



Marshallian theory of regional agglomeration

Michael Beenstock¹, Daniel Felsenstein²

¹ Hebrew University of Jerusalem, Department of Economics, Mount Scopus, Jerusalem 91905, Israel
(e-mail: msbin@mscc.huji.ac.il)

² Hebrew University of Jerusalem, Department of Geography, Mount Scopus, Jerusalem 91905, Israel
(e-mail: msdfels@mscc.huji.ac.il)

Received: 25 May 2008 / Accepted: 20 March 2009

Abstract. Most models of regional agglomeration are based on the new economic geography (NEG) model in which returns to scale are pecuniary. We investigate the implications for regional agglomeration of a ‘Marshallian’ model in which returns to scale derive from technological externalities. Workers are assumed to have heterogeneous ‘home region’ preferences. The model is designed to explain how ‘second nature’ determines regional wage inequality and the regional distribution of economic activity. We show that agglomeration is not a necessary outcome of Marshallian externalities. However, if centrifugal or positive externalities are sufficiently strong relative to their centripetal or negative counterparts, the model generates multiple agglomerating equilibria. These equilibria multiply if, in addition, there are scale economies in amenities. A dynamic version of the model is developed in which external economies and inter-regional labour mobility grow over time. Regional wage inequality overshoots its long run equilibrium and, there is more agglomeration in the long run.

JEL classification: J61, O18, R11, R13

Key words: Agglomeration, Marshallian externalities, wage disparities

1 Introduction

Empirical research in a growing number of countries¹ reveals persistent differences in regional wages. Moreover, these differences are inherent and cannot be explained away by regional differences in the cost-of-living, the characteristics of the labour force and amenities. Identical workers in terms of education, experience and other measures of human capital are paid differently depending upon the region in which they work (Beenstock and Felsenstein 2008). This body of interregional research more or less parallels convergence failure at the international level (Barro and Sala-I-Martin 1991).

¹ See for example: Duranton and Monastiriotis (2002) on Britain; Azzoni and Servo (2002) on Brazil; Maier and Weiss (1986) on Austria; and Beenstock and Felsenstein (2008) on Israel.

One might have thought that the prospects of discovering regional wage convergence were much greater than the prospects of discovering international wage convergence because capital is likely to be more mobile within countries than between them, because labour has the same characteristics, and because trade is likely to be freer within countries than between them. However, empirical results suggest that just as there is convergence failure between countries, there is convergence failure within them.

The discovery of convergence failure at the international level spawned new theories, such as endogenous growth theory (e.g. Grossman and Helpman 1993) to explain the empirical facts. The same has happened in the case of regional convergence failure; economists and regional scientists are developing theories to explain the facts. An important example of this is new economic geography (NEG) in which agglomerating forces may induce regional disparities.² Returns to scale in NEG are entirely pecuniary, and are induced by the 'market size effect', which enables producers to lower costs. NEG has emerged as the dominant paradigm in the economics of agglomeration.

An alternative paradigm to NEG is based upon Marshallian or technologically induced scale economies, which play no role in NEG. Although much of endogenous growth theory derives from technological scale economies and pecuniary economies have no role, the opposite applies to agglomeration theory. This dichotomy is puzzling because agglomeration is essentially growth across space rather than over time. The puzzle is all the greater since Marshallian externalities have formed the foundation of much of urban economics (Henderson 1974; David and Rosenbloom 1990; Abdel-Rahman and Fujita 2000; Duranton and Puga 2004; Rosenthal and Strange 2004). We see no intrinsic reason why Marshallian externalities should be confined to cities and not regions. Nor do we see any reason why pecuniary economies induced by market size effects should be confined to regions and not cities. Indeed, urban economic theory embraces both technological as well as pecuniary scale economies (Fujita and Thisse 2002, chapter 4). Various papers explore the role of technological externalities in promoting agglomeration. This is usually done within the urban economics paradigm of understanding the interactions between localisation economies, city size and urban productivity (Abdel-Rahman 2000). Another approach examines the common determinants of different modes of agglomeration. These include the home market effect, urbanisation or localisation effects, etc. (LaFountain 2005).

In this paper we extend this symmetry to the regional context.³ We develop a theory of regional agglomeration that is exclusively Marshallian. Although this theory is proposed as an alternative to the dominant NEG paradigm, we do not wish to detract from the importance of pecuniary economies in regional agglomeration.

To the best of our knowledge, only Michel et al. (1996) discuss regional agglomeration in a Marshallian context.⁴ However, as explained below, their assumption that scale economies are exponential induces corner solutions with complete agglomeration. To prevent this from happening, they borrowed from NEG the assumption that unskilled labour is completely immobile.

We generalise and extend their model in several ways. First, we introduce perfectly mobile capital into the model, and all labour is allowed to be mobile. Indeed, there are no immobile factors in our model that arbitrarily prevent complete agglomeration. Second, we show that the way in which scale economies are specified plays a crucial role in the Marshallian theory of agglomeration, and determines the existence or otherwise of multiple equilibria. Third, we introduce heterogeneous tastes into the model. We follow Murata (2003), Tabuchi and Thisse

² NEG refers to Krugman (1991). See also Fujita et al. (1999); Puga (1999); Brakman et al. (2001); Fujita and Thisse (2002, chapter 9); and Baldwin et al. (2003).

³ The main difference between urban and regional economics lies in the importance of land use in the former. See e.g. Fujita and Thisse (2002, p. 18).

⁴ Chapter 8 in Fujita and Thisse (2002) presents an identical theory to Michel et al. (1996).

(2002) and Nocco (2009) in assuming that workers have heterogeneous regional preferences, which are characterised by ‘home region’ preference.⁵ This means that workers are imperfectly mobile between regions, as originally suggested by Hicks (1932) and supported by subsequent empirical work (DaVanzo and Morrison 1981; Evans and McCormick 1994; Long 1991). However, unlike, Murata and Tabuchi and Thisse, we introduce amenities into the utility function.

Our model generates interior solutions for structural rather than technical reasons and has a rich taxonomy of cases. In some cases agglomeration raises a region’s wage differential while in other cases it decreases it. Indeed, the relationship between agglomeration and regional wage inequality depends critically on the nature of scale economies, for amenities as well as for production.

The Marshallian theory of agglomeration takes its inspiration from the following remarks:⁶

so great are the advantages which people following the same skilled trade get from near neighbourhood to one another. The mysteries of the trade become no mystery; but as it were, in the air, and children learn many of them unconsciously. Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organisation of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further ideas Marshall (1920, p. 271).

Again an increase in the scale of production of the industry as a whole . . . tends to open each business in the industry, whether large or small, access to improved plant, improved methods, and a variety of other ‘external’ economies Marshall (1919, p. 187).

Marshall did not develop these ideas into a fully-fledged theory of agglomeration. Nor did he propose a proper theory of industrial location. Rather, he was speculating on the causes of spatial clustering in economic activity. Foremost among these he identified knowledge spillovers, linkages between suppliers and producers and labour-market interactions.⁷ In doing so he distinguished between what today are referred to as the roles of first and second nature. He saw in knowledge spillovers and standardisation key ‘second nature’ determinants of external returns to scale, which accounted for spatial concentration in industry.

