FROM SECTORAL TO FUNCTIONAL URBAN SPECIALIZATION

Gilles Duranton
Diego Puga

From Sectoral to Functional Urban Specialization
Gilles Duranton and Diego Puga
NBER Working Paper No. 9112
August 2002
JEL No. R30, L23, R12

ABSTRACT

Striking evidence is presented of a previously unremarked transformation of urban structure from mainly sectoral to mainly functional specialization. We offer an explanation showing that this transformation is inextricably interrelated with changes in firms' organization. A greater variety of business services for headquarters and of sector-specific intermediates for production plants within a city reduces costs, while congestion increases with city size. A fall in the costs of remote management leads to a transformation of the equilibrium urban and industrial structure. Cities shift from specializing by sector -- with integrated headquarters and plants -- to specializing mainly by function -- with headquarters and business services clustered in larger cities, and plants clustered in smaller cities.

Gilles Duranton
Dept. of Geography and Environment
London School of Economics
Houghton Street
London WC2A 2AE
United Kingdom
g.duranton@lse.ac.uk

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

1. Introduction

Over the last few decades there has been a *shift in the main dimension along which cities specialise, from a specialisation by sector to a specialisation by function*. Urban economists have traditionally paid much attention to the specialisation of individual cities in a small number of sectors (see, e.g., Henderson, 1988). While specialisation continues to be an important feature of the urban system of the United States, cities are increasingly distinguished by their functional specialisation rather than by their sectoral specialisation. This transformation of urban structure has so far been unremarked. We provide striking evidence of it in Table 1.

The left-hand side of the table shows that sectoral specialisation within manufacturing, as measured by a Gini index, declined steadily for cities of all sizes between 1977 and 1997.[1] The median US metro area saw its Gini index of sectoral specialisation decline from .440 to .385 between 1977 and 1997. This is a very large change for such an index.[2]

The right-hand side of Table 1 shows that this falling sectoral specialisation has been mirrored by an increasing functional specialisation. In 1950 cities were not too different in terms of their proportions of managers and production workers. And although the largest cities already housed relatively more managers, there was no clear ranking by city size. By 1980 differences across cities had increased substantially and a clear ranking by size had emerged: larger cities had become specialised in management functions whereas smaller cities had become specialised in production. This pattern became even more marked over the following decade. In 1990 the nine largest US metro areas had 39% more managers per production worker than the national average. For metro areas with a current population between 1.5 and 5 million the figure was 25.7% above the national average. At the other extreme, non-metro Areas and metro areas with less than a quarter of a million people had

---

[1]Alternative measures of specialisation show a similar decline over time as well as a greater specialisation of smaller cities. Kim (1995) looks at US Census Regions instead of cities over a longer time period and finds that these have experienced a similar decline in their sectoral specialisation since the 1930s.

[2]Just to get a better idea of the size of this change, note that it is of the same order of magnitude — but of opposite sign — as the rise in a Gini index measuring US household income inequalities from .365 to .425 between 1979 and 1996, which has given rise to a large literature (Burtless, 1999).

Local population[a]	Sectoral specialisation[b]			Functional specialisation in management against production[c]			
	1977	1987	1997	1950	1970	1980	1990
5,000,000−19,397,717	.375	.369	.348	+10.2%	+22.4%	+30.8%	+39.0%
1,500,000− 4,999,999	.287	.275	.257	+ 0.3%	+16.7%	+21.7%	+25.7%
500,000− 1,499,999	.352	.338	.324	−10.9%	−10.0%	− 5.0%	− 2.1%
250,000− 499,999	.450	.409	.381	− 9.2%	− 9.7%	−10.9%	−14.2%
75,000− 249,999	.499	.467	.432	− 2.1%	− 6.6%	−12.7%	−20.7%
67− 75,000	.708	.692	.661	− 4.0%	−33.7%	−40.4%	−49.5%

Source: Authors' calculations based on data from County Business Patterns and Decennial Census of Population and Housing.

[a]Population by Metropolitan Statistical Area/Consolidated Metropolitan Statistical Area (New England County Metropolitan Area in New England), or Non-metro Area. The same areas are included in each population class throughout the table, on the basis of area definitions and population data from the Decennial Census of 2000.

[b]Median value for each population class of a Gini index comparing the local and national distributions of employment shares across 2-digit SIC manufacturing sectors. If s_h and $\bar{s}_h$ are respectively the local and national shares of employment in sector h, the Gini specialisation index is $\frac{1}{2} \sum_h |s_h - \bar{s}_h|$. Its value is close to one if a city is fully specialised in a sector that is very small at the national level and is equal to zero if local employment is dispersed across sectors in the same way as national employment.

[c]Percentage difference from the national average in the number of executives and managers per production worker (occupied in precision production, fabrication, or assembly).

Table 1. The diminishing sectoral specialisation and increasing functional specialisation of US cities

respectively 49.5% and 20.7% fewer managers per production worker than the national average. Together with management functions, larger cities have also absorbed business service employment as well as headquarters.[3]

In this paper we develop a formal framework that explains this urban transformation and suggests that it is inextricably interrelated with changes in firms' organisation. Specifically, to the *increasing separation of the management and production facilities of individual firms*. This separation is one of the most significant trends in the internal organisation of

---

[3]According to Kolko (1999), in 1910 manufacturing accounted for about six times as much employment as business services in all size-classes of cities. Over time this similarity across cities of different sizes has disappeared. There has been a growing concentration of manufacturing in small and medium-sized cities and of business services in larger cities, to the extent that the ratio of manufacturing to business service employment is now is now 4.5 times higher in non-metro areas than in metro areas with over 2.5 million people. Regarding headquarters, Shilton and Stanley (1999) study the location of over 5,000 headquarters in the US, and find that 40% of them are clustered in just 20 major urban counties.

firms of the last century (Chandler, 1977).[4] Kim (1999) carefully documents this process in his account of the rise of multi-location firms in the United States. The percentage of US manufacturing workers employed in multi-location firms increased from 51.4% in 1939 to 73.1% in 1987. A crucial component of this rise of multi-location firms has been the growing importance of separate establishments acting as headquarters. The number of stand-alone headquarters and their employment rose by 78.9% and 68.9% respectively between 1958 and 1987 — more than twice as fast as overall multi-location employment in manufacturing over the same period.

The decisions made by firms regarding their organisation (such as whether to be spatially integrated or to split management and production across multiple locations) are not made in a void. They are affected by the environment in which firms operate. And this environment is in turn shaped by the cumulation of firms' organisational choices. This paper develops a simple general equilibrium model in which firms' organisational choices are endogenous and interrelated and studies the links between firms' organisation and the urban system.

Our model (set up in Section 2) considers multiple final sectors, each made up of firms with both headquarter and production activities. Each firm gains from integrating headquarter and production in a single location because this saves in management costs. However, depending on the urban structure, there may also be gains from becoming a multi-location firm with headquarter and production establishments in different cities. This is because cities with a wider range of business service suppliers are less costly places in which to operate a headquarter. Similarly, the sharing of intermediate suppliers by production plants reduces productions costs in cities with more same-sector suppliers. At the same time, larger cities are more congested places which results in higher living costs. Workers are mobile across cities and occupations, and the number and composition of cities are endogenous.

---

[4]Another important trend has been firms' increasing reliance on outside suppliers. This growth in outsourcing has received wide attention in the literature (see Perry, 1989, for a survey) and is outside the scope of this paper, in which we take the boundaries of the firm as given and focus on the location of activities within those boundaries.

We begin to solve the model by determining the types of cities that can exist in equilibrium (Section 3). Then we derive both the equilibrium choice of organisational form by firms and the equilibrium urban structure (Section 4). When the additional costs associated with managing production from a remote headquarter are high, firms remain integrated. Given the benefits of sharing intermediate suppliers and urban congestion costs, cities host headquarters and production plants but specialise by sector. However, when the additional costs associated with managing production remotely fall below a certain level, for which we provide a closed-form solution, both the organisation of firms and the urban structure undergo profound changes. Firms previously organised as single units become multi-unit organisations. In choosing where to locate their now separate establishments, firms aim to place their headquarter in cities where business service employment is abundant and their production plant in cities with a greater same-sector specialisation in final production. Since similar organisational and location choices are made by a large number of firms, this in turn affects the employment patterns of cities, which is itself what makes the organisational choice worthwhile. This leads to a shift in the main dimension along which cities specialise, from a specialisation by sector to a specialisation by function. As a result of the benefits of sharing business service suppliers across firms and sectors, headquarters from different sectors and business services cluster in a few large cities. Manufacturing plants, for which localisation economies are weaker empirically as well as in the model, cannot afford to pay high enough wages to compensate for the high congestion costs. They move away and cluster in smaller separate cities. We conclude the paper with some final remarks and suggestions for further work (Section 5).

