities across locations of different size. Suppose we have K locations. Let us denote the population of location i by N_i and order locations from largest to smallest in terms of population: $N_1 > N_2 > \cdots > N_{K-1} > N_K$.

In the case of product-market competition, we can introduce transport costs for differentiated goods. Suppose that markets are segmented and that selling outside the location where a firm is located involves iceberg transport costs so that τ (> 1) units need to be shipped for one unit to arrive at the destination. Since firms now potentially sell in all locations, the free entry condition of equation (9) becomes

$$\frac{N_i}{4\gamma}\int_0^{\bar{h}_i}(\bar{h}_i - h)^2 g(h)\mathrm{d}h + \sum_{k\neq i}\frac{N_k}{4\gamma}\int_0^{\bar{h}_k/\tau}(\bar{h}_k - \tau h)^2 g(h)\mathrm{d}h = s\,, \tag{14}$$

for location i. The first term on the left-hand side captures operational profits from local sales and the second-term summation the operational profits from out-of-location sales. Note that only location i firms with marginal costs $h < \bar{h}_k/\tau$ sell in location k, where $\bar{h}_k$ is the cutoff for local firms in k, since location i firms must be able to cover not just production but also trade costs. Note that the cases of purely local or purely global product-market competition discussed above can still be captured as particular cases. The case of local product-market competition corresponds to $\tau = \infty$, which turns equation (14) into equation (9) with $C = N_i$. The case of global product-market competition corresponds to $\tau = 1$, which turns equation (14) into equation (9) with $C = \sum_{k=1}^{K} N_k$. In addition, we can now also consider intermediate cases where $1 < \tau < \infty$.

Regarding interactions, we can think of these as being subject to some spatial decay. Specifically, let us redefine the relevant argument for the interactions function $a(\cdot)$ as the sum of local population and outside population, with the latter adjusted by some decay factor as in Fujita and Ogawa (1982) and Lucas and Rossi-Hansberg (2002). This implies that the effective labour

[8]The absence of interactions between selection and agglomeration mechanisms is a consequence of having kept the assumption of quasi-linear preferences of Melitz and Ottaviano (2005), which eliminates income effects in the market for differentiated goods. The introduction of income effects would create an interaction between agglomeration and firm selection that would result in further left-truncation of the large-location log productivity distribution. This is because, with income effects, the log productivity advantages of agglomeration would translate into a larger market for differentiated goods in the large location. This would reinforce the increase in local product-market competition caused by the larger population, and strengthen firm selection. Thus, with income effects, agglomeration would appear as a right shift in the log productivity distribution, while selection as well as interactions between selection and agglomeration would appear as a left truncation. More complicated interactions between selection and agglomeration mechanisms would appear if the benefits from agglomeration for each worker varied depending on which firm they were working for.

supplied by an individual worker in location i is $a(N_i + \delta \sum_{k \neq i} N_k)$, where the decay parameter δ measures the strength of across-location relative to within-location interactions ($0 < \delta < 1$). From equation (10), the log productivity of a firm with marginal cost h in location i is given by $\phi = \ln\left[a(N_i + \delta \sum_{k \neq i} N_k)\right] - \ln(h)$. Thus, the gain in log productivity across the board due to interactions in location i (a local measure of the strength of agglomeration) of equation (12) can be redefined as

$$A_i \equiv \ln\left[a(N_i + \delta \sum_{k \neq i} N_k)\right] . \tag{15}$$

The case of local interactions discussed above corresponds to $\delta = 0$, which implies $A_i = \ln[a(N_i)]$. The case of global interactions discussed above corresponds to $\delta = 1$, which implies that $A_i = \ln[a(\sum_{k=1}^{K} N_k)]$. In addition, we can now also consider intermediate cases where $0 < \delta < 1$.

