Nursery Cities:
Urban diversity, process innovation, and the life-cycle of products

Gilles Duranton*‡
London School of Economics

Diego Puga*§
University of Toronto

Centre for Economic Performance Discussion Paper 445
February 2000

ABSTRACT: A simple model of process innovation is proposed, where firms learn about their ideal production process by making prototypes. We build around this a dynamic general equilibrium model, and derive conditions under which diversified and specialised cities coexist. New products are developed in diversified cities, trying processes borrowed from different activities. On finding their ideal process, firms switch to mass-production and relocate to specialised cities with lower costs. When in equilibrium, this configuration welfare-dominates those with only diversified or only specialised cities. We find strong evidence of this relocation pattern in establishment relocations across French employment areas 1993–1996.

Key words: cities, diversity, specialisation, innovation, learning, life-cycle.
JEL classification: R30, O31, D83.

*We are particularly grateful to Frédéric Lainé and Pierre Philippe Combes for kindly providing us with the data on, respectively, plant relocation and sectoral composition of employment for French employment areas; also to Vernon Henderson, Martin Osborne, and Dan Trefler, for very helpful discussions. We have also benefitted from comments and suggestions by Richard Baldwin, Jim Brander, Rikard Forslid, Elhanan Helpman, Maureen Kilkenny, Tomoya Mori, Nate Rosenberg, Barbara Spencer, Manuel Trajtenberg, as well as other participants at various seminars.

‡Also affiliated with the Centre for Economic Performance at the London School of Economics, and the Centre for Economic Policy Research.

§Also affiliated with the Canadian Institute for Advanced Research, the Centre for Economic Performance at the London School of Economics, the Centre for Economic Policy Research, and the Norges Handelshøyskole.
Funding from the Social Sciences and Humanities Research Council of Canada, and from the Connaught Fund of the University of Toronto is gratefully acknowledged.

Correspondence addresses:

Gilles Duranton
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

Diego Puga
Department of Economics
University of Toronto
150 St. George Street
Toronto, Ontario M5S 3G7
Canada
d.puga@utoronto.ca
http://dpuga.economics.utoronto.ca

1. Introduction

What makes a city an attractive place in which to produce? Is it a diversity of industries and people? Or is it the presence of many other producers engaged in similar activities and many other people with similar skills? There is a long-standing interest in this question (see Duranton and Puga, 2000, for a survey).

The study of the advantages of urban specialisation, often called 'localisation economies', is commonly traced back to Marshall (1890). Localisation economies have been formalised in a manner very close to the spirit of Marshall's work, based either on firms' ability to produce more efficiently when using a wider range of intermediates (Fujita, 1988; Abdel-Rahman and Fujita, 1990), on the benefits of specialisation for labour market matching (Helsley and Strange, 1990), or on the presence information spill-overs (Fujita and Ogawa, 1982). They have also been the subject of detailed empirical work (for example, Dumais, Ellison, and Glaeser, 1997, and Henderson, 1999).

The advantages of urban diversity, often called 'urbanisation economies', are frequently linked to the work of Jacobs (1969). She describes several examples where diversity facilitated innovation through the borrowing of processes from other sectors (such as 3M's development of masking tape, which took advantage of improvements in adhesives from the sandpaper industry). More recently, several empirical studies have found that diversity fosters growth in cities (Glaeser, Kallal, Scheinkman, and Schleifer, 1992), or at least in their most innovative sectors (Henderson, Kuncoro, and Turner, 1995). Looking more directly at innovation, Feldman and Audretsch (1999), Harrison, Kelley, and Gant (1996), and Kelley and Helper (1999) show that diversity fosters innovation in cities, while narrow specialisation hinders it. Surprisingly, however, there is no appropriate theoretical framework deriving urbanisation economies à la Jacobs starting from microeconomic foundations. And, in consequence, there is no framework in which to compare the relative advantages and disadvantages of urban diversity and specialisation, and to study their role in shaping urban systems. This paper develops such a framework.

This exercise is important for three main reasons. First, to understand urban patterns, it is essential to consider both localisation and urbanisation economies. If only localisation economies exist, and congestion costs increase with city size, all cities are fully specialised (as in Henderson, 1987, and Becker and Henderson, 2000).[1] Yet urban systems are characterised by the coexistence of diversified and specialised cities.

Second, the answer to the question of whether diversity or specialisation makes a location more attractive is unlikely to be universal. In some circumstances diversity is more important, while in others specialisation matters more. Modelling localisation and urbanisation economies starting from micro-foundations can help us understand the circumstances under which each characteristic is more important. The model we develop suggests that diversity and specialisation matter more in different periods of the life-cycle of products. This finding has implications for understanding not just urban systems, but also process innovation, and firm location and relocation patterns.

[1] A few papers introduce advantages from diversity in this kind of framework. However, they either assume those advantages, or give them static foundations that are quite distinct from the dynamic role assigned by Jacobs to diversity in fostering innovation (see Abdel-Rahman, 2000, Duranton and Puga, 2000, and Quigley, 1998, for detailed references and discussion).

Third, there is a large empirical literature on diversity and specialisation in cities that reaches seemingly contradictory conclusions. Papers examining the evolution of urban employment patterns (Glaeser *et al.*, 1992; Henderson *et al.*, 1995; Combes, 2000) find marked advantages of diversity in fostering urban employment growth and in attracting newer and more innovative activities. By contrast, papers looking at the evolution of urban productivity levels (Henderson, 1999) find weak effects of diversity and very strong effects of specialisation on productivity. A theoretical framework with both endogenous location decisions and productivity changes can help reconcile those conclusions.

Our model builds on two standard static ingredients. First, due to localisation economies, the cost of using a given production process diminishes as more local firms use the same type of process. Second, urban crowding places a limit on city size. This combination of localisation economies and congestion costs creates *static advantages to urban specialisation*.

The main novelty of the model is the simple model of process innovation that we develop and combine with those two more traditional ingredients.[2] We start from the assumption that a firm making a new product does not know how best to produce it. Nevertheless, it can make prototypes with any one of the types of production process already used locally (a firm does not want to be the only one using a given type of process locally because the absence of localisation economies makes this prohibitively costly). Once a firm produces a prototype with its ideal production process, it recognises it as such and is then able to begin mass-production. The combination of this learning process that draws from local types of production processes with costly firm relocation creates *dynamic advantages to urban diversity*.

We also incorporate firm turnover, by having some firms randomly close down each period. Optimal investment then ensures they are replaced by new firms producing new products. Finally, migration ensures that workers in all cities are equally well-off.

In our setting, three possible types of steady-states can occur: those with only diversified cities, those with only specialised cities, and mixed configurations with both diversified and specialised cities. We begin by studying the mixed configuration (Section 3). When both diversified and specialised cities coexist, it is because each firm finds it in its best interest to locate in a diversified city while searching for its ideal process, and later to relocate to a specialised city where all firms are using the same type of process.

Location in a diversified city during a firm's learning stage can be seen as an investment. It is costly because all firms impose congestion costs on each other, but only those using the same type of process create cost-reducing localisation economies. This results in comparatively higher production costs in diversified than in specialised cities. However, bearing these higher costs can be worthwhile for firms in search of their ideal process because they expect to have to try a variety of processes before finding their ideal one, and a diversified city allows them to do so without costly relocation after each trial. In this sense, diversified cities act as a 'nursery' for firms. Once a

[2] A significant literature addresses firms' learning about their technology (see, in particular, Jovanovic, 1982, and Jovanovic and MacDonald, 1994). However previous modelling approaches cannot be easily embedded in a general equilibrium model of a system of cities. Furthermore, they focus on firms learning in isolation or, in the strategic learning literature, on interactions based on imitations. The focus of this paper is instead on how the urban environment affects learning, and how best firms can choose their environment.

firm finds its ideal production process, it no longer benefits from being in a diverse environment. At this stage, if relocation is not too costly, the firm avoids the congestion imposed by the presence of other sectors by relocating to a city where all other firms share its specialisation.

We also study alternative steady-states with only diversified or only specialised cities (Section 4). Then we turn to stability issues, derive optimal city size, and welfare-rank the steady-states (Section 5). Or main result regarding welfare is that, whenever a steady-state with only diversified and/or a steady-state with only specialised cities exist for the same parameter values as a nursery steady-state with both diversified and specialised cities, the nursery steady-state provides a higher level of welfare. When a nursery configuration is a steady-state, everyone prefers to have both diversified and specialised cities. However, if there happen to be only diversified or only specialised cities, localisation economies prevent a mixed configuration from arising in the absence of coordination. Once we introduce some mechanism for city creation, such as perfectly competitive land developers, if a nursery configuration with cities of optimal size is a steady-state, it becomes the unique equilibrium configuration.

Turning to empirical issues, when both specialised and diversified cities co-exist in steady-state, the model predicts that cities are stable in their size and sectoral composition, but there is a constant turnover of firms. Some existing firms close down every period, and new firms enter to replace them with new products. This turnover follows a life-cycle pattern. New products are created and developed in diversified cities, but production eventually relocates to specialised cities. All of this is consistent with the existing evidence (reviewed in Section 6). However, while some of the relevant stylised facts are well established, the evidence on firm relocation patterns is very indirect. For this reason, we make use of a new data set that records the origin, destination, and sector of all establishment relocations across French employment areas, for the period 1993–1996. We find that indeed most relocating establishments move from particularly diverse cities to cities with a particularly strong specialisation in the relevant sector. Also, as suggested by the model, sectors that are both more innovative and more geographically concentrated relocate more and follow this diversified-to-specialised relocation pattern more often.

2. The model

There are N cities[3] and a continuum L of infinitely lived workers in the economy. Each worker has one of m possible discrete aptitudes. Let us index worker aptitudes by superscript j and cities by subscript i so that L_i^j denotes the gross amount of labour with aptitude j in city i. There are equal proportions of workers with each aptitude in the economy, but their distribution across cities is endogenous. We make assumptions (detailed in Section 3) to ensure that migration equalises utility.

Time is discrete and indexed by t (but to make notation less cumbersome, we only index variables by time when adding over different time periods). Each worker supplies one unit of labour in each period. There are congestion costs in each city, which result in the loss of a fraction of working time proportional to city size. The amount of labour net of congestion costs with aptitude

[3] N is assumed to be a continuous variable, but for simplicity we shall refer to it loosely as the 'number' of cities. For the moment, we shall take N as exogenously given. Later, in Section 5, we determine it endogenously.

j in city i is

$$l_i^j = L_i^j \left(1 - \tau \sum_{j=1}^{m} L_i^j \right) \ , \ \ \tau > 0 \ . \tag{1}$$

This expression corresponds to a situation in which workers live spread along linear cities in land plots of unit length, work at the city centre, and lose in commuting a fraction of their labour equal to 2τ times the distance travelled (one half of each city's population lives on each side of the centre). When workers receive the same wage, they are equally well off regardless of their location within the city: differences in land rents offset differences in commuting costs. Further assuming common ownership of the land, the expected wage income of a worker with aptitude j in city i is $(1 - \tau\sum_{j=1}^{m} L_i^j)w_i^j$, where w_i^j denotes the wage per unit of net labour of aptitude j in city i (see Fujita, 1989, for details and several generalisations).

Setting up a firm involves a one-off start-up cost, which enables firms to start making trial products, referred to as prototypes. Perfectly competitive and frictionless capital markets provide firms with finance for their start-up cost and remunerate workers' savings. Firms can eventually engage in mass-production, with lower production costs, but this involves using a certain 'ideal' production process. This ideal process is randomly drawn when the firm is created from m possible discrete processes, each with equal probability. There is a one-to-one mapping between each firm's possible production processes and workers' aptitudes, so that each of the m possible processes for each firm requires workers of a different aptitude. We say that two production processes for different firms are of the same type if they require workers with the same aptitude.

A newly created firm does not know its ideal production process, but can find it by trying different processes in the production of prototypes. After producing a prototype with a certain process, the firm knows whether this process is its ideal one or not. Thus, in order to switch from prototype to mass-production a firm needs to have produced a prototype with its ideal process first, or to have tried all of its m possible processes except one. However, the precise time at which a firms quits trying different processes is the result of an optimal stopping rule. Thus, we allow for the possibility that a firm decides to stop searching before learning its ideal process.

Firms have an exogenous probability δ of closing down each period (we can think of this as being due to the death of a shadow entrepreneur). Firms also lose a period of production whenever they relocate from one city to another. Thus, the cost of firm relocation increases with the exogenous probability of closure δ, a higher value of which makes firms discount future profits more relative to the profits foregone in the period lost in relocation.

We index varieties of goods (both prototypes and mass-produced goods) by h. We also distinguish variables corresponding to prototypes from those corresponding to mass-produced goods by an accent in the form of a question mark, ? (firms that can only produce prototypes are still wondering about their ideal production process).

Technology in prototypes is summarised by the following cost function:

$$\overset{?}{C}{}_i^j(h) = Q_i^j \overset{?}{x}{}_i^j(h) \ , \tag{2}$$

$$\text{where } \ Q_i^j = \left(l_i^j\right)^{-\epsilon} w_i^j \ , \ \ \epsilon > 0 \ . \tag{3}$$

Output of prototype h, made with a process of type j, in city i is denoted by $\overset{?}{x}^j_i(h)$. Q^j_i is the unit cost for firms producing prototypes using a production process of type j in city i. Note that Q^j_i decreases as l^j_i increases: there are localisation economies that reduce unit costs when there is a larger amount of labour, net of congestion costs, with the relevant aptitude in the same city (which also implies more firms using the same type of process in the same city). The Appendix derives this cost function from first principles, and shows that Q^j_i is the appropriate price index of a monopolistically competitive intermediate sector à la Ethier (1982). As in Fujita (1988) and Abdel-Rahman and Fujita (1990), each such intermediate sector hires workers of aptitude j and sells process-specific non-tradable services to final-good firms using a process of type j. These differentiated services enter the production function of final-good producers with the same constant elasticity of substitution $\frac{\epsilon+1}{\epsilon}$.