Of the large literature dealing with the organisational and location choices of firms, three particular streams are more closely related to our work. The theoretical literature on multinational enterprises has endogenised firms' choice of location for multiple facilities (Helpman, 1984, and Markusen, 1984, are two early examples, while Markusen, 1995, provides a survey). However, this line of work focuses on the location decisions of firms

taking the environment in which they operate as exogenous, whereas we are interested in determining firms' operating environment endogenously as a general equilibrium outcome shaped by firms' decisions. More recently, Grossman and Helpman (2002) have developed a general equilibrium framework where both the boundaries of the firm and the market structure in which they operate are endogenous. However, location choices are ignored so that there is unique environment faced by all firms. Finally, in the urban economics literature, Davis (2001) studies the location of production relative to distribution facilities and natural resources. However, his focus is on the relative magnitude of transport costs for final goods and for natural resources. More closely related to this paper is the work of Ota and Fujita (1993), who propose a model where firms make decisions about the location of their facilities and these decisions in turn affect the urban environment. However, their model deals only with location within a city, and is thus more useful to understand differences between a city's centre and its suburbs than differences across cities in an urban system.

2. The model

Preferences

There are *m* sectors producing goods for final consumption. Consumers have Cobb-Douglas preferences with equal shares of expenditure allocated to each of these sectors. We use superscripts to index sectors as well as worker occupations (described below) and subscripts to index cities. Thus, the indirect utility of a consumer working in occupation *h* in city *i* is

$$V_i = \frac{e_i^h}{P} ,$$ (1)

where e_i^h denotes individual consumption expenditure,

$$P = \prod_{h=1}^{m} \left(P^h \right)^{1/m}$$ (2)

is the global consumer price index, and P^h denotes the price of final goods in sector h. The latter are freely tradeable across cities, hence their prices are common to all cities. All workers in the economy are fully mobile across cities and occupations.

Technology

The input-output structure of each final sector is represented schematically in Figure 1. Final-goods firms have two facilities: a headquarter and a production plant. Each firm can adopt one of two organisational forms: spatially integrated or multi-location.[5] Spatially integrated firms have their headquarter and their production plant in the same city, while multi-location firms have their headquarter and their production plant in different cities. In either case, headquarters use labour (managers) and business services as inputs to produce headquarter services. Production plants then combine these headquarter services with sector-specific intermediate inputs to produce final goods. Both business services and sector-specific intermediate inputs are produced by outside suppliers with labour, and are non-tradeable across cities.[6] However, while the same business services are used by headquarters in all sectors, the intermediates used by production plants differ across sectors. For instance, the headquarters of apparel and automobile firms may work with the same banks, lawyers and advertising agencies. But the car parts purchased by automobile factories, while perhaps also used by other automobile producers, are very different inputs from the fabrics purchased by apparel factories.

It follows from this input-output structure that there are $2m + 1$ occupations for workers. We assign index 0 to workers employed in business services, and by extension to firms

---

[5]Note that throughout this paper integration is defined by the range of activities performed in the same location by individual firms (spatial integration) and not by the range of activities performed inside as opposed to outside the firm (organisational integration).

[6]Business services and sector-specific intermediate inputs are assumed non-tradeable for analytical tractability, but the qualitative conclusions of the model would be unaffected if they were instead tradeable subject to iceberg costs (this issue is further discussed in Section 4). Empirically, the proximity between business service suppliers and the headquarters of their customers appears to be very important. Kolko (1999) documents the increasing concentration of business services in the largest cities, which also attract a disproportionate share of headquarters (Shilton and Stanley, 1999). Furthermore Dekle and Eaton (1999) and Adserà (2000) provide evidence of localisation economies for business services in Japan and the US respectively.
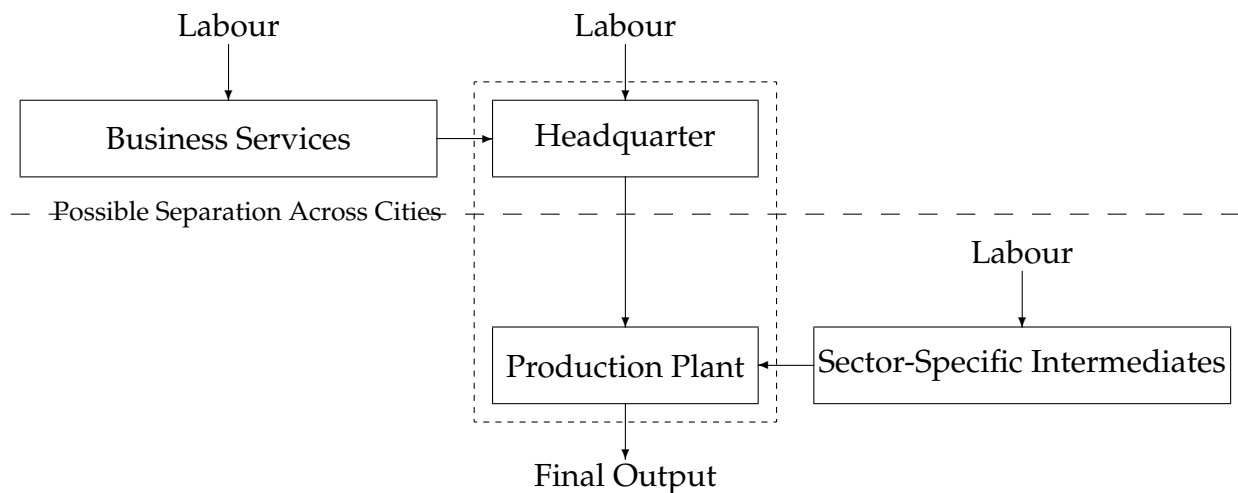
Figure 1. Input-output structure of each sector

in the business service sector; index h to workers directly employed by headquarters in sector h for $h = 1,...,m$ and to variables related to final production in that sector; and index $m + h$ to workers employed in the production of intermediate inputs specific to final sector h for $h = 1,...,m$, as well as to firms producing those intermediates.

Using this notation, we can now specify the details of firms' technology. Final-goods firms are perfectly competitive and have Cobb-Douglas technology, with cost shares η for the headquarter and $1 - \eta$ for the production plant. Thus, a sector h firm with its headquarter in city i and its production plant in city j (where j might be equal to or different from i) has a cost function given by

$$C_{i,j}^h = c_{i,j}^h \, x_{i,j}^h \, , \tag{3}$$

where $x_{i,j}^h$ denotes output,

$$c_{i,j}^h = \left(H_{i,j}^h \right)^{\eta} \left(Q_j^{m+h} \right)^{1-\eta} \, , \tag{4}$$

is the unit production cost, $H_{i,j}^h$ is the headquarter sub-cost, and Q_j^{m+h} is the production plant sub-cost.

Headquarter sub-costs are in turn split in a Cobb-Douglas fashion between labour directly employed by the headquarter and business services purchased from external suppliers, with sub-shares μ and $1 - \mu$ respectively. There is an endogenously determined mass s_i^0 of business service varieties produced in city i. All varieties enter into

7

headquarters' technology with constant elasticity of substitution $\frac{\theta+1}{\theta}$, where $\theta > 0$. Thus, for an integrated firm with its headquarter and its production plant in the same city i the headquarter sub-cost is

$$H_{i,i}^h = H_i^h = \left(w_i^h\right)^\mu \left(Q_i^0\right)^{1-\mu}, \tag{5}$$

where w_i^h denotes the wage per unit of labour for a worker employed in occupation h in city i,

$$Q_i^0 = \left\{ \int \left[q_i^0(k)\right]^{-1/\theta} \mathrm{d}k \right\}^{-\theta} \tag{6}$$

is the relevant price index of business services in city i, $q_i^0(k)$ is the price of variety k of business services produced in city i.

Each firm needs to transmit headquarter services (such as organisational decisions, marketing, or finance) to its production plant. The transmission of a firm's headquarter services to a production plant located in a different city involves iceberg costs (for instance, a fraction of managers' time is lost in travelling to visit the plant). Consequently, for a multi-location firm with its headquarter in city i and its production plant in another city j the headquarter input requirements are higher by a factor $\rho > 1$ than if it was an integrated firm with both facilities in city i :

$$H_{i,j}^h = \rho H_i^h. \tag{7}$$

This is meant to capture that, as Kim (1999) shows using US Census data, headquarter costs are significantly higher for a multi-location firm than for a similarly-sized firm with integrated headquarter and production.

Production plants for final goods use sector-specific intermediate inputs. There is an endogenously determined mass s_i^{m+h} of intermediate inputs specific to sector h produced in city i. All intermediate inputs specific to sector h enter into plants' technology with constant elasticity of substitution $\frac{\epsilon+1}{\epsilon}$, where $\epsilon > 0$. Thus, the relevant price index of intermediate inputs specific to sector h in city i is

$$Q_i^{m+h} = \left\{ \int \left[q_i^{m+h}(k)\right]^{-1/\epsilon} \mathrm{d}k \right\}^{-\epsilon}, \tag{8}$$

where $q_i^{m+h}(k)$ is the price of intermediate input k specific to sector h in city i.

