



1-2011

Is the Division of Labour Limited by the Extent of the Market? Evidence From French Cities

Gilles Duranton
University of Pennsylvania

Hubert Jayet

Follow this and additional works at: https://repository.upenn.edu/real-estate_papers

 Part of the [Real Estate Commons](#)

Recommended Citation

Duranton, G., & Jayet, H. (2011). Is the Division of Labour Limited by the Extent of the Market? Evidence From French Cities. *Journal of Urban Economics*, 69 (1), 56-71. <http://dx.doi.org/10.1016/j.jue.2010.08.004>

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/real-estate_papers/62
For more information, please contact repository@pobox.upenn.edu.

Is the Division of Labour Limited by the Extent of the Market? Evidence From French Cities

Abstract

This paper provides supportive evidence to the notion that the division of labour is limited by the extent of the (local) market. We first propose a theoretical model. Its main prediction is that scarce specialists occupations are over-represented in large cities. Using census data for French cities, we find strong empirical support for this prediction.

Keywords

division of labour, specialisation, extent of the market

Disciplines

Real Estate

Is the Division of Labour Limited by the Extent of the Market?

Evidence from French Cities

Gilles Duranton^{*‡}

London School of Economics

Hubert Jayet^{*†}

MÉDÉE, Université de Lille 1

Work in progress, 5 August 2004

ABSTRACT: This paper provides some evidence that the division of labour is limited by the extent of the (local) market. We first propose a theoretical model. Its main prediction is that scarce occupations are over-represented in large cities. Using census data for French cities, we then provide strong empirical support for this prediction.

Key words: Division of labour, specialisation, extent of the market.

JEL classification: R23, J24, J44.

*We are grateful to Yasushi Asami, Pierre-Philippe Combes, Masa Fujita, Laurent Gobillon, Vassilis Hajivassiliou, Yoshi Kanemoto, Tomoya Mori, Taka Tabuchi, and conference and seminar participants at Aberdeen, Berkeley, Kyoto, Lille, Paris, Philadelphia, and Tokyo for comments.

‡Department of Geography and Environment, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom, g.duranton@lse.ac.uk, <http://cep.lse.ac.uk/~duranton>. Also affiliated with the Centre for Economic Policy Research, and the Centre for Economic Performance at the London School of Economics. Financial support from the Leverhulme Trust is gratefully acknowledged.

†MÉDÉE, Université de Lille 1, Faculté des Sciences Economiques et Sociales, 59655 Villeneuve d'Ascq CEDEX, France, Hubert.Jayet@univ-lille1.fr.

1. Introduction

Cities offer considerable efficiency advantages for production. This empirical observation already figured prominently in Adam Smith's *Wealth of Nations* (1776) and in Alfred Marshall's *Principles* (1890). It was confirmed by modern econometric studies, which offer strong support for it. A doubling of city population is generally acknowledged to increase output per worker by 3 to 8% (see Rosenthal and Strange, 2004, for a comprehensive survey). To explain this fact, Smith (1776) put forward one specific mechanism that builds on three propositions. First, there are some efficiency benefits to specialisation and the division of labour. Second, the division of labour is limited by the extent of the market. Third, the extent of the market is in turn determined by transportation efficiency, making cities the 'natural' markets wherein the division of labour takes place. However, the division of labour is only one possible mechanism accounting for the greater efficiency of cities.

Identifying the exact sources of urban increasing returns is important for both theory and policy. Theorists need to be provided with empirical guidance regarding the relevance of their assumptions and which mechanisms to investigate, whereas policy makers need to know which market failures must be addressed. In this paper, we take a step towards resolving this identification problem by providing evidence about the division of labour and its relation to city size using data for French cities in 1990.¹ Our main result is that the division of labour strongly increases with city population.

Despite being a central tenet of economic analysis, there is remarkably little empirical work on the division of labour. The chief reason behind this paucity of work is that formidable difficulties arise to measure the division of labour. Indeed, the idea behind the division of labour is that the production process can be finely divided into smaller 'tasks' as workers become more specialised. This does not lend itself well to empirical analysis because industries divide their production differently and most tasks are industry-specific (if not firm-specific). The range of procedures performed by, say, physicians has obviously very little to do with the specialties of lawyers. Hence, a 'standard classification of tasks' would be very difficult to set up. Then, collecting information from workers would obviously be very costly. As a result, there is no systematic data on the tasks performed by workers in large cross-sections.

Hence it is no surprise that existing empirical work on the division of labour is scarce and tends to focus on particular industries: US physicians in Baumgardner (1988a) or lawyers in Garicano and Hubbard (2003). These detailed case-study analyses are important because they circumvent the main measurement problem. They can thus precisely document the main determinants of the division of labour. Their results support the idea that the division of labour increases with the size of the market. However, the size of the market may not be the sole determinant of the division of labour (Becker and Murphy, 1992). In the case of lawyers, Garicano and Hubbard (2003) also highlight the importance of subtle information and incentive problems. Asymmetric information

¹In this respect, our approach is in the spirit of Jaffe, Trajtenberg, and Henderson (1993) who provide evidence about technological spill-overs but do not quantify their effects. See also Holmes (1999) who provides evidence about the greater use of out-sourcing by firms in larger cities. Put differently, we provide a joint-test of Smith's second and third proposition leaving for further work his first proposition about measuring the gains from the division of labour.

implies for instance that lawyers have an incentive to retain customers, even when the requests of the latter do not correspond to their specialties.

The danger of case-studies is of course that such examples may not be representative of the wider economy and may offer a distorted perspective on the division of labour. An analysis covering the whole of manufacturing, or even better the whole economy, is obviously needed. Our starting point is that, although there is no standard classification of tasks, in most countries there is a standard occupational classification (SOC). These classifications are usually designed to reflect the division of labour. In particular, at their finest level occupations are defined according to job duties. We argue below that SOCs can provide important information about the division of labour. The key difficulty with SOCs is that their labels are not directly informative about the division of labour except for a few specific cases (e.g., specialised physician). Put differently, although SOCs are a natural source of information about the division of labour, they do not provide any immediate metric by which to measure it.

To solve this conceptual problem and to guide our empirical analysis, we start by developing a simple model of the division of labour. In his famous description of the farmers in the Scottish Highlands, Adam Smith (1776, p. 122) notes that "*In the lone houses and very small villages which are scattered about in so desert a country as the Highlands of Scotland, every farmer must be a butcher, baker, and brewer for his own family.*" He also underscores (p. 121) that "*There are some sorts of industry, even of the lowest kind, which can be carried on nowhere but in a great town. A porter, for example, can find employment and subsistence in no other place.*"² Building on this insight, we assume a production structure in which production in an industry consists in conducting a series of task. There is a 'central' task that defines the industry and a collection of 'peripheral tasks'. In the wood industry for instance, the central task is to chop trees whereas the monitoring of healthy trees, the transport of cut wood, its storage, the cleaning of the factory are all peripheral tasks. Taking another example from services, the central task in the educative sector is arguably teaching whereas scheduling courses or registering students or even possibly grading and tutoring students can all be viewed as peripheral tasks.

Workers choose their occupations ex-ante. Generalists can perform all tasks, but they are most productive at the central task. Specialists can only perform their specialist task, for which they have an advantage over generalists. The other key feature of the model is that specialist labour is not divisible. There are two main results in this model. First, on average the proportion of specialists for any task increases with market size. In very small markets, generalists will do everything just like Adam Smith's isolated farmers. In larger markets, peripheral tasks are taken over by specialists. Put differently, the division of labour increases with the extent of the market so that scarce occupations (like Smith's porter) are relatively over-represented in larger cities. The main intuition behind this result is that the indivisibility of specialist labour becomes less of an issue 'on average' as ever larger markets are considered. The second main property of the model follows

²A modern version of this is depicted in the 1983 feature movie *Local Hero*. The main character is an oil-field buyer from Houston — arguably an occupation reflecting a high level of division of labour — who is sent to a small village on the Northern coast of Scotland where oil has been discovered. The first local he meets runs the sole hotel-restaurant singlehandedly. He is also the local accountant, village leader and serves occasionally as taxi driver. As for the fisherman-pub-owner-baker, he asks somewhat incredulously to the American: "You only got one job?"

directly from the first one: As a result of the deepening of the division of labour and the subsequent replacement of generalists by specialists, output per worker increases on average with market size. This model constitutes a contribution of independent interest to the theoretical analysis of agglomeration economies.

Given that the second prediction is not specific to the model, our empirical work concentrates on the first prediction. In line with our theoretical model, the basic idea of our empirical test is to compare the distribution of workers across occupations and urban areas. Formally, we take a logit approach and consider that the observed distribution of workers across occupations in every industry and every urban area is the outcome of random draws. Our null hypothesis is that the division of labour is unrelated to city size. We test it against the prediction of the model of a positive effect of city size on the probability of receiving scarcer occupations. The scarcity of an occupation in a sector is defined empirically by its level of employment relative to other occupations in the same sector. In conducting our test, we control for the fact that industries are not equally represented in cities and that occupations are not evenly distributed across sectors.³

Our analysis uses a large extract from the 1990 French census with more than 5 million observations. We have 111 three-digit French sectors, 360 urban areas grouped in seven size classes, and 454 occupations grouped in four scarcity quartiles in each sector. Our null hypothesis is strongly rejected in favour of the alternative hypothesis suggested by our model. The proportion of workers employed in the scarcest occupations (i.e., occupations in the first scarcity quartile) is on average 70% larger in Paris than in the smallest French cities.

The quality of our results depends on how well we can identify the (scarce) specialists as opposed to the (more common) generalists. There are several possible measurement problems. First, although in the French standard occupational classification, occupations are finely defined on the basis of the tasks performed, there may be some division of labour taking place within occupations, which we may not be able to observe directly. Second, some occupational entries may be more precisely defined than others, which may lead us to identify as common generalist some occupations that are aggregates of specialists. Third, given the importance of out-sourcing, it may be that the division of labour takes place across sectors rather than within sectors. We devote considerable attention to these measurement issues and the potential pitfalls raised by our definition of scarcity. To deal with them, we successfully perform a series of robustness tests using alternative definitions of scarcity and alternative identification strategies for specialist workers.

Identifying what is the 'market' raises another set of issues. In the Smithian tradition, our baseline analysis identifies the market with the city. However, the relevant spatial scale wherein the division of labour takes place might be smaller than the city. Alternatively it may also be larger if tasks can be traded across cities. It is also the case that some tasks may be tradable across cities

³More precisely, in our logit we estimate the probability of a worker located in a given urban area to fill a job in a specific occupation. The latent variable of the logit model depends on city size crossed with the scarcity of the occupation. When the null hypothesis is true, the estimated cross-effects do not differ significantly from zero: scarce occupations occur with the same frequency in all cities. When the alternative hypothesis is true and using large cities and core occupations as references, the estimated cross-effects are negative: scarce occupations occur with a lower frequency in smaller cities. Furthermore, these cross-effects should become more negative as smaller cities and scarcer occupations are considered.

whereas others are not. In the last part of the paper, we devote some attention to the problems raised by the relevant spatial scale for the division of labour.

The final caveat is that the division of labour may take place not at the level of the local labour market but within product markets instead. Adam Smith's celebrated pin factory is one where a larger market leads workers to specialise on an ever more narrow range of tasks. Becker and Murphy (1992) argue that as the market for pins gets larger, there is only so far that workers can specialise to produce standard pins. Instead, the division of labour is more likely to take place between firms leading to the production of different varieties of pins. Unfortunately, we are not able to detect such division of labour within the product market (a.k.a. product differentiation) with our data.

As a theoretical contribution, our model relates to the small literature on the division of labour, which is surveyed in Xiang and Ng (1998) and Duranton and Puga (2004). Following, Baumgardner (1988*b*), a first strand of literature uses a continuous framework wherein labour specialisation can become ever more narrow as market size increases. In this type of framework, the division of labour increases with market size in a smooth and continuous fashion. However, the positive effect of worker specialisation on productivity is directly assumed as a reduced-form and has not received strong micro-economic foundations.

Following Stigler (1951), a second (and still underdeveloped) strand of literature views the basic tasks of economic activity as discrete. The source of the benefits to the division of labour are small indivisibilities at the worker level. Then, the division of labour proceeds discontinuously as market size increases. Our model belong to this second strand. The main difference with previous work is that we put more structure in the production process to generate some testable implications regarding the comparative statics on market size.

As an empirical contribution looking at a large number of sectors, our work is to our knowledge unique. More generally, we are not aware of theoretically informed empirical work focusing on the effect of the extent of the market on division of labour apart from the two case-studies already discussed above (Baumgardner, 1988*b*; Garicano and Hubbard, 2003).

The rest of the paper proceeds as follows. The next section presents our theoretical framework and derives its key prediction. Section 3 presents our data and some basic descriptive evidence about the division of labour. Section 4 presents our estimation methods and our main results. Section 5 performs a series of robustness tests. Finally Section 6 concludes.

2. Theory: a model of the pin factory

To motivate and guide the empirical work conducted below, we consider in this section a model of division of labour taking place across occupations. Our main mechanism relies on the existence of indivisibilities at the level of individual workers in some occupations. Our two main results are that scarcer occupations tend to be relatively more represented in larger markets and that output per worker increases with market size.