When a firm finds its ideal production process, it can engage in mass-production at a fraction ρ $(0 < \rho < 1)$ of the cost of producing a prototype. Thus the cost function for a firm engaged in mass-production is

$$C^j_i(h) = \rho\, Q^j_i x^j_i(h)\ , \tag{4}$$

where $x^j_i(h)$ denotes the output of mass-produced good h, made with a process of type j, in city i.

Turning to consumers, we assume that they have a zero rate of time preference.[4] Each period consumers allocate a fraction μ of their expenditure to prototypes and a fraction $1-\mu$ to mass-produced goods.[5] The instantaneous indirect utility of a consumer in city i is

$$V_i = \overset{?}{P}^{-\mu} P^{-(1-\mu)} e^j_i\ , \tag{5}$$

where e_i denotes individual expenditure,

$$\overset{?}{P} = \left\{ \sum_{j=1}^{m} \iint \left[\overset{?}{p}^j_i(h)\right]^{1-\sigma} \mathrm{d}h\, \mathrm{d}i \right\}^{1/(1-\sigma)}\ , \tag{6}$$

$$P = \left\{ \sum_{j=1}^{m} \iint \left[p^j_i(h)\right]^{1-\sigma} \mathrm{d}h\, \mathrm{d}i \right\}^{1/(1-\sigma)} \tag{7}$$

are the appropriate price indices of prototypes and mass-produced goods respectively, and $\overset{?}{p}^j_i(h)$ and $p^j_i(h)$ denote the prices of individual varieties of prototypes and mass-produced goods respectively. Double integration over h and i and summation over j include in the price indices all varieties produced with any type of process in any city. These price indices are equal in all cities because all final goods, whether prototypes or mass-produced, are freely tradable across cities. All

[4]Note that there is no form of accumulation in this model. All of consumers'savings are invested in financing firms' start-up costs. Each firms' expected profit stream must be sufficient to recover its start-up cost, and this limits investment at every period. Given that, when calculating this expected profit stream, single-period profits are already discounted by the probability that a firm closes down at any period, δ, introducing an additional discount rate through intertemporal consumer preferences would only obscure expressions without changing the nature of our results.

[5]It is common practice for Japanese electronics firms to sell prototypes of their goods to consumers before producing them at mass scale. However, these prototypes account only for a small fraction of sales. Similarly, some consumers are willing to purchase a β-version of Microsoft's latest operating system for US$60, whereas others are more than happy to wait until after it is released. This separation between the prototype and mass-production markets greatly helps us to obtain closed-form solutions.

prototypes enter consumer preferences with the same elasticity of substitution σ (> 2), and so do all mass-produced goods.[6]

Total income, Y, is the sum of expenditure, E, and investment:

$$Y = E + \overset{?}{P}^{\mu} P^{1-\mu} F \, \mathring{n} \; . \tag{8}$$

Investment, $\overset{?}{P}^{\mu} P^{1-\mu} F \, \mathring{n}$, comes from the aggregation of the start-up costs incurred by newly created firms (this cost is incurred only once: firms pay it when they are first created and never again — not even if they relocate their establishment, which simply involves the loss of one period of production). To come up with a new product (but not with the ideal way to produce it) firms must spend F on market research, purchasing the same combination of goods bought by the representative consumer (hence the presence of the price indices in this expression). $\mathring{n}$ denotes the total 'number'(mass) of new firms. Total expenditure is

$$E = \sum_{j=1}^{m} \int L_i^j e_i^j \, \mathrm{d}i = \sum_{j=1}^{m} \int l_i^j w_i^j \, \mathrm{d}i \; + \; \sum_{j=1}^{m} \iint \overset{?}{\pi}_i^j(h) \, \mathrm{d}h \, \mathrm{d}i \; + \; \sum_{j=1}^{m} \iint \pi_i^j(h) \, \mathrm{d}h \, \mathrm{d}i \; - \; \overset{?}{P}^{\mu} P^{1-\mu} F \, \mathring{n} \; . \tag{9}$$

The first term on the right-hand-side of (9) is total wage income, the second and third terms are total income from firm profits (where $\overset{?}{\pi}_i^j(h)$ and $\pi_i^j(h)$ denote the operational profits of each prototype and mass-producer respectively). Subtracting total investment from the sum of these incomes yields expenditure.

As a first step, it is insightful to take as given the number of new firms using a process of each type in each city (which determines the period allocation of income between consumption and savings) and also urban structure. By urban structure we mean the number of cities, the number of prototype and mass-producers using a process of each type in each city (denoted by $\overset{?}{n}_i^j$ and n_i^j respectively), and the number of workers of each aptitude in each city. This allows us to solve for a short-run static equilibrium in which consumers allocate expenditure optimally, firms maximise operational profits, and markets clear.

Lemma 1 (Output per worker) *In equilibrium, output per worker by firms using processes of type j in city i in a given period is*

$$\frac{\overset{?}{n}_i^j \overset{?}{x}_i^j + \rho \, n_i^j x_i^j}{L_i^j} = \left(L_i^j\right)^{\epsilon} \left(1 - \tau \sum_{j=1}^{m} L_i^j\right)^{\epsilon+1} .$$

Proof Total demand for each variety is the sum of consumer demand, obtained by application of Roy's identity to (5) and integration over all consumers, and demand by newly created firms, obtained by application of Shephard's lemma to their one-off start-up cost, $\overset{?}{P}^{\mu} P^{1-\mu} F$, and multiplication by the number of new firms. The product market clearing conditions for, respectively, prototypes and mass-produced goods made with a process of type j in city i are

$$\overset{?}{x}_i^j = \mu (\overset{?}{p}_i^j)^{-\sigma} \overset{?}{P}^{\sigma-1} Y \; , \tag{10}$$

$$x_i^j = (1-\mu)(p_i^j)^{-\sigma} P^{\sigma-1} Y \; . \tag{11}$$

[6]If $\sigma \leqslant 2$, the attractiveness of firm entry increases with the number of firms (see the Proof of Proposition 5 in the Appendix for details).

Note that we have dropped index h, since short-run equilibrium values may vary by city and type of process/aptitude, but do not vary by variety. Using (2) and (4), single-period operational profits can be written as

$$\overset{?}{\pi}{}_i^j = \left(\overset{?}{p}{}_i^j - Q_i^j\right) \overset{?}{x}{}_i^j\,, \tag{12}$$

$$\pi_i^j = \left(p_i^j - \rho\, Q_i^j\right) x_i^j\,. \tag{13}$$

Maximising (12) and (13) with respect to prices, and using (10) and (11), gives the profit-maximising prices for each prototype and for each mass-produced good firm. They are fixed relative markups over marginal costs:

$$\overset{?}{p}{}_i^j = \frac{\sigma}{\sigma-1}\, Q_i^j\,, \tag{14}$$

$$p_i^j = \rho\,\frac{\sigma}{\sigma-1}\, Q_i^j\,. \tag{15}$$

Substituting (10), (11), (14), and (15) into (12) and (13) yields maximised operational profits for prototype and mass-produced good firms:

$$\overset{?}{\pi}{}_i^j = \mu\,\frac{1}{\sigma}\left[\frac{\sigma-1}{\sigma}\,\frac{\overset{?}{P}}{Q_i^j}\right]^{\sigma-1} Y\,, \tag{16}$$

$$\pi_i^j = (1-\mu)\,\frac{1}{\sigma}\left[\frac{1}{\rho}\,\frac{\sigma-1}{\sigma}\,\frac{P}{Q_i^j}\right]^{\sigma-1} Y\,. \tag{17}$$

Demand for labour can be obtained by application of Shephard's lemma to (2)–(4) and integration over varieties. The labour market clearing condition for workers with aptitude j in city i is then

$$l_i^j = \overset{?}{n}{}_i^j\,\frac{\partial\,\overset{?}{C}{}_i^j}{\partial w_i^j} + n_i^j\,\frac{\partial\,C_i^j}{\partial w_i^j} = \left(l_i^j\right)^{-\epsilon}\left(\overset{?}{n}{}_i^j\overset{?}{x}{}_i^j + \rho\, n_i^j x_i^j\right)\,. \tag{18}$$

Substituting (1) into (18), rearranging, and dividing by L_i^j yields the result. □

Corollary 1 *With no need to learn and no market for prototypes ($\overset{?}{n}{}_i^j = 0$), output per worker is maximised when all workers in each city have the same aptitude (hence all firms use the same type of process) and $L_i = \frac{\epsilon}{(2\epsilon+1)\tau}$.*

In the absence of a learning stage, the optimal urban configuration involves only fully specialised cities, as in most models of systems of cities. This is also the equilibrium configuration if there is some mechanism for city creation, such as land developers (see, for instance, Henderson, 1987, and Becker and Henderson, 2000). Learning changes this, by creating dynamic advantages to diversity: diversity allows learning firms to produce a sequence of prototypes with different processes without costly relocations. This is a crucial innovation of this model, and will fundamentally affect the equilibrium urban system, by providing a motivation for the coexistence of diversified cities with specialised ones, and for production to change location over the life-cycle.

3. Mixed steady-states: Nursery cities

A steady-state equilibrium in this model is a configuration such that all of the following are true. Each consumer/worker allocates her income between consumption and savings, allocates her expenditure across goods, and takes her migration decisions so as to maximise expected utility. Each firm chooses a location/production strategy and prices so as to maximise its expected lifetime profits, given the location/production strategies and prices of all other firms. All profit opportunities are exploited, and the urban structure ($\mathring{n}^j_i$, n^j_i, L^j_i, and $\mathring{n}$) is constant over time.

To the keep matters simple, we make assumptions about migration to ensure that we have at most the following two kinds of cities.

Definition 1 (Specialised city) *A city is said to be (fully) specialised if all its workers have the same aptitude, so that all local firms use the same type of production process.*

Definition 2 (Diversified city) *A city is said to be (fully) diversified if it has the same proportion of workers with each of the m aptitudes, so that there are equal proportions of firms using each of the m types of production process.*

We assume that each worker can migrate only every once in a while.[7] Otherwise, all workers could relocate simultaneously and there would be nothing by which to identify a city. This once-in-a-while possibility of relocation is enough to ensure all workers are equally well-off in equilibrium. Workers know the population in each city. However, they have imperfect information about the distribution of each city's workforce across aptitudes. Specifically, they know only whether the largest group of workers with a common aptitude in each city is above some threshold (sufficiently larger than $\frac{1}{m}$), and if so which is this dominant specialisation. Further, workers form their expectations about wages in each city as if cities with a dominant specialisation were fully specialised, and as if cities with no dominant specialisation were fully diversified. With all workers forming their expectations in this way, their expectations turn out to be rational. A city with a dominant specialisation attracts only workers with the dominant aptitude, and so in steady-state is fully specialised. A city with no dominant specialisation seems equally attractive for workers of all aptitudes, and so in steady-state is fully diversified.

These assumptions about migration simplify matters considerably by reducing the set of possible steady-states to three types of configurations, all of them with a large degree of symmetry: those with only diversified cities, those with only specialised cities, and mixed configurations with both diversified and specialised cities. We start by looking at the mixed configurations in this section, and leave for the next section configurations with only diversified cities or only specialised cities.

In all cases, the procedure we follow to characterise the steady-state is the same. We start by specifying location/production strategies for all firms supporting the candidate steady-state.[8] For

[7]This includes migrations within each city, so as to avoid issues related to endogenous neighbourhood formation which are not the focus of this paper.

[8]We make explicit below all details about firms' strategies except for the order in which each firm tries different processes. The only concern here is that, for the sake of symmetry, in steady-state there must be the same proportion of firms trying each type of process in each diversified city at any period. Let us suppose that each firm chooses its order randomly. Since there is a continuum of firms, by the law of large numbers, that symmetry will be attained.

simplicity, we restrict ourselves to symmetric equilibria. Then we establish conditions under which no firm finds it profitable to deviate, so that the candidate steady-state is in fact a steady-state.

Definition 3 (Nursery configuration) *A nursery configuration is one that satisfies all of the following. Diversified and specialised cities coexist. There is the same proportion of cities specialised in each type of process. Each new firm locates in a diversified city and produces prototypes using a different type of production process each period. As soon as a firm finds its ideal production process, and only then, it relocates to a city specialised in that particular type of process, and commences mass-production.*

The remainder of this section studies when a nursery configuration is a steady-state. Let us start by simplifying notation, and replace subindex i with subindex D for diversified city variables and with subindex S for specialised city variables. Denote by N_D the number of diversified cities, and by $\overset{?}{n}_D$ the number of prototype producers using each of the m types of production processes in each diversified city. Since in a nursery configuration firms relocate to a city of the relevant specialisation as soon as they find their ideal process, there are no mass-producers in diversified cities. Thus, the total number of firms in each diversified city is $m\,\overset{?}{n}_D$. Denote by N_S the number of cities specialised in each of the m types of production processes (thus, the total number of specialised cities is mN_S), and by n_S the number of mass-producers (and since there are no prototype producers in specialised cities, also the number of firms) in each. This notation implicitly assumes that all diversified cities are identical, and that all specialised cities are identical except in their specialisation. In Section 5 we show these are necessary conditions for a steady-state to be stable with respect to small perturbations in the distribution of workers.[9]

The conditions for a nursery configuration to be a steady-state depend on three elements: relative production costs in diversified and specialised cities, the relative number of prototype to mass-producers, and the expected duration of the prototype and mass-production stages. Let us derive these as a function of parameters in the following three lemmas.