Both business services and sector-specific intermediate inputs are produced by monopolistically competitive firms à la Dixit and Stiglitz (1977). The cost function of a business service firm producing variety k in city i is

$$C_i^0(k) = \left[\alpha^0 + \beta^0 y_i^0(k)\right] w_i^0 , \tag{9}$$

where $y_i^0(k)$ denotes the firm's output. The expression in brackets 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 business services. Similarly, the cost function of an intermediate firm producing variety k specific to sector h in city i is

$$C_i^{m+h}(k) = \left[\alpha + \beta y_i^{m+h}(k)\right] w_i^{m+h} . \tag{10}$$

Internal urban structure

We model the internal spatial structure of cities in the simplest standard fashion. Production in each city takes place at a central point, defined as the Central Business District (CBD). Surrounding this, there is a line with residences of unit length. Residents commute from their residence to the CBD and back, losing in the process a fraction of their unit labour supply equal to 2τ times the distanced travelled. Each consumer chooses her place of residence so as to maximise utility, given her occupation and income, and the bid-rent curve in the city. Because of fixed lot size, this is equivalent to choosing residence so as to maximise net income. Thus, a consumer working in occupation h in city i maximises $w_i^h(1 - 2\tau z) - R_i(z)$ with respect to z, where z is the distance to the CBD and $R_i(z)$ is the differential land rent in city i for a residence located at distance z from the CBD.

The possibility of arbitrage across residential locations both within and across occupations ensures that at the residential equilibrium the sum of commuting cost and land rent expenditures is the same for all residents with the same wage; that workers sort themselves according to their wage, with higher-paid workers (who have a higher opportunity cost of commuting time) living closer to the CBD; that the city is symmetric and the city edges are at a distance $L_i/2$ of the CBD (where L_i is total population in city i); and that the

bid-rent curve is continuous, concave, and piece-wise linear. Land rent at the city edges is normalised to zero. Integrating land rent over the city yields total land rent:

$$R_i = \int_{-L_i/2}^{L_i/2} R_i(z)dz. \tag{11}$$

Let us denote net labour available at the CBD of city i in occupation h by l_i^h. This is equal to L_i^h minus the total commuting time incurred by workers in this occupation. Summing across occupations immediately implies:

$$\sum_{h=0}^{2m} l_i^h = L_i(1 - \tau L_i). \tag{12}$$

The cost of headquarters and production plants

Before we proceed to describe the mechanism for city creation, let us derive in two Lemmas simplified expressions for the cost of production plants and headquarters in final-goods sectors.

Lemma 1 (The cost of production plants) *In equilibrium the production plant sub-cost of a sector h integrated firm in city i is*

$$Q_i^{m+h} = \left(l_i^{m+h}\right)^{-\epsilon} w_i^{m+h} .$$

Proof By Shephard's Lemma, demand for each variety of intermediates is

$$y_i^{m+h} = \int \frac{\partial C_{j,i}^h}{\partial q_i^{m+h}} dj = (1-\eta)\left(\frac{H_i^h}{Q_i^{m+h}}\right)^\eta \left(\frac{q_i^{m+h}}{Q_i^{m+h}}\right)^{-(\epsilon+1)/\epsilon} \int x_{j,i}^h \, dj , \tag{13}$$

where we have dropped index k since all variables take identical values for all intermediate suppliers in the same sector and city. It follows from (13) that the elasticity of demand for each intermediate with respect to a firm's own price is $-\frac{\epsilon+1}{\epsilon}$, so that the profit-maximising price for each intermediate is a fixed relative markup over marginal cost:

$$q_i^{m+h} = (\epsilon + 1)\,\beta\,w_i^{m+h} . \tag{14}$$

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

is

$$y_i^{m+h} = \frac{\alpha}{\beta\,\epsilon} \, . \tag{15}$$

Demand for labour by intermediate producers can be obtained by application of Shephard's Lemma to (10) and integration over varieties as

$$l_i^{m+h} = s_i^{m+h} \frac{\partial C_i^h}{\partial w_i^{m+h}} = s_i^{m+h} \left[\alpha + \beta y_i^{m+h} \right] = s_i^{m+h} \, \alpha \, \frac{\epsilon+1}{\epsilon} \, , \tag{16}$$

where (15) has been substituted in to yield the final expression. By choice of units of intermediate output, we can set $\beta = \left(\frac{\epsilon}{\alpha}\right)^\epsilon (\epsilon+1)^{-(\epsilon+1)}$. Using (16) and (14), the price indices of (8) simplify into

$$Q_i^{m+h} = \left[s_i^{m+h} \left(q_i^{m+h} \right)^{-1/\epsilon} \right]^{-\epsilon} = \left(l_i^{m+h} \right)^{-\epsilon} w_i^{m+h} \, . \tag{17}$$

$\square$

As in Fujita (1988) and Abdel-Rahman and Fujita (1990), in our model there is a monopolistically competitive intermediate sector that hires workers and sells non-tradeable intermediates to final-goods producers. A larger workforce employed in any given intermediate sector $m+h$ of a city leads to a wider range of intermediates being produced for use in final-good sector h. Since these intermediates enter the production function of final-good producers in sector h with the same constant elasticity of substitution $\frac{\epsilon+1}{\epsilon}$, a wider range of available intermediates results in lower production costs in this final-good sector. Hence, despite constant returns to scale at the firm level in final production, there are aggregate increasing returns at the sector-city level, also known as localisation economies. The strength of these increases with ϵ.

Lemma 2 (The cost of headquarters) *In equilibrium the headquarter sub-cost of a sector h integrated firm in city i is*

$$H_i^h = \left(w_i^h \right)^\mu \left(Q_i^0 \right)^{1-\mu} ,$$

where

$$Q_i^0 = \left(l_i^0 \right)^{-\theta} w_i^0 \, .$$

The proof of this Lemma mirrors closely that of Lemma 1 and is thus relegated to Appendix A. The sharing of business services by headquarters, like the the sharing of intermediate suppliers by plants, gives rise to localisation economies, whereby a larger workforce employed in a city's business service sector leads to a wider range of business services being produced, which in turn results in lower costs of operating headquarters.[7] Since business services account for a fraction $1 - \mu$ of headquarter costs, these localisation economies increase with $\theta(1 - \mu)$. Furthermore, since by (7) multi-location firms face the same price index for business services as integrated firms, these localisation economies apply under either organisational form. Note that, since business service suppliers are shared by firms in different sectors, an increase in business service employment in a city reduces costs for local headquarters from all sectors. An increase in employment by intermediate suppliers, on the other hand, only reduces costs for local production plants in the sector using those intermediates.

City formation

To complete the model, we need to specify a mechanism for city formation. The simplest alternative is to consider perfectly competitive land development companies or land developers.[8] At the same time, active municipal governments with tax raising powers can play an equivalent role that would lead to identical results.[9] In fact, all we require to obtain our results is some mechanism that prevents urban configurations that are sustainable only under a coordination failure from arising in equilibrium — for instance, situations in which no-one moves away from inefficiently large cities for fear of no-one else following.

---

[7]Headquarters and business service producers may face lower costs in environments where there is a large presence of both activities for a variety of other reasons, such as drawing in part from a common pool of workers (Helsley and Strange, 1990) or human capital interactions (Jovanovic and Rob, 1989; Glaeser, 1999). Nevertheless, we specify the same micro-foundations for localisation economies in business services/headquarters as in intermediates/production plants to keep the model as simple as possible.

[8]There is ample evidence regarding the pervasive role played by 'large' private agents in the United States land market. Most of Garreau's (1991) examples of edge-cities were originally the product of decisions by a single large agent (see Henderson and Mitra, 1996, for complementary evidence on this).

[9]See Fujita (1989) and Henderson and Becker (2000) for discussions of this issue and for an equivalence result between these two types of institutions.

For simplicity, we also want to avoid lumpiness problems (such as sectors being smaller than the smallest city).

We thus assume a large economy in which there is a continuum of land developers, not all of which will be active in equilibrium. Each developer controls all land in several potential sites for a city. Land developers can give occupation-specific subsidies to workers.[10] Let us denote by T_i^h the subsidy given by the developer in city i to workers with occupation h. There is free entry and perfect competition amongst land developers. When active, each land developer seeks to maximise total land rent in its city, as given by (11), net of any transfers paid to workers in each occupation:

$$\max_{\{T_i^h\}} \Pi_i , \qquad \Pi_i = R_i - \sum_{h=0}^{2m} T_i^h L_i^h , \tag{18}$$

subject to workers in each active occupation in the city obtaining the highest consumption income available elsewhere, $\bar{e}$:

$$L_i^h e_i^h = L_i^h \bar{e} , \qquad L_i^h \geqslant 0 , \tag{19}$$

and subject also to firms with only their headquarter, only their production plant, or both facilities, in the city breaking even:

$$\left(\int_{i \neq j} x_{i,j}^h \, dj \right) \left(P^h - \left(\rho H_i^h \right)^\eta \left(\underline{Q}^{m+h} \right)^{1-\eta} \right) = 0 , \quad P^h - \left(\rho H_i^h \right)^\eta \left(\underline{Q}^{m+h} \right)^{1-\eta} \leqslant 0 , \quad \int_{i \neq j} x_{i,j}^h \, dj \geqslant 0 , \tag{20}$$

$$\left(\int_{i \neq j} x_{j,i}^h \, dj \right) \left(P^h - \left(\rho \underline{H}^h \right)^\eta \left(Q_i^{m+h} \right)^{1-\eta} \right) = 0 , \quad P^h - \left(\rho \underline{H}^h \right)^\eta \left(Q_i^{m+h} \right)^{1-\eta} \leqslant 0 , \quad \int_{i \neq j} x_{j,i}^h \, dj \geqslant 0 , \tag{21}$$

$$x_{i,i}^h \left(P^h - \left(H_i^h \right)^\eta \left(Q_i^{m+h} \right)^{1-\eta} \right) = 0 , \quad P^h - \left(H_i^h \right)^\eta \left(Q_i^{m+h} \right)^{1-\eta} \leqslant 0 , \qquad x_{i,i}^h \geqslant 0 , \tag{22}$$

where the integral $\int_{i \neq j} x_{i,j}^h \, dj$ aggregates the output of multi-location firms across all cities, and $\underline{Q}^{m+h}$ denotes the lowest sub-cost for a final-sector stand-alone plant elsewhere and

---

[10]See Helsley and Strange (1997) and Henderson and Becker (2000) for discussions of the instruments of control and the extent of ownership of land developers.