The distribution of firms' log productivities still has its probability density function given by equation (11), which using subindex i to specify the location becomes

$$f_i(\phi) = \begin{cases} 0 & \text{for } \phi < A_i - \ln(\bar{h}_i) \ , \\ \frac{e^{A_i - \phi} g\left(e^{A_i - \phi}\right)}{G(\bar{h}_i)} & \text{for } \phi \geqslant A_i - \ln(\bar{h}_i) \ , \end{cases} \tag{16}$$

In anticipation of the econometric approach developed in the next section, it will be useful to also write the corresponding cumulative density function, $F_i(\phi)$. To do that compactly, we need to introduce some additional notations. Let

$$S_i \equiv 1 - G(\bar{h}_i) \tag{17}$$

denote the proportion of firms that fail to survive product-market competition in location i (a local measure of the strength of selection). To further simplify notation, let us define

$$\tilde{F}(\phi) \equiv 1 - G\left(e^{-\phi}\right) \tag{18}$$

as the underlying cumulative density function of log productivities we would observe in all locations in the absence of any selection ($\bar{h}_i = \infty$, $\forall i$) and in the absence of any agglomeration ($A_i = 0$, $\forall i$). Without selection ($\bar{h}_i = \infty$, $\forall i$) all entrants survive regardless of their draw of h. Without agglomeration ($A_i = 0$, $\forall i$), $\phi = -\ln(h)$. Equivalently, $h = e^{-\phi}$. Using the change of variables theorem then yields (18) above. We can then write the cumulative density function of the distribution of log productivities for active firms in location i as

$$F_i(\phi) = \max\left\{0, \frac{\tilde{F}(\phi - A_i) - S_i}{1 - S_i}\right\} . \tag{19}$$

Relative to the underlying distribution we would observe in any location in the absence of agglomeration and selection, given by (18), selection eliminates a share S_i of entrants (those with lower productivity values) while agglomeration shifts the distribution rightwards by A_i. The next section will develop an econometric approach to estimate the relative magnitude across locations of agglomeration, as measured by A_i, and selection, as measured by S_i. The following proposition contains our main theoretical result, with predictions for how these expressions vary across locations of different sizes.

Proposition 1. Suppose there are K locations ranked from largest to smallest in terms of population: $N_1 > N_2 > \cdots > N_{K-1} > N_K$, there are positive trade costs τ across locations, where $1 < \tau < \infty$, and a decay δ for interactions across locations, where $0 < \delta < 1$. Firm selection left truncates the distribution of firm log productivities, and this truncation is

greater the larger a location's population: $\bar{h}_1 < \bar{h}_2 < \ldots < \bar{h}_{K-1} < \bar{h}_K$ and thus $S_1 > S_2 > \ldots > S_{K-1} > S_K$. Agglomeration leads to the distribution of firm log productivities (once truncated) being right shifted, and this right shift is greater the larger a location's population: $A_1 > A_2 > \ldots > A_{K-1} > A_K$.

Proof Consider any two areas i and j such that $i < j$ (and thus $N_i > N_j$). Writing the free entry condition of equation (14) for locations i and j gives:

$$\frac{N_i}{4\gamma}\int_0^{\bar{h}_i}(\bar{h}_i - h)^2 g(h)\mathrm{d}h + \frac{N_j}{4\gamma}\int_0^{\bar{h}_j/\tau}(\bar{h}_j - \tau h)^2 g(h)\mathrm{d}h + \sum_{k\neq i,k\neq j}\frac{N_k}{4\gamma}\int_0^{\bar{h}_k/\tau}(\bar{h}_k - \tau h)^2 g(h)\mathrm{d}h = s\ , \tag{20}$$

$$\frac{N_j}{4\gamma}\int_0^{\bar{h}_j}(\bar{h}_j - h)^2 g(h)\mathrm{d}h + \frac{N_i}{4\gamma}\int_0^{\bar{h}_i/\tau}(\bar{h}_i - \tau h)^2 g(h)\mathrm{d}h + \sum_{k\neq i,k\neq j}\frac{N_k}{4\gamma}\int_0^{\bar{h}_k/\tau}(\bar{h}_k - \tau h)^2 g(h)\mathrm{d}h = s\ . \tag{21}$$

Subtracting equation (21) from (20) and simplifying yields:

$$N_i\,\nu(\bar{h}_i,\tau) = N_j\,\nu(\bar{h}_j,\tau)\ . \tag{22}$$

where

$$\nu(z,\tau) \equiv \int_0^{z}(z-h)^2 g(h)\mathrm{d}h - \int_0^{z/\tau}(z-\tau h)^2 g(h)\mathrm{d}h\ . \tag{23}$$