We consider an industry with a three stage production process. At the first stage, generalist and specialist labour produce different tasks. These tasks are then used to produce an homogeneous

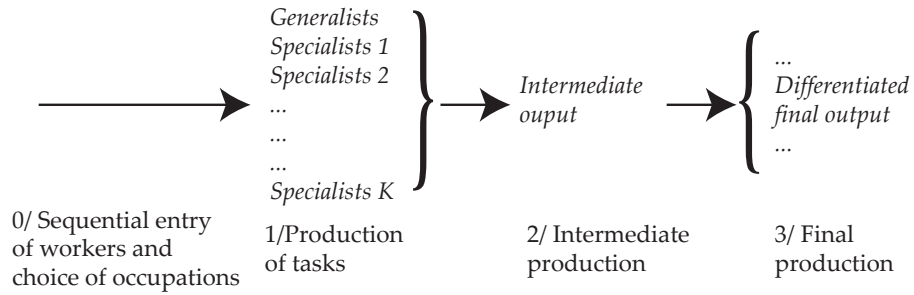


Figure 1. Production structure and timing

intermediate good at the second stage. Finally, this homogenous intermediate good is used as input by differentiated final producers at the third stage. This production process is represented in Figure 1. In what follows we first describe the details of the production of tasks and intermediate goods before turning to the description of the production of final goods.

The intermediate output z is produced by assembling $K + 1$ tasks indexed $0, 1, \dots, K$. These tasks are produced directly from generalist and specialist labour. We view these tasks as a discrete collection rather than a continuum wherein they could become ever more thinly defined (as for instance in Baumgardner, 1988*b*, or Duranton, 1998). Although we do not dispute that the division of labour may take place in this continuous fashion, our empirical work relies on a discrete classification of occupations.⁴

Task 0 is the task that defines the industry and we refer to it as the ‘central task’. For instance in the road transport industry, the central task is to drive trucks. The other tasks are referred to as the ‘peripheral tasks’. In this example, the peripheral tasks are (among others) the maintenance of trucks, the loading and unloading of cargo, taking orders, etc.

At the initial period 0, an endogenously determined number of workers sequentially chooses an occupation, taking final demand in the industry as given.⁵ This choice, made before tasks are produced, is irreversible. Each worker can either become a generalist or specialist for a given task. Generalist can produce task 0 as well as any other task. The labour supply of each generalist is perfectly divisible across tasks. For the central task, one unit of generalist labour produces one unit of this task. For any peripheral task k (with $k = 1, \dots, K$), one unit of generalist labour produces γ_k units of that task. Generalists are more efficient at the central task than at any peripheral task: $\gamma_k < 1$.

The unit labour supply of any specialist is indivisible. Specialists can only produce the task corresponding to their speciality.⁶ We refer to a specialist of task k as a k -specialist. For any peripheral task k , one unit of k -specialist labour produces σ_k units of that task. We assume that

⁴Whether the division of labour has a discrete or continuous nature is an empirical question beyond the scope of this paper. Nonetheless, we discuss its potential effects on our results in further detail below.

⁵Because of small indivisibilities, the workers are not in general symmetric ex-post. A simultaneous entry would lead to multiple generic equilibria and difficult co-ordination issues. Sequential entry avoids these problems and ensures the existence of a unique equilibrium.

⁶This assumption greatly simplifies the analysis since any task is restricted to be performed by at most two occupations.

specialists are more efficient at their task than generalists. We also assume that specialists are more efficient at their task than generalists at the central task so that $\sigma_k > 1$ for all $k = 1, \dots, K$.

According to these assumptions, the supply of tasks is thus given by:

$$z_0 = g_0, \quad (1)$$

and

$$z_k = \gamma_k g_k + \sigma_k s_k, \quad (2)$$

for $k = 1, \dots, K$, where g_k is the supply of (divisible) generalist labour on task k and s_k is the discrete number of k -specialists. We assume perfect complementarity across tasks to produce the intermediate good z .⁷ We have:

$$z = \text{Min}_{k \in \{0, 1, \dots, K\}} (z_k). \quad (3)$$

In equilibrium, cost minimisation by competitive intermediate producers implies $z_k = z$ for all tasks. For the central task, competition among (perfectly divisible) generalists implies that they are paid at their marginal product. Without loss of generality, we can normalise the price of the central task to unity so that the generalist wage is $w_0 = p_0 = 1$. For simplicity assume the wage outside the industry is inelastic and also equal to unity.

There is price competition among producers of task k (k -specialists and any generalist labour employed at this task).⁸ Following this, no equilibrium can be such that there is excess supply of k -specialists: $\sigma_k s_k > z$ when $s_k \geq 2$. In this case, price competition would drive the price of this task to zero: $p_k = 0$. Consequently the equilibrium must be such that there is insufficient supply of k -specialists to meet demand: $\sigma_k s_k \leq z$. Put differently, in equilibrium, any peripheral task is in general supplied by generalists as well as specialists. Under such circumstances, price competition implies limit pricing by specialists: $p_k = \frac{1}{\gamma_k}$.

To understand entry into the different occupations, it is useful to compare markets with different levels of intermediate demand. For arbitrarily small markets, no specialist of any kind finds it worthwhile to enter. A first k -specialist enters when her income is at least equal to that of a generalist: $\text{Min}(z, \sigma_k) p_k = 1$. Since $p_k = \frac{1}{\gamma_k}$, it follows that the first k -specialist entry takes place when: $z = \gamma_k$. No other k -specialist enters as long as it leads to excess specialist supply since her entry would drive prices (and thus her income) to zero. Hence, a second k -specialist enters for $z = 2\sigma_k$. More generally, we have:

$$\begin{cases} s_k = 0 & \text{if } z \leq \gamma_k, \\ s_k = 1 & \text{if } \gamma_k < z < 2\sigma_k \\ s_k = \lfloor \frac{z}{\sigma_k} \rfloor & \text{if } 2\sigma_k \leq z \end{cases} \quad (4)$$

where $\lfloor \cdot \rfloor$ denotes the floor function (or integer part since only positive numbers are considered).⁹

⁷Perfect complementarity is assumed for simplicity. The presence of different specialists receiving non-negative rents in equilibrium only requires imperfect substitutability across tasks.

⁸Qualitatively similar results are obtained under most alternative forms of competition, the crucial assumptions here are the indivisibility of specialist labour and the imperfect substitutability among tasks.

⁹Note that the first specialist entry for a task is entirely driven by the productivity of generalists at this task whereas the second and all subsequent entries are driven only by the productivity of specialists. Note also that in equilibrium the entry of the first specialist is efficient but subsequent are in general suboptimal. Indeed, efficient specialist entry requires the s^{th} k -specialist to enter when $z = (s-1)\sigma_k + \gamma_k$ instead of $z = s\sigma_k$ at the market equilibrium.

In equilibrium, the sequential entry process implies the following. In a market where there is room for multiple specialists for each task, the first worker to enter chooses the specialist occupation offering the highest labour income, $w_k = \frac{\sigma_k}{\gamma_k}$. Subsequent entrants choose the same occupation up to the point where further entry would lead to an excess supply of specialist labour on this task and drive its price to zero. At this point, the next entrant chooses the specialist occupation with the second highest $\frac{\sigma_k}{\gamma_k}$, etc. When no further specialist can enter without driving the price of a task to zero, the remaining entrants choose the generalist occupation until supply is equal to z for each task. Generalists perform the central task and satisfy the residual demand for the peripheral tasks.

Using (2) and the equilibrium condition $z_k = z$, we find that market equilibrium for peripheral task k implies:

$$g_k = \frac{z - \sigma_k s_k}{\gamma_k}. \quad (5)$$

Then labour market clearing implies $n = g_0 + \sum_{k=1}^K (g_k + s_k)$ where n is total labour supply in the industry. By (1) and (4), $g_0 = z$. After inserting this and (5) into the clearing condition and simplifying, we find:

$$n = z \left(1 + \sum_{k=1}^K \frac{1}{\gamma_k} \right) - \sum_{k=1}^K \left(\frac{\sigma_k}{\gamma_k} - 1 \right) s_k, \quad (6)$$

where s_k is given by (4). When $z \geq \underset{k}{\text{Max}}(\sigma_k)$, equations (4) and (6) imply that the proportion of k -specialists in the industry workforce is given by:

$$\frac{s_k}{n} = \frac{\lfloor z/\sigma_k \rfloor}{pz - \sum_{i=1}^K \left(\frac{\sigma_i}{\gamma_i} - 1 \right) \lfloor z/\sigma_i \rfloor}, \quad (7)$$

where $p = \sum_{k=0}^K p_k = 1 + \sum_{k=1}^K \frac{1}{\gamma_k}$ is the equilibrium price of the intermediate good.

Given the highly discontinuous nature of the floor function, $\lfloor z/\sigma_i \rfloor$, intervening in both the numerator and denominator of (7) it is not possible to derive general results regarding the behaviour of $\frac{s_k}{n}$. To see this consider three market sizes z_1, z_2 and z_3 such that $z_1 = \sigma_k s - \epsilon$, $z_2 = \sigma_k s$ and $z_3 = \sigma_k s + \epsilon$ where s is an integer and ϵ is a small quantity. Then from (7) it follows that the proportion of k -specialists first increases as market sizes increases from z_1 to z_2 and then decreases as market sizes increases from z_2 to z_3 .

To go further, it is useful to think of market size z as the outcome of an ex-ante draw from a uniform distribution over $[0, 2Z]$. In this case, we can derive some comparative results about $E\left(\frac{s_k}{n}\right)$, the expected proportion of k -specialists as a function of total employment.¹⁰ To derive these results, we need to assume additionally that the Z is large relative to the σ_k and that for $k \neq k'$, the ratio σ_k to $\sigma_{k'}$ cannot be written as the ratio of two integers.

We can now derive our first key result:

Proposition 1 *The expected employment share of k -specialists in the industry workforce increases with the expected demand for intermediates, Z .*

¹⁰This could correspond empirically to cases where market size is measured imprecisely by the statistician or when cities are lumped together in size classes as we do below.

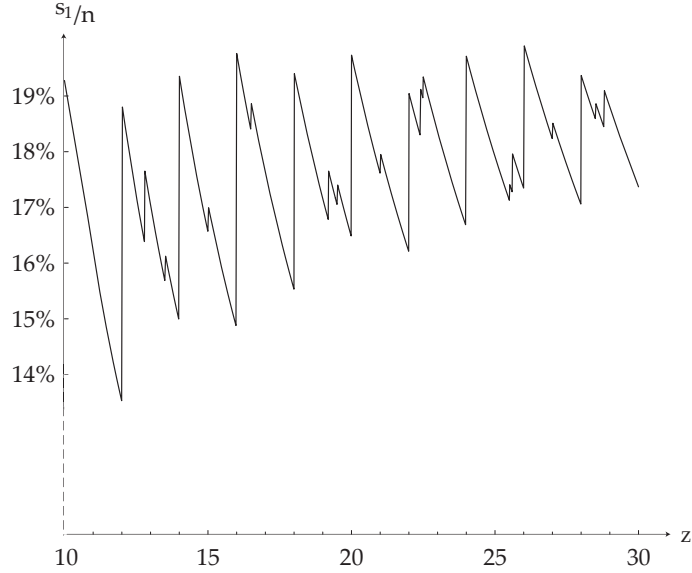


Figure 2. Share of task 1 specialists: An example

We considered three peripheral tasks and the following parameter values: $\sigma_1 = 2, \sigma_2 = 1.5, \sigma_3 = 3.2, \gamma_1 = 0.4, \gamma_2 = 0.68,$ and $\gamma_3 = 0.86$. The graph represent the share of task 1 specialists (s_1) as a function of the demand for intermediates (z) varying between 10 and 30.

The proof of this proposition is relegated to Appendix A.

The intuition for this proposition is the following. Each specialist is a discrete lump of unit labour so that it is only when there is a sufficiently large gap in the market for task k that a new k -specialist can enter. Otherwise the gap between demand and specialist output is seized by generalists. Note that the size of the gap for a new k -specialist to enter, σ_k , is independent of market size. Consequently the amount of generalist labour fluctuates between 0 and σ_k/γ_k . In a small but growing market, the *proportion* of specialists in the labour employed at a task declines considerably before a new specialist enters. By contrast, when the market is large, the gap needed for new k -specialist to enter is *relatively* small.¹¹ Hence on average, larger markets have a tendency to have a greater proportion of k -specialists.

Figure 2 represents a typical example for the evolution of the proportion of 1-specialists, $\frac{s_1}{n}$, as a function of the demand for intermediates, z . The main peaks represent the different integer thresholds for s_1 , $\lfloor \frac{z}{\sigma_1} \rfloor$, as the demand for the intermediate good increases. The smaller peaks between the main peaks represent the integer thresholds for the other tasks. These smaller peaks occur for the following reasons. When a new specialist enters to produce, say, task 2, a quantity σ_2 of generalist labour is released and there is an increase in efficiency. Hence when a new 2-specialist enters, the workforce needed to produce at this threshold z is reduced by $\sigma_2 - 1$. This decline in the denominator of (6) implies an increase in the share of 1-specialists in the workforce. This example also shows that, although there is no closed-form solution to the model in the general case, solutions can be easily computed through a simple algorithm when the distributions of γ_k

¹¹Formally the proportion of specialists fluctuates between 1 and $\frac{s_k}{s_k + \sigma_k/\gamma_k}$. The latter quantity is small when s_k is small whereas it is close to unity when s_k is large.

and σ_k are specified.