Marshallian externalities have not proved popular among regional agglomeration theorists for esthetic reasons. Fujita and Thisse (2002, p. 299) mention that this is because Marshallian externalities can be conceived as a theoretical ‘black box’, whereas by contrast pecuniary scale economies in NEG have clearly annunciated microfoundations. The latter stem from a marriage between the model of Dixit and Stiglitz (1977) of imperfect competition and iceberg assumptions about trade costs.

However, Marshallian technological externalities are not entirely a ‘black box’, and have not been conceived as such in urban economics. Glaeser (1999) suggests a theory of Marshallian externalities in which there are centrifugal and centripetal forces in the diffusion of skills. Knowledge spillovers have been subject to particularly rigorous treatment. Storper and Venables (2004) develop some of the microeconomic foundations for face-to-face contact among economic agents. They show how these improve co-ordination, increase productivity and mitigate the incentives problem in the creation of collaboration between agents. Helsley and Strange (2004) highlight the existence of endogenous knowledge spillovers in cities, induced for example by information barter. In contrast to standard exogenous knowledge transfers, they show that endogenous spillovers can give rise to an efficient equilibrium based on mutual

⁵ Tabuchi and Thisse (2002), Murata (2003) and Nocco (2009) all introduce regional preferences into NEG.

⁶ It is also mentioned in earlier editions of Marshall (1920) and arises in the first edition published in 1880. See also Marshall (1919).

⁷ Duranton and Puga (2004, p. 2066) review the micro-economic foundations of each of these processes under the headings of ‘sharing, matching and learning mechanisms’. Other (non-Marshallian) micro-foundations of agglomeration, such as home market effects, consumption effects and rent-seeking are reviewed by Rosenthal and Strange (2004).

reciprocity between agents. Berliant et al. (2006) model the interaction between agents with heterogeneous stocks of knowledge and the searching and matching process in which they engage. In addition, several papers have shown⁸ that the ‘cafeteria effect’ in knowledge spillovers is empirically important. Also several papers testify to the empirical importance of Marshallian externalities in regional agglomeration at the county level.⁹

We propose a simple model in which regions produce a single good under conditions of perfect competition.¹⁰ Returns to scale are constant at the level of the firm, but total factor productivity depends upon scale driven by Marshallian externalities in the region. These externalities are assumed to be completely localised so that they do not diffuse to other regions. This assumption is consistent with other studies such as Jaffe et al. (1993) and Anselin et al. (1997) who use patent and innovations data, to demonstrate that knowledge spillovers are mainly localised. There are no transport costs in the model so that inter-regional trade in the single good takes place freely and without friction, and its price is the same everywhere. However, because total factor productivity (TFP) depends upon scale, the cost of production varies by region.

In our model all workers are homogeneous in terms of skill. Therefore skill mix plays no role in our model. Workers may be homogeneous in terms of skill but they are heterogeneous in terms of ideas. When they meet in the proverbial cafe they have different ideas on how TFP may be raised, so that knowledge about TFP grows and percolates along lines suggested by Jovanovic and Rob (1989). The more populated a region the greater is the probability of meeting someone with new ideas, so that TFP varies directly with scale. On the other hand, scale may also increase the cost of communication; the proverbial café becomes crowded, so that TFP may also vary inversely with scale.¹¹

Our model explains the regional distribution of wages, population and output. First nature plays no role in our model since all regions are a priori identical physically. Therefore if agglomeration occurs it is entirely induced by second nature.

The paper concludes with a discussion of agglomeration dynamics. Marshall (1919) states that external returns to scale accumulate gradually, which implies that there should be more agglomeration in the long run than in the short run. Also, home region preference is stronger within generations than between generations. We propose an overlapping generations model in which individuals are raised in their home region, and work either in their home region or in another region. In the model there is more interregional mobility between generations than within generations. Since Marshallian externalities accumulate over time and interregional labour mobility increases over time, the model predicts that if agglomeration occurs, it occurs gradually over decades. Our model also shows that in the short run, regional wage gaps overshoot their long run equilibrium.

2 Regional supply and demand for labour

We begin this section by proposing a theory for regional labour supply in which individuals have regional preferences. Since each individual is assumed to supply one unit of labour, labour supply in a given region is equal to the population in that region. This is followed by a theory of labour demand by region in which Marshallian externalities play a key role. In Section 3 we discuss the equilibrium implications implied by our theory of regional labour markets.

⁸ e.g. Charlot and Duranton (2004); Glaeser and Mare (2001); and Fu (2007).

⁹ Henderson (2003) and Rosenthal and Strange (2004).

¹⁰ Since there is a single good there can be no ‘Jacobs externalities’ in our model.

¹¹ Similar scale related effects leading to either congestion or successful interaction are noted in both Helsley and Strange (2004) and Berliant et al. (2006).

2.1 Migration and residential location

People are assumed to have ‘status quo’ preferences in the sense that they are attached to their current place of residence. In the literature on discrete choice (Hartman et al. 1991) status quo bias is usually regarded as irrational, but in the case of migration decisions it has a rational justification. People have family and social attachments, and attachments to local cultural norms that have grown over time. Young people are naturally endowed with a location, their home region, because they have no say in where they were born and raised. Accident of birth most probably explains where most adults spend their lives.

We denote by U_{is} the utility of individual i raised in region r from living in region s . We assume that utility is generally higher in the home region, namely, when $s = r$, due to proximity to family and friends. We refer to this as home region preference. D_{rs} denotes the social distance between regions r and s . Due to home region preference individuals prefer regions that are socially closer to home.¹² Utility is hypothesised to depend upon four main factors in region s , the level of income (Y_s), the availability of amenities (H_s), social distance from the home region (D_{rs}) and unobserved heterogeneity. The random utility of individual i raised in region r from living in region s is assumed to be:

$$U_{is} = \delta H_s + \beta \log Y_s - \phi D_{sr} + \varepsilon_{is} \tag{1}$$

where ε captures unobserved heterogeneity. In Equation (1) $\log Y$ is specified rather than Y to conform with the analysis in section 2.2. Individuals are assumed to choose where to live by maximising U_{is} over all s . Because of home region preference, which is captured by ϕ , region r will be naturally over-represented in residential choice because $D_{rr} = 0$. Due to unobserved heterogeneity there will be individuals who choose to migrate from region r , their home region. The model allows for two-way migration. Given everything else, region s is less attractive to people from region r the greater the social distance between them.

If ε is independent and identically distributed with cumulative distribution $F(\varepsilon) = \exp(-e^{-\varepsilon})$ the proportion of people raised in region r choosing to live in region s will be determined by the following conditional logit model:

$$P_{rs} = \frac{\exp(\delta H_s + \beta \log Y_s - \phi D_{sr})}{\sum_{q=1}^R \exp(\delta H_q + \beta \log Y_q - \phi D_{qr})} \tag{2}$$

Equation (2) states that this proportion varies directly with income and amenities in s and inversely with income and amenities elsewhere.¹³

$$\frac{\partial P_{rs}}{\partial \log Y_s} = \beta P_{rs} (1 - P_{rs}) > 0 \tag{3}$$

$$\frac{\partial P_{rs}}{\partial H_s} = \delta P_{rs} (1 - P_{rs}) > 0 \tag{3a}$$

¹² With zero transportation costs it is always possible to visit loved ones costlessly. However, to ‘feel at home’ one has to be in physical contact with them; virtual or internet relationships are not enough.