Lemma 2 (Relative costs) *Unit production costs in diversified cities relative to those specialised cities in a nursery configuration are*

$$\frac{Q_D}{Q_S} = \left(\frac{L_S}{L_D/m}\right)^{\epsilon}\left(\frac{1-\tau L_S}{1-\tau L_D}\right)^{\epsilon+1} = \left(\frac{1-\mu}{\mu}\,\frac{N_D}{N_S}\right)^{\epsilon}\left(\frac{1-\tau\frac{(1-\mu)L}{mN_S}}{1-\tau\frac{\mu L}{N_D}}\right)^{\epsilon+1}.$$

Proof By Definition 3, in a nursery configuration there are L_D workers in each diversified city earning a wage w_D, and L_S workers in each specialised city earning a wage w_S. Hence the unit production costs of (3) become

$$Q_D = [(L_D/m)\,(1-\tau L_D)]^{-\epsilon} w_D \qquad (19)$$

[9]Note that there are other possible steady-states in which diversified and specialised cities coexist (in particular, configurations in which there are different numbers of cities with each specialisations). However, these are not robust with respect to small perturbations in the distribution of workers and the introduction of land developers, which we undertake in Section 5. Furthermore, they retain the main characteristics of the nursery configuration. The same comment applies to the configuration with only specialised cities studied in Section 4.

in diversified cities, and

$$Q_S = \left[L_S\,(1-\tau L_S)\right]^{-\epsilon} w_S \tag{20}$$

in specialised cities. Since each period some workers have the opportunity to move, and only wage income depends on location, in steady-state wages net of commuting costs must be equal across all locations:

$$(1-\tau L_D)\,w_D = (1-\tau L_S)\,w_S \;. \tag{21}$$

There are $m\,\overset{?}{n}_D$ prototype producers and no mass-producers in each of the N_D diversified cities, and there are n_S mass-producers and no prototype producers in each of the mN_S specialised cities. Hence, using (14) and (15), the price indices of (6) and (7) become

$$\overset{?}{P} = \frac{\sigma}{\sigma-1}\left(N_D m\overset{?}{n}_D\right)^{1/(1-\sigma)} Q_D \;, \tag{22}$$

$$P = \rho\,\frac{\sigma}{\sigma-1}\,(mN_S n_S)^{1/(1-\sigma)} Q_S \;. \tag{23}$$

Substituting (22) into (16) and valuing this in a diversified city, and substituting (23) into (17) and valuing this in a specialised city, yields operational profits in a nursery configuration for, respectively, prototype and mass-producers:

$$\overset{?}{\pi}_D = \frac{\mu Y}{\sigma\left(N_D m\overset{?}{n}_D\right)} \;, \tag{24}$$

$$\pi_S = \frac{(1-\mu)Y}{\sigma\,(mN_S n_S)} \;. \tag{25}$$

From (24), we see that total operational profits for all prototype producers are a fraction $\frac{1}{\sigma}$ of expenditure on prototypes, which in turn is a share μ of income. Similarly, from (25), total operational profits for mass-producers are a fraction $\frac{1}{\sigma}$ of expenditure on mass-produced goods, which in turn is a share $1-\mu$ of income. Since operational profits and costs must add up to expenditure, it follows that costs for all prototype producers (equal to the sum of the wage bill in all diversified cities) are a fraction $\frac{\sigma-1}{\sigma}$ of expenditure on prototypes. By the same reasoning, total costs for all mass-producers (equal to the sum of the wage bill in all specialised cities) are a fraction $\frac{\sigma-1}{\sigma}$ of expenditure on mass-produced goods. The wage bill in all diversified cities relative to the wage bill in all specialised cities is therefore $\frac{\mu}{1-\mu}$. With wages net of commuting costs equalised across cities, the relative populations of cities are equal to the relative wage bills. It follows that, in a nursery configuration, diversified cities account for a share μ of total population, L, and specialised cities for a fraction $1-\mu$:

$$N_D L_D = \mu L \;, \quad mN_S L_S = (1-\mu)L \;. \tag{26}$$

Dividing (19) by (20), and using (21) and (26) yields the result. □

An increase in the size of each sector present in each city has a cost-reducing effect, by strengthening localisation economies, but also a cost-increasing effect, by increasing city size and worsening congestion which raises labour costs. In specialised cities, all firms use the same type of production process and contribute to both effects. In diversified cities, however, only firms using the same type of process at the same time (a fraction $\frac{1}{m}$ of the total) contribute to localisation economies, while all firms impose on each other congestion costs. Thus, a diversified city is a more costly place to produce than a specialised city of the same size.

Lemma 3 (Relative number of firms) *The ratio of the total number of prototype producers to the total number of mass-producers in a nursery configuration is*

$$\Omega \equiv \frac{N_D \overset{?}{n}_D}{N_S n_S} = \frac{\delta(m+1) - 1 + (1-\delta)^{m-1}(1-2\delta)}{(1-\delta)^2\left[1-(1-\delta)^{m-2}(1-2\delta)\right]} \,.$$

Proof See the Appendix. □

Ω can be seen as a measure of how unlikely a firm is to find its ideal production process. Since firms can engage in mass-production only once they learn about their ideal process, whenever firms are unlikely to find their ideal process, the number of prototype producers is large relative to the number of mass-producers. Ω is a function of only two parameters, the number of types of production process, m, and the probability of a firm closing down at any period, δ. The larger either of these two parameters, the less likely that a firm will see itself through to the mass-production stage ($\frac{\partial\Omega}{\partial m} > 0$, $\frac{\partial\Omega}{\partial\delta} > 0$). Intuitively, if there are many possibilities for a firm's ideal production process and the closure rate is high, there is a large chance that a firm will close down before it can find its ideal process.

Lemma 4 (Expected duration of each stage) *A firm that follows the nursery configuration strategy expects to spend*

$$\overset{?}{\Delta} = \frac{\delta(m+1) - 1 + (1-\delta)^{m-1}(1-2\delta)}{m\delta^2}$$

periods producing prototypes in a diversified city, and

$$\Delta = \frac{(1-\delta)^2 - (1-\delta)^m(1-2\delta)}{m\delta^2}$$

periods engaged in mass-production in a city where all workers have the aptitude that corresponds to its ideal process. A firm that instead locates first in a specialised city, relocates across specialised cities to try different production processes, and on finding its ideal one fixes its location, expects to spend

$$\overset{?}{\Delta}_{OSC} = \frac{(1+m)(2-\delta)\delta - 1 + (1-\delta)^{2(m-1)}\left[1-2(2-\delta)\delta\right]}{m(2-\delta)^2\delta^2}$$

periods producing prototypes in different specialised cities, and

$$\Delta_{OSC} = \frac{(1-\delta) + (1-\delta)^{2(m-1)}\left[(3-\delta)\delta - 1\right]}{m(2-\delta)\delta^2}$$

periods engaged in mass-production in a city where all workers have the aptitude that corresponds to its ideal process.

Proof See the Appendix. □

To close the model we need to solve the general equilibrium level of investment, which yields the number of new firms created each period. It is particularly convenient to use Tobin's q approach (Tobin, 1969), in a way close to that in which it has recently been applied to endogenous growth models with monopolistic competition (see, in particular Baldwin and Forslid, 2000). Tobin's q is the ratio of the value of one unit of capital to its replacement cost. In steady-state, the general equilibrium level of investment is that for which $q = 1$. The asset value of a new firm is equal to its expected stream of operational profits. The cost of its replacement is the start-up cost. In this context, Tobin's $q = 1$ condition is therefore equivalent to a condition of zero expected net profits for firms:

$$q = \frac{\overset{?}{\Delta}\overset{?}{\pi}_D + \Delta\pi_S}{\overset{?}{P}^{\mu} P^{1-\mu} F} = 1 \; . \tag{27}$$

From (8), (9), and (22)–(27) we can calculate the number of firms using each type of process in each diversified city, $\overset{?}{n}_D$, and the number of firms in each specialised city, n_S. In steady-state, investment therefore keeps this distribution of firms constant by supporting the creation each period of just enough firms to replace those that have closed down.[10] We now have everything we need to derive necessary and sufficient conditions for the existence of a nursery steady-state.

Proposition 1 (Nursery steady-state) *A nursery configuration with N_S specialised cities of each type and N_D diversified cities is a steady-state if and only if the following five conditions are satisfied.*[11]

Condition 1.1 *(Firms relocate to a specialised city once they find their ideal process)*

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \geqslant \frac{1}{1-\delta}$$

Condition 1.2 *(Firms switch to mass-production once they find their ideal process)*

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \leqslant \frac{1-\mu}{\mu}\Omega$$

Condition 1.3 *(Firms stay in diversified cities until they find their ideal process)*

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \leqslant \frac{1}{1-\delta} + \frac{1-\delta}{2}\,\frac{1-\mu}{\mu}\Omega$$

Condition 1.4 *(Firms do not give up the search for their ideal process)*

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \leqslant \frac{m\delta}{m-1+\delta}\left[\overset{?}{\Delta} + \left(\Delta - \frac{1-\delta}{m\delta}\right)\frac{1-\mu}{\mu}\Omega\right]$$

Condition 1.5 *(Firms do not search for their ideal process by relocating across specialised cities)*

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \leqslant \frac{\overset{?}{\Delta}}{\overset{?}{\Delta}_{OSC}} + \frac{\Delta - \Delta_{OSC}}{\overset{?}{\Delta}_{OSC}}\,\frac{1-\mu}{\mu}\Omega$$

[10] By the law of large numbers, the proportion of firms closing down or finding their ideal process is constant over time. So is the flow of new firms.

[11] Note that while unit production costs in diversified cities and specialised cities are variables, Lemma 2 solves their ratio, $\frac{Q_D}{Q_S}$, solely as a function of parameters. We have not substituted this ratio into Conditions 1.1–1.5 to ease the intuition.

Proof A nursery configuration is a steady-state if and only if, with all firms following the nursery strategy, no firm finds it profitable to deviate from this strategy. Some deviations involve firms engaging in mass-production in diversified cities. Substituting (23) into (17) and valuing this in a diversified city yields operational profits for a firm doing this while all other firms follow the nursery strategy:

$$\pi_D = \left(\frac{Q_S}{Q_D}\right)^{\sigma-1} \frac{(1-\mu)Y}{\sigma\,(mN_S n_S)} \,. \tag{28}$$

Some deviations involve firms producing prototypes in specialised cities. Substituting (22) into (16) and valuing this in a specialised city, yields operational profits for a firm doing this while all other firms follow the nursery strategy:

$$\overset{?}{\pi}_S = \left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \frac{\mu Y}{\sigma\left(N_D m \overset{?}{n}_D\right)} \,. \tag{29}$$

From (24), (25), (28), and (29) we obtain the following profitability ratios:

$$\frac{\overset{?}{\pi}_S}{\overset{?}{\pi}_D} = \frac{\pi_S}{\pi_D} = \left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \,, \tag{30}$$

$$\frac{\pi_S}{\overset{?}{\pi}_D} = \frac{1-\mu}{\mu}\,\Omega \,. \tag{31}$$

Let us start ruling out possible deviations from the end. Consider a firm in a diversified city that knows its ideal process. Following the nursery strategy, it can relocate to a city of the relevant specialisation and, if it survives the relocation period (which happens with probability $1-\delta$), engage in mass-production there for an expected $\frac{1}{\delta}$ periods. Alternatively, it could engage in mass-production in the diversified city. The latter option is not a profitable deviation if and only if

$$\frac{\pi_D}{\delta} \leqslant (1-\delta)\frac{\pi_S}{\delta} \,. \tag{32}$$

Substituting (30) into (32) and rearranging yields Condition 1.1. For a firm to want to relocate once it finds its ideal process, unit production costs need to be sufficiently higher in diversified cities (as measured by $\frac{Q_D}{Q_S}$) to make the relocation cost (a probability δ to close down during relocation) worth incurring.

Another possible deviation for a firm that knows its ideal process is to nevertheless keep producing prototypes. But this is not more profitable than engaging in mass-production in the same type of city, provided that

$$\frac{\overset{?}{\pi}_S}{\pi_S} = \frac{\overset{?}{\pi}_D}{\pi_D} \leqslant 1 \,. \tag{33}$$

Using (30) and (31), (33) becomes Condition 1.2. For firms to want to switch to mass-production as soon as they can, the market for mass-produced goods needs to be large relative to the market for prototypes (as measured by $\frac{1-\mu}{\mu}$), and served by relatively few firms (as measured by $\Omega = \frac{N_D m \overset{?}{n}_D}{N_S m n_S}$), with relatively high operational profits per unit of output.

From Condition 1.2, a firm that knows its ideal process wants to engage in mass-production. From Condition 1.1, it prefers to do so in a city of the relevant specialisation rather in a diversified city, even if that involves relocating. Thus, Conditions 1.1 and 1.2 jointly guarantee that any firm in a diversified city that knows its ideal process wants relocate to a city of the relevant specialisation and engage in mass-production there.