$\rho \underline{H}^h$ denotes the lowest sub-cost for a final-sector stand-alone headquarter elsewhere. Note that the worker participation constraint of (19) is binding because of free worker mobility, which ensures that in equilibrium workers collect the same net returns everywhere. Similarly, the firm participation constraints of (20)–(22) are also binding.

Consumers are Arrow-Debreu shareholders in all firms and in all land development companies. However, free entry by firms and land developers drives their profits to zero in equilibrium. Thus, the equilibrium consumption expenditure of a worker with occupation h in city i living at a distance z from the CBD is the sum of wage income and any transfers, minus land rent expenditures:

$$e_i^h(z) = w_i^h(1 - 2\tau z) + T_i^h - R_i(z) . \tag{23}$$

To solve the model, it is useful to note that the developer's programme is equivalent to a situation in which the developer maximises the gap between the pre-transfer consumption expenditure of workers in the city and the post-transfer income which they could obtain elsewhere, subject to zero profit by firms and a feasibility constraint.

Lemma 3 (Equivalent developer's programme) *The developer's programme is equivalent to*

$$\max_{\{l_i^h\}} \Pi_i , \qquad \Pi_i = \sum_{h=0}^{2m} w_i^h l_i^h - L_i \bar{e} ,$$

subject to (20)–(22), $l_i^h \geqslant 0$ and (12).

Proof Integrating (23) over population in the city and using (11) and (12) yields $L_i^h e_i^h = \sum_{h=0}^{2m} w_i^h l_i^h + \sum_{h=0}^{2m} T_i^h L_i^h - R_i$. Substituting the worker participation constraint of the original programme, $L_i^h e_i^h = L_i^h \bar{e}$, into this expression implies transfers such that $\sum_{h=0}^{2m} T_i^h L_i^h = R_i - \sum_{h=0}^{2m} w_i^h l_i^h + L_i^h \bar{e}$. Inserting this into the original programme (18) yields the result. $\quad\square$

By lowering transfers down to the point where every worker gets $\bar{e}$, a developer can extract any surplus consumption income of workers in its city over that available elsewhere. Hence it is in the developer's interest to design the transfer structure so as to achieve an urban configuration such that workers obtain the highest possible pre-transfer consumption income, thus maximising the total wage bill in the city.

3. Equilibrium types and sizes of cities

In this section we show that in equilibrium there can be at most three types of cities: cities with only headquarters and their business service suppliers, cities with only production plants in one sector and their intermediate suppliers, and cities with only integrated firms in one sector and their business service and intermediate suppliers. This is done in three steps. First, we show that in equilibrium no city has both headquarters and plants belonging to multi-location firms (Lemma 4). Second, no city hosts production plants in more than one sector (Lemma 5). And third, a city with integrated firms has neither headquarters nor production plants belonging to multi-location firms (Lemma 6). We then proceed in the following section to study which of these three possible types of cities exist in equilibrium depending on parameters.

Lemma 4 (Multi-location headquarter and plant separation) *In equilibrium, headquarters and plants belonging to multi-location firms do not coexist in the same city.*

Proof Consider a city with stand-alone headquarters and production plants in the same sector. Then, from (20) and (21), $H_i^h = \underline{H}^h$ and $Q_i^{m+h} = \underline{Q}^{m+h} = \left(P^h\right)^{\frac{1}{1-\eta}}\left(\rho\underline{H}^h\right)^{-\frac{\eta}{1-\eta}}$. But then an integrated firm operating in the same city and sector would make profits equal to $P^h - \left(H_i^h\right)^{\eta}\left(Q_i^{m+h}\right)^{1-\eta} = \left(1 - \frac{1}{\rho^{\eta}}\right)P^h > 0$, which contradicts (22). Consider now a city with stand-alone production plants in sector h and stand-alone headquarters in a different sector h'. By Lemmas 1, 2, and 3, a developer can transfer headquarter resources from sector h' to sector h without affecting localisation economies in the production of either business services or sector-specific intermediates, nor its own profits. This will create the same profit opportunity from firm integration, which the developer can appropriate by lowering transfers. □

The headquarters of multi-location firms only locate in a city if it provides the lowest headquarter cost available anywhere. Similarly, the production plants of multi-location firms only locate in a city if it provides the corresponding intermediates at the lowest cost. If a city were to provide both the lowest headquarter cost and the lowest costs for

15

production plants for multi-location firms in the same sector and they broke even, then multi-location firms with either a headquarter or a production plant in the city would make positive profits by becoming integrated firms — thus saving on the transmission cost for headquarter services. The local developer could then capture the rents from this unexploited profit opportunity by lowering transfers, thus forcing firms to pay higher wages so that worker income stays at $\bar{e}$. Therefore, in equilibrium, headquarters and plants belonging to multi-location firms in the same sector do not coexist in the same city. Since resources used by headquarters, either directly or indirectly, can be shifted across sectors without changing any wages, headquarters and plants belonging to multi-location firms do not coexist in the same city even for different sectors. In effect, multi-location firms exist to exploit the advantages of different types of cities for headquarters and production plants, so the headquarters and production plants of multi-location firms are always located in different cities.

Lemma 5 (Production specialisation) *In equilibrium, each city hosts production plants in at most one sector.*

Proof Consider a developer choosing how to allocate across sectors the local amount of labour engaged in intermediate production. By Lemma 3, this involves maximising $\sum_{h=1}^{m} w_i^{m+h} l_i^{m+h}$ subject to (20)–(22) and keeping $\sum_{h=1}^{m} l_i^{m+h}$ constant. By Lemma 1, $w_i^{m+h} = Q_i^{m+h} \left(l_i^{m+h}\right)^{\epsilon}$. Changing the allocation of labour across occupations m to $2m$ in any city does not change any price, P^h, nor any headquarter cost, H_i^h. Consequently, by (20)–(22), it does not affect the Q_i^{m+h} consistent with zero profits. Hence, in choosing how to allocate across sectors the local amount of labour engaged in intermediate production, the developer maximises

$$\sum_{h=1}^{m} w_i^{m+h} l_i^{m+h} = \sum_{h=1}^{m} Q_i^{m+h} \left(l_i^{m+h}\right)^{1+\epsilon} \tag{24}$$

for given Q_i^{m+h} and $\sum_{h=1}^{m} l_i^{m+h}$. The convexity of this expression implies that all labour producing intermediates in any given city is allocated to a single sector. $\square$

This result is reminiscent of Henderson (1974). Recall that, because of its ability to extract any surplus consumption income of workers in its city, each developer behaves as if maximising the total wage bill in the city. With no equilibrium firm profits this implies maximising the value of production in the city. In the presence of localisation economies, final output in any given city and sector rises more than proportionately with employment in that city and sector. Maximisation of the value of production thus implies a corner solution for the allocation across sectors of labour employed in intermediate production. Consequently, any city that hosts plants hosts them in only one sector.[11]

Lemma 6 (Specialisation by type of firms) *In equilibrium, each city hosts either only final-sector headquarters and business service suppliers, only final-sector manufacturing plants and intermediate suppliers, or only integrated firms plus business services and intermediate suppliers.*

Proof By Lemma 2, $\left(w_i^h\right)^\mu \left(w_i^0\right)^{1-\mu} = H_i^h \left(l_i^0\right)^{\theta(1-\mu)}$. By Shephards's Lemma, $\sum_{h=1}^m w_i^h l_i^h = \frac{\mu}{1-\mu} w_i^0 l_i^0$. Combining these two equations yields $w_i^0 l_i^0 = \left(\frac{\mu}{1-\mu}\right)^{-\mu} \left(l_i^0\right)^{(1+\theta)(1-\mu)} \left[\sum_{h=1}^m \left(H_i^h\right)^{1/\mu} l_i^h\right]^\mu$. By inspection of this expression we can see that profit maximisation by land developers requires $H_i^h = H_i$ for all h such that $l_i^h > 0$. Hence

$$\sum_{h=0}^m w_i^h l_i^h = \mu^{-\mu}(1-\mu)^{-(1-\mu)} H_i \left(l_i^0\right)^{(1+\theta)(1-\mu)} \left(\sum_{h=1}^m l_i^h\right)^\mu. \tag{25}$$

Maximisation of this component of the developer's programme for given $\sum_{h=0}^m l_i^h$ shows that

$$\sum_{h=1}^m l_i^h = \frac{\mu}{(1+\theta)(1-\mu)} l_i^0. \tag{26}$$

Substituting (24)–(26) and Lemma 5 into the developer's programme of Lemma 3, we can rewrite this as

$$\max_{\{l_i^0, l_i^h, l_i^{m+h}\}} \Pi_i, \qquad \Pi_i = \frac{(1+\theta)^{-\mu}}{1-\mu} H_i \left(l_i^0\right)^{1+\theta(1-\mu)} + Q_i^{m+h} \left(l_i^{m+h}\right)^{1+\epsilon} - L_i \bar{e}, \tag{27}$$

subject to (20)–(22), $l_i^h \geqslant 0$ and (12).