The second proposition, although very simple, is crucial for the empirical identification of the model. It states that:

Proposition 2 *Specialists are scarce occupations relative to generalists.*

This result stems directly from equation (4), which can be used to show that the ratio of generalists to k -specialists is no smaller than $\sigma_k > 1$.

Proposition 2 relies on two main assumptions. First, all tasks are symmetric (in the sense that all coefficients in the production function of intermediates are equal to one). This assumption could be relaxed easily. Some tasks could be more demanding than others (in the sense that they could have a coefficient above one in the production function (3)). The only requirement for our prediction to hold in a statistical sense in this setting is that peripheral tasks are no more demanding (in labour time) on average than the central task. Arguably (and by a natural convention), the central task can be defined as the most demanding one (i.e., that with the highest coefficient in the production function).

The second key assumption underpinning this result is that $\sigma_k > 1 > \gamma_k$ (i.e., specialists are more efficient at their task than generalists at the central task). One could justify this productive advantage of specialists as a compensation for being able to execute fewer tasks. Furthermore, in a more complete model, this assumption could also be justified by the large human capital investment needed for specialists occupations. If instead of $\sigma_k > 1 > \gamma_k$, we had $1 > \sigma_k > \gamma_k$ (i.e., specialists are less efficient at their task than generalists at the central task), the proposition above would hold only for markets below a given size threshold (recall that for a small enough market, there is no specialist).

Together, Propositions 1 and 2 imply the following testable empirical implication:

Empirical Implication 1 *The employment share of scarce occupations increases with market size.*

The next result is also important since it leads to further testable empirical predictions.

Proposition 3 *The expected relative number of k -specialists $\left(\frac{s_k}{z/\sigma_k}\right)$ decreases with their productivity (σ_k).*

The proof is relegated in Appendix B. It is a consequence of the fact that the integer problem becomes on average relatively more important with the size of the indivisibility (σ_k). This proposition implies that, for a given market size, we expect an increase in the productivity of k -specialists to imply a more than proportionate decline in their number.

Together, Propositions 1 and 3 imply that:

Empirical Implication 2 *The ratio of the employment share of a scarcer occupation to the share of a less scarce occupation increases with market size.*

Finally, since specialists have a higher output than generalists on their task, a greater proportion of specialists obviously implies some efficiency gains.

Proposition 4 *On average output per worker increases with the demand for intermediates, z .*

Turning to the production of final goods, we could assume a competitive framework and close the model straightaway. To check the robustness and the limits of the mechanism developed so far, we embed it instead in the standard monopolistic competition framework developed by Salop (1979).

There is an endogenously determined number of differentiated final producers, N , in the industry. The product space is the unit circle. Each firm produces one variety of the industry good, and we assume that varieties are evenly spaced around the unit circle.¹² The production of the final good entails a fixed cost, ϕ , paid in units of intermediate goods and a constant requirement, λ , of intermediate goods for each marginal unit. The profit of final producer i is given by

$$\Pi(i) = q(i)y(i) - \lambda p y(i) - \phi p, \quad (8)$$

where $q(i)$ is the price of the variety produced by final producer i and $y(i)$ its output. There is a continuum of D consumers whose tastes are uniformly distributed around the unit circle. Each consumer consumes one unit of final good. If a consumer buys a good that is less than a perfect match for her own taste, she has to bear a cost of mismatch. When the distance between the characteristics of the good and the consumer's taste is x , the cost of mismatch expressed in terms of numéraire is equal to μx . Each firm sets the price of its variety so as to maximise its profits. Each consumer buys the variety with the lowest delivery price (i.e., the sum of the price and mismatch cost).

Suppose N firms have entered the market. Because firms and consumers are symmetrically located in the characteristic space, it is natural to look for a symmetric equilibrium in which they all set the same price, q . Each firm effectively competes with two competitors whose variety characteristics are at a distance $\frac{1}{N}$ to its right and to its left. A consumer whose taste is located at a distance x from firm i is indifferent between buying her unit of final good from firm i selling at price $q(i)$ or from i 's closest competitor when

$$q(i) + \mu x = q + \mu \left(\frac{1}{N} - x \right). \quad (9)$$

Firm i will thus sell its goods to all consumers with tastes located within a distance x of its characteristics and have total sales equal to:

$$y(i) = 2Dx = \frac{D}{N} + [q - q(i)] \frac{D}{\mu}. \quad (10)$$

By lowering its price $q(i)$, firm i can increase its market area beyond its proportional market share ($\frac{D}{N}$). The effect of a lower price on total sales becomes stronger as the cost of mismatch declines or the number of consumers increases. Substituting (10) into the expression giving profits (8), differentiating the resulting concave function with respect to $q(i)$, and then substituting $q(i) = q$ yields the equilibrium price of any final variety. It is given by:

$$q = \lambda p + \frac{\mu}{N}. \quad (11)$$

¹²Maximal differentiation is usually imposed exogenously in this class of model. See Duranton and Puga (2004), whose presentation we follow, for a discussion of this issue.

Prices differ from marginal cost because final producers have some market power. However this market power decreases with the number of firms in the market (N) and increases with the cost of mismatch (μ). There is thus a competition effect whereby a larger number of firms forces them to lower their prices. We can now substitute (10) and (11) into the profit (8). After simplification, free-entry implies:

$$N = \sqrt{\frac{\mu D}{\phi p}}. \quad (12)$$

The total demand for intermediate goods is then equal to

$$z = \left(\frac{D}{N} \lambda + \phi \right) N = D \lambda + \sqrt{\frac{\phi \mu D}{p}}. \quad (13)$$

There are obviously increasing returns in the production of the final goods. From (13), an increase in final demand, D , leads to a less than proportional increase in the demand for intermediates. The sources of increasing returns lie in the competition effect highlighted above. A larger market increases firms entry but it also reduces their profit margins. This implies that the number of firms, N , grows less than proportionately with demand, D , as made clear by equation (12). Hence, as the market grows, so does output per firm. Consequently, as the market grows, the fixed cost is spread over a larger output which reduces average intermediate requirement. We can also note that consumers will not only benefit from this competition effect but also from a matching externality. The average distance between a consumer's taste and the characteristics of the variety she buys is equal to $\frac{1}{4N}$. Hence using (12) the average mismatch cost is equal to $\sqrt{\phi p \mu / D} / 4$. It also decreases with demand, D .

From (13), an increase in final demand, D , leads to an increase in the demand for intermediates, z . Consequently all the propositions and empirical predictions derived above apply not only to the size of the market for intermediates, z , but also to final demand, D , which we identify with city population in our baseline estimation. In a nutshell:

Empirical Implication 3 *The employment share of scarce occupations increases as larger cities are considered. So does the ratio of the employment share of a scarcer occupation to the share of a less scarce occupation.*

A few comments are in order. Note first that our results are to a large extent independent from the market structure for final goods. We chose a market structure based on Salop (1979), but our results also directly apply to any market structure for final goods wherein there is a positive relation between final demand and intermediate consumption.¹³

Second, although in our model the production of the intermediate goods is out-sourced to competitive suppliers, we could easily consider an extension in which intermediate goods would be specific to the firm and be produced in-house. In such a case, the division of labour would be

¹³Note nonetheless that we cannot explore the aspects of the division of labour related to product differentiation between firms. To understand this, consider the limit case where the division of labour takes place within firms, which produce different varieties of the same good using the same technology. When a larger market only implies more firms rather than larger firms (Dixit and Stiglitz, 1977), the occupational structure of the industry will remain invariant with market size.

limited by the size of the firm and no longer by the size of the market. However, our main results would carry through nonetheless. This is because the output of each firm increases (see equations 12 and 13) as the size of the market increases. This increase in output increases the demand by each firm for each firm-specific task. In turn, this justifies the replacement of generalists by specialists. Hence, with firm-specific intermediates the division of labour would still be *indirectly* limited by the extent of the market.¹⁴ Put differently, our main predictions are robust to the exact form taken by the division of labour.¹⁵

Third, the division of labour is not the sole source of increasing returns in this model when all the stages of production are considered. Competition between producers facing a fixed cost and able to exploit it more intensively as well as gains from better matches also play a role when market size increases. This implies that we cannot expect to identify the gains from the division of labour of labour separately from those of the other sources of increasing returns by investigating the effects of market size on output per worker. Our identification strategy is instead to test the other predictions of our model regarding the evolution of the employment shares of scarce occupations as larger cities are considered.

Fourth, in this model, the division of labour has a discrete rather than continuous nature. The main justification for this is that the key indivisibility, which matters for the division of labour, is that of the worker themselves. Another important justification is that our empirical work relies on a discrete occupational classification. In this respect, a model with discrete choices is more appropriate. The problem is of course that the standard occupational classification we use may not capture perfectly the occupations defined in the model. This important issue is discussed at length below.

Fifth, despite their different settings, our argument bears some resemblance with that put forward by Stigler (1951).¹⁶ In his informal argument, he assumes that the operations of the firms can be divided into a series of tasks for which there is a U-shaped average cost curve. Initially for a small market, a unique firm produces all the tasks. For some tasks, the firm is rather far from the efficient size. As the market grows, some tasks can be out-sourced to specialist producers operating closer to the efficient size. On average in a larger market, specialist suppliers of a given task will be operating closer to the efficient size. Although our model cannot pretend to be a direct formalisation of Stigler's argument, the two approaches are strongly related in the sense that in both cases market expansion allows producers to operate closer to the efficient size.

Finally, note that the model represents a closed economy. Put differently, it assumes that workers, tasks and intermediate/final goods are not mobile across markets. We then view the

¹⁴Note that in this case, the efficiency results would be different. Conditional on the number of firms and their output, the division of labour would be constrained-efficient in this setting. This is because in-house specialists would be allocated through a command equilibrium rather than a market equilibrium.

¹⁵The downside is that in absence of occupation data at the firm level, we cannot distinguish between these different aspects of the division of labour in our empirical work. Being able to distinguish between these different aspects of the division of labour should be a priority for future research given that such exploration may provide a window on the spatial organisation of firms.

¹⁶Stigler (1951) considers that the division of labour takes place between firms, rather than at the worker level, with an argument relying on smooth cost functions.

city as the natural spatial scale at which the division of labour takes place.¹⁷ Regarding workers, the justification for their lack of mobility is the importance of commuting frictions and moving costs. Regarding tasks, it is difficult to view most tasks as mobile (especially because the division of labour itself reinforces the need for close co-ordination). An exception may be some very specialised consumer services (like those provided by brain surgeons) for which consumers are willing to travel between cities to acquire them. We think that such tasks are the exception, rather than the rule.

The lack of mobility of intermediate and final goods is empirically more questionable. When tasks are immobile but the intermediate and final goods can be traded freely across cities, the city is not the relevant market for the local industry. Instead, the relevant market may be national or international in scope. But even when the goods are mobile, the division of labour remains limited by the size of the industry locally (to repeat, this holds provided tasks are immobile). When goods are mobile, we can still examine whether the division of labour is limited by the local size of the industry. Our empirical work below pays some attention to these issues.

3. Data and some basic evidence

Data

We use a random sample of the 1990 French census, covering 25% of the French population in full-time employment. It contained initially 5,518,704 observations. For each observation, we know the occupation, workplace location at the municipality level, and sector.

For occupations, we use the French 1982 official standard occupational classification ('nomenclature des professions'). At the one-digit level, this classification distinguishes workers on the basis of their work status and general skills: farmers, self-employed, professionals, 'intermediate' occupations, other white-collar employees, and blue-collar workers. At the two-digit level, the classification retains a skill dimension (e.g., skilled and unskilled blue-collars). Work status also keeps playing a role (e.g., private vs. public sector). It also starts distinguishing workers by functions (e.g., administration vs. trade). At the three-digit level, functions and the type of product (e.g., chemical vs. metal) become the dominant criteria. Finally, at the four-digit level occupations are nearly exclusively defined by their job tasks. For instance, among IT technicians, it distinguishes between development, production, and installation/maintenance. At the four digit level, the French SOC contains 454 entries for active workers. It is worth noting that, at the four-digit level, this SOC was explicitly designed to capture the division of labour between occupations (INSEE and DARES, 2003). This information is reliable because occupational entries are used for wage-setting purpose in sector-wide bargaining between employers and unions.

The spatial units are the French urban areas ('aires urbaines'). Urban areas are aggregates of the elementary French geographical units, the municipalities (36,500 of them approximatively). Each urban area aggregates municipalities around a built urban centre, so that all municipalities within

¹⁷Although there is nothing specifically urban about our model, it could be easily extended to incorporate commuting frictions (see Duranton and Puga, 2004, regarding the possible ways to embed models with increasing returns in an urban setting).

a given urban area send more than 40% of their employed inhabitants to other municipalities within the urban area. There are 360 urban areas with employment above 4000. We only consider the workers whose employment is located in one of these urban areas (4,319,480 of them). For computational tractability reasons, we must group urban areas into classes of different sizes. We define 7 classes. Some summary statistics about this grouping are reported in Table 1. The issues raised by our choice of spatial units with respect to our estimation are discussed below.