It follows from (22) and $N_i > N_j$ that

$$\nu(\bar{h}_i,\tau) < \nu(\bar{h}_j,\tau)\ . \tag{24}$$

We wish to show that $\bar{h}_i < \bar{h}_j$. It is therefore enough to prove that $\nu(z,\tau)$ is an increasing function of its first argument. This can be shown easily by differentiating (23) with respect to z:

$$\begin{aligned}\frac{\partial\nu(z,\tau)}{\partial z} &= 2\left[\int_0^{z}(z-h)g(h)\mathrm{d}h - \int_0^{z/\tau}(z-\tau h)g(h)\mathrm{d}h\right]\\ &= 2\left[(\tau-1)\int_0^{z/\tau}hg(h)\mathrm{d}h + \int_{z/\tau}^{z}(z-h)g(h)\mathrm{d}h\right] > 0\ .\end{aligned} \tag{25}$$

By equation (19), the distribution of firm log productivities is left truncated relative to the distribution $\tilde{F}$ we would observe without firm selection or agglomeration. The proportion of truncated values of $\tilde{F}$ is S_i in location i and S_j in location j. Since $\bar{h}_i < \bar{h}_j$, by equation (17), $S_i > S_j$. Incorporating $\bar{h}_i < \bar{h}_j$ into (19) implies that the distribution of firm log productivities (once left truncated) is right-shifted relative to $\tilde{F}$. The extent of this right shift is A_i in locations i and A_j in location j (with the shift applied to the respective left truncated distributions). By equation (15), $A_i > A_j$. □

3. Econometric specification

We now develop an econometric approach to estimate the parameters that quantify the importance of selection and agglomeration in the theoretical model for locations of different market sizes. Denote $\tilde{\lambda}(y) \equiv \tilde{F}^{-1}(y)$ the y^{th} quantile of $\tilde{F}$ and $\lambda_i(y) \equiv F_i^{-1}(y)$ the y^{th} quantile of F_i. From (19), we can write

$$\lambda_i\,(y) = \tilde{\lambda}\,(S_i + (1 - S_i)y) + A_i \text{ , for } y \in [0{,}1] \text{ ,} \tag{26}$$

which, with the change of variables $y \to \frac{y - S_i}{1 - S_i}$, turns into

$$\lambda_i \left(\frac{y - S_i}{1 - S_i} \right) = \tilde{\lambda}(y) + A_i \text{ , for } y \in [S_i{,}1] \text{ .} \tag{27}$$

We cannot work with equation (27) directly since we do not observe the quantiles of the underlying distribution, $\tilde{\lambda}(y)$, in the data. However, we can subtract from (27) the corresponding equation for another location j of smaller population size to get:

$$\lambda_i \left(\frac{y - S_i}{1 - S_i} \right) - \lambda_j \left(\frac{y - S_j}{1 - S_j} \right) = A_i - A_j \text{ , for } y \in \left[\max(S_i{,}S_j){,}1\right] \text{ .} \tag{28}$$

Note that, while Proposition 1 predicts that $S_i > S_j$, our estimation must allow for the possibility that $S_i \leqslant S_j$, and hence the max operator in equation (28).

Since quantiles are measured with error, equation (28) does not hold exactly. We wish to look for parameter values that minimise the mean squared error of (28) for all allowed values of y. Because of the max operator in (28), it is convenient to consider two cases separately: $S_i > S_j$ and $S_i \leqslant S_j$.

Let us start with the case where $S_i > S_j$. With the change of variables $\frac{y - S_i}{1 - S_i} \to y$, equation (28) becomes

$$\lambda_i\,(y) - \lambda_j \left(\frac{S_i - S_j}{1 - S_j} + \frac{1 - S_i}{1 - S_j} y \right) = A_i - A_j \text{ , for } y \in [0{,}1] \text{ .} \tag{29}$$

This can be written more compactly by defining

$$S \equiv \frac{1 - S_i}{1 - S_j} \text{ ,} \tag{30}$$

as the share of the underlying log productivity distribution that is not truncated in location i relative to the share that is not truncated in location j (i.e., the relative shares of surviving entrants in the theoretical model), and

$$A \equiv A_i - A_j \text{ ,} \tag{31}$$

as the difference in shift between the (already truncated) underlying log productivity distribution in location i and the one in location j (i.e., the difference in the strength of agglomeration-induced productivity gains in the theoretical model). Using (30) and (31), equation (29) can be rewritten as

$$\lambda_i\,(y) = \lambda_j\,((1 - S) + Sy) + A \text{ , for } y \in [0{,}1] \text{ .} \tag{32}$$