The next issue is whether a firm located in a diversified city stays there until it finds its ideal process. Alternatively, it could relocate from a diversified to a specialised city after making $m-2$ prototypes without finding its ideal process, not yet knowing which of the two remaining processes is its ideal one. Such a deviant firm relocates one period earlier than under the nursery strategy, and so has a lower probability, $1-\delta$, of making its next prototype. On the other hand, if it makes this next prototype, it will do so at a lower cost (if Condition 1.1 is satisfied), getting the higher operational profits associated with specialised cities. However, it may turn out (with probability $\frac{1}{2}$) that its ideal process is not this next one but the one it left to try last, in which case it has to relocate once more than under the nursery strategy, delaying the mass-production stage in which it gets operational profits π_S per period. Thus a firm does not find it profitable to deviate from the nursery strategy by relocating from a diversified to a specialised city not yet knowing which of two remaining processes is its ideal one if and only if

$$(1-\delta)\overset{?}{\pi}_S + \frac{1}{2}\frac{(1-\delta)^3}{\delta}\pi_S \leqslant \overset{?}{\pi}_D + \frac{1}{2}\frac{(1-\delta)^2}{\delta}\pi_S \,. \tag{34}$$

Substituting (30) and (31) into (34) and rearranging yields Condition 1.3. This condition is satisfied when the market for mass-produced goods is large relative to the market for prototypes, when the number of firms serving this market is small, when the cost advantage of specialised cities is not too large, and when (if Conditions 1.1 and 1.4 are satisfied) relocation costs (as measured by δ) are not too low.

There are other possible deviations from the nursery strategy for a firm that initially locates in a diversified city. A firm could relocate from a diversified to a specialised city, not just with two processes left to try, but with any number of untried processes between 2 and $m-1$. If it did so and its ideal process did not correspond to this city's specialisation, it could either keep trying to find its ideal process by relocating further to other specialised cities, or it could at some point give up the search for its ideal process and remain in a specialised city producing prototypes. Alternatively, it could also decide at some point to return to a diversified city. The Appendix shows that none of these deviations impose additional parameter constraints on a nursery steady-state beyond those of Conditions 1.1–1.5.

The final step is to show that a firm does not find it profitable to locate initially in a specialised rather than in a diversified city. The Appendix rules out deviations involving relocation from a specialised to a diversified city, or a firm giving up the search for its ideal process after producing more than one prototype. This leaves only two possibilities for a firm that deviates from the nursery strategy by locating initially in a specialised city. First, it could remain in this specialised city regardless of the outcome of the first trial. If it manages to survive the first period in a specialised city (which happens with probability $1-\delta$), it will be able to engage in mass production with probability $\frac{1}{m}$ or else will keep producing prototypes. This is not a profitable deviation from the nursery strategy if and only if

$$\overset{?}{\pi}_S + (1-\delta)\left[\frac{1}{m}\frac{\pi_S}{\delta} + \frac{m-1}{m}\frac{\overset{?}{\pi}_S}{\delta}\right] \leqslant \overset{?}{\Delta}\overset{?}{\pi}_D + \Delta\pi_S \; . \tag{35}$$

Substituting (30) and (31) into (35) and rearranging yields Condition 1.4. A firm is deterred from giving up the search for its ideal process by a lower cost disadvantage of diversified cities, $\frac{Q_D}{Q_S}$, a larger market for mass-produced goods relative to the market for prototypes, $\frac{1-\mu}{\mu}$, and a smaller relative number of firms serving this market, Ω. It is also more likely to stick to the nursery strategy when the expected number of periods producing prototypes, $\overset{?}{\Delta}$, and the expected additional periods engaged in mass-production under the nursery strategy, $\Delta - \frac{1-\delta}{m\delta}$, are large relative to the expected number of periods producing prototypes if it only tries one process, $\frac{m-1+\delta}{m\delta}$.

The remaining alternative for a firm locating initially in a specialised city is to search for its ideal process solely in specialised cities, which would mean relocating from one specialised city to another between prototypes in order to try different production processes until finding the ideal one, and then stay in a city of the relevant specialisation engaged in mass-production. This is not a profitable deviation from the nursery strategy if and only if

$$\overset{?}{\Delta}_{OSC}\overset{?}{\pi}_S + \Delta_{OSC}\pi_S \leqslant \overset{?}{\Delta}\overset{?}{\pi}_D + \Delta\pi_S \; . \tag{36}$$

Substituting (30) and (31) into (36) and rearranging yields Condition 1.5. A firm takes into account the relative expected duration of the prototype and mass-production stages under each strategy (note that $\overset{?}{\Delta}_{OSC}$ only includes the expected periods producing prototypes, but not the periods of relocation from one specialised city to another one during the prototype production stage). A firm is also deterred from searching for its ideal process in specialised cities by a low cost advantage of specialised cities, a large market for mass-produced goods relative to the market for prototypes, and a small number of firms serving this market.

Therefore, if Conditions 1.1–1.5 are satisfied, with all firms following the nursery strategy, no firm finds it profitable to deviate from this strategy. □

The nursery strategy can be seen as a risky investment. Whether it is worthwhile or not depends on its cost, on the payoff if successful, and on the likelihood of success. The nursery strategy is costly because in a diversified city all firms impose congestion costs on each other, but only those using the same type of process create cost-reducing localisation economies, and this results in comparatively higher production costs (from Condition 1.1, $\frac{Q_D}{Q_S} > 1$). If the cost advantage of specialised cities is too large, a firm may find it worthwhile to produce in specialised cities before finding its ideal process (Conditions 1.3-1.5). On the other hand, if the cost advantage is too small, a firm may never want to incur the cost of moving away from a diversified city (Condition 1.1). The payoff to learning is also important, and this increases with the size of the market for mass-produced goods relative to the market for prototypes ($\frac{1-\mu}{\mu}$). It also depends on how crowded each market is, as measured by the relative number of firms (Ω, which in turn depends on m and δ). Finally, the likelihood of a firm finding its ideal process depends on the number of alternatives (m), and the chances of closure in any period (δ). Figure 1 illustrates the dependence of Conditions 1.1–1.5 on m and δ. This is plotted for parameter values given in the Appendix, and ignores that m

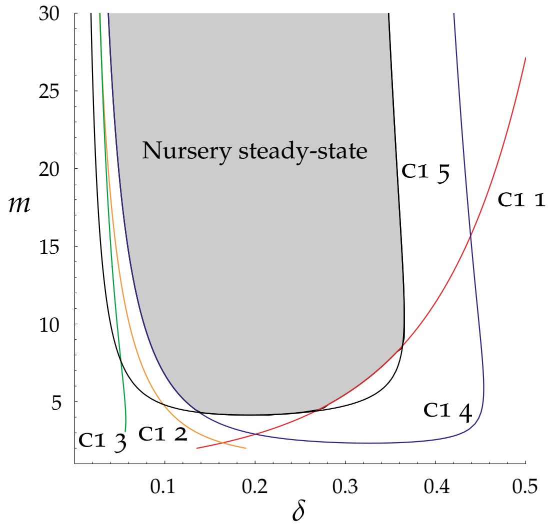

Figure 1. Dependence of Conditions 1.1–1.5 on m and δ

is an integer for visual clarity. The area where this nursery configuration is a steady-state is shaded in gray.

A larger value of m makes finding the ideal production process more difficult for firms. Consequently, the nursery strategy becomes more attractive than other strategies that involve relocations while producing prototypes.[12]

Regarding δ, a low value of this parameter makes searching for the ideal process across specialised cities a less costly alternative to the nursery strategy (Condition 1.3 and downward sloping portion of Condition 1.5). It also implies that, with all firms following the nursery strategy, a higher proportion of them will get to the mass-production stage. This makes it more attractive for a firm to deviate and stop looking for its ideal process (downward sloping portion of Condition 1.4), or to keep on producing prototypes even if it finds its ideal process (Condition 1.2). On the other hand, a high value of δ makes it unlikely that a firm makes it to the mass-production stage. This increases the importance of getting higher operational profits while producing prototypes, encouraging firms to search for their ideal process across specialised cities (upward sloping portion of Condition 1.5) or discouraging them from searching altogether (upward sloping portion of Condition 1.4). It is therefore for intermediate values of δ that the nursery configuration is a steady-state.

4. Steady-states with only diversified or only specialised cities

Consider now two alternative configurations: only diversified cities and only specialised cities.

Definition 4 (Configuration with only diversified cities) *A configuration with only diversified cities is one that satisfies all the following. Each firm produces prototypes using a different type of production*

[12]There are also circumstances (in particular, a very large share of demand being allocated to prototypes) under which the increased uncertainty associated with a larger value of m can deter a firm from trying to find its ideal process (a violation of Condition 1.4). This is, however, mostly an artifact of the convenient functional form chosen for consumer preferences, which guarantees a share of expenditure to prototype producers.

process each period, and as soon as it finds its ideal production process, it commences mass-production. Firms never relocate.

Proposition 2 (Steady-state with only diversified cities) *A configuration with only diversified cities is a steady-state if and only if the following condition is satisfied.*

Condition 2.1 *(Firms switch to mass-production once they find their ideal process)*

$$\frac{1-\mu}{\mu}(1-\delta)\Omega \geqslant 1 \ .$$

Proof Without specialised cities and with symmetric diversified cities, the only possible deviation is for a firm to keep producing prototypes in the same city after it finds its ideal production production process. This deviation is not profitable whenever $\pi_D \geqslant \overset{?}{\pi}_D$. From (6), (7), (14) and (15), operational profits in the configuration with only diversified cities are a share $\frac{1}{\sigma}$ of revenue:

$$\overset{?}{\pi}_D = \frac{\mu Y}{\sigma \left(N_D m \overset{?}{n}_D \right)} \ , \tag{37}$$

$$\pi_D = \frac{(1-\mu)Y}{\sigma \left(N_D m n_D \right)} \ . \tag{38}$$

The strategy that each firm follows in a configuration with only diversified cities is identical to the nursery strategy up until it finds its ideal process. Then, instead of relocating to a specialised city (of which there are not any), the firm switches to mass production without relocating. Thus, the expected duration of the prototype production stage is the same as under the nursery strategy. And, since there is no relocation, the expected duration of the mass-production stage is $\frac{1}{(1-\delta)}$ times that under the nursery strategy. Consequently the number of prototype producers relative to the number of mass-producers is $(1-\delta)$ times that under the nursery strategy: $\frac{\overset{?}{n}_D}{n_D} = (1-\delta)\Omega$. Using this, (37), and (38), the inequality $\pi_D \geqslant \overset{?}{\pi}_D$ becomes Condition 2.1. □

Definition 5 (Configuration with only specialised cities) *A configuration with only specialised cities is one that satisfies all of the following. There are only specialised cities, and there is the same proportion of cities specialised in each type of process. Each firms searches for its ideal production process by relocating across specialised cities to produce different prototypes. As soon as a firm finds its ideal production process, it commences mass-production.*

Proposition 3 (Steady-state with only specialised cities) *A configuration with only specialised cities is a steady-state if and only if the following condition is satisfied.*

Condition 3.1 *(Firms do not give up the search for their ideal process)*

$$\frac{1-\mu}{\mu}\Omega_{OSC} \geqslant \frac{m(1-\delta\overset{?}{\Delta}_{OSC})-1+\delta}{m\delta\Delta_{OSC}-1+\delta} \ ,$$

$$\textit{where} \ \ \Omega_{OSC} \equiv \frac{\overset{?}{n}_S}{n_S} = \frac{\delta(m+1)(2-\delta)-1+(1-\delta)^{2(m-1)}[1-2\delta(2-\delta)]}{(2-\delta)\{1-\delta-(1-\delta)^{2(m-1)}[1-(3-\delta)\delta]\}} \ .$$

Proof See the Appendix. □

In steady-states with only diversified or only specialised cities, firms want to engage in mass production when the market for mass-produced goods is large relative to the market for prototypes (as measured by $\frac{1-\mu}{\mu}$), and when the proportion of firms engaged in mass-production is low (as measured by $(1-\delta)\Omega$ with only diversified cities, and by Ω_{OSC} with only specialised cities). Furthermore, in steady-states with only specialised cities, learning involves frequent relocations. Consequently, a lower probability of closing down during relocation (δ) deters firms from giving up.

5. City size and welfare

Since three possible configurations can be in steady-state, it is important to compare their robustness and welfare properties. To do this, we first derive optimal city size, and establish that a steady-state can be stable with respect to small perturbations in the distribution of workers only when all cities are no smaller than their optimal size. We then show that when a nursery configuration is a steady-state, it provides a higher welfare level than the alternative configurations, for given sizes of specialised and diversified cities. Finally, instead of treating the number of cities parametrically, we open up a market for the development of cities through the activity of perfectly competitive land developers. In this case, if the nursery configuration with cities of optimal size is a steady-state then this becomes the unique equilibrium configuration.

Lemma 5 (Optimal city size) *Optimal city size is* $\frac{\epsilon}{(2\epsilon+1)\tau}$.

Proof This follows from maximising output per worker in each city, as derived in Lemma 1. □

The size of a city affects its efficiency by changing the balance between the economies of localisation and congestion costs. Optimal city size increases with localisation economies, as measured by ϵ, and diminishes with the congestion costs parameter, τ. Note that optimal size for any city is independent of the composition of its population. This result stems directly from the assumption of homothetic production function.[13] Note also that, for any configuration, total welfare reaches its restricted maximum when all cities are of this size.

Equilibrium stability is closely related to optimal city size. While mobility ensures that workers in all cities are equally well-off, there might be configurations where this equality can be broken by a small perturbation in the spatial distribution of workers. The three steady-states studied above are such that all specialised cities (if any) are of the same size and all diversified cities (if any) are also of the same size. Proposition 4 implies that no steady-state can be stable otherwise.

Proposition 4 (Stability) *A steady-state is stable with respect to small perturbations in the spatial distribution of workers if and only if all diversified cities (if any) are of the same size, all specialised cities (if any) of the same type are of the same size, and no city is smaller than* $\frac{\epsilon}{(2\epsilon+1)\tau}$.