---

[11] In practice, while many cities are narrowly specialised, some are quite diversified. However, this has much to do with the relative advantages of diversity and specialisation for innovation and production, an issue which is outside the scope of this paper (see Duranton and Puga, 2001).

From Lemmas 4 and 5, only five types of cities are possible: the three mentioned in the Lemma, plus cities with both production plants of multi-location firms and integrated firms, and cities with both headquarters of multi-location firms and integrated firms. To prove the Lemma, it is sufficient to show that the last two types of cities are incompatible with profit maximisation by the developers. By (20)–(22), in a city with both integrated firms and production plants of multi-location firms, $\left(\rho \underline{H}^h\right)^\eta \left(Q_i^{m+h}\right)^{1-\eta} = \left(H_i\right)^\eta \left(Q_i^{m+h}\right)^{1-\eta} = P^h$, and hence $H_i = \rho \underline{H}^h$ and $Q_i^{m+h} = \left(P^h\right)^{\frac{1}{1-\eta}} \left(\rho \underline{H}^h\right)^{-\frac{\eta}{1-\eta}} = \underline{Q}^{m+h}$. Similarly, in a city with both integrated firms and headquarters of multi-location firms, $\left(\rho H_i\right)^\eta \left(\underline{Q}^{m+h}\right)^{1-\eta} = \left(H_i\right)^\eta \left(Q_i^{m+h}\right)^{1-\eta} = P^h$, and hence $H_i = \frac{1}{\rho}\left(P^h\right)^{\frac{1}{\eta}} \left(\underline{Q}^{m+h}\right)^{-\frac{1-\eta}{\eta}} = \underline{H}^h$ and $Q_i^{m+h} = \rho^{\frac{\eta}{1-\eta}} \underline{Q}^{m+h}$. In either case, H_i and Q_i^{m+h} are pinned down by the firm participation constraints, and hence they are taken as given by the developer when implementing the programme of (27). By inspection of (27), given that $\theta > 0$ and $\epsilon > 0$, a configuration with both integrated firms and production plants of multi-location firms in the same city cannot be an equilibrium, since the developer would do better either by having only production plants or by having no production plants except those needed to have integrated firms. Similarly, a configuration with both integrated firms and headquarters of multi-location firms in the same city cannot be an equilibrium, since the developer would do better either by having only headquarters or by having no headquarters except those needed to have integrated firms. □

Localisation economies in the production of both business services and sector-specific intermediates imply that a profit-maximising developer will avoid mixing headquarters and production plants in a city, except possibly if these belong to integrated firms — due to the lower headquarter requirements associated with integrated firms. Having reduced the possible city types to three, we can now derive the equilibrium size of the different types of cities.

Lemma 7 (Equilibrium city sizes) *In equilibrium, cities hosting only final-sector production plants and intermediate suppliers (if any) have a population*

$$L_M = \frac{\epsilon}{\tau(1 + 2\epsilon)} \; ;$$

cities hosting only final-sector headquarters and business service suppliers (if any) have a population

$$L_H = \frac{\sigma}{\tau(1 + 2\sigma)} \; ;$$

and cities hosting only integrated firms plus business services and intermediate suppliers (if any) have a population

$$L_I = \frac{\gamma}{\tau(1 + 2\gamma)} \; ;$$

where $\sigma \equiv \theta(1 - \mu)$, and $\gamma \equiv \eta\theta(1 - \mu) + (1 - \eta)\epsilon$.

Proof Denote by subscript M variables corresponding to a city hosting only production plants for multi-location firms. From (12),

$$l_M^{m+h} = L_M^{m+h}(1 - \tau L_M^{m+h}) \; . \tag{28}$$

The developer's programme of (27), for the case of a city hosting only production plants for multi-location firms in sector h, can thus be expressed as

$$\max_{\{L_M^{m+h}\}} \Pi_M^{m+h} \; , \qquad \Pi_M^{m+h} = \underline{Q}^{m+h} \left[L_M^{m+h}(1 - \tau L_M^{m+h}) \right]^{1+\epsilon} - L_M^{m+h} \bar{e} \; . \tag{29}$$

The first order condition for (29) implies that workers consumption expenditure equals the value of their social marginal product

$$\bar{e} = (1 + \epsilon)\underline{Q}^{m+h} \left[L_M^{m+h}(1 - \tau L_M^{m+h}) \right]^{\epsilon} (1 - 2\tau L_M^{m+h}) \; . \tag{30}$$

Free entry of land developers exhausts their profits in equilibrium. Substituting (30) into (29) and using $\Pi_M^{m+h} = 0$ yields equilibrium population for any city hosting only production plants as

$$L_M = L_M^{m+h} = \frac{\epsilon}{\tau(1 + 2\epsilon)} \; . \tag{31}$$

Similarly, denote by subscript H variables corresponding to a city hosting only headquarters for multi-location firms. From (12) and (26),

$$l_H^0 = \frac{1 + \sigma - \mu}{1 + \sigma} L_H^h (1 - \tau L_H^h) , \tag{32}$$

$$l_H^h = \frac{\mu}{1 + \sigma} L_H^h (1 - \tau L_H^h) . \tag{33}$$

By symmetry, $\underline{H}^h = \underline{H}$ for all sectors. Substituting (32) into (27), the programme for the developer of a city hosting only headquarters for multi-location firms is

$$\max_{\{L_H\}} \Pi_H , \qquad \Pi_H = \frac{(1 + \theta)^{1 + \sigma - \mu}(1 - \mu)^\sigma}{(1 + \sigma)^{1 + \sigma}} \underline{H}[L_H(1 - \tau L_H)]^{1 + \sigma} - L_H \bar{e} . \tag{34}$$

The first order condition for (34) together with $\Pi_H = 0$ yields equilibrium population for any city hosting only headquarters as

$$L_H = \frac{\sigma}{\tau(1 + 2\sigma)} . \tag{35}$$

Finally, denote by subscript I variables corresponding to a city hosting only integrated firms. From (4), and Lemmas 2 and 1, the equilibrium unit production cost for integrated firms in sector h is

$$c_I^h = (w_I^h)^{\eta\mu}(w_I^0)^{\eta(1-\mu)}(w_I^{m+h})^{1-\eta}\left(l_I^0\right)^{-\eta\sigma}\left(l_I^{m+h}\right)^{-(1-\eta)\epsilon} . \tag{36}$$

By Shephard's Lemma,

$$w_I^0 l_I^0 = \frac{1 - \mu}{\mu} w_I^h l_I^h = \frac{\eta(1 - \mu)}{1 - \eta} w_I^{m+h} l_I^{m+h} . \tag{37}$$

From (26), Lemma 5, and (37),

$$\frac{w_I^h}{w_I^0} = 1 + \theta , \qquad \frac{w_I^{m+h}}{w_I^0} = \frac{(1 - \eta)l_I^0}{\eta(1 - \mu)l_I^{m+h}} . \tag{38}$$

Substituting (38) into (36) yields

$$c_I^h = w_I^0 l_I^0 (1 + \theta)^{\eta\mu} \left(\frac{1 - \eta}{\eta(1 - \mu)}\right)^{1-\eta} \left(l_I^0\right)^{-\eta(1+\sigma)} \left(l_I^{m+h}\right)^{-(1-\eta)(1+\epsilon)} . \tag{39}$$

Zero profits for final firms imply that, if in equilibrium any city has only integrated firms in sector h, $c_I^h = P^h$. We can thus use this and (39) to solve for $w_I^0 l_I^0$ as a function of l_I^0, l_I^{m+h},

P^h, and parameters. From (37), we obtain

$$w_I^0 l_I^0 + w_I^h l_I^h + w_I^{m+h} l_I^{m+h} = \frac{1}{\eta(1-\mu)} w_I^0 l_I^0 = \frac{P^h \left(l_I^0\right)^{\eta(1+\sigma)} \left(l_I^{m+h}\right)^{(1-\eta)(1+\epsilon)}}{\eta^\eta (1-\eta)^{1-\eta}(1+\theta)^{\eta\mu}(1-\mu)^\eta} . \qquad (40)$$