Table 1. Summary statistics

Population size	Number of urban areas	Number of employees
2,000,000 at least (Paris)	1	1,232,759
300,000 to 1,999,999	5	567,198
150,000 to 299,999	8	395,302
80,000 to 149,999	22	611,381
40,000 to 79,999	39	529,258
20,000 to 39,999	63	439,459
less than 20,000	222	544,123

Finally for sectors, we use the French classification of activities at the three-digit levels with 111 non-agricultural sectors.

We can tabulate employment by occupation and sector. Among the 50,394 cells (454 occupations \times 111 sectors), 24,275 are non-empty. Within each sector, we only consider occupations with sufficient employment. This is because we are not interested in anecdotal cases like gardeners employed in financial intermediation or metal workers in wholesale trade. Therefore, we select only the 6,733 sector-occupation cells with at least 30 workers in the census extract we use.¹⁸ The corresponding number of workers is 4,209,995. Among the 111 sectors, 2 disappear, so we now have 109 sectors. The list of urban areas is unchanged.

We now turn to the definition of what constitutes a scarce occupation. In line with our model, our main definition of scarcity is based on absolute employment in each sector. We say that occupation k is scarcer than occupation l in sector j if there are fewer workers employed by sector j in occupation k than in occupation l at the national level: $n_k^j < n_l^j$. For tractability reasons, we define scarcity groups of occupations. In our baseline, we define 4 groups. We first sort occupations by increasing employment within each sector. Then, for the K occupations represented within sector j , the first $\frac{K}{4}$ occupations are in group 1. The following $\frac{K}{4}$ are in group 2, and so on. Therefore, a higher rank for a group of occupations corresponds to a higher level of scarcity.

The measurement problems associated with the French SOC and the problems relating to our definitions of scarcity are discussed in the next section after the presentation of the estimation equation.

¹⁸Note that this censoring is likely to bias our results downwards since the very scarcest occupations (which are likely to be in the largest cities) are ignored.

Basic descriptive data and preliminary evidence

Recall that a key aspect of our model is the distinction between common generalist occupations and scarce specialists. In the sample we use, the five most common occupations are secretaries (3.57%), office workers in public service (1.74%), accounting employees (1.63%), cleaners (1.58%), and secondary school teachers (1.50%). The five scarcest occupations are private sector archivists, crew masters for fishing boats, modelers, civil planning engineers, and glass blowers (each representing less than 0.01% of our sample). Consistent with our model, the five most common occupations are typically core occupations in a sector (school teachers, office workers for civil service) or occupations reflecting a low degree of division of labour (secretaries and cleaners), whereas the five scarcest occupations are highly specialised and reflect a high degree of division of labour.

Within each sector, the most common occupation employs on average 18% of the sector workforce. There is of course a lot of variation across sectors with a minimum at 5% in naval construction and maxima at 62% for fishing and 52% for road transport (salaried fishermen and truck drivers respectively). The average share of sectoral employment of the five most common occupation is 49%. The other 51% of employment are on average spread across more than 50 occupations. The average Gini coefficient for the distributions of occupations in a sector is 0.89. This strongly suggests that most sectors are organised around a few core occupations and many scarce peripheral occupations.

It is also useful to look at the distribution of occupations within cities and how it varies with city size. First, we can compute a relative diversity index $RDI(u)$ for each urban area u :

$$RDI(u) = \frac{1}{\sum_k \left| \frac{n_{u,k}}{n_{u,\cdot}} - \frac{n_{\cdot,k}}{n_{\cdot,\cdot}} \right|}, \quad (14)$$

where $n_{u,k}^j$ is employment in city u , sector j , and occupation k and a dot stands for a sum over the missing index. As cities grow larger (and in absence of division of labour), we may expect their occupational composition to reflect more accurately the distribution of occupations in the economy. Instead, our theory predicts that both large and small cities should be different from the national average. This is because scarce occupations are over-represented in large cities, whereas more ubiquitous occupations are over-represented in small cities.

The linear correlation between employment size in a city and its relative diversity index is positive. However there is some evidence of a bell-shaped relationship. RDI is highest for mid-size French cities such as Bordeaux, Nantes, Tours, Valence or Rouen whereas the largest city, Paris is 26th.

Our model also predicts that an absolute diversity index for occupations like:

$$ADI(u) = 1 / \sum_k \left(\frac{n_{u,k}}{n_{u,\cdot}} \right)^2 \quad (15)$$

should increase strongly with city size. When we Spearman-rank correlate this index with city size, the correlation is strong at 0.71 and is highly significant. This clearly points at a stylised fact whereby larger cities have more occupations being represented. This is confirmed by the rank

correlation between the number of occupations with no employment and city population, which is equal to -0.77 .¹⁹

These broad features about the distributions of occupations across sectors and cities are consistent with the basic assumptions and the results of our model. It remains however to test our results more precisely and more directly.

4. Estimation and main results

Specification

Within every sector j , the $n_{u,k}^j$ workers in occupation k located in city u result from random drawings from a population of $n_{..}^j$ workers with probabilities $\pi_{u,k}^j$, $u = 1, \dots, U$, $k = 1, \dots, K$. For every j , $\pi_{u,k}^j$ is supposed to have the following logit form:

$$\pi_{u,k}^j = \frac{\exp(Y_{u,k}^j)}{\sum_{\substack{i=1,\dots,U \\ l=1,\dots,K}} \exp(Y_{i,l}^j)}, \quad (16)$$

with

$$Y_{u,k}^j = \alpha_u^j + \beta_k^j + \zeta_{m(u),r(j,k)}, \quad (17)$$

where $m(u)$ is the class of city u (with $m = 1 \dots M$) and $r(j,k)$ is the scarcity level of occupation k in sector j (with $r = 1 \dots R$). The three coefficients in (17) are α_u^j , the fixed-effect for each city and sector, β_k^j , the fixed-effect for each occupation and sector, and $\zeta_{m(u),r(j,k)}$, the effect of being an occupation of a given level of scarcity in a sector for each class of city size. The occupation-sector fixed-effects control for the fact that occupations are unevenly distributed across sectors. The city-sector fixed-effects control for the fact that sectors are not evenly spread across cities.²⁰ The coefficients of interest are of course the cross-effects $\zeta_{m(u),r(j,k)}$. They capture the relative distribution of occupations across classes of city sizes according to their level of scarcity.

A natural null hypothesis is that within sector j the distribution of occupation k across cities is independent from the scarcity of the occupation k , $r(j,k)$. Such null hypothesis implies that the conditional probabilities $\frac{\pi_{u,k}^j}{\pi_{u..}^j}$ are the same for all cities u .²¹ When this null hypothesis holds, all $\zeta_{m(u),r(j,k)}$ are equal to zero.

Instead our model suggest an alternative hypothesis whereby (i) the employment share of scarce occupations increases with city size and (ii) the ratio of employment shares of a very scarce occupation to a less scarce occupation also increases with city size. More formally, if r corresponds to scarcer occupations than r' then the sequence $\zeta_{1,r} \dots \zeta_{M,r}$ should decrease more rapidly (or increases less rapidly) than the sequence $\zeta_{1,r'} \dots \zeta_{M,r'}$. Equivalently, if m corresponds

¹⁹Note that this estimate takes into account the fact that, even in absence of spatial difference, larger cities are more likely to receive a positive employment in any occupation.

²⁰See Mori, Nishikimi, and Smith (2003) for a detailed analysis of this issue.

²¹This is equivalent to the conditional probabilities $\pi_{u,k}^j / \pi_{..,k}^j$, $u = 1, \dots, U$ not depending on the occupation k .

to smaller cities than m' , then the sequence $\zeta_{m,1}, \dots, \zeta_{m,R}$ should decrease more rapidly (or increase less rapidly) than sequence $\zeta_{m',1}, \dots, \zeta_{m',R}$.

Equation (17) implies the estimation of a very large number of fixed-effects: one for each sector-occupation and one for each sector-city. Even though the model is made conditional on the sectors, we would have $(109 - 1) \times 360 = 38,880$ city-sector fixed effects, $(109 - 1) \times 454 = 49,032$ occupation-sector fixed-effects, and $(7 - 1) \times (4 - 1) = 18$ ζ coefficients. This is well beyond our computational power. Besides, among the 87,930 coefficients, only 18 are of interest. In Appendix C, we show that the likelihood of (16) can be separated into three parts, which can be estimated separately as in nested logit models. The first part estimates the same fixed-effects but for each group of cities instead of each city (7 instead of 360) each scarcity level of occupation instead of each occupation (4 instead of 454) and the ζ coefficients. The second and third components of the likelihood estimate conditional probabilities for the choices across cities within a given class and across occupations for a given scarcity level. Note that this procedure reduces the number of coefficients to be estimated fairly dramatically. To estimate the first part of the likelihood, we only need to estimate $(109 - 1) \times 7 + (109 - 1) \times 4 + 6 \times 3 = 1,204$ coefficients instead of 87,930.²² However, 1,204 is still a fairly large number for a direct estimation. Instead, we use an iterative three-stage estimation procedure, which in case of convergence, yields the same results as the direct estimation.

Estimation issues

Several issues regarding the data and our definition of scarcity can now be discussed in relation to our estimating equation. First, the main drawback of the French SOC (and of all existing SOCs) is that, although it is very detailed, it may not be able to capture subtle aspects of the division of labour. A fairly obvious case regards university professors: they are all part of the same entry. Economists are lumped together with scholars doing comparative literature or chemistry. Among economists, field distinctions are of course also ignored. We expect the inability of the SOC to capture subtle differences between jobs to be more of an issue in places where the division of labour is the most extensive. Thus, this measurement problem implies that our results are likely to underestimate the importance of the effect of market size on the division of labour.

Furthermore, this measurement problem may not occur evenly. Some occupations in the SOC may be more finely defined than others. Imagine an occupational entry that lumps together several different specialist occupations. As a consequence, this entry may not be 'scarce' when we measure it. This measurement problem should again bias our results downwards.

To deal with this measurement problem, two (partial) solutions can be envisioned. We note first that there is a systematic sectoral dimension to it. The level of detail of the French SOC is much greater for 'manufacturing' jobs than for 'service' jobs. For instance, blue-collar workers (mostly found in manufacturing) represent less than 25% of French workers, but they account for

²²Another positive implication is that at this level of aggregation the number of zeroes is small (and much smaller than if the estimation was performed directly at the individual level).

more than half the entries in the French SOC.²³ Hence, to assess the impact of mis-measurement, we can perform our analysis for large groups of sectors separately (manufacturing, business and consumer services). Second, we expect differences in how finely defined occupations are to be unrelated to sectors.²⁴ We can exploit this feature in the following way. Rather than looking at the absolute employment in each occupation and sector, we can look instead at the employment in each occupation and sector cell relative to national employment in the occupation. Put differently, if the breadth of occupational entries was random, we could learn something by looking within each sector at the employment in each occupation relative to national employment in this occupation.

Second, some occupations related to some particular functions may be scarcer than others and tend to agglomerate in larger cities. For instance, managers may agglomerate in large cities for reasons unrelated to the division of labour (Duranton and Puga, 2001). In sectors where managerial occupations are scarce, this will bias our results and lead us to over-estimate the effect of market size on the division of labour. To avoid this bias, we can run our analysis on groups of occupations within sectors, rather than all occupations in sectors. For instance, instead of looking all the occupations in the steel industry, we will be looking at all the managerial occupations in this industry and then all then intermediate professions in the same industry, etc. Related to this is the concern that many tasks may be out-sourced by final producers to suppliers that belong to different sectors. Put differently, all occupations in narrowly defined sectors may not be the appropriate units of observation when looking at the division of labour. It may be worth instead looking at groups of occupations within more broadly defined sectors.

Third, note that it is possible to avoid the problems raised by our identification of scarcity with specialisation by taking a completely different approach. Rather than identifying specialised occupations from their number of workers, we can use the characteristics of the workers employed. More specifically, specialised occupations usually require a specific training. To a large extent, this training is acquired before entry on the labour market. Hence, it is possible to use the homogeneity of the educational achievements of workers in an occupation to measure how specialised this occupation is. Obviously, very broad measures of educational achievements are not going to be very helpful. For instance, most executives have gone through higher educations. However, detailed educational achievements containing, not only the length of the training, but also its type (general vs. vocational) and its speciality are more useful for our purpose. In this respect, we can use the information contained in the French Labour Force Survey (Enquête Emploi). For about half a million French workers in the 1990s, we know both their detailed occupations and their detailed educational achievement. This survey distinguishes between general and vocational education, and for both types of education it gives the type of degree obtained and its subject (i.e. medical, engineering, humanities, etc). Obviously, this alternative definition has drawbacks of its own but it nonetheless provides a good robustness test.