[13] In particular, it might be more realistic to assume that the intensity of increasing returns is decreasing with net local employment above a given threshold for each type of process. This would yield a larger optimal city size for diversified cities than for specialised cities.

Proof Steady state nominal wages are equal to output per worker, given by Lemma 1. Wage equalisation then implies that cities of the same type (diversified cities, or cities with the same specialisation) can be at most of two different sizes. One of these two possible sizes is no greater than the optimal size, $\frac{\epsilon}{(2\epsilon+1)\tau}$, while the other one is no smaller than this. Taking the derivative of output per worker, as derived in Lemma 1, with respect to the number of workers of each type in each city, shows that a small positive perturbation increases output per worker and hence wages in a city of suboptimal size (by (19) and (20) it also lowers costs). This makes this city more attractive relative to other cities with a similar composition. As workers and firms move in, the size of this city increases. A negative perturbation has the opposite effect of pushing more and more workers and firms away from such a city. By contrast, in a city no smaller than the optimal size, a positive perturbation makes the city less attractive, while a negative perturbation makes the city more attractive to workers and firms. Consequently, stability requires that all cities of the same type are of the same size, and that no city is smaller than $\frac{\epsilon}{(2\epsilon+1)\tau}$. □

Whenever a nursery configuration is a steady-state, a configuration with only diversified cities is also a steady-state (from Conditions 1.1, 1.2 and 2.1), and a configuration with only specialised cities can also be a steady-state. It is thus important to compare their welfare properties, for comparable city sizes (diversified cities of the same size under the nursery configuration as under the configuration with only diversified cities, and specialised cities of the same size under the nursery configuration as under the configuration with only specialised cities).

Proposition 5 (Welfare ranking of the steady-states) *Whenever a nursery steady-state exists, it provides a higher level of welfare than a steady-state with only diversified or only specialised cities, for comparable city sizes.*

Proof See the Appendix. □

This result is easy to understand. Consider first the comparison of a nursery steady-state and a steady-state with only diversified cities, assuming diversified cities are of the same size in both configurations (which makes nominal wages equal). If the nursery configuration is a steady-state, by Condition 1.1, each firm would like to relocate to a specialised city once it finds its ideal process. There production costs are sufficiently lower as to offset the cost of relocation, increasing firms' profits with respect to not relocating. As free entry and exit of firms exhausts profits net of the start-up cost, this also gets translated into higher welfare by means of lower price indices. Since without specialised cities a firm cannot exploit this opportunity, the nursery steady-state provides higher welfare than the steady-state with only diversified cities. A similar argument can be made to compare a nursery steady-state and a steady-state with only specialised cities, assuming specialised cities are of similar size in both configurations. If the nursery configuration is a steady-state, by Condition 1.5, each firm prefers to search for its ideal process in a diversified city rather than by relocating across specialised cities, with all other firms following the nursery strategy. If all other firms are searching for their ideal process across specialised cities, doing the same is even less profitable. Again, free entry and exit of firms translates this into higher welfare in the nursery steady-state relative to the steady-state with only specialised cities. Thus, when a nursery configuration is a steady-state, everyone prefers to have both diversified and specialised cities. However, if there happen to be only diversified or only specialised cities, localisation economies prevent a mixed configuration from arising in the absence of coordination. All of this highlights the importance of looking at city formation as a way to resolve the coordination problem that results in a multiplicity of equilibria.

Thus far, the number and type of cities has been treated parametrically. Nonetheless new cities are being created, not just by the autonomous decisions of small agents, but also by private bodies acting as land development companies ('developers') or by local governments pursuing active development policies (see Fujita, 1989, and Becker and Henderson, 2000, for a discussion of this issue and for an equivalence result between these two types of institutions).

Suppose that instead of there being public ownership of land, each potential site for a city is owned by a different developer, and there is an unexhausted supply of such sites, all of which are identical. Each developer seeks competitively to maximise land rents in the city net of all costs. Each developer offers workers and firms a contract that cannot be renegociated over time. These contracts specify the size of the city, whether it has a dominant sector, and any per worker or per firm subsidies to those who locate there.[14]

Proposition 6 (Land developers) *With competitive land developers, all cities achieve optimal size, and all local land rents are transferred to local firms, filling the gap between the private and the public marginal product of firms.*

Proof The proof follows from the analysis of Becker and Henderson (2000). □

It is immediate that this result, which is also known as the 'Henry George Theorem', also implies in our model that if the nursery configuration with cities of optimal size is a steady-state, then this is the unique equilibrium configuration.

6. Empirical implications

The scarcity of theoretical work on the relative advantages and disadvantages of urban diversity and specialisation is in contrast with the wealth of empirical work on the topic. This section discusses existing empirical evidence that relates to the implications of the model. It also presents new evidence on establishment relocations across France in 1993–1996.

Coexistence of diversified and specialised cities

The model stresses the advantages of urban configurations in which diversified and specialised cities coexist. This coexistence is a pervasive fact (see Henderson, 1988, for evidence for the United States, as well as Brazil and India; Black and Henderson, 1998, and Duranton and Puga, 2000, for for further US evidence, and Lainé and Rieu, 1999, for evidence for France).

[14]It may be nonetheless more realistic to assume that developers are 'constrained' (Helsley and Strange, 1997), and in particular that they may develop specialised cities but not diversified cities. In this case, specialised cities reach optimal size whereas diversified cities are too large.

One interpretation of the coexistence of diverse and specialised cities is that patterns of specialisation and diversity are merely random outcomes. However, in their careful study of the agglomeration of industries in the US, Ellison and Glaeser (1997) show that most four-digit sectors are too concentrated geographically for their distribution to be the result of a random allocation. A second interpretation is that patterns of specialisation and diversification in cities merely mirror the spatial distribution of resources. However, the results of Ellison and Glaeser (1999) attribute only about one-fifth of the concentration reported in their earlier paper to observable natural advantages. Henderson (1997*a*) also looks at the role of natural advantages, using panel data to separate externalities from persistent comparative advantage. He shows that, even when local fixed effects are accounted for, externalities remain important. Patterns of specialisation and diversity are thus to a large extent the result of economic interactions taking place both within sectors and across sectors.

The relative advantages of diversity and specialisation

The relative advantages of diversity and specialisation for innovation and production are the basis of our model. Glaeser *et al.* (1992) assess the importance of these advantages by examining the evolution of urban employment patterns. They find that diversity fostered urban employment growth in US cities between 1956 and 1987. Pursuing this line of research, Henderson *et al.* (1995) look at differences across sectors. They find that, while urban diversity is indeed important for attracting new and innovative sectors, a history of similar past specialisation appears to matter more to retain mature industries. Combes (2000) finds similar results for France between 1984 and 1993, with service sectors as well as more innovative manufacturing sectors benefitting from diversity.

Fujita and Ishii (1998) study the location of trial plants producing prototypes and mass-production plants for nine major Japanese electronic firms. They show that trial plants are overwhelmingly located in more diverse metropolitan areas, whereas mass-production plants are almost always located in more specialised cities. Henderson (1997*b*) provides further evidence for the US. He shows that the production of less standardised or non-traditional items tends to be more concentrated in diversified metropolitan areas. On the other hand, the production of more standardised or traditional items tends to be more heavily concentrated in cities which are often quite specialised in terms of their exports to other cities.

Henderson (1999) takes a different approach, and looks at the evolution of productivity in US manufacturing plants for nine industries (high-tech and machinery) over thirty years, conditioning on the local composition of economic activities. He finds that local same-sector specialisation tends to have a positive effect on productivity in both types of industries. The effects are stronger in high-tech industries, consistent with these being more agglomerated. Regarding diversity, he finds only weak effects of diversity on productivity. However, when he looks at employment changes, his findings are in line with the previous literature, indicating that diversity matters for the spatial allocation of activities.

All of this is fully consistent with the predictions of the model, where observed productivity is higher in specialised cities. By contrast, diversity allows firms to find production processes that will help them attain higher future productivity. But since firms relocate to specialised cities to exploit such processes, the positive effects of diversity cannot be observed in productivity levels in diversified cities. They do, however, manifest in the presence of newer or less standardised activities in diversified cities and of older or more standardised activities in specialised cities. The advantages of diversity become particularly clear when one looks explicitly at process innovation.

Urban diversity and process innovation

The microeconomic foundations of the model specifically emphasise the benefits of diversity for process innovation. Detailed evidence on this particular aspect is provided by Harrison *et al.* (1996) and Kelley and Helper (1999).[15] They study the adoption of new production processes, involving the switch from traditional tools to numerically controlled or computer numerically controlled machines, by individual establishments making machine tools. Kelley and Helper (1999) find that a diversity of local employment contributes significantly towards the adoption of new production processes, even after controlling for other geographical, technical and organisational characteristics. Their sample includes US establishments from 21 different three-digit machine-making industries (ranging from heating equipment and plumbing fixtures to guided missiles and aircraft). A strong local presence of the aggregated 21 industries also has a positive effect on adoption, albeit weaker than that of overall diversity. At the same time, when they look at local employment in individual machine-making industries, Harrison *et al.* (1996) find that same-sector specialisation hinders the adoption of the new processes. They also suggest that those plants that relocated away from diverse urban areas to nearby suburban counties are amongst the most likely adopters of the new processes. This link between process innovation and firms' location and relocation decisions is particularly important because it is at the core of our model.

Establishment turnover and location patterns

When both specialised and diversified cities co-exist in steady-state, the model predicts that cities are stable in their size and sectoral composition, but there is a constant turnover of firms. Some existing firms close down every period, and are replaced by new firms. The entrants prefer to locate in diversified cities because they provide a better environment for process innovation, but eventually relocate to specialised cities.

As predicted by the model, there is a great stability in the relative sizes of cities. The systematic analysis of all US cities between 1900 and 1990 by Black and Henderson (1998) shows that, with few exceptions (such as Phoenix, Detroit, or Pittsburgh), the relative sizes of US cities changed little over the course of the last century. According to Eaton and Eckstein (1997), this pattern of overall stability is even stronger in France and Japan. The economic structure of cities has also been shown to be extremely stable over time. Over the long period between 1860 and 1987, Kim (1995) finds a correlation of 0.64 for the coefficient of regional localisation at the State level for two-digit US

[15] Feldman and Audretsch (1999) look at product, rather than process, innovation. They study the location of establishments responsible for new products reported by trade journals in the US. Diversity across industries with a common science base has a large positive effect on innovative output, whereas same-industry specialisation has a negative effect on innovative output. The effect of city size is also positive but much weaker.

industries. Dumais *et al.* (1997) provide further evidence, showing that for most US three-digit industries the stability in their concentration across US States over 1972-1992 is striking. The same picture of stability emerges from Henderson (1997*b*, 1999) who looks at some specific sectors across US cities.

This stability of city sizes and sectoral composition is in contrast with the high rate of establishment turnover. According to Dumais *et al.* (1997), nearly three fourths of the plants existing in 1972 were closed by 1992, and more than one half of all US manufacturing employees in 1992 worked in plants that did not exist in 1972. The same authors then show that plant openings and closures are spatially biased. The opening of new plants tends to reduce the degree of agglomeration of particular sectors, suggesting that new plants are created in locations with below-average specialisation in the corresponding sector. On the other hand, the closure of existing plants tends to increase the degree of agglomeration of particular sectors. This is consistent with our nursery steady-state, where the larger rate of plant openings in diversified cities (due to new plants locating in diversified cities) makes sectoral concentration lower than it would otherwise be, and the corresponding larger rate of plant closures (due to relocations departing from diversified cities) makes concentration higher than it would otherwise be.[16]

Unfortunately, the evidence on relocation patterns in these and most other papers is very indirect, since the data rarely separates openings due to the creation of new plants from openings due to the relocation of existing plants. A new exhaustive data set on establishment relocations across France makes such distinction clear. We conclude this section by using that data to show that, at least for France, there is strong direct evidence of the relationship implied by the model between establishment relocations and urban economic structures.

Evidence from establishment relocations across France 1993–1996

The data, extracted from the SIRENE database of the Institut National de la Statistique et des Etudes Economiques (INSEE), contains the geographical origin and destination and the sectoral classification of every single establishment relocation that took place in France between 1993 and 1996 (see Lainé, 1998, for a detailed description). Only complete relocations are included in the data (that is to say, episodes in which the complete closure of an establishment is followed by the opening in a different location of an establishment owned by the same firm and performing the full same range of activities).

The geographical origin and destination of relocating establishments is identified at the level of employment areas (zones d'emplois). Continental France is fully covered by 341 employment areas, whose boundaries are defined on the basis of daily commuting patterns. Relocating establishments are classified by sector according to level 36 of the Nomenclature d'Activités Française (NAF) classification of the INSEE. The 18 sectors we study cover all of manufacturing and business-services, with the exception of postal services.

[16]From equation (A 11) in the Appendix, per period plant openings and closures in diversified cities represent a fraction $\delta(1+\frac{1}{(1-\delta).\Omega})$ of the local stock, whereas per period plant openings and closures in specialised cities represent a lower fraction δ of the local stock.

The model predicts that establishments will tend to relocate from particularly diversified areas to particularly specialised areas. To characterise French employment areas in terms of diversity and specialisation, we use sectoral employment data for each employment area from the Enquête Structure des Emplois (ESE) for December 1993. We measure the specialisation of employment area i in sector j by the share of the corresponding NAF36 sector in local manufacturing and business-service employment, s_i^j. We measure the diversity of employment area i by the inverse of a Herfindahl index of sectoral concentration of local employment, $1/\sum_{j=1}^{85}(s_i^j)^2$, calculated in this case at a higher level of sectoral disaggregation given by the NAF85 classification. In order to identify employment areas which are particularly specialised in a given sector or particularly diversified, we normalise both measures by their median value for all employment areas. By these measures, Lyon and Nantes are amongst France's most diversified areas, Chateaudun has the median diversity, while Lavelanet is both the least diversified and one of the most specialised areas (in textiles, which in 1993 accounted for 84% of local manufacturing and business service employment).