Notice the similarity between equations (32) and (26). This tells us that if the log productivity distributions in locations i and j, F_i and F_j, are the result of left-truncating and shifting to different extents the same underlying log productivity distribution $\tilde{F}$, and if the extent of left-truncation is greater in location i than in location j ($S_i > S_j$), then F_i can also be obtained by left-truncating and then shifting F_j. In particular, by left-truncating a share S of values of F_j and then shifting the truncated distribution by A. This also implies that in the data we will not be able to separately identify S_i, S_j, A_i and A_j, but only $S = (1 - S_i)/(1 - S_j)$ and $A = A_i - A_j$.

Since quantiles are measured with error, equation (32) does not hold exactly. Denote $\hat{\lambda}_i(y)$ the y^{th} sample quantile of F_i which, if there is no specification error and only a sampling error, is a consistent estimator of $\lambda_i(y)$. In the case where $S_i > S_j$ ($S \in [0,1)$), we are going to look for values of S and A that minimise the mean square error for equation (32) in the sample, by solving the following minimisation programme

$$\min_{S \in [0,1), A} D_i(S, A) , \tag{33}$$

with

$$D_i(S, A) = \int_0^1 \left[\hat{\lambda}_i\left(y\right) - \hat{\lambda}_j\left((1 - S) + Sy\right) - A\right]^2 \mathrm{d}y . \tag{34}$$

Consider now the case where $S_i \leqslant S_j$. With the change of variables $\frac{y - S_j}{1 - S_j} \to y$, equation (28) becomes

$$\lambda_i\left(\frac{S_j - S_i}{1 - S_i} + \frac{1 - S_j}{1 - S_i} y\right) - \lambda_j\left(y\right) = A_i - A_j \text{ , for } y \in [0,1] . \tag{35}$$

Using (30) and (31), equation (35) can be rewritten as

$$\lambda_j\left(y\right) = \lambda_i\left(\left(1 - \frac{1}{S}\right) + \frac{1}{S} y\right) - A \text{ , for } y \in [0,1] . \tag{36}$$

In the case where $S_i \leqslant S_j$ ($S \geqslant 1$), we are going to look for values of S and A that minimise the mean square error for equation (36) in the sample, by solving the following minimisation programme

$$\min_{S \geqslant 1, A} D_j(S, A) , \tag{37}$$

with

$$D_j(S, A) = \int_0^1 \left[\hat{\lambda}_j\left(y\right) - \hat{\lambda}_i\left(\left(1 - \frac{1}{S}\right) + \frac{1}{S} y\right) + A\right]^2 \mathrm{d}y . \tag{38}$$

Finally, we combine the minimisation programmes of (33) (for $S \in [0,1)$) and (37) (for $S \geqslant 1$) into the following global minimisation programme:

$$\min_{S > 0, A} \left[1_{\{S \leqslant 1\}} D_i(S, A) + 1_{\{S > 1\}} D_j(S, A)\right] . \tag{39}$$

If $S \in [0,1)$ (the distribution in location i is more truncated), the value of A that solves (39) for any given S is the average of the differences between the quantiles of the distribution in location i and quantiles of the distribution in location j after a share S of values have been truncated from the latter. This can be seen by minimising $D_i(S, A)$, as given by equation (34), with respect to A to obtain

$$\hat{A}(S) = \int_0^1 \left[\hat{\lambda}_i\left(y\right) - \hat{\lambda}_j\left((1 - S) + Sy\right)\right] \mathrm{d}y \text{ , if } S \in [0,1) . \tag{40}$$

If $S > 1$ (the distribution in location j is more truncated), the value of A that solves (39) for any given S is the average of the differences between the quantiles of the distribution in location i and quantiles of the distribution in location j after a share $1/S$ of values have been truncated from the former. This can be seen by minimising $D_j(S, A)$, as given by equation (38), with respect to A to obtain

$$\hat{A}(S) = \int_0^1 \left[\hat{\lambda}_i \left(\left(1 - \frac{1}{S}\right) + \frac{1}{S} y\right) - \hat{\lambda}_j\left(y\right)\right] \mathrm{d}y \text{ , if } S > 1 \text{ .} \tag{41}$$

Note that if quantile functions are continuous, $\hat{A}(S)$ is a continuous function of S, especially at $S = 1$. The value of S that minimises (39), denoted $\hat{S}$, is found by replacing the expression of $\hat{A}(S)$ given by (40) or (41) into (39) and using a grid search. We deduce the value of A that minimises (39) as $\hat{A} = \hat{A}(\hat{S})$.