¹³ Note that in the logit formulation variables such as Y , H and D do not appear directly in Equations (3) and (4); their effect is expressed indirectly via P_{rs} and P_{rq} .

$$\frac{\partial P_{rs}}{\partial \log Y_q} = -\beta P_{rs} P_{rq} < 0 \quad (4)$$

$$\frac{\partial P_{rs}}{\partial H_q} = -\delta P_{rs} P_{rq} < 0 \quad (4a)$$

Given their home region preference, people from r are therefore imperfectly mobile with respect to wage differentials unless $\beta = \infty$, in which case they are perfectly mobile. If $\beta = 0$, they are completely immobile. If, home region preference is absolute so that $\phi = \infty$, people will be immobile because $P_{rs} = 0$. Note that there are two separate forces in Equation (2). Home region preference is captured by ϕ , which raises P_{rr} relative to P_{rs} because $D_{rr} = 0$. Equation (2) implies that P_{rr} varies directly with ϕ and D_{rs} . A quite separate force is heterogeneity, which implies that even if $\phi = 0$ (no home region preference), individuals do not regard regions as perfect substitutes.¹⁴

In Section 3 we show that in general economic equilibrium Y_s , H_s and P_{rs} are jointly determined. If more people choose to live in region s the regional distribution of income is likely to be affected. In the meanwhile Equation (2) holds in partial equilibrium. The same applies to amenities. If more people choose to live in region s there may be a positive ‘conviviality effect’ since scope for social interaction is greater, and there may be a negative ‘crowding effect’ due to congestion, in which case the regional distribution of amenities¹⁵ will be jointly determined with P_{rs} . We denote by θ_{rs} the marginal effect of P_{rs} on amenities in region s (H_s). Although we have no explicit market for housing in the model, the ‘crowding effect’ may also be regarded as an expression of the positive correlation between housing costs and agglomeration (Beenstock and Felsenstein 2008).

What does the migration model with endogenous amenities imply about the relationship between regional population shares and the wage gap? From Equation (2) we obtain:

$$dP_{rs} = \frac{\partial P_{rs}}{\partial \log Y_s} d \log Y_s + \frac{\partial P_{rs}}{\partial H_s} dH_s$$

Because amenities are endogenous dH_s depends on $d \log Y_s$ as follows:

$$dH_s = \theta_{rs} \frac{\partial P_{rs}}{\partial \log Y_s} d \log Y_s$$

Substituting the latter into the former and Equations (3) and (3a) into the result gives the relationship between wage gaps and the population share when amenities are endogenous:¹⁶

$$\frac{dP_{rs}}{d \log Y_s} = \beta P_{rs} (1 - P_{rs}) [1 + \delta \theta_{rs} P_{rs} (1 - P_{rs})] \quad (5)$$

¹⁴ The second of these forces forms the basis of Tabuchi and Thisse (2002). Tabuchi and Thisse (2002) and Murata (2003) ignore the role of amenities.

¹⁵ More generally, H should represent the quality of life. Blomquist (2006) estimates the effect of first and second nature on the quality of life. By definition first nature plays no role in our model. The effects of second nature on the quality of life are captured by what we term amenity effects. Our specification assumes that amenities and the quality of life are scale dependent. Empirical evidence suggests that crime rates, pollution, congestion, etc. are scale dependent, but not exclusively so.

¹⁶ Footnote 13 explains why variables Y , H and D do not feature directly in Equation (5).

Equation (5) states that if θ_{rs} is positive, population shares vary directly with the wage gap. Indeed, they increase by more than they would had amenities not been scale dependent. If, however, θ_{rs} is sufficiently negative Equation (5) indicates that when amenities are endogenous population shares may vary inversely with the wage gap. More generally the relationship between population shares and wage gaps may not be monotonic.

If there are only two regions ($R = 2$) A and B, Equation (2) simplifies to the logit case. In the initial equilibrium half the population is assumed to live in each region and wages and amenities are the same in both regions. Let P_{AA} denote the probability that a person raised in A chooses to live in A and let P_{BA} denote the probability that a person raised in B chooses to live in A. A's share of the population is $P = P_0 P_{AA} + (1 - P_0) P_{BA}$. Since $P_0 = 1/2$, $P = 1/2(P_{AA} + P_{BA})$, or:

$$P = \frac{1}{2} \left(\frac{1}{1 + \exp(\beta y + \delta h - \phi D)} + \frac{1}{1 + \exp(\beta y + \delta h + \phi D)} \right) \tag{6}$$

where $y = \log(Y_B/Y_A)$, $h = H_B - H_A$ and $D_{AB} = D_{BA} = D$. P_{AA} varies directly with D because the greater the distance from A to B the more likely residents of A will prefer to remain there. P_{BA} varies inversely with D because it reduces the incentive to move from B to A. Note that in the initial equilibrium $y = h = 0$ in which case $P = 1/2$ according to Equation (6) and $P_{AA} = 1/[1 + \exp(-\phi D)]$ and $P_{BA} = 1/[1 + \exp(\phi D)]$. The same would apply in the general case where population shares equal $1/R$.

Since H may be increasing or decreasing at the margin we assume that $H_A = G(P)$. If $G' > 0$ amenities vary directly with scale (P), and if $G'' > 0$ the effect of agglomeration on amenities is divergent. Because of crowding G'' may be negative in which case $G(P)$ is not monotonic. For reasons of convenience, flexibility and analytical tractability we follow Michel et al. (1996) who assumed a quadratic relationship between amenities and population size:

$$H_A = \bar{\omega} + \nu P - \sigma P^2 \tag{7}$$

If $\sigma < 0$, H_A is divergent. Since P is bounded by 0 and 1, H_A is bounded by $\bar{\omega}$ and $\bar{\omega} + \nu - \sigma$. If $(\bar{\omega} + \nu)/2\sigma < 1$, H_A is non-monotonic. Therefore, Equation (7) caters for a broad class of scale dependencies.

Due to symmetry, H_B is obtained by substituting $1 - P$ for P in Equation (7). Therefore $h = (\nu - \sigma)(1 - 2P)$ so that h varies inversely with P when $\nu > \sigma$ and directly with P when $\nu < \sigma$. Note that when $P = 1/2$ $h = 0$. Equation (6) further implies that the relationship between A's population share and the log wage gap is:

$$\frac{\partial P}{\partial y} = - \frac{\beta}{2} \frac{P_{AA} P_{AB} + P_{BA} P_{BB}}{1 + \delta(\sigma - \nu)(P_{AA} P_{AB} + P_{BA} P_{BB})} \tag{8}$$

If $\sigma > \nu$, Equation (8) is negative, implying that A's population share varies directly with its wage gap. If, however, ν is sufficiently larger than σ , A's population share may vary inversely with the wage gap. Because P_{AA} , etc. are endogenous, Equation (9) implies that the relationship between P and y in this case is not necessarily monotonic.