²³In the baseline, these differences between manufacturing and services will lead us to further underestimate the true extent of the division of labour in cities which are relatively specialised in services, that is the largest cities.

²⁴In part this is because the French SOC was constructed independently from the French SIC. Furthermore, we expect that it will be mostly small peripheral occupations in all sectors that are lumped together.

Finally, nothing guarantees that the urban area is the right level of spatial aggregation when looking at the division of labour. When mobility is very low, the municipality may be more relevant whereas when mobility is high, the regional level may instead be more relevant. When a sector exports most of its goods to other cities, looking at how the division of labour increases with local employment across sectors may not make very much sense. Instead, one would like to look at how the division of labour increases with the size of the city-*and*-sector.

The estimations for these alternative definitions of specialisation/scarcity and the exploration of the proper spatial scope for the division of labour are performed in Section 5.

Estimation results

Table 2. Estimation results for the base model (all sectors)

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.184 (0.00002)	-0.258 (0.00003)	-0.272 (0.00002)	-0.343 (0.00002)	-0.355 (0.00002)	-0.383 (0.00002)
High	0	-0.214 (0.00004)	-0.284 (0.00006)	-0.342 (0.00004)	-0.405 (0.00004)	-0.429 (0.00005)	-0.439 (0.00004)
Very high	0	-0.208 (0.00004)	-0.312 (0.00006)	-0.414 (0.00003)	-0.463 (0.00004)	-0.511 (0.00004)	-0.524 (0.00003)

3,052 observations. Population in thousands for city sizes. Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets. The null hypothesis of all coefficients being equal to zero is rejected.

The estimation results for the full model are reported in Table 2. The null hypothesis of null effects is strongly rejected. Instead, the estimation results conform to the predictions of our model. Scarce occupations are more represented as cities get larger and this over-representation increases with the scarcity of the occupation. Taking Paris and the most common occupations as references, the coefficients $\xi_{..}$ become more negative as smaller cities and scarcer occupations are considered. The only exception to this regular ordering involves the coefficient for high and very high scarcity in the second class of cities. Note however that this exception is very minor.

Although we built our estimation so that it under-estimates the effects of city size on the division of labour, our results points at relatively large effects. The scarcest occupations are about 70% more frequent in Paris than in the smallest French cities.

5. Robustness checks

In this section, we perform a series of robustness checks.

Alternative thresholds for scarcity

Because in some sectors the distribution of employment across occupations is highly skewed, a grouping in quartiles to measure scarcity may not be appropriate. On average, the 14th largest

occupation in a sector (which typically corresponds to the first quartile) contains only 1.5% of the sector workforce. Hence in most sectors, we are already beyond 'core' occupations. Since in our baseline we use these occupations in our reference, our results may be biased downwards.

It is worth conducting our analysis again with different cut-off points for the scarcity groups. In Table 3 we report the results when we take the 15th, 45th, and 75th percentile of the employment distribution rather than the quartiles to construct four scarcity groups.

Table 3. Estimation results for alternative scarcity thresholds (all sectors)

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.067 (0.00001)	-0.126 (0.00001)	-0.155 (0.00001)	-0.179 (0.00001)	-0.206 (0.00001)	-0.223 (0.00001)
High	0	-0.154 (0.00001)	-0.215 (0.00002)	-0.259 (0.00001)	-0.310 (0.00001)	-0.339 (0.00001)	-0.346 (0.00001)
Very high	0	-0.219 (0.00001)	-0.327 (0.00001)	-0.380 (0.00001)	-0.475 (0.00001)	-0.506 (0.00001)	-0.544 (0.00001)

3,052 observations. Population in thousands for city sizes. Within each of the 109 sectors, occupations are grouped by their distribution by employment size with threshold at 15% (scarcest), 45%, and 75%. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

A straightforward comparison between Tables 3 and 2 shows that the results are very similar. Hence, the relative over-representation of scarce occupations is robust to the exact thresholds chosen to define scarcity.

Separate estimations by groups of sectors

As noted above, the occupations predominantly found in manufacturing are more narrowly defined than the occupations found in services. Given that services are found predominantly in large cities, this may bias our results and lead us to underestimate the effects of city size on the division of labour. As an alternative to the analysis above and using the same definitions of urban areas and occupations, we can run separate estimations for three groups of sectors (manufacturing, business services, and consumer services).

The results are reported in Tables 4(a-c). The null hypothesis of null effects is again strongly rejected for all three groups of sectors. Instead the results again conform to the predictions of the model. The coefficients are in general more negative for smaller urban areas and scarcer occupations. For manufacturing activities, the estimated effects are even stronger than for the joint estimation of all sectors. The results are somewhat weaker for business services and consumer services. However, they still work in the same direction.

From this second robustness test, our conclusion is the following. We note first that the level of detail of the SOC may affect the magnitude of the effects. However our evidence regarding the

Table 4. Estimation results for the base model by groups of sectors

(a) Manufacturing sectors

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.283 (0.00005)	-0.334 (0.00007)	-0.350 (0.00005)	-0.426 (0.00005)	-0.400 (0.00005)	-0.431 (0.00004)
High	0	-0.296 (0.00009)	-0.371 (0.00014)	-0.456 (0.00008)	-0.487 (0.00009)	-0.527 (0.00009)	-0.512 (0.00008)
Very high	0	-0.307 (0.00010)	-0.426 (0.00013)	-0.491 (0.00008)	-0.568 (0.00008)	-0.607 (0.00008)	-0.635 (0.00006)

1,260 observations. Population in thousands for city sizes. Within each of the 45 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets. The null hypothesis of all coefficients being equal to zero is rejected.

(b) Business service sectors

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.102 (0.0001)	-0.164 (0.0002)	-0.168 (0.0002)	-0.208 (0.0002)	-0.221 (0.0003)	-0.317 (0.0002)
High	0	-0.203 (0.0003)	-0.249 (0.0004)	-0.265 (0.0004)	-0.309 (0.0005)	-0.286 (0.0006)	-0.451 (0.0004)
Very high	0	-0.223 (0.002)	-0.307 (0.0003)	-0.408 (0.0002)	-0.405 (0.0003)	-0.464 (0.0003)	-0.417 (0.0005)

420 observations. Population in thousands for city sizes. Within each of the 15 sectors (financial and business oriented services), occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

(c) Consumer service sectors

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.115 (0.0001)	-0.214 (0.0001)	-0.179 (0.0001)	-0.265 (0.0001)	-0.285 (0.0001)	-0.338 (0.0001)
High	0	-0.091 (0.0001)	-0.185 (0.0002)	-0.165 (0.0002)	-0.296 (0.0002)	-0.331 (0.0002)	-0.328 (0.0001)
Very high	0	-0.025 (0.0004)	-0.141 (0.0003)	-0.244 (0.0001)	-0.316 (0.0002)	-0.315 (0.0002)	-0.348 (0.0001)

360 observations. Population in thousands for city sizes. Within each of the 14 sectors (real estate, household services, health and education), occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

existence of a pervasive increase in division of labour with city size does not depend on how finely defined the SOC is.²⁵

Estimation results for relative scarcity

As highlighted above, some occupational entries may be more finely defined than others. In this case, a definition of scarcity based on total employment may be misleading since many ubiquitous occupations may be just constellations of scarce occupations. Their share may increase with city size and thus bias our results downwards. This measurement problem should in principle affect all sectors since occupations are defined independently from sectors. Consequently, instead of defining scarcity as a function of total employment, we can define it as a function of employment in the sector relative to national employment for the occupation. Put differently, we use a relative measure of scarcity (i.e., specificity) instead of an absolute one (total employment).

More formally, occupation k is scarcer than occupation l in sector j if the share of sector j in occupation k is lower than its share in occupation l at the national level: $\frac{n_k^j}{n_k} < \frac{n_l^j}{n_l}$. To define scarcity groups of occupations for this alternative definition, we again sort occupations by increasing specificity. Then, as with the previous definition for scarcity, we define 4 groups with (approximately) the same number of occupations.

Table 5. Estimation results with relative scarcity thresholds (all sectors)

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.071 (0.00001)	-0.172 (0.00002)	-0.168 (0.00001)	-0.229 (0.00001)	-0.234 (0.00001)	-0.284 (0.00001)
High	0	-0.187 (0.00002)	-0.269 (0.00003)	-0.333 (0.00002)	-0.389 (0.00002)	-0.372 (0.00002)	-0.400 (0.00002)
Very high	0	-0.174 (0.00002)	-0.258 (0.00003)	-0.335 (0.00002)	-0.393 (0.00002)	-0.392 (0.00002)	-0.400 (0.00002)

3,052 observations. Population in thousands for city sizes. Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets. The null hypothesis of all coefficients being equal to zero is rejected.

As shown by the results reported in Table 5, the null hypothesis of null effects is again strongly rejected. The results are qualitatively similar to previous estimations. They also conform to the predictions of our model. For every scarcity level, the smaller the urban areas, the more negative is the effect. And the higher the scarcity level, the more negative is the coefficient. Then, the scarcer an occupation within an activity and the smaller an urban area, the lower is the employment share of that area compared to its employment share in Paris. We can note that the effects are slightly weaker than with the baseline definition. This suggests that our initial definition of scarcity (based on absolute employment) is more adequate than this alternative (based on relative employment).

²⁵These findings are difficult to interpret further since they may result from the fact that the division of labour can be observed more easily when occupations are more finely defined or, alternatively, the division of labour may be more pervasive in manufacturing than services.

Controlling for the functional specialisation of cities

As described above, functionally related occupations are usually lumped together in the same one-digit occupational categories in the French SOC. These occupational categories are not evenly distributed across cities. Managers are typically over-represented in large cities, whereas blue-collar workers are over-represented in small cities. If managerial occupations have on average less employment than blue-collar occupations, we will observe that scarce occupations (the managers) are indeed relatively more represented in large cities. This is consistent with the predictions of our model. However, this feature need not be an outcome of the division of labour. A model of functional urban specialisation in which (scarce) managers locate in larger cities may generate this result without the division of labour playing any role (Duranton and Puga, 2001). Our baseline analysis may thus over-estimate the effect of city size on the division of labour.

To check for the robustness of our results we can run our baseline on groups of occupations for each sector rather than each sector (i.e., instead of considering 99 sectors we consider 5×99 groups of occupations within sectors — self-employed, managers, intermediate occupations, employees, blue-collar workers). Except for the fact that we consider only three scarcity groups (with a third of the occupations each to avoid scarcity groups having too few occupations), the analysis is then the same as in our baseline. The results are reported in Table 6. They show that introducing controls for broad occupational groups only have a small impact on the baseline results.

Table 6. Estimation results for sectors and occupational groups (all sectors)

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Low	0	0	0	0	0	0	0
Intermediate	0	-0.134 (0.00002)	-0.187 (0.00003)	-0.186 (0.00003)	-0.235 (0.00004)	-0.262 (0.0001)	-0.296 (0.0006)
High	0	-0.151 (0.00007)	-0.224 (0.0001)	-0.297 (0.00007)	-0.335 (0.00008)	-0.389 (0.0002)	-0.363 (0.0005)

3,212 observations. Population in thousands for city sizes. Within each sector and group of occupation, occupations are grouped by terciles of their distribution by employment size. Standard errors are between brackets. The null hypothesis of all coefficients being equal to zero is rejected.

In Table 7, we report additional results for 5 groups of occupations within sectors. This time, we are using a more aggregated sectoral classification with 16 sectors. This should allow us to capture any deepening of the division of labour taking place within broad sectoral branches instead of narrowly defined sectors. Although we still observe some relationship between the division of labour and market size, the results are far from being as strong as in the baseline. This suggests that the effects of market size we evidenced are unlikely to be driven by any mechanism leading large cities to specialise functionally into occupations which happen to be scarce nationally.

Division of labour within cities vs. within cities-and-sectors

When intermediate and final goods are freely tradable, the size of the city in which the tasks are assembled is not the relevant market size to look at since most of the output is exported to other

Table 7. Estimation results for sectors and occupational groups (aggregate sectors)

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.057 (0.00004)	-0.035 (0.00005)	-0.077 (0.00004)	-0.084 (0.00004)	-0.114 (0.00005)	-0.109 (0.00004)
High	0	-0.170 (0.0001)	-0.186 (0.0001)	-0.186 (0.0001)	-0.273 (0.0001)	-0.317 (0.0001)	-0.317 (0.0001)
Very high	0	-0.173 (0.0003)	-0.173 (0.0004)	-0.257 (0.0003)	-0.383 (0.0004)	-0.429 (0.0004)	-0.344 (0.0003)

448 observations. Population in thousands for city sizes. Within each of the 16 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

cities, if not other countries. On the other hand, the mobility of intermediate and final goods does not change anything to the first part of the model. It is only the way the market is defined that changes. Put differently, the mobility of goods implies that only the third of Smith's 1776 proposition need to be revised (provided of course that tasks remain immobile). Transport frictions no longer make the city the natural market to look at. However, the division of labour may still take place within each sector locally. In this case, it is the size of the local sector that should matter.