In the results below, we report $\hat{S}$ and $\hat{A}$, as well as the corresponding criterion value $\left[1_{\{\hat{S}\leqslant 1\}} D_i(\hat{S}, \hat{A}) + 1_{\{\hat{S}>1\}} D_j(\hat{S}, \hat{A})\right]$. We also calculate standard errors and confidence intervals for the estimated parameters by bootstrapping in the distribution of log productivity[9]

The advantages of this approach are the following. First, it plainly fits with the theory. We have an underlying log productivity distribution that we do not observe, and the distributions in locations of different sizes are transformations of this distribution (with left truncation and right shift). Second, we do not need to choose *a priori* a distribution of reference from which the other distribution would be left truncated and right shifted. Third, we rely on all the moments of the distributions and yet we do not perform any smoothing of the observed distributions (which would introduce an extraneous source of variation that would be difficult to deal with).

4. Data and TFP estimations

Data

To construct our data for 1994-2002, we merge together four large-scale, French, administrative data sets from the French statistical institute (INSEE).

The first is the BRN ('Bénéfices Réels Normaux') which contains the balance sheet of all large firms in the traded sectors. For tax purpose, firms with a turnover above 730,000 euros report detailed information about their output, intermediate good consumption and materials, productive and financial assets, and their wage bill. This allows us to construct a reliable measure of value added for each firm and each year. We also retain information about total employment and the value of assets. To estimate TFP (see below), we use a measure of capital stock based on the sum of the reported book values.[10] In some TFP estimations we use the cost of capital rather than assets values. We do this using the detailed procedure developed by Boutin and Quantin (2006).

Our second data set is the RSI ('Régime Social des Indépendants') which contains the balance sheet of all small firms (i.e., those with a turnover below 730,000 euros). Although the details of the reporting differs from that of BRN, for our purpose RSI contains essentially the same information as the previous data and we thus treat it in the same way. Note that the union of BRN and RSI

[9]In a future version, we plan to bootstrap in the sample of observations used to compute the log productivities.

[10]Evaluating assets at their historical costs is not without problems. We minimise them by estimating TFP at the sector level, using sometimes highly disaggregated sectors. We also use time dummies. An alternative would be to deflate assets using economic criteria. However, our panel is rather short which makes it difficult to trace the original investments. Our procedure also differs from that of Olley and Pakes (1996) who use a permanent inventory method.

provides detailed information about nearly all French firms. Because the information about very small firms tend to be noisy, we retain only firms with more than 6 employees.

This firm level data contains more information than is usually available. For instance, US based research needs to rely either on sectoral surveys or on five-yearly censuses for which value added is difficult to compute. Despite their richness, BRN and RSI are not enough for our purpose because our model requires us to work at the establishment level. To comply with this requirement, we use two further data sets, SIREN ('Système d'Identification du Répertoire des ENtreprises') and DADS ('Déclarations Annuelles de Données Sociales'). SIREN is an exhaustive registry of all establishments in the traded sector. For each establishment and year, SIREN reports a firm identifier, a municipality code, and employment by two-digit occupational category. To avoid estimating too many coefficients for different types of labour, we aggregate two-digit occupational categories into 3 groups: high-, intermediate- and low-skill workers following the procedure of Burnod and Chenu (2001).

Since BRN and RSI only report total employment and not hours worked, we prefer to use employee level data from DADS, a matched employer-employee data set, which is exhaustive for 1994-2002. In particular, the DADS data contains the number of paid hours for each employee in each establishment.

To merge these four data sets, we extended the procedure of Aubert and Crépon (2003) for BRN and DADS to include also RSI and SIREN. The total number of observations for 1994 is 942,506. The number of observations rises slowly over the period. Finally, we also allocate each municipality to its metropolitan area ('Aire Urbaine') when there is one or classify it as rural.

To sum up, for each firm between 1994 and 2002, we know its value added, the value of its assets (with and without financial assets), total employment, and detailed cost share information. For each establishment within each firm, we know its detailed location, its wage bill, and its number of paid hours by skill level.