Results regarding relocation patterns are presented in Table 1. Looking first at the aggregate figures in the bottom row, we see that complete establishment relocations across French employment areas represented 4.7% of the average stock over this period (29,358 relocations from an average stock of 624,772 establishments). We find this number surprisingly high, given the restrictive definition of relocations in the data. Firms can reallocate their productive activities in other ways. For instance, they can can transfer to the plant being opened only some of the activities previously carried out in the plant being closed, or (as in the study by Fujita and Ishii, 1998, for Japanese electronics firms) they can maintain prototype and mass production plants and change over time the products made in each. None of these partial relocations are included in our data.

An even more striking result is that 72% of all complete establishment relocations across employment areas were from an area with above median diversity to an area with above median specialisation in the corresponding sector. The model not only predicts this pattern of relocations, it also suggests that innovative sectors with strong localisation economies will have a stronger tendency to relocate from particularly diversified to particularly specialised areas. More traditional sectors, by contrast, will tend to experience fewer relocations and not necessarily with this 'nursery' pattern. That is precisely what comes out of Table 1 when we split relocations by sector. R&D, pharmaceuticals and cosmetics, IT and consultancy services, and business services have relocation rates of between 5 and 8.1%. And between 75.8 and 93% of relocations in these sectors are from an area with above median diversity to an area with above median specialisation. They also appear to be some of the sectors where firms benefit more from being in the same locations, as reflected in the geographic concentration indices on the right-most column. On the other hand, food and beverages, furniture and fixtures, wood, lumber, pulp and paper, primary metals, and non-metallic mineral products have relocation rates of only between 0.8 and 2.7%, and less than 35% of those relocations are from an area with above median diversity to an area with above median specialisation. They are also not particularly agglomerated sectors.

One might not have expected this pattern had it not been for the model. More agglomerated sectors are likely to be those where firms benefit most from being close to each other or from being

	% of relocations from diversified to specialised areas[a]	Relocations as a % of the stock[b]	Geographic concentration[c]
R&D	**93.0**	8.1	0.023
Pharmaceuticals and cosmetics	**88.3**	6.4	0.020
IT and consultancy services	**82.1**	7.3	0.030
Business services	**75.8**	5.0	0.015
Printing and publishing	**73.3**	5.4	0.026
Aerospace, rail and naval equipment	**71.6**	3.3	0.026
Electrical and electronic equipment	**69.1**	4.2	0.011
Motor vehicles	**62.5**	2.7	0.020
Electrical and electronic components	**60.9**	5.9	0.007
Textiles	**46.4**	2.5	0.024
Chemical, rubber and plastic products	**38.3**	3.9	0.009
Metal products and machinery	**37.6**	3.2	0.005
Clothing and leather	**36.3**	3.4	0.013
Food and beverages	**34.6**	0.8	0.007
Furniture and fixtures	**32.6**	2.7	0.008
Wood, lumber, pulp and paper	**30.6**	1.7	0.009
Primary metals	**30.0**	2.5	0.009
Non-metallic mineral products	**27.3**	2.0	0.012
Aggregate	**72.0**	4.7	

[a]Percentage of all establishments relocating across employment areas that move from an area with above median diversity to an area with above median specialisation.

[b]Establishment relocations across employment areas relative to the average number of establishments.

[c]Ellison and Glaeser (1997) geographic concentration index.

Source: Authors' calculations based on the SIRENE and ESE data sets.

Table 1. Establishment relocations across French employment areas 1993–1996

at certain locations. This could make establishments in those sectors relocate less, and when they do relocate to move to similar places. By contrast, the model suggests that firms in more innovative and agglomerated activities benefit most from the advantages that diversity and specialisation offer at different stages of the product-cycle. Therefore, they relocate more and tend to move from particularly diversified to particularly specialised areas.

7. Concluding remarks

In the empirical literature and economic policy discussions about urban economic structures, the debate has been mostly framed in terms of diversity versus specialisation, as if the answer was one or the other. This paper suggests instead that both diversified and specialised urban environments are important in systems of cities. There is a role for each type of local economic environment but at different stages of a firm's life-cycle. Diversified cities are more suited to the early stages of a product's life-cycle whereas more specialised places are better to conduct mass-production of fully developed products. For manufacturing and services, unlike for agriculture, 'sowing' and 'reaping' can take place in different locations.

A 'balanced' urban system may thus not be one where all cities are equally specialised or equally diversified but one where both diversified and specialised cities co-exist. In such a system, some cities specialise in churning new ideas and new products (which requires a diversified base), whereas other cities specialise in more standardised production (which, in turn, is better carried out in a more specialised environment).

The usefulness of a model in relation to the facts can come from helping make better sense of facts we already know about, or from pointing to new facts to look for. In this paper we have tried to do both. The distribution of economic activities emerging from the model and its microeconomic foundations have been shown to be consistent with seemingly contradictory findings in the empirical literature attempting to measure spatial externalities and their relationship with urban diversity and specialisation. At the same time, the model suggests that the varying relative importance of diversity and specialisation at different stages of the product-cycle will lead many firms (especially those in more innovative and agglomerated sectors) to relocate from particularly diversified to particularly specialised cities. We have used data on complete plant relocations across France to take a first cut at this prediction, and show that the pattern is strongly there. However, casual observation and case studies suggest that the complete relocation of establishments may only be a small component in the mobility of production. Production relocations across plants within the same firm seem to play an important role in many industries, from electronics to automobiles. Multi-establishment firms, by locating their facilities in different places, may be able to take advantage of different types of environment at the same time. Thus, what our empirical findings reflect may be just the tip of the iceberg in the relocation of production over the life-cycle.

References

Abdel-Rahman, Hesham M. 2000. City systems: General equilibrium approaches. In Jean-Marie Huriot and Jacques-François Thisse (eds.), *Economics of cities*. Cambridge: Cambridge University Press.

Abdel-Rahman, Hesham M. and Masahisa Fujita. 1990. Product variety, Marshallian externalities, and city sizes. *Journal of Regional Science* 30(2): 165–183.

Baldwin, Richard E. and Rikard Forslid. 2000. Trade liberalisation and endogenous growth: A q-theory approach. *Journal of International Economics* (forthcoming).

Becker, Randy and J. Vernon Henderson. 2000. Intra-industry specialisation and urban development. In Jean-Marie Huriot and Jacques-François Thisse (eds.), *Economics of cities*. Cambridge: Cambridge University Press.

Black, Duncan and J. Vernon Henderson. 1998. Urban evolution in the US. Working Paper 98–21, Brown University.

Combes, Pierre Phillipe. 2000. Economic structure and local growth: France 1984–1993. *Journal of Urban Economics* (forthcoming).

Dumais, Guy, Glenn Ellison, and Edward L. Glaeser. 1997. Geographic concentration as a dynamic process. Working Paper 6270, National Bureau of Economic Research. URL http://www.nber.org/.

Duranton, Gilles and Diego Puga. 2000. Diversity and specialisation in cities: Why, where and when does it matter? *Urban Studies* 37(3): 533–555.

Eaton, Jonathan and Zvi Eckstein. 1997. Cities and growth: Theory and evidence from France and Japan. *Regional Science and Urban Economics* 27(4-5): 443–474.

Ellison, Glenn and Edward L. Glaeser. 1997. Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy* 105(5): 889–927.

Ellison, Glenn and Edward L. Glaeser. 1999. The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review Papers and Proceedings* 89(2): 311–316.

Ethier, Wilfred J. 1982. National and international returns to scale in the modern theory of international trade. *American Economic Review* 72(3): 389–405.

Feldman, Maryann P. and David B. Audretsch. 1999. Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review* 43(2): 409–429.

Fujita, Masahisa. 1988. A monopolistic competition model of spatial agglomeration: A differentiated product approach. *Regional Science and Urban Economics* 18(1): 87–124.

Fujita, Masahisa. 1989. *Urban economic theory: Land use and city size*. Cambridge: Cambridge University Press.

Fujita, Masahisa and Ryoichi Ishii. 1998. Global location behavior and organizational dynamics of Japanese electronics firms and their impact on regional economies. In Alfred D. Chandler Jr., Peter Hagström, and Örjan Sölvell (eds.), *The Dynamic Firm: The Role of Technology, Strategy, Organization and Regions*. Oxford: Oxford University Press, 343–383.

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

Glaeser, Edward L., Heidi Kallal, José A. Scheinkman, and Andrei Schleifer. 1992. Growth in cities. *Journal of Political Economy* 100(6): 1126–1152.

Harrison, Bennett, Maryellen R. Kelley, and Jon Gant. 1996. Specialization versus diversity in local economies: The implications for innovative private-sector behavior. *Cityscape* 2(2): 61–93.

Helsley, Robert W. and William C. Strange. 1990. Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics* 20(2): 189–212.

Helsley, Robert W. and William C. Strange. 1997. Limited developers. *Canadian Journal of Economics* 30(2): 329–348.

Henderson, J. Vernon. 1987. General equilibrium modelling of systems of cities. In Edwin S. Mills (ed.), *Handbook of Regional and Urban Economics*, volume 2. Amsterdam: Elsevier.

Henderson, J. Vernon. 1988. *Urban Development: Theory, Fact and Illusion*. Oxford: Oxford University Press.

Henderson, J. Vernon. 1997*a*. Externalities and industrial development. *Journal of Urban Economics* 42(3): 449–470.

Henderson, J. Vernon. 1997*b*. Medium size cities. *Regional Science and Urban Economics* 27(6): 583–612.

Henderson, J. Vernon. 1999. Marshall's economies. Working Paper 7358, National Bureau of Economic Research. URL http://www.nber.org/.

Henderson, J. Vernon, Ari Kuncoro, and Matt Turner. 1995. Industrial development in cities. *Journal of Political Economy* 103(5): 1067–1090.

Jacobs, Jane. 1969. *The Economy of Cities*. New York: Random House.

Jovanovic, Boyan. 1982. Selection and the evolution of industry. *Econometrica* 50(3): 649–70.

Jovanovic, Boyan and Glenn M. MacDonald. 1994. Competitive diffusion. *Journal of Political Economy* 102(1): 24–52.

Kelley, Maryellen R. and Susan Helper. 1999. Firm size and capabilities, regional agglomeration, and the adoption of new technology. *Economics of Innovation and New Technology* 8(1–2): 79–103.

Kim, Sukkoo. 1995. Expansion of markets and the geographic distribution of economic activities: The trends in US regional manufacturing structure, 1860-1987. *Quarterly Journal of Economics* 110(4): 881–908.

Lainé, Frédéric. 1998. Mobilité des établissements et déconcentration urbaine. In Denise Pumain and Francis Godard (eds.), *Données urbaines*, volume 2. Paris: Anthropos, 263–272.

Lainé, Frédéric and Carole Rieu. 1999. La diversité industrielle des territoires. INSEE Premiere 650, Institut National de la Statistique et des Etudes Economiques.

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

Quigley, John M. 1998. Urban diversity and economic growth. *Journal of Economic Perspectives* 12(2): 127–138.

Tobin, James. 1969. A general equilibrium approach to monetary theory. *Journal of Money Credit and Banking* 1(1): 15–29.

Appendix

Units costs as intermediate prices Here we derive from first principles the reduced-form cost functions presented in the text. Suppose there are m monopolistically competitive intermediate service sectors. Each of these sectors employs workers with one of the m possible aptitudes to produce differentiated varieties. The cost function of a sector j service firm producing variety g in city i is

$$Cs^j_i(g) = \left[\alpha + \beta y^j_i(g)\right] w^j_i \,, \tag{A 1}$$

where $y^j_i(g)$ denotes the firm's output. The expression in parenthesis is the unit labour requirement, which has both a fixed and a variable component. Thus, there are increasing returns to scale in the production of each variety of services.

Each of the m types of production process for final-good firms (whether prototype or mass-producers) corresponds to the use as inputs of services from one of the m intermediate service sectors. Services are non-tradeable across cities. All potential varieties (only some of which will be produced in equilibrium) enter symmetrically into the technology of final-good firms with a constant elasticity of substitution $\frac{\epsilon+1}{\epsilon}$. The cost function of a final-good firm producing prototype h, using a process of type j, in city i is

$$\overset{?}{C}{}^j_i(h) = Q^j_i \overset{?}{x}{}^j_i(h) \ , \tag{A 2}$$

and that of a final-good firm mass-producing good h, which uses a process of type j, in city i is

$$C^j_i(h) = \rho \ Q^j_i x^j_i(h) \ , \tag{A 3}$$

where

$$Q^j_i = \left\{ \int \left[q^j_i(g) \right]^{-1/\epsilon} \mathrm{d}g \right\}^{-\epsilon} \ , \tag{A 4}$$

$0 < \rho < 1$, $\epsilon > 0$, and $q^j_i(g)$ is the price of variety g of service sector j produced in city i.