And by (12) and (26), $l_I^0 + l_I^h + l_I^{m+h} = \frac{1+\sigma}{1+\sigma-\mu} l_I^0 + l_I^{m+h} = L_I^h(1-\tau L_I^h)$, where L_I^h denotes total population in this city. Hence, the developer's programme of (27) for the case of a city hosting only integrated firms can be expressed as

$$\max_{\{l_I^0, l_I^{m+h}\}} \Pi_I^h, \qquad \Pi_I^h = \frac{P^h \left(l_I^0\right)^{\eta(1+\sigma)} \left(l_I^{m+h}\right)^{(1-\eta)(1+\epsilon)}}{\eta^\eta (1-\eta)^{1-\eta}(1+\theta)^{\eta\mu}(1-\mu)^\eta} - L_I^h \bar{e}, \qquad (41)$$

subject to

$$\frac{1+\sigma}{1+\sigma-\mu} l_I^0 + l_I^{m+h} = L_I^h(1-\tau L_I^h) . \qquad (42)$$

The first order conditions for (41) together with (41), (42), and $\Pi_I^h = 0$ yield equilibrium population for any city hosting only headquarters as

$$L_I = L_I^h = \frac{\gamma}{\tau(1+2\gamma)} , \qquad (43)$$

and its split between occupations as

$$l_I^0 = \frac{\eta(1+\sigma-\mu)}{1+\gamma} L_I(1-\tau L_I) , \qquad (44)$$

$$l_I^h = \frac{\eta\mu}{1+\gamma} L_I(1-\tau L_I) , \qquad (45)$$

$$l_I^{m+h} = \frac{(1-\eta)(1+\epsilon)}{1+\gamma} L_I(1-\tau L_I) . \qquad (46)$$

□

As is standard in the literature, there is a tradeoff between commuting costs (τ) and the extent of aggregate increasing returns in the city (ϵ, σ, or γ). Here aggregate increasing returns arise despite constant-returns technology in final production due to the sharing of intermediate suppliers in cities with production plants of multi-location firms, due to the sharing of business service suppliers in cities with headquarters of multi-location firms, and due to the sharing of both types of suppliers in cities with integrated firms. The ability of competitive developers to control the composition of the workforce in their city

and to extract any surplus makes them efficient. At the competitive equilibrium, cities will thus be of optimal size conditional on their type.[12] This size increases with the intensity of increasing returns and decreases with the cost of commuting. To achieve optimal city size, developers must make transfers to local workers so as to cover the gap between the value of their private and social marginal products — as can be seen from (30). With zero profits for developers, total land rents equal total transfers, and thus are just enough to cover that gap. This is known as the Henry George Theorem and is a standard result in urban economics, both in the context of production subject to local aggregate increasing returns (Serck-Hanssen, 1969; Starrett, 1974; Vickrey, 1977) and in the context of local public goods (Flatters, Henderson, and Mieszkowski, 1974; Stiglitz, 1977; Arnott and Stiglitz, 1979).

Lemma 8 (The Henry George Theorem) *In equilibrium, all cities achieve optimal size, and developers transfer all land rents in their city to local workers, filling the gap between the private and the public marginal product of labour.*

4. From sectoral to functional urban specialisation

It is now possible to derive our main result. The choice of organisational form by firms as well as the urban structure depend on how much higher is the cost of providing headquarter services to a plant in a different city as compared to a plant in the same city (ρ).

---

[12]It is only because the economy is large that cities can reach their optimal size regardless of the urban configuration (i.e., the number of cities of each type) . Having instead a small economy creates a lumpiness problem whereby cities in equilibrium are typically larger than their optimal size (since any size below optimum city size is unstable) and the relative inefficiency of equilibrium city sizes depends on the configuration. Thus, in a small economy the main result derived in Section 4 below would not depend on the comparison of the optimal city sizes but on the comparison of the optimal sizes subject to the lumpiness constraint (which converge to the the optimal sizes as the economy grows arbitrarily large). In the more pathological case of very small sectors consuming a large share of business services, Lemma 5 may not even hold as some cities may find it worthwhile to host production plants of integrated firms in two or more sectors so that they can share the costs of business services. For further discussion of these issues, see Papageorgiou and Pines (2000).

Proposition 1 (Equilibrium firm organisation and urban structure) *If $\rho > \hat{\rho}$, where*

$$\hat{\rho} \equiv \left(\frac{(1+2\gamma)^{1+2\gamma}\sigma^{\eta\sigma}\epsilon^{(1-\eta)\epsilon}}{\gamma^{\gamma}[\eta^{\sigma}(1+2\sigma)^{1+2\sigma}]^{\eta}[(1-\eta)^{\epsilon}(1+2\epsilon)^{1+2\epsilon}]^{1-\eta}} \right)^{\frac{1}{\eta}},$$

then all firms adopt an integrated organisational form, and all cities specialise by sector, hosting headquarters and production plants plus business service and intermediate suppliers. If instead $\rho < \hat{\rho}$ then all firms adopt a multi-location organisational form, and all cities specialise by function, with a share $\frac{\eta(1+2\sigma)\epsilon}{\eta(1+2\sigma)\epsilon+(1-\eta)(1+2\epsilon)\sigma}$ of cities hosting headquarters from any sector plus business service suppliers, and the remaining cities hosting production plants plus intermediate suppliers from the same sector, with equal proportions of cities specialised in production in each of the m sectors.

Proof From (4), (7), and Lemmas 2 and 1, the equilibrium unit production cost for multi-location firms is

$$c_M^h = \rho^{\eta}(w_H^h)^{\eta\mu}(w_H^0)^{\eta(1-\mu)}(w_M^{m+h})^{1-\eta}\left(l_H^0\right)^{-\eta\sigma}\left(l_M^{m+h}\right)^{-(1-\eta)\epsilon}. \tag{47}$$

Dividing this by the unit production cost for integrated firms of (36) yields

$$\frac{c_M^h}{c_I^h} = \left(\frac{\rho}{\hat{\rho}}\right)^{\eta}, \tag{48}$$

where

$$\hat{\rho} \equiv \left[\left(\frac{w_I^h}{w_H^h}\right)^{\eta\mu}\left(\frac{w_I^0}{w_H^0}\right)^{\eta(1-\mu)}\left(\frac{w_I^{m+h}}{w_M^{m+h}}\right)^{1-\eta}\left(\frac{l_I^0}{l_H^0}\right)^{-\eta\sigma}\left(\frac{l_I^{m+h}}{l_M^{m+h}}\right)^{-(1-\eta)\epsilon}\right]^{\frac{1}{\eta}}. \tag{49}$$

The firm participation constraints of (20)–(22) imply that to be sustainable cities must offer firms the possibility of producing at the lowest available unit cost. If $\rho > \hat{\rho}$ then $c_M^h > c_I^h$, all firms adopt an integrated organisational form, and every city hosts headquarters and production plants, which by Lemma 5 belong to the same sector, plus their business service and intermediate suppliers. If instead $\rho < \hat{\rho}$ then $c_M^h < c_I^h$, all firms adopt a multi-location organisational form, and every city either hosts headquarters from any sector plus their business service suppliers or hosts production plants, which by Lemma 5 belong

23

to the same sector, plus their intermediate suppliers. By Shephard's Lemma, $w_H^0 l_H^0 = \frac{1-\mu}{\mu} w_H^h l_H^h$. From this, (26), and Lemma 5,

$$\frac{w_H^h}{w_H^0} = 1 + \theta \, . \tag{50}$$

The developer's programme of Lemma 3, together with the exhaustion of developers' profits due to free entry imply

$$\frac{1}{\eta(1-\mu)} w_I^0 l_I^0 = L_I \bar{e} \, , \qquad \frac{1}{1-\mu} w_H^0 l_H^0 = L_H \bar{e} \, , \qquad w_M^{m+h} l_M^{m+h} = L_M \bar{e} \, . \tag{51}$$

From (38), (50), and (51),

$$\frac{w_I^h}{w_H^h} = \frac{w_I^0}{w_H^0} = \frac{\eta L_I \, l_H^0}{L_H \, l_I^0} \, , \qquad \frac{w_I^{m+h}}{w_M^{m+h}} = \frac{(1-\eta) L_I \, l_M^{m+h}}{L_M \, l_I^{m+h}} \, . \tag{52}$$

Substituting (28), (32), (44), (46), (52), and Lemma 7, into (49), and simplifying yields the value of $\hat{\rho}$ given in the Proposition. Denote by N_H the mass of cities that host headquarters of multi-location firms and by N_M the mass of cities that host production plants of multi-location firms when $\rho < \hat{\rho}$. By Shephard's Lemma, $\frac{w_H^0 l_H^0 N_H}{w_M^{m+h} l_M^{m+h} N_M} = \frac{\eta(1-\mu)}{1-\eta}$. Combining this with (51) yields the share of cities that host only headquarters when $\rho < \hat{\rho}$ as $\frac{\eta(1+2\sigma)\epsilon}{\eta(1+2\sigma)\epsilon + (1-\eta)(1+2\epsilon)\sigma}$. The remaining cities specialise in production and, by symmetry, there are equal proportions of cities specialised in each of the m sectors. $\qquad \square$

Discussion

In our model, firms face a trade-off between the gains from integrating their headquarter and production plant and the gains from separating them. The gains from integration stem from lower headquarter requirements. The gains from separation arise from the lower costs of operating headquarters in cities with a higher specialisation in business services, and the lower costs of operating production plants in cities with a higher specialisation in the corresponding intermediates.