The issue is thus whether the relevant unit of observation for the market is total employment in the city or the employment in the city-sector. Note that this opposition is reminiscent of the traditional distinction in urban economics between localisation economies and urbanisation economies. When localisation economies prevail, the interactions take place mostly within each sector. It is thus the size of the local sector that matters. With urbanisation economies, the interactions take place mostly across sectors. In this latter case, it is the size of the entire local economy that matters.

To investigate this question, we can run the same analysis as before but use classes of city-sector sizes instead of classes of city sizes. These classes could be defined using either absolute employment thresholds or employment thresholds relative to the size of each sector. Given that some sectors (such as those in consumer services) are much larger than the others and fairly ubiquitous, using absolute employment thresholds is unlikely to yield meaningful results. Instead, we defined relative thresholds for seven classes. The distribution of employment across cities in most sectors is very skewed. Thus, trying to have the same number of workers in each class would lead the lowest class to have a very large number of city-sectors (more than 80% of them), whereas the highest class would have only a very small number of large sectors. Conversely, trying to have the same number of city-sectors in each class would lead to large imbalances in terms of employment. Hence we decided for a compromise solution trying to balance in each class the number of city-sectors and total employment.

The results are reported in Table 8. There is some evidence that scarce occupations are more represented where the local sector is relatively larger. The effects are, however, less strong than for the baseline. The ordering of coefficients by decreasing relative size and increasing scarcity is also less clear-cut. This suggests that the city is a more adequate spatial unit to look at than the city-sector.

Table 8. Estimation results for employment in city-sectors (all sectors)

Scarcity	Relative employment in city-sectors						
	1	2	3	4	5	6	7
Very low	0	0	0	0	0	0	0
Low	0	-0.171 (0.00001)	-0.194 (0.00001)	-0.230 (0.00002)	-0.250 (0.00002)	-0.232 (0.00003)	-0.218 (0.00003)
High	0	-0.243 (0.00002)	-0.263 (0.00003)	-0.309 (0.00003)	-0.293 (0.00004)	-0.313 (0.00006)	-0.322 (0.00006)
Very high	0	-0.232 (0.00003)	-0.290 (0.00003)	-0.383 (0.00003)	-0.378 (0.00003)	-0.396 (0.00004)	-0.328 (0.00007)

448 observations. The size classes of relative employment are as follows:

Class 1: top 1.1% of relative size (298 city-sectors representing 34.5% of national employment.

Class 2: the following 2.7% (736 city-sectors representing 16.5% of national employment.

Class 3: the following 4.8% (1303 city-sectors representing 14.5% of national employment.

Class 4: the following 8.2% (2248 city-sectors representing 12.8% of national employment.

Class 5: the following 12.5% (3409 city-sectors representing 9.8% of national employment.

Class 6: the following 18.5% (5058 city-sectors representing 6.9% of national employment.

Class 7: the following 52.2% (14239 city-sectors representing 5.0% of national employment.

Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

Alternative choices of spatial units

We now turn to the issue of the spatial scale at which the division of labour takes place. To check further the robustness of our baseline results, we run our baseline estimation using French employment areas rather than French urban areas. Employment areas correspond to local labour markets. The two key differences with urban areas is that the 341 French employment areas cover the whole French territory and that large French cities may comprise more than one employment area, whereas neighbouring small urban areas may be lumped together in the same employment area.

The results are reported in Table 9 (a). The first interesting feature is that for the last two size classes, the coefficients are even more negative than for the baseline. This is because these last two sizes classes are composed of rural areas (absent in the baseline), where the division of labour is even less extensive than in small French urban areas. The second main feature of this table is that the results for the first five size classes are very close to the baseline. This is because there is a strong correlation between the size of an urban area and that of its employment area. For instance, although Paris is divided into several employment areas, six of them are in the top 10 largest employment areas in the country.

In Table 9 (b), we report results for much larger spatial units, the 22 French regions. Unfortunately, there is a strong correlation between the population of a region and the population of its main city since French regions are typically defined around a major urban centre. Despite this, the results are much weaker than when cities or employment areas are used. Between the third and the sixth class, there is no clear ordering of the coefficient like in the baseline. These results suggest that the division of labour takes place at a spatial scale smaller than the region.

Finally in Table 9 (c), we report results for the smallest spatial units for which we have data: the

Table 9. Estimation results for alternative spatial units (all sectors)

(a) Employment areas

Scarcity	Class of size						
	2000 to 400	400 to 250	250 to 140	140 to 80	80 to 50	50 to 30	30 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.139 (0.00001)	-0.242 (0.00001)	-0.315 (0.00001)	-0.398 (0.00001)	-0.433 (0.00001)	-0.499 (0.00001)
High	0	-0.180 (0.00003)	-0.279 (0.00003)	-0.378 (0.00003)	-0.492 (0.00003)	-0.508 (0.00003)	-0.572 (0.00003)
Very high	0	-0.218 (0.00002)	-0.325 (0.00003)	-0.439 (0.00003)	-0.545 (0.00003)	-0.526 (0.00003)	-0.632 (0.00003)

3,052 observations. Population in thousands for employment area sizes. The reference class is composed of Central Paris, Nanterre (in the West of Paris) and Lyon. The employment size cut-offs were chosen so that all classes contain approximately the same overall employment. Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

(b) Regions

Scarcity	Class of region size						
	10 to 4	4 to 1.2	1.2 to 1	1 to 0.75	0.75 to 0.6	0.6 to 0.48	0.48 to 0
Very low	0	0	0	0	0	0	0
Low	0	-0.283 (0.00001)	-0.390 (0.00002)	-0.339 (0.00002)	-0.351 (0.00002)	-0.428 (0.00002)	-0.502 (0.00003)
High	0	-0.337 (0.00002)	-0.454 (0.00004)	-0.407 (0.00003)	-0.412 (0.00004)	-0.504 (0.00005)	-0.610 (0.00007)
Very high	0	-0.378 (0.00002)	-0.473 (0.00004)	-0.449 (0.00003)	-0.418 (0.00004)	-0.557 (0.00005)	-0.614 (0.00007)

3052 observations. Population in million for region sizes. The employment size cut-offs were chosen so that all classes contain approximately the same overall employment. The reference region is Ile de France which comprises Paris. Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

(c) Municipalities

Scarcity	Class of municipality size						
	Paris	135 to 35.6	35.6 to 15.3	15.3 to 6.2	6.2 to 2.16	2.16 to 0.54	0.54 to 0
Very low	0	0	0	0	0	0	0
Low	0	-0.127 (0.00002)	-0.163 (0.00001)	-0.248 (0.00001)	-0.283 (0.00001)	-0.382 (0.00001)	-0.625 (0.00001)
High	0	-0.164 (0.00004)	-0.226 (0.00003)	-0.295 (0.00003)	-0.347 (0.00003)	-0.453 (0.00003)	-0.623 (0.00002)
Very high	0	-0.247 (0.00002)	-0.263 (0.00003)	-0.371 (0.00002)	-0.424 (0.00002)	-0.533 (0.00002)	-0.648 (0.00003)

3052 observations. Population in thousands for municipality sizes. The employment size cut-offs were chosen so that all classes contain approximately the same overall employment. The reference is Central Paris. Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets.

The null hypothesis of all coefficients being equal to zero is rejected.

36,500 French municipalities. The results are strong with a clear ordering of the coefficients and large differences between Paris and the smallest French municipalities. Our ability to draw sharp conclusions here is limited by the fact that there is again a strong correlation between the size of an urban area and the size of its municipalities. Nonetheless the strength of these results hints at the fact that the division of labour may be more extensive in central cities (the largest municipalities) than in their suburbs. This hypothesis should be explored more in depth in future work.

Defining specialisation by the ex-ante education

To assess whether an occupation is specialised or not, we can use the details of the educational achievement of the workers in this occupation. The identifying assumption is that specialised occupations require a specific training. Hence, we can measure the degree of specialisation for an occupation by the share of the dominant degree rather than by scarcity. Denoting by $v(d,k)$ the share of degree d in occupation k , we can compute $\bar{v}(k) = \underset{d}{\text{Argmax}} v(d,k)$, the share of the main degree for occupation k .

We can then proceed as previously and rank all existing occupations within each sector according to $\bar{v}(k)$. Again we can define four quartiles of increasing scarcity and perform the same estimation as in the baseline.²⁶

[[To be done]]

We can also see how much correspondence there is between this definition of specialist occupation and the baseline definition based on scarcity.

[[To be done]]

Model selection

Finally, it is important to look at the issue of model selection. Our empirical results are consistent with the predictions of our model and support the idea that the division of labour is limited by the extent of the market. However it might be the case that another theory may generate similar predictions.

The crux of our model is that small (i.e., worker level) indivisibilities are easier to spread in larger markets. Hence our model belongs to the broader family of sharing mechanisms that provide micro-foundations for urban increasing returns (Duranton and Puga, 2004). Two other broad class of mechanisms can also generate urban increasing returns: matching and learning. Standard matching models typically take the structure of occupations needed for production as given. To generate some clustering, a matching function with increasing returns is needed. If occupations within a sector cannot be spatially separated, clustering should take place by sector

²⁶Another possible definition could be based on the importance of workers' turnover within the occupation relative to the turnover in and out of the occupation. The lower the turnover in and out of the occupation relative to the turnover within the occupation, the more specialised is the occupation. A first problem is that differences in the definitions of occupations may bias directly the results since, all else equal, a more narrowly defined occupation will exhibit more turnover in and out of the occupation. A second and even more serious problem is that in many specialised occupations, within occupation turnover is extremely low because of the prevalence of self-employment (like highly specialised consultants). The only turnover observed may then be in and out of that occupation. This would run against what this alternative definition is trying to capture.

irrespective of the scarcity of the occupation. If instead occupations can be spatially separated, there is no obvious reason why unrelated occupations would want to cluster in the largest cities and bear the cost of congestion instead of locating in small cities specialised in the same sector and set of occupations. Put differently, it is not immediately clear why all the scarcest occupations should cluster in the largest cities in simple matching models.

Learning mechanisms provide a more appealing alternative. Imagine that workers are able to occupy more specialised occupations as they learn. When learning is faster in larger cities, we may expect specialised occupations to be over-represented there.²⁷ Hence such learning models deserve further empirical investigation.

Table 10. Estimation results for young workers (below 30 years of age, all sectors)

Scarcity	Class of city size						
	Paris	2000 to 300	300 to 150	150 to 80	80 to 40	40 to 20	20 to 4
Very low	0	0	0	0	0	0	0
Low	0	-0.135 (0.00006)	-0.201 (0.00008)	-0.252 (0.00005)	-0.284 (0.00006)	-0.324 (0.00006)	-0.346 (0.00005)
High	0	-0.141 (0.0001)	-0.289 (0.0001)	-0.316 (0.0001)	-0.332 (0.0001)	-0.371 (0.0001)	-0.394 (0.0001)
Very high	0	-0.186 (0.00008)	-0.293 (0.0001)	-0.339 (0.0001)	-0.435 (0.00008)	-0.455 (0.0001)	-0.445 (0.00009)

3,052 observations. Population in thousands for city sizes. Within each of the 109 sectors, occupations are grouped by quartiles of their distribution by employment size. Standard errors are between brackets. Note that we selected occupation-sector cells with employment above 20 (instead of 30 in the baseline).

The null hypothesis of all coefficients being equal to zero is rejected.

To distinguish between a learning explanation and our model of the division of labour, note that learning models predict that there should be no difference across locations for ‘young’ workers who have not learnt yet. In Table 10 we report estimation results for workers below 30. Although the differences across classes of city size are slightly weaker than in the baseline (the coefficients in Table 10 are about 10% below those in the baseline), there is no indication that learning could account for more than a small fraction of the relative over-representation of scarce occupations in larger cities.

6. Conclusions

Our paper starts from the observation that there is no immediate way to assess empirically how far the division of labour is limited by the extent of the market. Instead, we develop a model-based approach to explore this issue. Our theoretical model builds on the existence of small indivisibilities at the worker level. Its key prediction is that scarce occupations should be over-represented in large cities. Using a large extract of the French census, we test this prediction successfully. We also

²⁷In this line of thought, a promising candidate is the model developed by Matouschek and Robert-Nicoud (2005), which analyses how cities can mitigate the hold-up problem in the acquisition of industry specific skills. In this model, the occupational structure is endogenous. It explains well why workers willing to specialise and acquire a given set of very specific skills want to cluster in the same city where more employers are available. However this model does not explain why more specialised occupations should be scarcer.

verify that it is robust to a number of estimation issues such as measurement errors, our definition of scarcity, alternative spatial units, etc.

We think that the two most important outstanding issues are the following. First, our analysis says nothing about the productivity benefits of a more extensive division of labour. Further work should address this key issue to assess how much of the productivity benefits of larger cities can be attributed to the division of labour as opposed to other sources of agglomeration economies. Second and because of the nature of our data, our analysis pays no attention to firms and the role they play in the division of labour. To gain more insight on the spatial organisation of economic activity we believe further work should pay more attention to firms and how they mediate the division of labour.