Demand for each intermediate variety is the sum of demand by prototype producers and demand by mass-producers, obtained by application of Shephard's lemma to (A 2) and (A 3) respectively and integration over final-good firms. The market clearing condition for each service variety is then:

$$y^j_i = \overset{?}{n}{}^j_i \frac{\partial \overset{?}{C}{}^j_i}{\partial q^j_i} + n^j_i \frac{\partial C^j_i}{\partial q^j_i} = \left(\frac{\overset{?}{q}{}^j_i}{\overset{?}{Q}_i} \right)^{-(\epsilon+1)/\epsilon} (\overset{?}{n}{}^j_i \overset{?}{x}{}^j_i + \rho n^j_i x^j_i) \ . \tag{A 5}$$

where prices $\overset{?}{q}{}^j_i$ still need to be replaced by their profit-maximising values, and we have dropped index g since all variables take identical values for all service firms in the same sector and city. The profit-maximising price for each service is a fixed relative markup over marginal cost:

$$q^j_i = (\epsilon + 1) \ \beta \ w^j_i \ . \tag{A 6}$$

Free entry and exit in services drives maximised profits to zero. From the zero profit condition, the only level of output in services consistent with zero profits is

$$y^j_i = \frac{\alpha}{\beta \epsilon} \ . \tag{A 7}$$

Demand for labour can be obtained by application of Shephard's lemma to (A 1) and integration over varieties. Using (A 7), the labour markets clearing condition becomes

$$l^j_i = s^j_i \frac{\partial {C_S}^j_i}{\partial w^j_i} = s^j_i \left[\alpha + \beta y^j_i(g) \right] = s^j_i \, \alpha \, \frac{\epsilon + 1}{\epsilon} \ , \tag{A 8}$$

where s^j_i is the equilibrium 'number' of sector j service firms in city i.

By choice of units of intermediate output, we can set $\beta = (\frac{\epsilon}{\alpha})^\epsilon (\epsilon+1)^{-(\epsilon+1)}$. Using (A 8) and (A 6), the price indices of (A 4) simplify into

$$Q^j_i = \left[s^j_i \left(q^j_i \right)^{-1/\epsilon} \right]^{-\epsilon} = \left(l^j_i \right)^{-\epsilon} w^j_i \ . \tag{A 9}$$

Equations (A 2), (A 3), and (A 9) are the reduced-form cost functions of equations (2)–(4) in the main text. □

Proof of Lemma 3 Think of firms as drawing in advance the order in which they are going to try the m types of processes. The process used to produce prototype number t has the same probability of being the ideal one, $\frac{1}{m}$, as any other one. But to calculate the probability of a firm that follows the nursery strategy finding its ideal process in period t, for $1 \leqslant t \leqslant m-2$, we have to weight this by the probability of such a firm's not closing down before or on that period, $(1-\delta)^t$. Finally, if a firm following the nursery strategy gets to produce $m-1$ prototypes and remains in operation (which happens with probability $\frac{2}{m}(1-\delta)^{m-1}$), it leaves the diversified city at that point, since it has either just found its ideal process or otherwise knows that the only process left to try must be its ideal one. The total number of firms relocating from diversified to specialised cities each period is therefore a fraction $[\sum_{t=1}^{m-2}\frac{1}{m}(1-\delta)^t + \frac{2}{m}(1-\delta)^{m-1}]$ of the number of firms starting up each period, $\mathring{n}$. After simplification, this becomes $\frac{1-\delta}{\delta m}[1-(1-\delta)^{m-2}(1-2\delta)]\,\mathring{n}$. The number of firms arriving each period in specialised cities is a fraction $(1-\delta)$ of those that relocated from diversified cities the previous period, since a fraction δ close down in the period of idleness that makes relocation costly. With a constant number of firms in each city, this quantity must also equal the number of firms closing down in specialised cities each period, $\delta m N_S n_S$:

$$\frac{(1-\delta)^2}{\delta m}\left[1-(1-\delta)^{m-2}(1-2\delta)\right]\,\mathring{n} = \delta\, m N_S n_S\,. \tag{A 10}$$

In steady-state, the number of firms created each period must equal the total number of closures, which is a fraction δ of all existing firms:

$$\mathring{n} = \delta\left(N_D m \overset{?}{n}_D + \frac{\delta}{1-\delta}\, m N_S n_S + m N_S n_S\right)\,. \tag{A 11}$$

Denote the ratio of the total number of prototype producers, $N_D m \overset{?}{n}_D$, and the total number of mass-producers, $m N_S n_S$, by Ω. Then, eliminating $\mathring{n}$ from (A 10) and (A 11) yields the result. □

Proof of Lemma 4 For $t = 1, ..., m-2$, a firm following the nursery strategy can stop producing prototypes after t periods for two reasons. First, it may find that its ideal process is the one it has decided to try in attempt number t (which has the same probability of being the ideal one, $\frac{1}{m}$, as any other process) if it gets to that point (this happens with probability $(1-\delta)^{t-1}$). Second, such a firm may stop producing prototypes after t trials because its ideal process is one of the $m-t$ it has decided to try later (this happens with probability $\frac{m-t}{m}$), but it closes down right after producing prototype number t (this happens with probability $(1-\delta)^{t-1}\delta$). Finally, a firm following the nursery strategy that gets to produce prototype number $m-1$ does so because its ideal process is either the one it decided to try in attempt number $m-1$ or the remaining one (this happens with probability $\frac{2}{m}$), and remains in operation after $m-2$ periods (this happens with probability $(1-\delta)^{m-2}$). The expected number of periods producing prototypes (equal to the expected stay in a diversified city) for a firm following the nursery strategy is therefore

$$\overset{?}{\Delta} = \sum_{t=1}^{m-2} t\left[\frac{1}{m}(1-\delta)^{t-1} + \frac{m-t}{m}\,\delta(1-\delta)^{t-1}\right] + (m-1)\frac{2}{m}(1-\delta)^{m-2}\,, \tag{A 12}$$

which simplifies to the expression in Lemma 4. Since firms have a probability δ of closing down at any period, their expected duration is $\frac{1}{\delta}$. The expected duration of a firm following the nursery strategy after finding its ideal production process (taking into account the period of idleness when relocating from a diversified to a specialised city) is $\frac{\Delta}{1-\delta}$, where Δ denotes the expected number of periods engaged in mass-production. Thus, $\overset{?}{\Delta} + \frac{\Delta}{1-\delta} = \frac{1}{\delta}$. Substituting $\overset{?}{\Delta}$ and solving for Δ yields the expression in Lemma 4.

Consider now the expected duration of each stage for a firm locating initially in a specialised city, relocating across specialised cities to try different production processes until it finds its ideal one, and then staying in a city of the relevant specialisation engaged in mass-production. In this case, for $t = 1, ..., m-2$, such a firm can cease production of prototypes after t prototypes for two reasons. First, it may find that its ideal process is the one it has decided to try in attempt number t if it gets to that point (which has the same probability of being the ideal one, $\frac{1}{m}$, as any other process). And we have to take into account that this firm will only sometimes make it to that point (this happens with probability $(1-\delta)^{2(t-1)}$, given that under this strategy every period producing a prototype is followed by a period of relocation). Second, such a firm may cease production of prototypes after t prototypes because its ideal process is one of the $m-t$ it has decided to try later (this happens with probability $\frac{m-t}{m}$), but it closes down in period $2t-1$ right after producing prototype number t or in period $2t$ while it is relocating (this happens with probability $(1-\delta)^{2(t-1)}[\delta + (1-\delta)\delta]$). Finally, a firm following this strategy that gets to produce prototype number $m-1$ does so because its ideal process is either the one it decided to try in attempt number $m-1$ or the remaining one (this happens with probability $\frac{2}{m}$), and remains in operation after $2(m-2)$ periods (this happens with probability $(1-\delta)^{2(m-2)}$). The expected number of periods producing prototypes for a firm sticking to specialised cities with this strategy is therefore

$$\overset{?}{\Delta}_{OSC} = \sum_{t=1}^{m-2} t \left[\frac{1}{m} (1-\delta)^{2(t-1)} + \frac{m-t}{m} (1-\delta)^{2(t-1)} \delta(2-\delta) \right] + (m-1) \frac{2}{m} (1-\delta)^{2(m-2)} , \quad \text{(A 13)}$$

which simplifies to the expression in Lemma 4. To calculate the expected number of periods engaged in mass production for a firm sticking to specialised cities with this strategy, first note that such a firm has a probability $\frac{1}{m}(1-\delta)^{2t-1}$ of finding its ideal process by producing its prototype number t with it, for $t = 1, ..., m-1$, and a probability $\frac{1}{m}(1-\delta)^{2(m-1)}$ of finding it by trying everything else and surviving relocation to a city specialised in the only type of process with which it never made a prototype. Conditional on starting, the expected duration of the mass-production stage is $\frac{1}{\delta}$. The unconditional expected duration of this stage for a firm following this strategy is then

$$\Delta_{OSC} = \left(\sum_{t=1}^{m-2} \frac{1}{m} (1-\delta)^{2t-1} + \frac{1}{m} (1-\delta)^{2(m-1)} \right) \frac{1}{\delta} , \quad \text{(A 14)}$$

which again simplifies to the expression in Lemma 4. □

Further deviations from the nursery strategy Besides the deviations from the nursery strategy that are directly ruled out by Conditions 1.1–1.5, there are other possible deviations that do not impose additional parameter constraints on a nursery steady-state.

Condition 1.3 guarantees that a firm will not relocate to a specialised city not knowing which of two remaining processes is its ideal one, but instead wait until it knows which of them is its ideal process. However, it could also relocate to a specialised city after producing $m - t$ prototypes without finding its ideal process, despite having another t processes left to try (where $2 \leqslant t \leqslant m - 1$). If it did so and its ideal process did not correspond to this city's specialisation, it could either keep trying to find its ideal process or it could give up.

Consider first deviations involving giving up the search for the ideal process. After producing $m - t$ prototypes in a diversified city and not finding its ideal process, a firm could relocate to a specialised city where all firms are using a type of process it has not yet tried, and stay there whatever happens (engaging in mass-production if it finds its ideal process by producing its prototype number $m - t + 1$, or producing prototypes with the same process thereafter if it does not). Such a firm has probability $1 - \delta$ of still being in operation after the period lost in relocation from the diversified to the specialised city, in which case it gets $\overset{?}{\pi}_S$ by producing its prototype number $m - t + 1$; and if it is still in operation one period later, it also gets for an expected $\frac{1}{\delta}$ periods π_S if it just found its ideal process (which happens with probability $\frac{1}{t}$) and $\overset{?}{\pi}_S$ if it did not. Thus, the expected profits over its remaining operating life for such a firm are $(1 - \delta) \left\{ \overset{?}{\pi}_S + (1 - \delta) \left[\frac{1}{t} \frac{\pi_S}{\delta} + \frac{t-1}{t} \frac{\overset{?}{\pi}_S}{\delta} \right] \right\}$. We need to compare this with the expected profits over its remaining operating life for a firm following the nursery strategy. As defined by Lemma 4, $\overset{?}{\Delta}$ is the expected number of periods of prototype production for a firm that has m processes to try and follows the nursery strategy. Thus, replacing m by t in $\overset{?}{\Delta}$ gives the expected remaining number of periods of prototype production for a firm that has t processes left to try and follows the nursery strategy. Let us denote this by $\overset{?}{\Delta}(t)$. Similarly, denote by $\Delta(t)$ the expected number of periods of mass-production for a firm that has t processes left to try and follows the nursery strategy, which results from replacing m by t in Δ. The expected profits over its remaining operating life for a firm that has t processes left to try and follows the nursery strategy is therefore $\overset{?}{\Delta}(t)\overset{?}{\pi}_D + \Delta(t)\pi_S$. A firm that has produced $m - t$ prototypes in a diversified city without finding its ideal process will therefore prefer to wait until it finds it before relocating to a specialised city, rather than relocate immediately to a specialised city and remain there, if and only if

$$(1 - \delta) \left\{ \overset{?}{\pi}_S + (1 - \delta) \left[\frac{1}{t} \frac{\pi_S}{\delta} + \frac{t-1}{t} \frac{\overset{?}{\pi}_S}{\delta} \right] \right\} \leqslant \overset{?}{\Delta}(t)\overset{?}{\pi}_D + \Delta(t)\pi_S \, . \tag{A 15}$$

Substituting (30) and (31) into (A 15) and rearranging yields

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \leqslant \frac{t\delta}{(1-\delta)(t-1+\delta)} \left[\overset{?}{\Delta}(t) + \left(\Delta(t) - \frac{(1-\delta)^2}{t\delta} \right) \frac{1-\mu}{\mu} \Omega \right] \, . \tag{A 16}$$

In principle, we would need to check that (A 16) is satisfied for $2 \leqslant t \leqslant m - 1$. However, if Conditions 1.1 and 1.2 are satisfied, the right-hand-side of (A 16) is decreasing in t: if a firm finds it profitable to deviate from the nursery strategy by giving up the search for its ideal process, it prefers to do so as early as possible. Hence, we only need to check that (A 16) is satisfied for $t = m - 1$, which involves a firm relocating to a specialised city immediately after producing the first prototype in a diversified city. But whenever this is a profitable deviation, locating initially in a specialised city and staying there whatever happens must be even better, as long as Condition 1.1 is satisfied. Consequently Condition 1.4 is the necessary and sufficient condition for this deviation not to be profitable.