2.2 Production

There is a single traded good (Q), which is freely traded between regions in a perfectly competitive market at zero transport cost. Capital (K) is perfectly mobile between regions so that the return to capital is the same in each region. The national economy is closed. The production

function is neoclassical, for simplicity Cobb-Douglas, so that output in region $r = A, B$ is $Q_r = A_r K_r^\alpha L_r^{1-\alpha}$. Note that for symmetry $0 < \alpha < 1$ does not vary by region so that internal returns to scale are therefore assumed to be constant and identical across regions. However, external returns to scale may vary across regions depending on their size.

The way scale should be represented depends upon the factors of production to which external returns to scale accrue. If they only accrue to capital it would be preferable to use capital (K) to represent scale. We assume that total factor productivity depends ‘neutrally’ upon scale in which case scale may be represented by either capital or population.¹⁷ In equilibrium the capital labour ratio is constant, hence it makes no analytical difference to which factors of production scale effects accrue, and how scale is represented.

Land is not an explicit factor of production in the model. In principle, we see no reason why the introduction of markets for land and housing would change Marshallian agglomeration theory differently from NEG. Helpman (1998) and Suedekum (2006) show, not surprisingly, that in the NEG model housing markets impede agglomeration because land becomes more expensive in agglomerating regions. The same would apply in our model; industrial rents would rise thereby impeding agglomeration in terms of production. Also, house prices would rise, thereby impeding agglomeration in terms of residential choice. In Section 2.1 we suggested that the effect of agglomeration on residential rents will be captured by negative amenities or crowding. Here too, the effect of agglomeration on industrial rents is captured by congestion effects on TFP .

Nor, in principle, do we think that extending the model beyond the single good case would substantively alter Marshallian agglomeration theory as long as the factor price frontier is downward sloping, namely, there is a negative relationship between the real wage and the rate of interest, and the labour intensity of production varies inversely with the real wage. When there is only one good the factor price frontier necessarily slopes downwards. However, when there is more than one good reswitching becomes a theoretical possibility¹⁸ in which event the factor price frontier may slope upwards. Barring reswitching, we see no reason why the Marshallian theory of agglomeration would be affected just as neoclassical theory remains unaffected as long as reswitching does not occur.

In principle, therefore, the specification of production in the model may be generalised to incorporate externalities that accrue non-neutrally with respect to factors of production, and land and housing markets may be introduced too. Finally, the model may be extended to more than a single good. However, in the interests of transparency and methodological minimalism we believe that a parsimonious treatment of Marshallian agglomeration theory is preferable to introducing further generalisations at this expository stage.

For simplicity we assume that there are two regions A and B so that $L = L_A + L_B$ with $\psi = L_A/L$. Since capital is mobile the marginal productivity of capital must be equated between regions ($MPK_A = MPK_B$), which implies:

$$\frac{k_A}{k_B} = \left[\frac{A_A}{A_B} \right]^{1/(1-\alpha)} \quad (9)$$

Equation (9) states that relative capital per head ($k = K/L$) varies directly with relative TFP. If workers are paid their marginal products, the ratio of wages in A relative to B is equal to:

¹⁷ Michel et al. (1996) assume that scale effects accrue to skilled labour.

¹⁸ See Ferguson (1969) and Harcourt (1972) on the ‘Cambridge Criticism’ of Neoclassical production theory and reswitching.

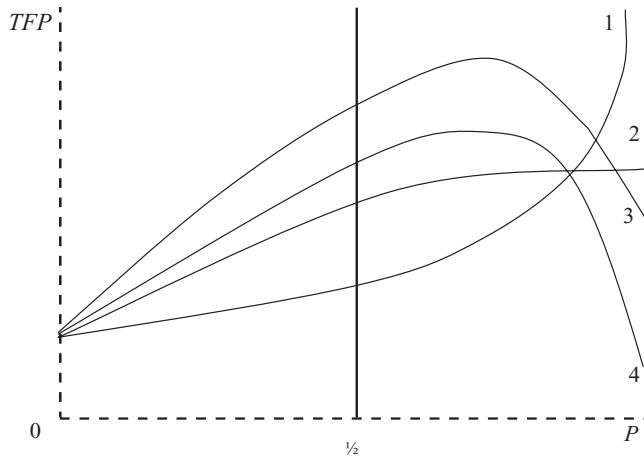


Fig. 1. The Relationship between scale and *TFP*

$$\frac{Y_A}{Y_B} = R = \frac{k_A^\alpha A_A}{k_B^\alpha A_B} = \left[\frac{A_A}{A_B} \right]^{1/\alpha} = \Theta^{1/\alpha} \tag{10}$$

Notice that in the absence of scale effects ($A_A = A_B = A$) Equation (10) states that regional wages are equated.

Whereas workers must live in a region to work there, profits are not region specific. Profits earned in region *A* may be distributed in either region. Therefore, profits unlike wages play no role in the model. Matters would be quite different, of course, if profits could not be transferred across regional boundaries.

Figure 1 illustrates a number of hypothetical relationships between *TFP* (*A*) and scale, where the latter is measured by the proportion of employment in the region. In schedules 1 and 2 external returns to scale are always increasing, however, in schedule 1 the relationship is divergent whereas in schedule 2 it is convergent. In schedules 3 and 4 returns to scale eventually decrease because if regions become too large congestion may generate external diseconomies of scale. In schedule 4 *TFP* eventually falls below its level at scale zero. The unknown relationship between *TFP* and scale in region *A* is denoted by $F(\psi)$. We refer to its partial derivative as F_A , which is the marginal agglomeration effect in *A*. Symmetry implies that *TFP* in region *B* is $F(1 - \psi)$, and F_B denotes the marginal agglomeration effect in *B*. Note that if F_A is negative as in schedule 3, F_B is positive by symmetry.

It may be shown that the effect of agglomeration on relative $TFP = \Theta$ is:

$$\frac{d\Theta}{d\psi} = \Theta' = \frac{TFP_B F_A + TFP_A F_B}{TFP_B^2} \tag{11}$$

Equation (11) is positive when marginal agglomeration effects are positive ($F_A > 0, F_B > 0$) as in schedules 1 and 2 in Figure 1. When schedule 3 applies F_A can be negative in which case F_B is positive and $\Theta > 1$. Equation (11) may still, however, be positive. But it will be negative when $F_A + \Theta F_B < 0$. Therefore relative *TFP* in *A* will vary directly with ψ when schedules 1 and 2 apply, but it may vary inversely with ψ when schedule 3 applies. This applies a fortiori in the case of schedule 4 since Θ may be less than one when *A* is over-agglomerated. When schedules 3 and 4 apply, Θ' may be positive when *A* is under-agglomerated and negative when it becomes over-agglomerated.

Differentiating Equation (11) with respect to ψ reveals whether the effect of agglomeration on relative TFP in A is divergent or convergent:

$$\Theta'' = \frac{TFP_B^2 F_{AA} + TFP_A TFP_B F_B + TFP_B (F_A F_B - TFP_A F_{BB}) + 2TFP_A F_B^2}{TFP_B^3} \quad (12)$$

where F_{AA} and F_{BB} denote the effects of agglomeration in A on the marginal agglomeration effects¹⁹ in regions A and B respectively. If $\Theta' > 0$ and $\Theta'' < 0$ agglomeration in A has a monotonic and convergent effect on relative TFP in A . If, however, $\Theta'' > 0$ agglomeration in A has a divergent effect. A further possibility is that Θ'' is positive over some range of ψ and then becomes negative. Since there may be sign reversal in Θ' and Θ'' there is a rich taxonomy in the relationship between relative TFP and scale.