TFP *estimation*

For a given sector (either at the two-digit level, at the three-digit level, or sometimes using even more disaggregated sectors), we wish compute TFP at the establishment level. Appendix ? [[to be added]] shows that, under some assumptions, we can write a firm production function as if production occurred at the firm level using the capital stock and labour aggregated over all establishments. Using an approximation to aggregate labor with different skills (see Hellerstein, Neumark, and Troske, 1999) yields the following specification for the production function:

$$\ln(VA_{kt}) = b_0 + b_1 \ln K_{kt} + b_2 \ln L_{kt} + \sum c_s SKILL_{skt} + \sum d_i GEO_{ikt} + b_3 SEC_{kt} + v_t + \phi_{kt} \qquad (42)$$

where k indexes the firm; VA_{kt} is the value added in year t; b_0 is a constant; K_{kt} is the capital stock measured by the value of assets; L_{kt} is employment measured by the total number of paid hours obtained from the aggregation of the number of paid hours across all of k's establishments; $SKILL_{skt}$ is the share of skill, low, intermediate, and high, in employment (also computed from the number of paid hours in all the establishments of the firm); GEO_{ikt} is the share of the firm's employment in establishments located in location i; SEC_{kt} is the share of k's labour not in the considered sector; v_t is a time fixed effect; and ϕ_{kt} is a residual. Finally, the estimation is performed for each sector separately, the skill reference group is low, and the reference location is the smallest.

Equation (42) can be estimated with OLS or with GMM (Blundell and Bond, 1998). This latter method uses two sets of IV equations: the endogenous explanatory variables in level regressed on their lagged differences, and the differences of the endogenous explanatory variables regressed on their lagged levels. The endogenous explanatory variables are: labour, capital, skill shares. There are also some explanatory variables that are considered to be exogenous: the labour shares in every type of locations, the labour share in other sectors than the one considered and time dummies. Note that specification tests (using Sargan statistics) are often rejected. Appendix **??** [[to be added]] describes a variety of other ways to estimate TFP in our context: data envelopment (with and without instrumenting) and the procedures developed by Olley and Pakes (1996) and Levinsohn and Petrin (2003).

We then construct establishment TFP in the following way. For single-plant firms, the most natural measure is $\widehat{d_i} + \widehat{\phi}_{kt}$ where i is the firm location, $\widehat{d_i}$ is estimated from (42) and $\widehat{\phi}_{kt}$ is the estimator of the residual of (42). For multi-plant firms, we construct as many measures of productivity as there are locations where the firm locates establishments. For a firm with two plants, one in location i and one in location j, we have two productivity measures: $\widehat{d_i} + \widehat{\phi}_{kt}$ and $\widehat{d_j} + \widehat{\phi}_{kt}$. The productivity measure $\widehat{d_i} + \widehat{\phi}_{kt}$ is used to compute the productivity distribution in location i, and $\widehat{d_j} + \widehat{\phi}_{kt}$ is used to compute the productivity distribution in location j.

There are a number of fundamental issues pertaining to the estimation of production functions. The literature has long been concerned by a possible correlation between the residual, that is the TFP, and factor usage. A number of solutions have been proposed for this problem (in particular, Olley and Pakes, 1996, Levinsohn and Petrin, 2003). Adding to this, a number of things such as factor quality and output prices are not observed. The level of aggregation at which the estimation is performed is also an issue. Finally there is some uncertainty about the particular functional form that should be estimated. Before going any further it is important to note that those issues matter to us only to the extent that they affect the estimation differently across locations. A geographically-uniform bias is differenced out by our estimation approach.

In the above list, three issues are of particular concern. First, labour quality varies across locations. As shown by Combes *et al.* (2008), more productive workers tend to locate in more productive locations, arguably working for more productive firms. To avoid confounding this sorting of workers with the effects of agglomeration, we need to control for workers' skills as precisely as possible. For this reason we consider several labour types in our estimating equation (42). The second issue is that land prices vary considerably across locations. Since a satisfactory modelling of the land market is beyond the scope of our model, we leave this issue for future work. Nonetheless, we note that we get very similar results when using factor quantities and factor costs in estimating TFP. This suggests that differences in unobserved factor quality or in unobserved factors do not play any significant role. The third major issue relates to the price of final output which differs across locations. [[More to be written about this]]