When the transmission cost for headquarter services is high, firms remain integrated. Because of localisation economies in production and urban congestions costs increasing

24

with size, this in turn leads cities to specialise by sector. In contrast, when the transmission cost for headquarter services is low, firms wish to locate their headquarter in cities where business service employment is abundant and their production plant in cities with a greater same-sector specialisation in final production. And it is in the interest of profit-maximising developers to make it feasible for firms to implement this separation. Since the same organisational choice is made by a large number of firms, it in turn affects the employment patterns of cities, which is itself what makes the organisational choice worthwhile. As a result, some cities specialise in headquarters and business services and others in final and intermediate production, the latter also being specialised by sector.

Many business and economic historians have argued that the extra costs of coordinating and monitoring multi-location firms relative to integrated firms has decreased very significantly following key technological developments in transport and communication technologies, as well as new management practices (Chandler, 1977; Kim, 1999). The cost of transporting goods, people, and ideas has declined dramatically over the last century. For instance, maritime freight in 1990 was only one third as costly as in 1920, while for air-travel the revenue per passenger-mile in 1990 was one sixth of what it was in 1930. Large as they may seem, these changes are dwarfed by the free fall in telecommunication costs: the cost of a three-minute telephone conversation between New York and London fell by 98.7% between 1930 and 1990 (Jones, 1997). According to Yates (1989; 1991), managerial and accounting innovations have played an even greater role than the sheer reduction in the cost of physical distance. Of particular importance has been the development of duplicators of various types, such as the typewriter together with the use of carbon copies, and later the photocopying and fax machines, and e-mail. Such duplicators have made the transmission of information along a managerial hierarchy and across space much faster and much cheaper. Other recent developments in information technologies, such as online stock monitoring and ordering have also facilitated management over large distances. On the whole, it seems that this reduction in the cost of coordination and monitoring across distant locations can to a very large extent be taken as exogenous to the rise of

multi-location firms.

This reduction in the additional costs related to remote management can be simply represented in the model as a reduction in ρ. Figure 2 plots the unit production cost of multi-location firms (c_M) relative to the unit production cost of integrated firms (c_I) as a function of ρ.[13] When ρ is above $\hat{\rho}$ any developers attempting to set up cities with either only headquarters or only production plants cannot attract firms in the face of the competition from developers in cities hosting only integrated firms, which offer lower unit production costs. Thus every firm adopts an integrated organisational structure and every city hosts headquarters and production plants from one sector. By Proposition 1, when ρ falls below $\hat{\rho}$ cities with either only headquarters or only production plants of multi-location firms offer a combined unit cost below that available in integrated cities. Firms previously organised as a single unit become multi-unit organisations. Multi-unit firms locate their headquarters separately from their own production plants, but close to other headquarters and business service firms. They also locate their plants close to other manufacturing facilities with which they can share intermediate suppliers. This leads to a change in urban structure, with a shift in the main dimension along which cities specialise, from a specialisation by sector to a specialisation by function.

In practice, localisation economies are likely to be stronger for business services than for manufacturing intermediates (see Markusen, 1989, for a conceptual discussion of this issue and Dekle and Eaton, 1999, for empirical evidence). In terms of the model, this simply requires that $\sigma > \epsilon$. In this case, it follows from Lemma 7 and Proposition 1 that when cities specialise by function, headquarters and business services becomes concentrated in larger cities and manufacturing production gets located in smaller cities

---

[13]The figure plots (48) for $\hat{\rho}$ given by Proposition 1 with parameters $\sigma = 0.07$ (a 1% increase in employment in a city hosting headquarters and business services increases local productivity by 1.07%), $\epsilon = 0.05$ (a 1% increase in employment in a city hosting only production plants and intermediate suppliers increases local productivity by 1.05%), and $\eta = 0.2$ (headquarters account for 20% of total costs). By symmetry, superscript h has been dropped from c_M^h and c_I^h. Note that, by (36) and (47), c_M varies with ρ but c_I does not. With these parameter values, the change in firms' organisation and in urban structure takes place when the additional costs of remote management fall below 17.3% of the management costs for an integrated firm ($\hat{\rho} = 1.173$).
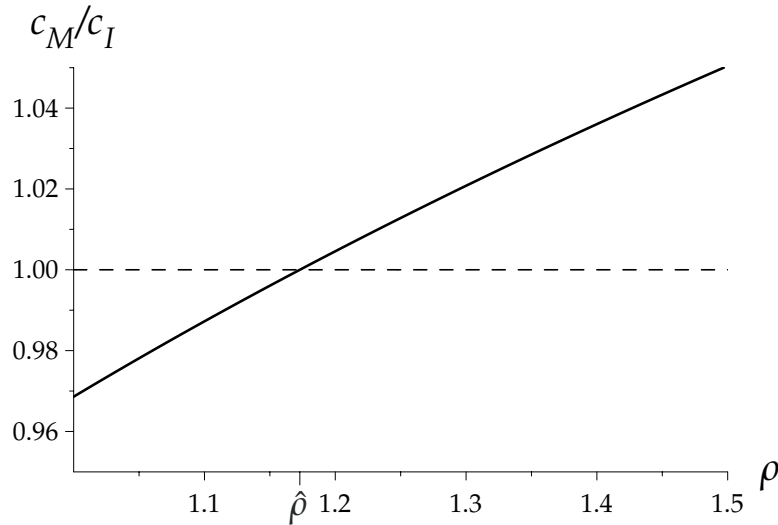
Figure 2. Unit costs of multi-location relative to integrated firms

(with $\sigma > \epsilon$, $L_H > L_I > L_M$).[14] All of this provides a consistent explanation for the stylised facts documented in the introduction.

While the model predicts recent changes in firms' organisation and in the urban structure, it presents these as happening suddenly as ρ falls below a certain threshold. The catastrophic aspect of these changes is due to the stylised nature of the model. A smoother transition could be obtained with a sector-specific cost of separating headquarters and production plant, ρ^h. Assume a gradual decline in this cost across all sectors with $\rho^1 < \rho^2 < ... < \rho^m$ at all times. Then as ρ^1 falls below $\hat{\rho}$, cities previously specialised in sector 1 undergo a shift from sectoral to functional specialisation. Some cities are left with only production plants in this sector, whereas their headquarters concentrate in business centres. The same happens again in sector 2,...,m as ρ^2,...,ρ^m fall below $\hat{\rho}$.

While improvements in physical transportation technologies have facilitated the dispersion of manufacturing activities, improvements in telecommunication technologies may have the opposite effect on business services and other knowledge-intensive activ-

---

[14]If we were to introduce the possibility of supplying business services and manufacturing intermediates to firms in a different city subject to some trade cost, then a greater difficulty in supplying business services at a distance would also result in higher localisation economies for business services than for manufacturing intermediates. Similarly, a greater impact of the congestion costs associated with larger cities on manufacturing than on business services, say because of their different land intensity, would have analogous implications for the relative magnitude of localisation economies.

ities (Leamer and Storper, 2001). Because of the complementarity between telecommunication and face-to-face contact (as evidenced for instance by Gaspar and Glaeser, 1998), cheaper and more sophisticated telecommunications increase the benefits of clustering in activities characterised by complex and timely interactions such as business services. In our framework, this implies a growing difference in the strength of localisation economies between business services and manufacturing intermediates.[15] To isolate the effects of such a change formally, consider an increase in localisation economies in business services, σ, and a corresponding reduction in localisation economies in manufacturing, ϵ, so as to keep average localisation economies, $\gamma \equiv \eta\sigma + (1 - \eta)\epsilon$, constant. In the absence of changes in the urban structure, such changes in localisation economies imply smaller manufacturing cities and a growth in the size of cities hosting headquarters and business services. However, theses changes in localisation economies may also trigger further changes in the organisation of production within firms and across cities. Regarding the threshold for the separation of headquarters and production ($\hat{\rho}$), there are two effects of opposite sign:

$$\left.\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}\sigma}\right|_{\mathrm{d}\epsilon=-\frac{\eta}{1-\eta}\mathrm{d}\sigma} = \log\left[\frac{(1-\eta)\sigma(1+2\epsilon)^2}{\eta\epsilon(1+2\sigma)^2}\right]\hat{\rho} \gtrless 0 . \tag{53}$$

There is a negative effect on $\hat{\rho}$ working through relative city sizes: if $\sigma \gg \epsilon$, $L_H \gg L_M$, so lumping production with headquarters and business services saves headquarter resources without much added congestion. On the other hands, there is a positive effect on $\hat{\rho}$ working through relative wages: if $\sigma \gg \epsilon$, $w_H \gg w_M$, so lumping production with headquarters and business services involves paying production workers much more than they would earn in separate cities. If business services generate much stronger localisation economies than manufacturing intermediates this second effect tends to dominate: urban concentrations of headquarters and business services tend to be very large, and the high wages that workers must be paid to compensate for high congestion costs drive away

---

[15] An increase in the share of headquarter costs accounted for by business services purchased from external suppliers, $1 - \mu$, as compared to directly employed labour would have the same effect (recall that $\sigma \equiv (1 - \mu)\theta$).

production plants and their suppliers, for which localisation economies are weaker.[16]

5. Concluding comments

Our model relates key stylised facts regarding changes in the organisation of firms and changes in the urban structure. Technological progress in transport and telecommunication technologies have made it less costly for firms to separate their production facilities from their headquarter and management facilities. Such a separation is worthwhile only if the firms can locate their production facilities in environments with greater same sector specialisation and their headquarters in business centres where business service employment is abundant. This in turn gives a strong incentive for cities to shift from a main specialisation along a sectoral dimension to a main specialisation along a functional dimension, leading to the emergence of separate business centres and manufacturing cities. With greater benefits from proximity for headquarters and business services than for manufacturing, the shift from sectoral to functional specialisation also implies that business centres are few and large whereas manufacturing centres are more numerous and smaller in size.