Michel et al. (1996) assume that $TFP_A = \exp(a\psi)$ in which case external returns to scale are always increasing and agglomeration induces divergence in relative TFP since:

$$\Theta' = 2a \frac{\exp(a)}{\exp[2a(1-\psi)]} > 0 \quad (13a)$$

$$\Theta'' = 2a\Theta' > 0 \quad (13b)$$

This choice of functional form naturally induces corner solutions and complete agglomeration. To generate interior solutions Michel et al. (1996) assume that unskilled labour is completely immobile. Together with the Inada condition this constitutes a *deus ex machina* that rules out complete agglomeration.

Another possible specification of $F(\psi)$ follows Equation (7) and assumes that TFP is quadratic in scale, namely, $F(\psi) = a + b\psi - c\psi^2$. Here too we choose this specification for reasons of convenience, flexibility and analytical tractability. Scale effects are beneficial in A and adverse in B when $\psi < b/2c$, and adverse in A and beneficial in B when $\psi > b/2c$. This specification implies schedule 3 in Figure 1 if $b > c$ and it implies schedule 4 if $b < c$. It also implies:

$$\Theta' = (b-c) \frac{2\left(a - c\psi(1-\psi) + \frac{b}{2}\right)}{TFP_B^2} \quad (14)$$

Since the maximum value of $\psi(1-\psi) = 1/4$ Equation (14) is positive when $b > c$ and negative otherwise.

In summary, the effect of agglomeration in A on relative TFP may be positive or negative, and it may be divergent or convergent. If negative Marshallian externalities are ruled out Θ' is positive, but the sign of Θ'' is ambiguous. Finally, from Equation (10) the marginal effect of agglomeration in region A on relative wages in A is equal to:

$$\frac{dR}{d\psi} = R' = \frac{1}{1-\alpha} \Theta^{\alpha/1-\alpha} \Theta' \quad (15)$$

The sign of Equation (15) depends directly on the sign of Θ' . This process is divergent when:

¹⁹ F_{AA} is the derivative of F_A with respect to ψ and F_{BB} is the derivative of F_B with respect to $1-\psi$.

$$R'' = \frac{\Theta^{\alpha/1-\alpha}}{1-\alpha} \left[\frac{\alpha}{1-\alpha} \frac{\Theta'}{\Theta} + \Theta'' \right] > 0 \tag{16}$$

Equation (16) shows that if $\Theta' > 0$ and $\Theta'' > 0$ then relative wages in *A* are divergent in agglomeration, which happens when agglomeration increases relative *TFP* divergently. Note that if agglomeration increases relative *TFP* convergently, namely, $\Theta' > 0$ and $\Theta'' < 0$, R'' may still be positive, in which case agglomeration has a positive and divergent effect on *A*'s relative wages. Therefore convergence in relative productivity does not guarantee convergence in relative wages. If $R'' < 0$ then relative wages are convergent.

3 Regional equilibrium

In this section we characterise the equilibrium that is implied by the model presented in Section 2. The residential choice model generated a potentially U-shaped relationship between *A*'s population share (denoted by *P*) and the log wage gap ($y = \log R$) between *A* and *B*. In Figure 2 we denote this relationship by schedule *M*, which is drawn under the 'standard' assumption that *A*'s wage gap has to improve to attract more people, namely, $\sigma > \nu$ in Equation (8). Note that symmetry requires the log wage gap be zero when half the population lives in *A*. By contrast, according to Equation (10) the production model generates a positive relationship between *A*'s population share (ψ) and the wage ratio (*R*) when $b > c$ so that $\Theta' > 0$, and a negative relationship when $b < c$ so that $\Theta' < 0$. In Figure 2 we denote this relationship by schedule *Q*, which slopes downwards because we initially assume that $b < c$ in Equation (14). Note that symmetry requires schedule *Q* intersect the horizontal axis when half the population lives in *A* and regional wages are equated ($R = 1$).

The equilibrium in Figure 2 is determined where schedules *M* and *Q* intersect, namely, when the supply of people choosing to reside in regions *A* and *B* is equal to firms' demand for labour in *A* and *B*. The equilibrium condition is $P = \psi$. Schedules *M* and *Q* have a single intersection at E_1 at which wages are equated, and due to symmetry, half the population is in region *A*. The equilibrium at E_1 is stable and unique. To see this, suppose that there was in-migration to *A* such that $P > 1/2$. This would drive down wages in *A* relative to *B*'s wages, which would induce reverse

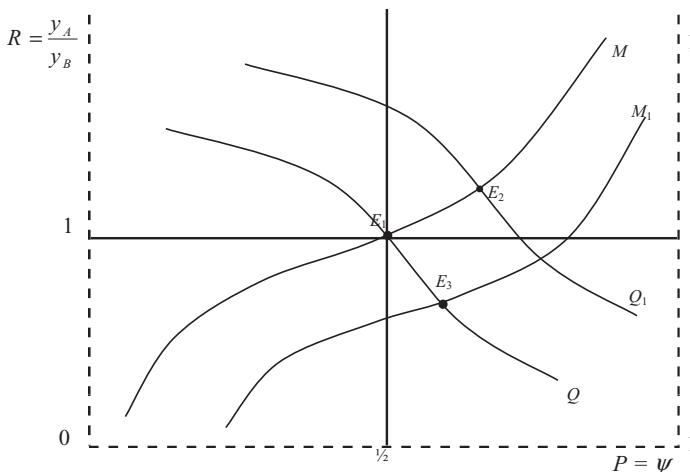


Fig. 2. Regional equilibrium: Unique case

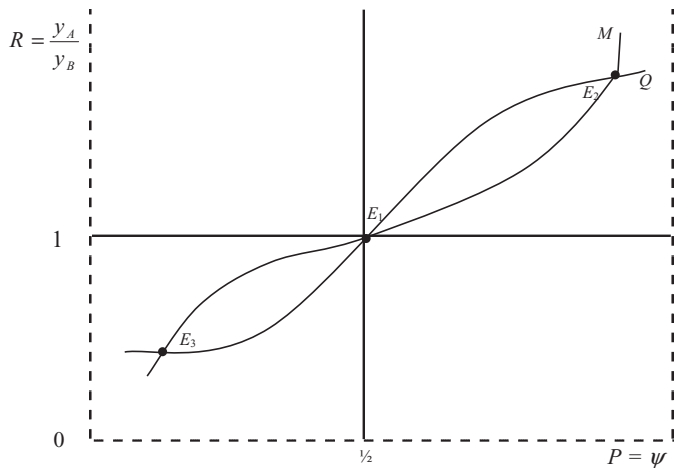


Fig. 3. Multiple regional equilibria

migration, thereby restoring the equilibrium to E_1 . If there is an autonomous increase in productivity in A, schedule Q would shift to the right (to Q_1) and the new equilibrium would be at a point such as E_2 , at which the wage gap opens in favour of A, and A's population share increases. If there is an autonomous increase in residential preference in favour of A, schedule M would shift to the right (to M_1) and the new equilibrium would be at a point such as E_3 , at which the wage gap opens in favour of B and A's population share increases.