Table 1. Estimation results (based on GMM TFP estimates)

Sector	Agglomeration and selection			Agglomeration only		Selection only		
	A	S	R^2	A	R^2	S	R^2	obs.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Food, beverages and tobacco	-0.11	0.98	0.50	-0.07	0.42	1.01	0.06	25853
Clothing and leather	0.39	1.10	0.87	0.25	0.72	0.92	0.09	5964
Publishing and printing	0.27	1.01	0.97	0.24	0.96	0.90	0.44	10493
Pharmaceuticals products	0.35	1.01	0.91	0.33	0.91	0.89	0.39	1831
Domestic equipment	0.20	1.01	0.99	0.18	0.98	0.96	0.32	6880
Motor vehicles	0.31	1.09	0.80	0.17	0.53	0.98	0.02	1816
Ships, locomotives and rolling stock	0.34	0.97	0.97	0.40	0.94	0.81	0.67	1143
Machinery and equipment	0.40	1.02	0.90	0.35	0.89	0.88	0.35	16332
Electric and electronic equipment	0.22	1.00	0.98	0.22	0.98	0.93	0.46	7735
Textiles	0.16	1.00	0.96	0.17	0.96	0.93	0.52	3718
Wood, pulp and paper	0.18	1.00	0.99	0.19	0.99	0.94	0.51	5985
Chemicals	0.26	1.01	0.98	0.24	0.97	0.92	0.40	6987
Basic metals and fabricated metal products	0.12	0.99	0.97	0.14	0.96	0.96	0.52	15161
Electric and electronic components	0.10	0.98	0.96	0.13	0.92	0.95	0.67	3337
Construction	0.28	1.01	0.91	0.26	0.91	0.94	0.36	78511
Vehicle sale and maintenance	0.22	1.00	0.96	0.22	0.96	0.93	0.41	32052
Wholesale and commission trade	0.30	1.02	0.98	0.27	0.96	0.91	0.39	70242
Transportation	0.26	1.01	0.86	0.24	0.85	0.94	0.42	39948
Consultancy and assistance activities	0.36	1.01	0.99	0.34	0.99	0.89	0.53	56129
Average	0.26	1.01	0.92	0.24	0.90	0.93	0.39	

5. Main results

Columns (1) and (2) in Table 1 show our baselines estimates for two-digit manufacturing and business service sectors of A and S, estimated by comparing the distribution of firm log-productivities in large locations (urban areas with over 200,000 people) with the distribution of firm log-productivities in small locations (urban areas with less than 200,000 people and rural areas). Firm log-productivities are TFP computed using GMM. Recall that $S \equiv \frac{1-S_i}{1-S_j}$, where i corresponds to large locations and j corresponds to small locations. If $S < 1$, then $S_i > S_j$, i.e., the strength of selection increases with location size. Recall also that $A \equiv A_i - A_j$. If $A > 0$, then $A_i > A_j$, i.e., the strength of agglomeration increases with location size. [[Standard errors are yet to be computed (this will require several more weeks of computation time)]]. Columns (3) in Table 1 reports a pseudo-R^2.

For most sectors, S is very close to 1, which suggests that there is not much difference in the strength of selection across locations of different size. Note that this does not imply that selection is not important, it simply suggests that its importance is similar in locations of different size. Our model showed that the extent to which selection varied across locations of different size was closely related to the extent to which product market competition is local or global (national in this case). Our results would be consistent with a situation where French firms compete with similar intensity on national markets regardless of their location.

For almost all sectors, A is above 0, which suggests that agglomeration economies are stronger in large locations than in small locations. A back-of-the-envelope calculation suggests that the average value of A we find (0.26) corresponds to a productivity increase across the board (after accounting for selection) of roughly 4% for a doubling of city size, which is in line with what is found using very different methods in the agglomeration literature (Rosenthal and Strange, 2004, Combes *et al.*, 2007). Our model showed that the extent to which agglomeration economies varied across locations of different size was closely related to the extent to which interaction are local or global (national in this case). Our results would be consistent with a situation where interactions are quite local, which matches the empirical literature looking at the spatial decay of different types of agglomeration economies (Rosenthal and Strange, 2004).

Column (4) in Table 1 reports our estimates of A when we impose the restriction $S = 1$ (no difference in the strength of selection between large and small areas), and column (5) reports the corresponding pseudo-R^2. The restriction does not change results too much.