Consider now deviations involving early relocation to specialised cities to continue searching for the ideal process. If a firm that leaves a diversified city before finding its ideal process does not eventually come back to a diversified city (and we show below that, with Conditions 1.1 and 1.3–1.5 satisfied, it never finds such a deviation from the nursery strategy to be profitable), then it can only try different processes by relocating from one specialised city to another between prototypes. As defined by Lemma 4, $\overset{?}{\Delta}_{OSC}$ is the expected number of periods of prototype production for a firm with m processes to try that is located in a specialised city, and intends to relocate across specialised cities to try different production processes until it finds its ideal one, and then stay in a city of the relevant specialisation engaged in mass-production. Thus, replacing m by t in $\overset{?}{\Delta}_{OSC}$ gives the expected remaining number of periods of prototype production for a firm that follows this same strategy with t processes left to try. Let us denote this by $\overset{?}{\Delta}_{OSC}(t)$. Similarly, denote by $\Delta_{OSC}(t)$ the expected number of periods of mass-production for this firm, which results from replacing m by t in Δ_{OSC}. Thus, the expected profits over its remaining operating life for a firm that relocates from a diversified to a specialised city not yet knowing which of the t processes it has not yet tried is its ideal one, with the intention of finding this by relocating across cities of different specialisation, is $(1-\delta)[\overset{?}{\Delta}_{OSC}(t)\overset{?}{\pi}_S + \Delta_{OSC}(t)\pi_S]$ (note that it only remains in operation after relocating to the first specialised city with probability $1-\delta$). This will not be a profitable deviation from the nursery strategy if and only if

$$(1-\delta)\left[\overset{?}{\Delta}_{OSC}(t)\overset{?}{\pi}_S + \Delta_{OSC}(t)\pi_S\right] \leqslant \overset{?}{\Delta}(t)\overset{?}{\pi}_D + \Delta(t)\pi_S \,. \qquad \text{(A 17)}$$

Substituting (30) and (31) into (A 17) and rearranging yields

$$\left(\frac{Q_D}{Q_S}\right)^{\sigma-1} \leqslant \frac{\overset{?}{\Delta}(t)}{(1-\delta)\,\overset{?}{\Delta}_{OSC}(t)} + \frac{\Delta(t)-(1-\delta)\,\Delta_{OSC}(t)}{(1-\delta)\,\overset{?}{\Delta}_{OSC}(t)}\,\frac{1-\mu}{\mu}\Omega \,. \qquad \text{(A 18)}$$

In principle, we would need to check that (A 18) is satisfied for $2 \leqslant t \leqslant m-1$. However, note first that that Condition 1.4 is never satisfied for $\delta \geqslant 0.5$, as long as Condition 1.1 is satisfied (for $\delta \geqslant 0.5$, since the right-hand-side of Condition 1.4 is no greater than 1, while Condition 1.1 requires the left-hand-side to be greater than 1). Then, for $\delta < 0.5$, dependence of the right hand side of (A 18) on t is such that whenever this condition is not satisfied for some t in $[2, m-1]$, it is not satisfied for one or both limits of this interval. Further, early relocation with $t = 2$ is already addressed by Condition 1.3. Hence, we only need to check that (A 18) is satisfied for $t = m-1$, which implies early relocation to a specialised city immediately after producing the first prototype in a diversified city. But note that whenever this is a profitable deviation, if Condition 1.1 is satisfied, a firm must prefer instead to locate initially in a specialised city, relocate across specialised cities to try different production processes until finding the ideal one, and then stay in a city of the relevant specialisation engaged in mass-production. Consequently, Condition 1.5 is the necessary and sufficient condition for this other deviation not to be profitable.

Thus, if Conditions 1.1–1.5 are satisfied, a firm that locates initially in a diversified city will not find it profitable to do anything else than remain there until it finds its ideal process, and then relocate to a city of the relevant specialisation to engage in mass-production. However, these conditions also guarantee that a firm will not find it profitable to locate initially in a specialised city either. If it did so, it could try a single process and, if this is not its ideal process, give up trying to find it altogether and produce prototypes with the same process thereafter. Condition 1.4 was derived as necessary and sufficient for this deviation not to be profitable. A firm locating initially in a specialised city could also search for its ideal process solely in specialised cities, which would mean relocating from one specialised city to another between prototypes in order to try different production processes. Condition 1.5 was derived as the necessary and sufficient condition for this deviation not to be profitable either.

Alternatively, a firm locating initially in a specialised city could search for its ideal process by relocating across specialised cities for up to its first t prototypes and, if it has not found its ideal process by then and is still in operation, relocate to a diversified city and follow the nursery strategy thereafter. The ideal process of a firm following this deviation from the nursery strategy is one of the t it intends to try in specialised cities with probability $\frac{t}{m}$, in which case it expects to spend in specialised cities $\frac{(1-\delta)[1-(1-\delta)^{2t}]}{t(2-\delta)\delta^2}$ periods producing prototypes, and $\frac{(1+t)(2-\delta)\delta-1+(1-\delta)^{2t}[1-(2-\delta)\delta]}{t(2-\delta)^2\delta^2}$ periods engaged in mass-production (the derivation of these expected durations follow those of $\overset{?}{\Delta}_{OSC}$ and Δ_{OSC} above, except for the lack of asymmetry in the last prototype). This firm's ideal process is instead one of the $m-t$ it intends to try in a diversified city with probability $\frac{m-t}{m}$, in which case it expects to spend $\frac{1-(1-\delta)^{2t}}{(2-\delta)\delta}$ periods producing prototypes in specialised cities, and if it gets though this stage and to a diversified city (which happens with probability $(1-\delta)^{2t}$) it expects to spend another $\overset{?}{\Delta}(m-t)$ periods producing prototypes in this diversified city, and $\Delta(m-t)$ periods engaged in mass-production in a specialised city. Thus, the expected lifetime profits of a firm following this deviation are

$$\frac{t}{m}\left(\frac{(1-\delta)\left[1-(1-\delta)^{2t}\right]}{t(2-\delta)\delta^2}\overset{?}{\pi}_S+\frac{(1+t)(2-\delta)\delta-1+(1-\delta)^{2t}\left[1-(2-\delta)\delta\right]}{t(2-\delta)^2\delta^2}\pi_S\right) \\ +\frac{m-t}{m}\left(\frac{1-(1-\delta)^{2t}}{(2-\delta)\delta}\overset{?}{\pi}_S+(1-\delta)^{2t}(\overset{?}{\Delta}(m-t)\overset{?}{\pi}_D+\Delta(m-t)\pi_S)\right) \quad \text{(A 19)}$$

With Conditions 1.1, 1.4 and 1.5 satisfied, this expression is no greater than the left-hand-side of (36), which in turn is no greater than the expected profits of a firm following the nursery strategy. This is therefore not a profitable deviation.

This brings us back to the deviation alluded to above but not formally discussed, in which a firm follows the nursery strategy for some periods, if unsuccessful continues the search for its ideal process in specialised cities for some periods, and if still unsuccessful comes back to the nursery strategy. The expected profits for such a firm from the point at which it first deviates from the nursery strategy are the result of replacing m in (A 19) by the number of processes the firms has not yet tried at that point, replacing t by the number of processes it intends to try if necessary in specialised cities, and multiplying the result by $1-\delta$. But, if Conditions 1.1 and 1.3–1.5 are satisfied, this is no greater than the left-hand-side of (A 17), which in turn we have already shown is no greater than the expected profits of continuing to follow the nursery strategy. So Conditions 1.1 and 1.3–1.5 guarantee that this deviation is not profitable either.

This also eliminates all other deviations involving location in diversified cities after some periods of prototype production in specialised cities. Given the strategy of other firms, the profitability of any strategy only depends on a firm's current location and, if it does not know its ideal process, on how many possibilities there are left. Deviations that result from replacing the nursery strategy for some other strategy after moving from a specialised to a diversified city are therefore covered by the conditions derived so far.

Therefore, if Conditions 1.1–1.5 are satisfied, with all firms following the nursery strategy, no firm finds it profitable to deviate from this strategy. □

Parameter values for Figure 1 If cities are of the same (possibly optimal) size, Conditions 1.1–1.5 depend only on three parameters other than m and δ: μ, ϵ, and σ. Figure 1 is plotted for $\mu = 0.2$ (prototypes represent 20% of the market, with mass-produced goods accounting for the remaining 80%), $\epsilon = 0.07$ (a 1% increase in the amount of labour with a certain aptitude net of commuting costs increases a city's output from that labour by 1.07%), and $\sigma = 4$ (firms mark-up marginal costs by $\frac{1}{3}$). □

Proof of Proposition 3 Let us start from the end and consider first a firm that knows its ideal production process and is located in a city of the relevant specialisation. This firm engages in mass production if and only if $\pi_S \geqslant \overset{?}{\pi}_S$. From (6), (7), (14) and (15), operational profits in the configuration with only specialised cities are:

$$\overset{?}{\pi}_S = \frac{\mu Y}{\sigma \left(m N_S \overset{?}{n}_S \right)} \, , \tag{A 20}$$

$$\pi_S = \frac{(1-\mu) Y}{\sigma \left(m N_S n_S \right)} \, . \tag{A 21}$$

By the same reasoning used for the nursery configuration, the number of prototype producers relative to the number of mass-producers in the configuration with only specialised cities is Ω_{OSC}, as given in the text. Substituting (A 20),(A 21), and Ω_{OSC} into $\pi_S \geqslant \overset{?}{\pi}_S$, this becomes

$$\frac{1-\mu}{\mu} \Omega_{OSC} \geqslant 1 \, . \tag{A 22}$$

The other possible deviations involve firms giving up the search for their ideal process. Consider a firm arriving in a specialised city after producing $m-t$ different prototypes and not finding its ideal process. This firm can stay there whatever happens. With probability $\frac{1}{t}$ the process tried by the firm is its ideal one. In this case, the firm gets $\overset{?}{\pi}_S$ for the current period and with probability $1-\delta$ survives this period and gets an expected intertemporal profit equal to $\frac{\pi_S}{\delta}$. With probability $\frac{t-1}{t}$, the process tried by the firm is not its ideal one and in that case, it gets an expected intertemporal profit equal to $\frac{\overset{?}{\pi}_S}{\delta}$. If instead the firm keeps searching for its ideal process it gets ${}_t\overset{?}{\Delta}_{OSC} \overset{?}{\pi}_S + \Delta_{OSC}(t) \pi_S$. Thus a firm does not want to give up the search for its ideal process after $m-t$ unsuccessful trials if and only if:

$$\frac{1}{t} \left(\overset{?}{\pi}_S + \frac{1-\delta}{\delta} \pi_S \right) + \frac{t-1}{t} \leqslant {}_t\overset{?}{\Delta}_{OSC} \overset{?}{\pi}_S + \Delta_{OSC}(t) \pi_S \, . \tag{A 23}$$

Substituting (A 20), (A 21) and Ω_{OSC} into (A 23) and rearranging yields:

$$\frac{1-\mu}{\mu}\Omega_{OSC} \geqslant \frac{t\left[1-\delta\overset{?}{\Delta}_{OSC}(t)\right]-1+\delta}{t\delta\Delta_{OSC}(t)-1+\delta} \; . \tag{A 24}$$

The right hand side of this equation is maximised for $t = m$, when it becomes Condition 3.1. Furthermore if this condition is satisfied, so is (A 22). □

Proof of Proposition 5 Consider first indirect utility in the nursery steady-state. Let us choose as numéraire the wage in a specialised city, w_S. Wage income is equalised across workers and is equal to expenditure so that $e = 1 - \tau L_S$. Inserting this expression and (27) into (5) yields:

$$V = \frac{F(1-\tau L_S)}{\overset{?}{\Delta}\overset{?}{\pi}_D + \Delta\pi_S} \; . \tag{A 25}$$

Substituting (25) and (31) into (A 25) and rearranging implies:

$$V = \frac{\sigma m \Omega F(1-\tau L_S)}{Y[\mu\overset{?}{\Delta} + (1-\mu)\Omega\Delta]} N_S n_S \; . \tag{A 26}$$

Then inserting (22), (23) and Ω (from Lemma 3) into (27) implies

$$\overset{?}{\Delta}\overset{?}{\pi}_D + \Delta\pi_S = F\rho^{1-\mu}\frac{\sigma}{\sigma-1} Q_D{}^{\mu} Q_S{}^{1-\mu} m^{\frac{1}{1-\sigma}} \Omega^{\frac{\mu}{1-\sigma}} (N_S n_S)^{\frac{1}{1-\sigma}} \; . \tag{A 27}$$

Substituting (25) and (31) into the previous expression yields

$$(N_S n_S)^{\frac{2-\sigma}{1-\sigma}} = \frac{Y(\sigma-1)}{F\sigma^2\rho^{1-\mu}}\left[\mu\overset{?}{\Delta} + (1-\mu)\Omega\Delta\right]\Omega^{-1-\frac{\mu}{1-\sigma}} m^{\frac{\sigma-2}{1-\sigma}} Q_D{}^{-\mu} Q_S{}^{-(1-\mu)} \; . \tag{A 28}$$

Note that this expression leads to a well-defined solution for n_S only when $\sigma > 2$. (If $\sigma < 2$, an increase in the number of firms reduces the start-up cost $\overset{?}{P}^{\mu}P^{1-\mu}F$ so much as to make further firm entry more profitable). Income in a nursery configuration can be calculated from (8), (9), (21), and (24)–(26):

$$Y = \frac{\sigma}{\sigma-1}\,(1-\tau L_S)\,L = \frac{\sigma}{\sigma-1}\left(1-\tau\frac{(1-\mu)L}{mN_S}\right)L \; . \tag{A 29}$$

Inserting (19)–(21), (A 28), (A 29), Ω (from Lemma 3), and $\overset{?}{\Delta}$ and Δ (from Lemma 4) into (A 26) gives V only as a function of parameters. Then the same exercise can be repeated for the other two configurations to obtain welfare levels under these, V_{ODC} and V_{OSC}. If Condition 1.1 is satisfied, $V \geq V_{ODC}$. If Condition 1.5 is satisfied, $V \geq V_{OSC}$.

□