References

- Baumgardner, James R. 1988a. Physicians' services and the division of labor across local markets. *Journal of Political Economy* 96(5):948–982.
- Baumgardner, James R. 1988b. The division of labor, local markets, and worker organization. *Journal of Political Economy* 96(3):509–527.
- Becker, Gary S. and Kevin M. Murphy. 1992. The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics* 107(4):1137–1160.
- Dixit, Avinash K. and Joseph E. Stiglitz. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67(3):297–308.
- Duranton, Gilles. 1998. Labor specialization, transport costs, and city size. *Journal of Regional Science* 38(4):553–573.
- Duranton, Gilles and Diego Puga. 2001. From sectoral to functional urban specialisation. Discussion Paper 2971, Centre for Economic Policy Research.
- 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.
- Garicano, Luis and Thomas Hubbard. 2003. Specialization, firms and markets: The division of labour between and within law firms. Discussion Paper 3699, Centre for Economic Policy Research. URL <http://www.cepr.org/>.
- Holmes, Thomas J. 1999. Localisation of industry and vertical disintegration. *Review of Economics and Statistics* 81(2):314–325.
- INSEE and DARES. 2003. *Nomenclature des Professions et Catégories Socioprofessionnelles des emplois salariés d'entreprises PCS-ESE 2003*. Paris: Institut National de la Statistique et des Etudes Economiques.
- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108(3):577–598.
- Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.
- Matouschek, Niko and Frédéric Robert-Nicoud. 2005. The role of human capital investments in the location decision of firms. *Regional Science and Urban Economics* (forthcoming).
- Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes. 1974. *Introduction to the Theory of Statistics*. Third edition. Singapore: McGraw Hill.
- Mori, Tomoya, Koji Nishikimi, and Tony E. Smith. 2003. Relative entropy index of industrial localization: A statistic for large sample analysis. Processed, Kyoto University.
- Rosenthal, Stuart S. and William C. 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.
- Salop, Steven C. 1979. Monopolistic competition with outside goods. *Bell Journal of Economics* 10(1):141–156.

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

Stigler, George J. 1951. The division of labor is limited by the extent of the market. *Journal of Political Economy* 59(3):185–193.

Xiang, Xiaokai and Siang Ng. 1998. Specialization and division of labor: A survey. In Kenneth J. Arrow, Yew-Kwang Ng, and Xiaokai Yang (eds.) *Increasing Returns and Economic Analysis*. Houndmills, Basingstoke: Macmillan.

Appendix A. Proof of Proposition 1

We consider that the size of the market, z is drawn ex-ante from a uniform distribution over $[0, 2Z]$. Following Mood, Graybill, and Boes (1974), a simple Taylor expansion shows that:

$$E\left(\frac{s_k}{n}\right) \approx \frac{E(s_k)}{E(n)} - \frac{\text{cov}(s_k, n)}{E(n)^2} + \frac{E(s_k)}{E(n)^3} \text{var}(n). \quad (\text{A } 1)$$

Note first that, because any real number x is the sum of its floor $\lfloor x \rfloor$ and its remainder (or decimal part) $\{x\}$, equation (4) implies:

$$s_k = \left\lfloor \frac{z}{\sigma_k} \right\rfloor = \frac{z}{\sigma_k} - \left\{ \frac{z}{\sigma_k} \right\}. \quad (\text{A } 2)$$

From equation (6), we can also write:

$$n = Az + \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1 \right) \left\{ \frac{z}{\sigma_j} \right\}, \quad (\text{A } 3)$$

where $A \equiv \left(1 + \sum_{j=1}^K \frac{1}{\sigma_j} \right)$.

Then, note that for any z drawn from a uniform distribution over $[0, 2Z]$, $E\left(\left\{ \frac{z}{\sigma_k} \right\}\right) = \frac{1}{2}$. Consequently equations (A 2) and (A 3) imply that the first term of (A 1) can be re-written as:

$$\frac{E(s_k)}{E(n)} = \frac{\frac{Z}{\sigma_k} - \frac{1}{2}}{AZ + \frac{1}{2} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1 \right)} = \frac{\frac{1}{\sigma_k} - \frac{1}{2Z}}{A + \frac{1}{2Z} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1 \right)}, \quad (\text{A } 4)$$

which is obviously increasing in Z .

The rest of the proof shows that the sum of the last two terms of (A 1) also increases with Z . Starting with the second term of (A 1), simple algebra yields:

$$\begin{aligned} \text{cov}(s_k, n) &= \text{cov} \left(\frac{z}{\sigma_k} - \left\{ \frac{z}{\sigma_k} \right\}, Az + \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1 \right) \left\{ \frac{z}{\sigma_j} \right\} \right) \\ &= \frac{A}{\sigma_k} \text{var}(z) - A \text{cov} \left(\left\{ \frac{z}{\sigma_k} \right\}, z \right) + \frac{1}{\sigma_k} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1 \right) \text{cov} \left(z, \left\{ \frac{z}{\sigma_j} \right\} \right) \\ &\quad - \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1 \right) \text{cov} \left(\left\{ \frac{z}{\sigma_k} \right\}, \left\{ \frac{z}{\sigma_j} \right\} \right). \end{aligned} \quad (\text{A } 5)$$

Then, it can be verified that:

$$\begin{aligned}
cov(z, \{\frac{z}{\sigma_i}\}) &= E\left(z \times \{\frac{z}{\sigma_i}\}\right) - E(z)E\left(\{\frac{z}{\sigma_i}\}\right) \\
&= \int_{x=0}^{2Z} \frac{x}{2Z} \{\frac{x}{\sigma_i}\} dx - \frac{Z}{2} \\
&= \sum_{j=0}^{\lfloor 2Z/\sigma_i - 1 \rfloor} \int_{x=0}^{\sigma_i} \frac{x + \sigma_i j}{2Z} \frac{x}{\sigma_i} dx - \frac{Z}{2} + o\left(\frac{1}{Z}\right) \\
&= \frac{\sigma_i}{12} + o\left(\frac{1}{Z}\right).
\end{aligned} \tag{A 6}$$

Note that the term in $1/Z$ (which can be neglected when Z is large) arises because $2Z/\sigma_i$ is not an integer in general. A similar reasoning implies that $\{z/\sigma_i\}$ and $\{z/\sigma_j\}$ are approximately independent from each other when Z is large and σ_i/σ_j cannot be written as the ratio of two integers. Consequently $cov\left(\{\frac{z}{\sigma_i}\}, \{\frac{z}{\sigma_j}\}\right) \approx 0$. This implies that (A 5) can be re-written as:

$$cov(s_k, n) = \frac{AZ^2}{3\sigma_k} - \frac{A\sigma_k}{12} + \frac{1}{\sigma_k} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) \frac{\sigma_j}{12} + o\left(\frac{1}{Z}\right) \tag{A 7}$$

As for the third term, using the same reasoning as above, note first that:

$$\begin{aligned}
var(n) &= var\left(Az + \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) \{\frac{z}{\sigma_j}\}\right) \\
&= A^2 var(z) + \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)^2 var\left(\{\frac{z}{\sigma_j}\}\right) \\
&\quad + 2 \sum \sum_{i < j} \left(\frac{\sigma_j}{\gamma_j} - 1\right) \left(\frac{\sigma_i}{\gamma_i} - 1\right) cov\left(\{\frac{z}{\sigma_i}\}, \{\frac{z}{\sigma_j}\}\right) + 2A \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) cov\left(z, \{\frac{z}{\sigma_j}\}\right) \\
&= \frac{AZ^2}{3} + \frac{1}{12} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)^2 + \frac{A}{6} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) \sigma_j + o\left(\frac{1}{Z}\right)
\end{aligned} \tag{A 8}$$

Given that the first term of (A 1) increases with Z , it is sufficient to have that the sum of the second and third term increases with Z for $E(s_k/n)$ to increase with Z . From (A 3), (A 4), (A 7), and (A 8), we obtain

$$\begin{aligned}
-\frac{cov(s_k, n)}{E(n)^2} + \frac{E(s_k)}{E(n)^3} var(n) &= \frac{E(s_k)var(n) - E(n)cov(s_k, n)}{E(n)^3} \\
&= \frac{\left(\frac{1}{\sigma_k} - \frac{1}{2Z}\right) \left(\frac{A^2}{3} + \frac{1}{12Z^2} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)^2 + \frac{A}{6Z^2} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) \sigma_j + o\left(\frac{1}{Z^3}\right)\right)}{\left(A + \frac{1}{2Z} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)\right)^3} \\
&\quad - \frac{\left(A + \frac{1}{2Z} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)\right) \left(\frac{A}{3\sigma_k} - \frac{A\sigma_k}{12Z^2} + \frac{1}{\sigma_k Z^2} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) \frac{\sigma_j}{12} + o\left(\frac{1}{Z^3}\right)\right)}{\left(A + \frac{1}{2Z} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)\right)^3} \\
&= \frac{\frac{-A^2}{6Z} - \frac{A}{6\sigma_k Z} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right) + o\left(\frac{1}{Z^2}\right)}{\left(A + \frac{1}{2Z} \sum_{j=1}^K \left(\frac{\sigma_j}{\gamma_j} - 1\right)\right)^3}
\end{aligned} \tag{A 9}$$

Tedious algebra then shows that for Z large enough, (A 9) increases in Z .

Appendix B. Proof of Proposition 3

Following the same approach as in the previous proof:

$$\begin{aligned}
 E\left(\frac{s_k}{\sigma_k}\right) &= \frac{E(s_k)}{E\left(\frac{z}{\sigma_k}\right)} - \frac{\text{cov}\left(s_k, \frac{z}{\sigma_k}\right)}{E\left(\frac{z}{\sigma_k}\right)^2} + \frac{E(s_k)}{E\left(\frac{z}{\sigma_k}\right)^3} \text{var}\left(\frac{z}{\sigma_k}\right) \\
 &= 1 - \frac{\sigma_k}{2Z} - \frac{\text{var}(z)}{Z^2} + \frac{\sigma_k \text{cov}\left(\left\{\frac{z}{\sigma_k}\right\}, z\right)}{Z^2} + \frac{Z - \frac{\sigma_k}{2}}{Z^3} \text{var}(z) \\
 &= 1 - \frac{2\sigma_k}{3Z} + \frac{1}{12Z^2} + o\left(\frac{\sigma_k}{Z^3}\right),
 \end{aligned} \tag{B 1}$$

which is obviously decreasing in σ_k for Z large enough.

Appendix C. Estimation of the model

The log-likelihood

The log-likelihood to be maximised is

$$L = \sum_{j,u,k} n_{u,k}^j \ln \pi_{u,k}^j. \tag{C 1}$$

Note first that

$$\begin{aligned}
 \sum_{\substack{u=1,\dots,U \\ k=1,\dots,K}} \exp(Y_{u,k}^j) &= \sum_{\substack{u=1,\dots,U \\ k=1,\dots,K}} \exp(\alpha_u^j + \beta_k^j + \xi_{m(u),r(j,k)}), \\
 &= \sum_{m,r} \exp(\xi_{m,r}) \sum_{u,m(u)=m} \exp(\alpha_u^j) \sum_{k,r(j,k)=r} \exp(\beta_k^j), \\
 &= \sum_{m,r} \exp(\tilde{\alpha}_m^j + \tilde{\beta}_r^j + \xi_{m,r}),
 \end{aligned} \tag{C 2}$$

with

$$\begin{aligned}
 \tilde{\alpha}_m^j &= \log \left[\sum_{u,m(u)=m} \exp(\alpha_u^j) \right], \\
 \tilde{\beta}_r^j &= \log \left[\sum_{k,r(j,k)=r} \exp(\beta_k^j) \right].
 \end{aligned} \tag{C 3}$$

Note also that $\pi_{u,k}^j$ may be written as

$$\pi_{u,k}^j = \tilde{\pi}_{m(u),r(j,k)}^j \tilde{\pi}_u^j \tilde{\pi}_k^j, \tag{C 4}$$

where

$$\tilde{\pi}_{m,r}^j = \frac{\exp(\tilde{\alpha}_m^j + \tilde{\beta}_r^j + \xi_{m,r})}{\sum_{\substack{i=1,\dots,M \\ l=1,\dots,R}} \exp(\tilde{\alpha}_i^j + \tilde{\beta}_l^j + \xi_{i,l})} \tag{C 5}$$

is within sector j the probability of a worker employed in an occupation of scarcity level r being located in a city of type m ,

$$\tilde{\pi}_u^j = \exp(\alpha_u^j - \tilde{\alpha}_{m(u)}^j) = \frac{\exp(\alpha_u^j)}{\sum_{i,m(i)=m(u)} \exp(\alpha_i^j)} \tag{C 6}$$

is within sector j the probability of a worker being located in city u , conditional on being located in a city of type $m(u)$, and

$$\tilde{\pi}_k^j = \exp(\beta_k^j - \tilde{\beta}_{r(j,k)}^j) = \frac{\exp(\beta_k^j)}{\sum_{i,r(j,i)=r(j,k)} \exp(\beta_i^j)} \quad (\text{C } 7)$$

is within sector j the probability of a worker being employed in occupation k , conditional on having an occupation being of scarcity level $r(j,k)$.