Column (6) in Table 1 reports our estimates of S when we impose the restriction $A = 0$ (no difference in the strength of agglomeration between large and small areas), and column (7) reports the corresponding pseudo-R^2. If we do not allow agglomeration to vary across locations of different size, then part of this is picked up as variation in selection (the estimates of S are now clearly below 1), but the fit deteriorates very substantially.

We have also worked with other estimation methods for TFP, such as data envelopment (with and without instrumenting) and the procedures developed by Olley and Pakes (1996) and Levinsohn and Petrin (2003) with similar conclusions.

[[Further results to be added]]

6. Conclusions

[[To be written]]

References

Amiti, Mary and Lisa Cameron. 2007. Economic geography and wages. *Review of Economics and Statistics* 89(1):15–29.

Aubert, Patrick and Bruno Crépon. 2003. La productivité des salariés âgés : une tentative d'estimation. *Economie et Statistique* 0(368):65–94.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. Plants and productivity in international trade. *American Economic Review* 93(4):1268–1290.

Blundell, Richard W. and Steve R. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87(1):115–143.

Boutin, Xavier and Simon Quantin. 2006. Une méthodologie d'évaluation comptable du coût du capital des entreprises françaises : 1984–2002. Working paper, INSEE-DESE G2006-09.

Burnod, Guillaume and Alain Chenu. 2001. Employés qualifiés et non-qualifiés: Uneproposition d'aménagement de la nomenclature des catégories socioprofessionnelles. *Travail et Emploi* 0(86):87–105.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63:forthcoming.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2007. Estimating agglomeration effects: Does playing with different instruments give a consistent tune? Processed, University of Toronto.

Del Gatto, Massimo, Giordano Mion, and Gianmarco I.P. Ottaviano. 2006. Trade integration, firm selection and thecosts of non-Europe. Discussion Paper 5730, Centre for Economic Policy Research.

Del Gatto, Massimo, Gianmarco I.P. Ottaviano, and Marcello Pagnini. 2008. Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms. *Journal of Regional Science* 48:forthcoming.

Dixit, Avinash K. 1979. A model of duopoly suggesting a theory of entry barriers. *Bell Journal of Economics* 10(1):20–32.

Duranton, Gilles and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2063–2117.

Fujita, Masahisa and Hideaki Ogawa. 1982. Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics* 12(2):161–196.

Head, Keith and Thierry Mayer. 2004. The empirics of agglomeration and trade. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2609–2669.

Heckman, James J. and Bo E. Honore. 1990. The empirical content of the Roy model. *Econometrica* 58(5):1121–1149.

Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1999. Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labour Economics* 17(3):409–446.

Henderson, J. Vernon. 2003. Marshall's scale economies. *Journal of Urban Economics* 53(1):1–28.

Hopenhayn, Hugo. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* 60(5):1127–1150.

Levinsohn, James and Amil Petrin. 2003. Estimating production functions using inputs to control for unobservables. *Review of Economic Studies* 70(2):317–342.

Lucas, Robert E., Jr. and Esteban Rossi-Hansberg. 2002. On the internal structure of cities. *Econometrica* 70(4):1445–1476.

Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.

Melitz, Marc and Gianmarco I. P. Ottaviano. 2005. Market size, trade and productivity. Working Paper 11393, National Bureau of Economic Research.

Melitz, Marc J. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71(6):1695–1725.

Olley, G. Steven and Ariel Pakes. 1996. The dynamics of productivity in the telecommunication equipment industry. *Econometrica* 64(6):1263–1297.

Ottaviano, Gianmarco I. P., Takatoshi Tabuchi, and Jacques-Franccois Thisse. 2002. Agglomeration and trade revisited. *International Economic Review* 43(2):409–436.

Redding, Stephen and Anthony J. Venables. 2004. Economic geography and international inequality. *Journal of International Economics* 62(1):53–82.

Rosenthal, Stuart S. and William Strange. 2004. Evidence on the nature and sources of agglomeration economies. In Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2119–2171.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Printed for W. Strahan, and T. Cadell.

Syverson, Chad. 2004. Market structure and productivity: A concrete example. *Journal of Political Economy* 112(6):1181–1222.

Vives, Xavier. 1990. Trade association disclosure rules, incentives to share information, and welfare. *Rand Journal of Economics* 21(3):409–430.