Had there been perfect inter-regional labour mobility schedule M would have been a horizontal line emanating from 1 on the vertical axis. In the absence of scale effects in production schedule Q would also have been a horizontal line emanating from 1 on the vertical axis.²⁰ In this case regional wages are always equated regardless of the population shares because capital mobility ensures that the marginal productivity of capital is equated across regions. If there are no scale effects in production but people have residential preferences, schedule M would be as drawn in Figure 2 and schedule Q would be horizontal at zero, so that the wages are equated and half the population lives in A.

What happens, however, if instead of sloping downwards as in Figure 2, schedule Q slopes upwards and convergently as in Figure 3 because $b > c$ and $R'' < 0$? In Figure 3 we assume that schedule Q cuts schedule M from below when there is no agglomeration ($P = 1/2$). According to the logit model schedule M is naturally flattest at this point.²¹ Schedules M and Q intersect at three points so that there are multiple equilibria. The first of these equilibria, denoted by E_1 , is non-agglomerating since $R = 1$ and $\psi = 1/2$. However, this equilibrium is unstable because if the population in A happened to increase, relative wages in A would increase according to schedule Q to such an extent that yet more inward migration in A's favour would take place. This happens because schedule Q lies above schedule M over the relevant range. This process converges to an agglomerating equilibrium at E_2 where schedules M and Q intersect. Its symmetrical counterpart is E_3 , which is also stable. Therefore, stable equilibria require that schedule M intersects schedule Q from below.

In Figure 4 we consider the case where schedule M is U-shaped over some range due to endogenous amenities and because $v > \sigma$ in Equation (8), but schedule Q slopes downwards as in Figure 2. There are three equilibria. At E_1 wages are necessarily equated and $P = \psi = 1/2$ as in

²⁰ Or $\Theta = 0$.

²¹ From Equation (11) the slope of schedule Q at this point is equal to $2F_A/TFP_A$.

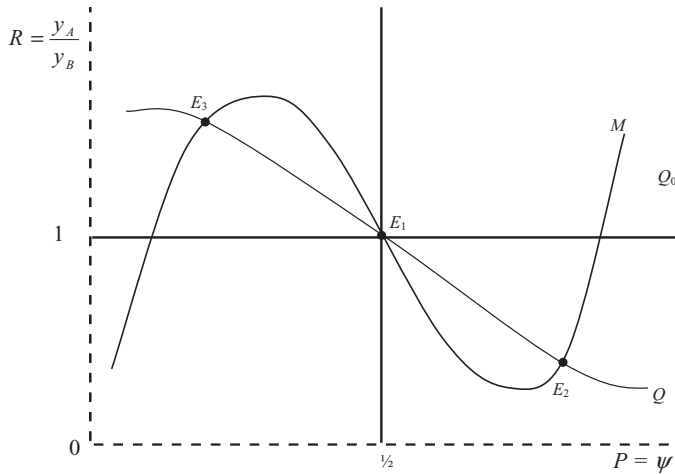


Fig. 4. Regional equilibria with amenities

Figures 2 and 3. However, the initial equilibrium E_1 is unstable whereas E_2 and E_3 are stable because schedule M cuts schedule Q from below. Had schedule Q sloped upwards (schedule Q') instead of downwards there would be two equilibria, one at E_1 and the other at E_4 . Here too, E_1 is unstable but E_4 is stable.

The model has been developed for the case where unobserved heterogeneity in regional preferences is assumed to have a logistical distribution. Similar qualitative results would have been obtained had we assumed instead that regional preferences are normally distributed. In this case the probit model would have applied in the two regions case and the conditional probit model in the multi-region case. Indeed, any parametric assumption that implies that schedule M is asymptotic would have produced broadly the same result. However, if regional preferences are uniformly distributed schedule M will be linear. This gives rise to two possibilities. In the first, the linear M schedule intersects the Q schedule three times, in which case the result is broadly the same as in Figure 3; the equilibrium is agglomerating. In the second, the linear M schedule intersects the Q schedule only once at $R = 1$ and $\psi = 1/2$, in which case there is no agglomeration. If the Q schedule is relatively flat because Marshallian externalities are weak, the second case will arise and there will be no agglomeration.

The cases in Figures 2–4 illustrate the interplay between Marshallian externalities as they are represented in schedule Q and regional preferences as they are represented in schedule M . Clearly, many other possibilities may arise.

4 Intergenerational implications of the model

The equilibria discussed in Section 3 are essentially short-term because they refer to a single generation. In this section we discuss the intergenerational, or long term implications of the model. The long term solution to the model differs from its short-term counterpart for two separate reasons. First, we show that the migration model presented in Section 2.1 implies more regional mobility between generations than within generations. However, asymptotic mobility remains imperfectly elastic. Second, if the Marshallian externalities discussed in Section 2.2 take time to develop, they will be more pronounced in the long run than in the short run.

Recall that the short run lasts one generation, which is still a considerable period of time. By contrast the long run may be a matter of decades or even centuries. This recognises the fact that

agglomeration is likely to be a long and protracted process. This should be contrasted with the speed of the agglomeration process in models that treat internal migration in an ad hoc fashion. For example, Michel et al. (1996) assume that migration is perfectly elastic in the long run but, due to frictions, is imperfectly elastic in the short run. This specification permits regional wage disparities in the short run but eliminates them in the long run. The speed of the agglomeration process depends in their case on the rate of internal migration as a function of regional wage disparities. In the absence of frictions, the agglomeration process would be infinitely rapid. The microfoundations underpinning these frictions are unclear, yet they play a key role in determining the speed of the agglomeration process, which in any case has been left vague.

By contrast, internal migration in our model is underpinned by microfoundations and the biological clock of the agglomeration process has a generational tick. This stems from the assumption that individuals make one major location decision during their lifetime. In our model internal migration is therefore an intergenerational phenomenon rather an intragenerational phenomenon as in Fujita and Thisse (2002). We could relax the restriction that individuals make only one major location decision during their lifetime in which case internal migration would be both an intergenerational and intragenerational phenomenon. But we would never expect labour to move as freely as capital since capital is footloose, but people have ties to their location. Since intergenerational migration is greater than intragenerational migration, there is more scope for agglomeration in our model than in models which are essentially designed for a single generation.

In Section 2.1 we pointed out that in the initial equilibrium at time (generation) $t = 0$ $P_{AA} = 1/[1 + \exp(-\phi D)]$ and $P_{BA} = 1/[1 + \exp(\phi D)]$. Consider what happens to P_t (A 's population share at $t = 1$) following an increase in P_{BA} (or P_{AA}) induced by an autonomous increase in income in A relative to B . Assuming, for expositional purposes, that P_{AA} and P_{BA} are fixed and therefore have no time indices, the model in Section 2.1 implies the following first-order difference equation for A 's population share:

$$P_t = (P_{AA} - P_{BA})P_{t-1} + P_{BA} \quad (17)$$

the general solution for which is:

$$P_t = P^* + (P_0 - P^*)\lambda^t \quad (18)$$

$$\lambda = P_{AA} - P_{BA}$$

$$P^* = \frac{P_{BA}}{P_{AB} + P_{BA}} \quad (19)$$

Since $P_{AA} > P_{BA}$ it follows that $0 < \lambda < 1$. P^* denotes A 's asymptotic or long run population share. Equation (19) states that A 's long run population share varies directly with P_{BA} and inversely with P_{AB} . Equation (18) states that A 's population share converges upon P^* .