The log-likelihood may then take the following form

$$L = \sum_{j,u,k} n_{u,k}^j \left[\ln \tilde{\pi}_{m(u),r(j,k)}^j + \ln \hat{\pi}_u^j + \ln \check{\pi}_k^j \right] = \tilde{L} + \hat{L} + \check{L}, \quad (\text{C } 8)$$

with

$$\begin{aligned} \tilde{L} &= \sum_{j,u,k} n_{u,k}^j \ln \tilde{\pi}_{m(u),r(j,k)}^j = \sum_{j,m,r} \tilde{n}_{m,r}^j \ln \tilde{\pi}_{m,r}^j, \\ \hat{L} &= \sum_{j,u,k} n_{u,k}^j \ln \hat{\pi}_u^j = \sum_{j,u} n_{u,\cdot}^j \ln \hat{\pi}_u^j, \\ \check{L} &= \sum_{j,u,k} n_{u,k}^j \ln \check{\pi}_k^j = \sum_{j,k} n_{\cdot,k}^j \ln \check{\pi}_k^j, \end{aligned} \quad (\text{C } 9)$$

where

$$\begin{aligned} \tilde{n}_{m,r}^j &= \sum_{\substack{k,r(j,k)=r \\ u,m(u)=m}} n_{u,k}^j \\ n_{u,\cdot}^j &= \sum_{k=1,\dots,K} n_{u,k}^j \\ n_{\cdot,k}^j &= \sum_{u=1,\dots,U} n_{u,k}^j. \end{aligned} \quad (\text{C } 10)$$

To estimate the $\xi_{m,r}$ coefficients, we need to maximise \tilde{L} only.

Identification

We have

$$\tilde{\pi}_{m,r}^j = \frac{\exp(\tilde{Y}_{m,r}^j)}{\sum_{\substack{i=1,\dots,M \\ l=1,\dots,R}} \exp(\tilde{Y}_{i,l}^j)}, \quad (\text{C } 11)$$

with

$$\tilde{Y}_{m,r}^j = \tilde{\alpha}_m^j + \tilde{\beta}_r^j + \xi_{m,r}. \quad (\text{C } 12)$$

Note first note that regardless of a^j for $j = 1 \dots J$, of b_m for $m = 1 \dots M$, and of c_r for $r = 1 \dots R$, we have

$$\tilde{\alpha}_m^j + \tilde{\beta}_r^j + \xi_{m,r} = (\tilde{\alpha}_m^j + a^j + b_m) + (\tilde{\beta}_r^j - a^j + c_r) + (\xi_{m,r} - b_m - c_r) \quad (\text{C } 13)$$

which implies that, once the $\tilde{Y}_{m,r}^j$ are known, the computation of $\tilde{\alpha}_m^j$, $\tilde{\beta}_r^j$ and $\xi_{m,r}$ requires $J + M + R$ identification conditions.

Moreover, as usual with logit models, for every sector j and for every values of the probabilities $(\tilde{\pi}_{m,r}^j)_{m=1,\dots,M,r=1,\dots,R}$, the terms $\tilde{Y}_{m,r}^j$ are identified up to a constant. Then we need J more restrictions. Therefore, we have to choose $2J + M + R$ identification conditions. Our restrictions are the following

$$\begin{aligned} \forall j, \quad \alpha_1^j = \beta_1^j = 0, \\ \forall m, \quad \zeta_{m,1} = 0, \\ \forall r, \quad \xi_{1,r} = 0. \end{aligned} \tag{C 14}$$

Likelihood derivatives

\tilde{L} may also be written as

$$\tilde{L} = \sum_r \tilde{L}_r, \quad \tilde{L}_r \equiv \sum_{m,r} \tilde{n}_{m,r}^j \ln \tilde{\pi}_{m,r}^j. \tag{C 15}$$

Knowing that

$$\frac{\partial \ln \tilde{\pi}_{i,v}^l}{\partial \tilde{\alpha}_m^j} = \delta_{l,j}(\delta_{m,i} - \tilde{\pi}_{m,\cdot}^j), \quad \frac{\partial \ln \tilde{\pi}_{i,v}^l}{\partial \tilde{\beta}_r^j} = \delta_{l,j}(\delta_{r,v} - \tilde{\pi}_{\cdot,r}^j), \quad \frac{\partial \ln \tilde{\pi}_{i,v}^l}{\partial \tilde{\zeta}_{m,r}} = \delta_{m,i}\delta_{r,v} - \tilde{\pi}_{m,r}^l \tag{C 16}$$

where $\delta_{l,j} = 1$ if $l = j$, zero otherwise and

$$\tilde{\pi}_{m,\cdot}^j = \sum_r \tilde{\pi}_{m,r}^j, \quad \tilde{\pi}_{\cdot,r}^j = \sum_m \tilde{\pi}_{m,r}^j, \tag{C 17}$$

the first-order derivatives of the log-likelihood are

$$\begin{aligned} \frac{\partial \tilde{L}}{\partial \tilde{\alpha}_m^j} &= \frac{\partial \tilde{L}_j}{\partial \tilde{\alpha}_m^j} = \sum_{i,v} \tilde{n}_{i,v}^j (\delta_{m,i} - \tilde{\pi}_{m,\cdot}^j) = \tilde{n}_{m,\cdot}^j - \tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{m,\cdot}^j, \\ \frac{\partial \tilde{L}}{\partial \tilde{\beta}_r^j} &= \frac{\partial \tilde{L}_j}{\partial \tilde{\beta}_r^j} = \sum_{i,v} \tilde{n}_{i,v}^j (\delta_{r,v} - \tilde{\pi}_{\cdot,r}^j) = \tilde{n}_{\cdot,r}^j - \tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{\cdot,r}^j, \\ \frac{\partial \tilde{L}}{\partial \tilde{\zeta}_{m,r}} &= \sum_{j,i,v} \tilde{n}_{i,v}^j (\delta_{m,i}\delta_{r,v} - \tilde{\pi}_{m,r}^j) = \sum_j (\tilde{n}_{m,r}^j - \tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{m,r}^j), \end{aligned} \tag{C 18}$$

while the non-zero second-order derivatives are

$$\begin{aligned} \frac{\partial^2 \tilde{L}}{\partial \tilde{\alpha}_m^j \partial \tilde{\alpha}_i^j} &= -\tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{i,\cdot}^j (\delta_{m,i} - \tilde{\pi}_{m,\cdot}^j) & \frac{\partial^2 \tilde{L}}{\partial \tilde{\beta}_r^j \partial \tilde{\beta}_v^j} &= -\tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{\cdot,r}^j (\delta_{r,v} - \tilde{\pi}_{\cdot,r}^j), \\ \frac{\partial^2 \tilde{L}}{\partial \tilde{\zeta}_{m,r} \partial \tilde{\zeta}_{i,v}} &= -\sum_j \tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{m,r}^j (\delta_{m,i}\delta_{r,v} - \tilde{\pi}_{i,v}^j) & \frac{\partial^2 \tilde{L}}{\partial \tilde{\alpha}_m^j \partial \tilde{\beta}_r^j} &= -\tilde{n}_{\cdot,\cdot}^j (\tilde{\pi}_{m,r}^j - \tilde{\pi}_{m,\cdot}^j \tilde{\pi}_{\cdot,r}^j), \\ \frac{\partial^2 \tilde{L}}{\partial \tilde{\zeta}_{m,r} \partial \tilde{\alpha}_i^j} &= -\tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{m,r}^j (\delta_{m,i} - \tilde{\pi}_{\cdot,i}^j) & \frac{\partial^2 \tilde{L}}{\partial \tilde{\zeta}_{m,r} \partial \tilde{\beta}_v^j} &= -\tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{m,r}^j (\delta_{r,v} - \tilde{\pi}_{\cdot,v}^j). \end{aligned} \tag{C 19}$$

Finally, after defining

$$h_{m,i,r,v}^j = -\tilde{n}_{\cdot,\cdot}^j \tilde{\pi}_{m,r}^j (\delta_{m,i}\delta_{r,v} - \tilde{\pi}_{i,v}^j), \tag{C 20}$$

one can easily prove that

$$\begin{aligned} \frac{\partial^2 \tilde{L}}{\partial \tilde{\alpha}_m^j \partial \tilde{\alpha}_i^j} &= \sum_{r,v} h_{m,i,r,v}^j & \frac{\partial^2 \tilde{L}}{\partial \tilde{\beta}_r^j \partial \tilde{\beta}_v^j} &= \sum_{m,i} h_{m,i,r,v}^j & \frac{\partial^2 \tilde{L}}{\partial \tilde{\zeta}_{m,r} \partial \tilde{\zeta}_{i,v}} &= \sum_j h_{m,i,r,v}^j, \\ \frac{\partial^2 \tilde{L}}{\partial \tilde{\alpha}_m^j \partial \tilde{\beta}_v^j} &= \sum_{i,r} h_{m,i,r,v}^j & \frac{\partial^2 \tilde{L}}{\partial \tilde{\zeta}_{m,r} \partial \tilde{\alpha}_i^j} &= \sum_v h_{m,i,r,v}^j & \frac{\partial^2 \tilde{L}}{\partial \tilde{\zeta}_{m,r} \partial \tilde{\beta}_v^j} &= \sum_i h_{m,i,r,v}^j. \end{aligned} \tag{C 21}$$

Coefficients estimation

Taking into account the identification restrictions given above, \tilde{L} must be maximised with respect to $J(M + R - 2) + (M - 1)(R - 1)$ parameters, which may be quite large. However, overall maximisation may be replaced by the following iterative procedure, easier to implement because it is less demanding in terms of computer resources:

1. Start: initialise $\xi_{m,r}$ coefficients at zero, for all $m = 2, \dots, M$ and $r = 2, \dots, R$.
2. Stage 1: For every $j = 1, \dots, J$, maximise \tilde{L}_j with respect to $\tilde{\alpha}_2^j, \dots, \tilde{\alpha}_M^j$ and $\tilde{\beta}_2^j, \dots, \tilde{\beta}_R^j$.
3. Stage 2: Maximise \tilde{L}_r with respect to $\xi_{m,r}$, $m = 2, \dots, M$ and $r = 2, \dots, R$.
4. Conclusion: iterate stages 1 and 2 until convergence.

Estimating the variance-covariance matrix

To compute the test statistics, we need to recover an estimate of the variance-covariance matrix, at least for the parameters of interest, $\xi_{m,r}$, $m = 2, \dots, M$ and $r = 2, \dots, R$ (recall that $\xi_{1,r} = \xi_{m,1} = 0$). For this, we need to recover a partial inverse of the Hessian. Let us introduce the following notations:

$$\begin{aligned}
 H_{\alpha,\alpha}^j &= \left[\sum_{r,v} h_{m,i,r,v}^j \right]_{\substack{m=2,\dots,M \\ i=2,\dots,M}} & H_{\beta,\beta}^j &= \left[\sum_{m,i} h_{m,i,r,v}^j \right]_{\substack{r=2,\dots,R \\ v=2,\dots,R}} \\
 H_{\xi,\xi}^j &= \left[\sum_j h_{m,i,r,v}^j \right]_{\substack{m=2,\dots,M, r=2,\dots,R \\ i=2,\dots,M, v=2,\dots,R}} & H_{\alpha,\beta}^j &= \left[\sum_{i,r} h_{m,i,r,v}^j \right]_{\substack{m=2,\dots,M \\ v=2,\dots,R}} \\
 H_{\alpha,\xi}^j &= \left[\sum_v h_{m,i,r,v}^j \right]_{\substack{m=2,\dots,M, r=2,\dots,R \\ i=2,\dots,M}} & H_{\beta,\xi}^j &= \left[\sum_i h_{m,i,r,v}^j \right]_{\substack{m=2,\dots,M, r=2,\dots,R \\ v=2,\dots,R}}
 \end{aligned} \tag{C 22}$$

and

$$H_{(\alpha,\beta)}^j = \begin{bmatrix} H_{\alpha,\alpha}^j & \left(H_{\alpha,\beta}^j \right)' \\ H_{\alpha,\beta}^j & H_{\beta,\beta}^j \end{bmatrix} \quad H_{(\alpha,\beta),\xi}^j = \left\{ H_{\alpha,\xi}^j \quad H_{\beta,\xi}^j \right\} \tag{C 23}$$

Then, the full Hessian matrix is

$$H = \begin{bmatrix} H_{(\alpha,\beta)}^1 & 0 & 0 & \left(H_{(\alpha,\beta),\xi}^1 \right)' \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & H_{(\alpha,\beta)}^J & \left(H_{(\alpha,\beta),\xi}^J \right)' \\ H_{(\alpha,\beta),\xi}^1 & \dots & H_{(\alpha,\beta),\xi}^J & H_{\xi,\xi} \end{bmatrix} \tag{C 24}$$

As our parameters of interest are the $\xi_{m,r}$, $m = 2, \dots, M$ and $r = 2, \dots, R$, we only need to recover the lower right corner of the inverse matrix of H . Using block-inversion formula, we find

$$V^{\beta,\beta} = \left(H_{\xi,\xi} - \sum_{j=1}^J \left(H_{(\alpha,\beta),\xi}^j \right) \left(H_{\alpha,\alpha}^j \right)^{-1} \left(H_{(\alpha,\beta),\xi}^j \right)' \right)^{-1} \tag{C 25}$$