Although outside the scope of this paper, two extensions of this model are worth considering. First, one could enrich the model by assuming different types of business services. Some business services (e.g., standard business banking or equipment leasing services) need to be provided on a frequent basis and benefit from a proximity between suppliers and customers. There are also sophisticated business services (e.g., specialised financial advice for mergers-and-acquisitions) that are not required as often and for which a proximity between similar business service firms is more important. Finally, there are routinised business services (e.g., call centres) that can be easily be provided from far-away. In this case, when business services and manufacturing become separated, we would expect day-to-day services to locate in relatively large cities close to the headquar-

---

[16]Dekle and Eaton (1999) and Adserà (2000) find evidence that this kind of process is taking place in Japan and the US respectively.

ters of their customers, more sophisticated services might cluster in a handful of 'world cities' (such as New York, London, and Tokyo) concentrating all the highest management functions (Hall, 1984), and the likes of call centres might be farmed-out to low-cost locations. Second, one could incorporate a richer geography. For instance, the cost of managing production remotely might increase in the distance from the headquarter. This could explain the existence of regional business centres that host headquarters whose plants are predominantly in close-by cities.

Appendix A. Proof of Lemma 2

By Shephard's Lemma, demand for each variety of business services is

$$
y_i^0 = \int \frac{\partial C_{i,j}^h}{\partial q_i^0} dj = \eta(1-\mu) \left(\frac{Q_i^{m+h}}{H_i^h} \right)^{1-\eta} \left(\frac{w_i^h}{Q_i^0} \right)^\mu \left(\frac{q_i^0}{Q_i^0} \right)^{-(\theta+1)/\theta} \left(x_{i,i} + \rho \int_{j \neq i} x_{i,j} \, dj \right). \quad \text{(A1)}
$$

It follows from (A1) that the elasticity of demand for each business service with respect to a firm's own price is $-\frac{\theta+1}{\theta}$, so that the profit-maximising price for each business service is a constant relative markup over marginal cost:

$$
q_i^0 = (\theta + 1) \beta^0 w_i^0 . \quad \text{(A2)}
$$

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

$$
y_i^0 = \frac{\alpha^0}{\beta^0 \theta} . \quad \text{(A3)}
$$

Demand for labour by business service suppliers can be obtained by application of Shephard's Lemma to (9) and integration over varieties as

$$
l_i^0 = s_i^0 \frac{\partial C_i^0}{\partial w_i^0} = s_i^0 \left[\alpha^0 + \beta^0 y_i^0 \right] = s_i^0 \alpha^0 \frac{\theta+1}{\theta} , \quad \text{(A4)}
$$

where (A3) has been substituted in to yield the final expression. By choice of units of business service output, we can set $\beta^0 = \left(\frac{\theta}{\alpha^0} \right)^\theta (\theta+1)^{-(\theta+1)}$. Using (A4) and (A2), the price indices of (6) simplify into

$$
Q_i^0 = \left[s_i^0 \left(q_i^0 \right)^{-1/\theta} \right]^{-\theta} = \left(l_i^0 \right)^{-\theta} w_i^0 . \quad \text{(A5)}
$$

□

References

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

Adserà, Alícia. 2000. Sectoral spillovers and the price of land: a cost analysis. *Regional Science and Urban Economics* 30(5):565–585.

Arnott, Richard J. and Joseph E. Stiglitz. 1979. Aggregate land rents, expenditure on public goods, and optimal city size. *Quarterly Journal of Economics* 93(4):471–500.

Burtless, Gary. 1999. Effects of growing wage disparities and changing family composition on the US income distribution. *European Economic Review* 43(3–4):853–865.

Chandler, Alfred D. 1977. *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, MA: Belknap.

Davis, James C. 2001. *Headquarters, Services and Mobility in the Urban System*. Unpublished PhD dissertation, Brown University.

Dekle, Robert and Jonathan Eaton. 1999. Agglomeration and land rents: Evidence from the prefectures. *Journal of Urban Economics* 46(2):200–214.

Dixit, Avinash K. and Joseph E. Stiglitz. 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67(3):297–308.

Duranton, Gilles and Diego Puga. 2001. Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review* 91(5):1454–1477.

Flatters, Frank, J. Vernon Henderson, and Peter Mieszkowski. 1974. Public goods, efficiency, and regional fiscal equalization. *Journal of Public Economics* 3(2):99–112.

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.

Garreau, Joel. 1991. *Edge City: Life on the New Frontier*. New York: Doubleday.

Gaspar, Jess and Edward L. Glaeser. 1998. Information technology and the future of cities. *Journal of Urban Economics* 43(1):136–156.

Glaeser, Edward L. 1999. Learning in cities. *Journal of Urban Economics* 46(2):254–277.

Grossman, Gene M. and Elhanan Helpman. 2002. Integration versus outsourcing in industry equilibrium. *Quarterly Journal of Economics* 117(1):85–120.

Hall, Peter G. 1984. *The World Cities*. London: Weidenfeld and Nicolson.

Helpman, Elhanan. 1984. A simple theory of international trade with multinational corporations. *Journal of Political Economy* 92(3):451–471.

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. 1974. The sizes and types of cities. *American Economic Review* 64(4):640–656.

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

Henderson, J. Vernon and Randy Becker. 2000. Political economy of city sizes and formation. *Journal of Urban Economics* 48(3):453–484.

Henderson, J. Vernon and Arindam Mitra. 1996. The new urban landscape: Developers and edge cities. *Regional Science and Urban Economics* 26(6):613–643.

Jones, Charles I. 1997. The upcoming slowdown in us economic growth. Working Paper 6284, National Bureau of Economic Research.

Jovanovic, Boyan and Rafael Rob. 1989. The growth and diffusion of knowledge. *Review of Economic Studies* 56(4):569–582.

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.

Kim, Sukkoo. 1999. The rise of multiunit firms in us manufacturing. *Explorations in Economic History* 36(4):360–386.

Kolko, Jed. 1999. Can I get some service here? Information technologies, service industries and the future of cities. Processed, Harvard University.

Leamer, Edward E. and Michael Storper. 2001. The economic geography of the internet age. *Journal of International Business Studies* 32(4):641–665.

Markusen, James R. 1984. Multinationals, multi-plant economies, and the gains from trade. *Journal of International Economics* 16(3–4):205–226.

Markusen, James R. 1989. Trade in producer services and in other specialized intermediate inputs. *American Economic Review* 79(1):88–95.

Markusen, James R. 1995. The boundaries of multinational enterprises and the theory of international trade. *Journal of Economic Perspectives* 9(2):169–189.

Ota, Mitsuru and Masahisa Fujita. 1993. Communication technologies and spatial organization of multi-unit firms in metropolitan areas. *Regional Science and Urban Economics* 23(6):695–729.

Papageorgiou, Yorgos Y. and David Pines. 2000. Externalities, indivisibility, nonreplicability, and agglomeration. *Journal of Urban Economics* 48(3):509–535.

Perry, Martin K. 1989. Vertical integration: Determinants and effects. In Richard Schmalensee and Robert D. Willig (eds.) *Handbook of Industrial Organization*. Amsterdam: North-Holland, 183–255.

Serck-Hanssen, Jan. 1969. The optimal number of factories in a spatial market. In Hendricus C. Bos (ed.) *Towards Balanced International Growth*. Amsterdam: North-Holland, 269–282.

Shilton, Leon and Craig Stanley. 1999. Spatial patterns of headquarters. *Journal of Real Estate Research* 17(3):341–364.

Starrett, David A. 1974. Principles of optimal location in a large homogeneous area. *Journal of Economic Theory* 9(4):418–448.

Stiglitz, Joseph E. 1977. The theory of local public goods. In Martin S. Feldstein and Robert P. Inman (eds.) *The Economics of Public Services*. London: MacMillan Press, 274–333.

Vickrey, William S. 1977. The city as a firm. In Martin S. Feldstein and Robert P. Inman (eds.) *The Economics of Public Services*. London: MacMillan Press, 334–343.

Yates, JoAnne. 1989. *Control Through Communication: The Rise of System in American Management*. Baltimore, MD: Johns Hopkins University Press.

Yates, JoAnne. 1991. Investing in information: Supply and demand forces in the use of information in american firms, 1850–1929. In Peter Temin (ed.) *Inside the Business Enterprise: Historical perspectives on the use of information*. Chicago, IL: University of Chicago Press, 117–154.