From Equation (17) the short run effect of an increase in P_{BA} is equal to $1 - P_0 = 1/2$. Therefore, if more B residents migrate to A , A 's population share increases by half the change in P_{BA} . To obtain the long run effect we differentiate Equation (19) with respect to P_{BA} :

$$\frac{\partial P^*}{\partial P_{BA}} = \frac{P_{AB}}{(P_{AB} + P_{BA})^2} \quad (20)$$

In the symmetric case when $P_{AB} = P_{BA}$ Equation (20) simplifies to $1/4P_{AB}$. Due to home region preference $P_{AB} < 1/2$, therefore Equation (20) exceeds a half in the symmetric case, as well as in

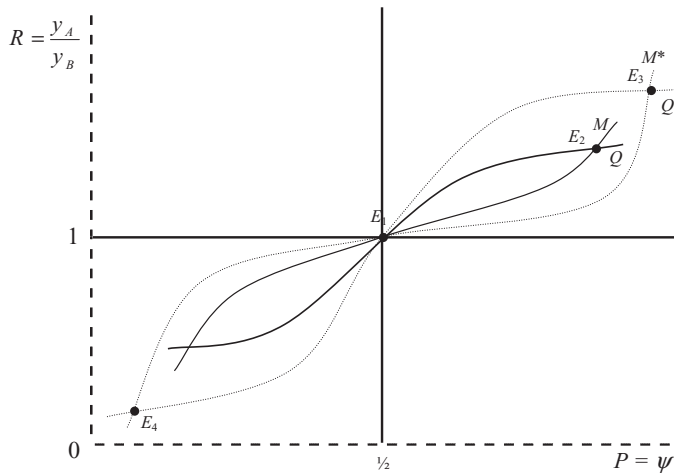


Fig. 5. Dynamic regional equilibrium

almost all asymmetric cases.²² Therefore changes in P_{BA} typically induce more migration in the long run than in the short run. A similar argument applies to changes in P_{AB} . This means that the M schedule in Figure 2 is flatter in the long run than in the short run as indicated in Figure 5 where M denotes the short run schedule and M^* is its long run counterpart. This means that if the income gap increases in favour of region A, A's population share will grow from one generation to the next until it eventually settles down at its new long run equilibrium determined by schedule M^* .

If Marshallian externalities grow over time and the Q schedule is upward sloping as in Figure 5, the short term Q schedule will be flatter than its long term counterpart.²³ This happens for any given population share because relative wages must grow over time if Marshallian externalities are dynamic. If this process is convergent the long run counterpart to schedule Q in Figure 5 is represented by Q^* .

To formalise these dynamic externalities is difficult because relative wages vary directly but nonlinearly with the population share according to Equation (10). In the simple first-order case R and P are related through a nonlinear difference equation such as:

$$R_t = F(R_{t-1}, P_t) \tag{21}$$

If both partial derivatives are positive, an increase in P has a greater effect on R in the long run than in the short run, so that schedule Q^* is steeper than schedule Q .

The short term agglomerating equilibria (E_1 and E_2), determined by the intersections between schedules M and Q , have already been discussed in Figure 3. The long run agglomerating equilibria are determined by the intersections between schedules M^* and Q^* (E_3 and E_4). There is more agglomeration in the long run than in the short run because schedule M^* lies below schedule M when $P > 1/2$ and schedule Q^* lies above schedule Q .

Not shown in Figure 5 is the trajectory between the short and long run equilibria, for example, between E_1 and E_3 . To solve the trajectory for state variables R and P it would be necessary to substitute Equation (21) into Equation (17) since P_{AA} and P_{BA} vary directly but nonlinearly with R . What would result is a set of second order nonlinear difference equations in

²² If, for example, $P_{AB} = 0.45$, P_{BA} must be less than 0.4987 for Equation (20) $> 1/2$.

²³ If schedule Q slopes downwards it will be flatter.

the state variables, which do not have analytical solutions. It is obvious that A 's population share and the relative wage may overshoot their long run equilibrium at E_3 . And even in the absence of overshooting the adjustment paths for the state variables may not be monotonic. Clearly many different adjustment paths are possible. What is important for our purposes is not the adjustment path, about which little can be said, but the fact that there must be more agglomeration and regional wage inequality in the long run than in the short run. In terms of Figure 5, E_3 must lie to the north east of E_2 .

When there are more than two regions people face greater migration choices so that the pull of home region preference naturally weakens.²⁴ Therefore migration in both the short and long runs is greater in the multi-region case. When there are only two regions intergenerational migration is forward and backwards; if children migrate from A to B , grandchildren migrate back to A where their grandparents live. By contrast, in a multi-region context grandchildren may migrate onwards to regions C and beyond. Nevertheless, while the two region case is restrictive, it has the virtues of simplicity, and it serves the purpose of demonstrating that there must be more migration intergenerationally than intragenerationally.

5 Conclusions

Although urban economists and growth theorists have attached importance to technological spillovers, regional agglomeration theorists have focused almost exclusively on pecuniary scale economies. We suggest that technological scale economies may also form the basis of a theory of regional agglomeration. Our model extends and generalises Michel et al. (1996) by focusing upon the interaction between scale dependence in amenities and total factor productivity, and by assuming that individuals have heterogeneous preferences regarding where they wish to live and work. In Michel et al. (1996) wage inequality depends crucially on the scale dependence of amenities. In our model regional wage inequality arises out of heterogeneous preferences as well as scale dependence in amenities. In the absence of scale dependence in amenities, Michel et al. (1996) are forced to assume that unskilled labour is immobile to prevent complete agglomeration. In our model such corner solutions are avoided more naturally by heterogeneity in residential choices. We show that relative wages in the agglomerating region may increase under some circumstances, but decrease under others. However, the existence of Marshallian scale externalities do not automatically lead to regional agglomeration.

If marginal externalities are always positive there will be multiple equilibria, and the agglomerating equilibria are stable. In the two region case it is a matter of chance or history whether agglomeration occurs in one region rather than another. If there is agglomeration, small shocks are self correcting, but large shocks may alter the equilibrium. In this way ghost regions may emerge, while other regions become 'hot spots'. If in addition, there are scale economies in amenities, the number of multiple equilibria may increase. This means that shocks may not have to be so large to create a new regional equilibrium.

We also investigated the dynamic implications of our model for agglomeration. In doing so, we distinguish between two forces. The first concerns dynamic effects in Marshallian externalities which may be larger in the longer run than in the short run. The second concerns labour mobility between regions, which according to our residential choice theory is greater between generations than within generations. These two forces induce more agglomeration between generations than within generation, but their implications for regional wage disparities are ambiguous.

²⁴ Variety naturally reduces the demand for individual items according to discrete choice theory since there are greater opportunities for substitution.

Finally, it might be asked why do we need another theory of regional agglomeration when NEG already exists? First, the theoretical routes taken by the two theories are very different. There is an obvious intellectual interest when phenomena such as regional agglomeration and regional inequality can be explained using different and even rival axioms. Second, as noted by Suedekum (2006) NEG predicts that real wages should be lower in regions that are more agglomerated. As noted in Section 2.2, this result changes when housing is introduced into the NEG model. Empirically, there is growing evidence that regional wages are positively correlated with agglomeration (Beenstock and Felsenstein 2008). Our model (Figure 3) predicts that real wages should be positively correlated with agglomeration even without the introduction of housing. However, under some circumstances this correlation may be negative (Figure 2). Third, the empirical contents of the two theories are very different. NEG directs empirical investigators to focus on pecuniary scale economies and product varieties. By contrast Marshallians will direct their empirical efforts to the determinants of regional *TFP* including networking and related externalities. In the final analysis, of course, the theories have to be tested empirically. Finally, the Marshallian model is simpler and more transparent than NEG. It has analytical solutions whereas NEG is usually solved numerically.

References

- Abdel-Rahman HM (2000) Multi-firm city versus company town: A micro-foundation model of localisation economies. *Journal of Regional Science* 40: 755–769
- Abdel-Rahman HM, Fujita M (2000) Product variety, Marshallian externalities and city size. *Journal of Regional Science* 30: 165–182
- Anselin L, Varga A, Acs ZJ (1997) Geographic spillovers between university research and high technology innovations. *Journal of Urban Economics* 42: 422–448
- Azzoni CR, Servo LMS (2002) Education, cost of living and regional wage inequality in Brazil. *Papers in Regional Science* 81: 157–175
- Baldwin R, Forslid R, Martin P, Ottaviano G, Robert-Nicoud F (2003) *Economic geography and public policy*. Princeton University Press, Princeton, NJ
- Barro RJ, Sala-I-Martin X (1991) Convergence across states and regions. *Brookings Papers in Economic Activity* 1: 107–182
- Beenstock M, Felsenstein D (2008) Regional heterogeneity, conditional convergence and regional inequality. *Regional Studies* 42: 475–488
- Berliant M, Reed RR, Wang P (2006) Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics* 60: 69–95
- Blomquist GC (2006) Measuring quality of life. In: Arnott RJ, McMillen DP (eds) *A companion to urban economics*. Blackwell, Oxford
- Brakman S, Garretsen H, van Marrewijk C (2001) *An introduction to geographical economics*. Cambridge University Press, Cambridge
- Charlot S, Duranton G (2004) Communications externalities in cities. *Journal of Urban Economics* 56: 581–613
- DaVanzo J, Morrison F (1981) Return and other sequences of migration in the US. *Demography* 18: 85–101
- David PA, Rosenbloom JL (1990) Marshallian factor market externalities and the dynamics of industrial location. *Journal of Urban Economics* 28: 349–370
- Dixit AK, Stiglitz JE (1977) Monopolistic competition and optimum product diversity. *American Economic Review* 67: 297–308
- Duranton G, Monastiriotis V (2002) Mind the gaps: The evolution of regional inequalities in the UK 1982–1997. *Journal of Regional Science* 42: 219–256
- Evans P, McCormick B (1994) The new pattern of regional unemployment: Causes and policy significance. *The Economic Journal* 104: 663–647
- Duranton G, Puga D (2004) Micro-foundations of urban agglomeration economies. In: Henderson JV, Thisse J-F (eds), *Handbook of regional and urban economics*, Vol. 4. North Holland, Amsterdam
- Ferguson CE (1969) *The neoclassical theory of production and distribution*. Cambridge University Press, Cambridge
- Fu S (2007) Smart café cities: Testing human capital externalities in the Boston Metropolitan Area. *Journal of Urban Economics* 61: 86–111
- Fujita M, Krugman P, Venables A (1999) *The spatial economy: Cities, regions and international trade*. MIT Press, Cambridge, MA

- Fujita M, Thisse J-F (2002) *Economics of agglomeration: Cities, industrial location and regional growth*. Cambridge University Press, Cambridge
- Glaeser EL (1999) Learning in cities. *Journal of Urban Economics* 46: 254–277
- Glaeser EL, Mare DC (2001) Cities and skills. *Journal of Labour Economics* 19: 316–342
- Grossman GE, Helpman E (1993) *Innovation and growth in the global economy*. MIT Press, Cambridge, MA
- Harcourt G (1972) *Some Cambridge controversies in the theory of capital*. Cambridge University Press, Cambridge
- Hartman S, Doanne MK, Woo CK (1991) Consumer rationality and the status quo. *Quarterly Journal of Economics* 106: 141–162
- Helpman E (1998) The size of regions. In: Pines D, Sadka E, Zilcha I (eds) *Topics in public economics: Theoretical and applied analysis*. Cambridge University Press, Cambridge
- Helsley RW, Strange WC (2004) Knowledge barter in cities. *Journal of Urban Economics* 56: 327–345
- Henderson JV (1974) The sizes and types of cities. *American Economic Review* 64: 640–656
- Henderson JV (2003) Marshall's scale economies. *Journal of Urban Economics* 53: 1–28
- Hicks JR (1932) *A theory of wages*. Macmillan, London
- Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localisation of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics* 108: 577–598
- Krugman PR (1991) Increasing returns and economic geography. *Journal of Political Economy* 99: 483–499
- La Fountain C (2005) Where do firms locate? Testing competing models of agglomeration. *Journal of Urban Economics* 58: 338–366
- Long L (1991) Residential mobility differences among developed countries. *International Regional Science Review* 14: 133–148
- Marshall A (1919) *Principles of economics*, 8th edition. Macmillan, London
- Marshall A (1920) *Industry and trade*. Macmillan, London
- Jovanovic B, Rob R (1989) The growth and diffusion of knowledge. *Review of Economic Studies* 56: 569–582
- Maier G, Weiss P (1986) The importance of regional factors in the determination of earnings: The case of Austria. *International Regional Science Review* 10: 211–220
- Michel P, Perrot A, Thisse J-F (1996) Interregional equilibrium with heterogeneous labour. *Journal of Population Economics* 9: 95–114
- Murata Y (2003) Product diversity, taste heterogeneity, and geographic distribution of economic activities: Market vs. non-market interactions. *Journal of Urban Economics* 53: 126–144
- Nocco A (2009) Preference, heterogeneity and economic geography. *Journal of Regional Science* 49: 33–56
- Puga D (1999) The rise and fall of regional inequalities. *European Economic Review* 43: 303–334
- Rosenthal S, Strange W (2004) Evidence on the nature and sources of agglomeration economies. In: Henderson JV, Thisse J-F (eds) *Handbook of regional and urban economics*, Vol. 4. North Holland, Amsterdam
- Storper M, Venables AJ (2004) Buzz: Face to face contact and the urban economy. *Journal of Economic Geography* 4: 351–370
- Suedekum J (2006) Agglomeration and regional living costs. *Journal of Regional Science* 46: 529–543
- Tabuchi T, Thisse J-F (2002) Taste heterogeneity, labour mobility and economic geography. *Journal of Development Economics* 69: 155–177