Revised: 8/10/2010

Productivity in cities:
self-selection and sorting*

Anthony J. Venables
Department of Economics
University of Oxford.

Abstract

Productivity is high in cities partly because the urban environment acts as a self-selection mechanism. If workers have imperfect information about the quality of workers with whom they match and matches take place within cities, then high-ability workers will choose to live and work in expensive cities. This self-selection improves the quality of matches in such cities. The mechanism may be reinforced by the development of informational networks in cities with a large proportion of high ability workers. As a consequence productivity in these cities is high for workers of all ability types.

Keywords: economic geography, productivity, city, urban, sorting, self-selection.

JEL codes: R0, R1

* Paper prepared for the 10th anniversary issue of the *Journal of Economic Geography*; it completes ideas first developed in Venables (2002). Thanks to Vince Crawford and referees for helpful comments.

Correspondence address:
Department of Economics,
University of Oxford,
Manor Road Building,
Oxford OX1 3UQ,
United Kingdom

Tony.venables@economics.ox.ac.uk

1. Introduction

Space is a major determinant of the efficiency of economic and social interactions; urban environments are good for some interactions, and proximity is good for most. These simple observations have profound implications and have spawned a decade of innovative papers in the *Journal of Economic Geography*. Why does space matter? What are the underlying mechanisms that create spatial variations in efficiency and quantitatively how important are they? Given the existence of these effects, what are their general equilibrium implications for cities, regional inequalities, and the distribution of economic activity across countries? And can improved understanding of these forces lead to better policies for urban and regional design, and also for the wider agenda of international trade and development? The insights provided by research into these questions fundamentally changes the way we see the world. The smoothing forces of diminishing returns are countered by forces of agglomeration. Spatial inequalities in economic activity and income arise endogenously and persistently, not just as transient phenomena. Policy issues are seen more clearly, although translating the insights of economic geography into policy has proved difficult, partly because of the pervasive nature of some of the market failures involved, and partly because policy instruments do not map directly into outcomes; multiple equilibria, hysteresis and path dependence make the effects of policy unpredictable.

This short paper contributes to the first set of issues outlined above; spatial productivity effects. We know that these can arise from 'technological' externalities that are spatially limited in range – knowledge spillovers, Marshall's 'secrets of the trade', and increased opportunities for learning. They can also arise from firm level increasing returns to scale and indivisibilities in production; these interact with transport costs to provide benefits from proximity to markets and suppliers.[1] The research frontier is now based on the idea that productivity gains arise as spatial organisation can mitigate existing market failures.[2] Thick markets are more competitive, less prone to hold up, allow for risk pooling, and may provide better information and matching possibilities. This paper develops an aspect of this based on imperfect information and the ability of cities to reduce market failures associated with informational asymmetries in the workforce. We show how the high cost of living in cities can induce self-selection by workers, so that more expensive cities have disproportionately many high-ability workers.

[1] This is the main mechanism that trade theory brought to economic geography, see Fujita et al (1999).
[2] Puga and Duranton (2004) and Puga (2010) offer good surveys

The context is one of partnerships formed between workers to undertake projects. Work in many professional activities takes this form. Co-authorship is one example; others are teams of architects and engineers; investment bankers and clients; film producers and directors.[3] The formalization we develop here is matching between pairs of individuals, but the ideas can be extended to matching between firms, such as the firms of architects, engineers, surveyors, lawyers and builders needed for a building project. Why might partnerships work better in some economic environments – in particular in cities – than others? One line of argument is to do with the value of face-to-face (F2F) contact and the role of economic density in facilitating such contacts.[4] F2F is an efficient mode of communication allowing high frequency exchange of ideas and 'complex discourse' (Searle 1969). It facilitates building trust, a point made by many authors (e.g. Putnam 2000) and now receiving support from behavioral economics. Valley et al (2002) conduct an experiment to investigate inefficiencies in a bargaining game and show how pre-play written communication has ambiguous effects, while F2F communication unambiguously reduces inefficiencies by enabling 'high levels of communication not compromised by deception' (Valley et al p148). Another line of argument comes from the matching literature. Larger and thicker labour markets can improve the quality of the match between firms with particular skill needs and workers with particular skill attributes, can increase competition in the matching process, and can increase the frequency of meetings (Helsley and Strange 1990, Amiti and Pissarides 2005, Glaeser 1999).

This paper takes a different line, arguing that cities provide a way of keeping low quality individuals out of matches. The argument rests on aspects of partnerships that have received attention in the literature on assortative matching (e.g. Shimer and Smith 2000). Workers are heterogenous (differing in ability) and types are private information, known to the worker but not observed by potential partners. Partnerships exhibit supermodularity so that types are productive complements and there are aggregate gains from positive assortative matching (matching similar types rather than mixing).[5] In our context, this means simply that the benefit of forming a partnership with someone of high ability is greater for those with high ability than for those with low ability. We outline a model in which these forces combine with cost of living differences between cities to induce self-selection by workers of different abilities. An expensive city will attract a higher proportion of high ability workers

[3] Standard references on the economics of partnerships include Farrell and Scotchmer (1988), Shimer and Smith (2000).

[4] Storper and Venables (2004) offer an overview of these issues.

than a city with a low cost of living. It will therefore have higher productivity both because of the direct effect of the ability mix of the population and also because, given ability, workers of all types are making better matches and therefore being more productive. We then go on to argue that these effects might be reinforced by informational networks that are likely to develop in cities with a high proportion of high ability workers, but not in cities where this proportion is low. The model we develop is kept simple, but provides a basis for some of the empirical findings about the role of sorting in explaining urban productivity.[6]

2. Cities induce self-selection

We initially focus on one sector or profession that has a fixed number of workers of whom H are high ability (type-H) and L low ability (type-L). There are two cities, 1 and 2, and workers choose to live and work in one or the other. The proportions of type-H and type-L workers who choose city 1 are θ_H and θ_L respectively, so the total numbers of workers in each city are $N^1 = H\theta_H + L\theta_L$, $N^2 = H(1-\theta_H) + L(1-\theta_L)$.

Output in the sector is produced by pairs of workers who form a partnership to undertake a project. If two high ability workers form a partnership the value of their output is $2q_{HH}$, two low-ability $2q_{LL}$, and one of each ability level $2q_{HL}$, $q_{HH} > q_{HL} > q_{LL}$. We assume that individuals know their own type but cannot directly observe that of the partner with whom they undertake the project. Since ability is unobserved when partnerships are formed matches take place randomly, and crucially *within* each city. Thus, the probability of matching with a high-ability worker depends on location. The probability in city *i* is denoted μ^i so

$$\mu^1 = \frac{H\theta_H}{H\theta_H + L\theta_L}, \qquad \mu^2 = \frac{H(1-\theta_H)}{H(1-\theta_H) + L(1-\theta_L)} \quad . \tag{1}$$

The total output of a partnership is divided equally between the two partners. The expected returns to a match made by type-H and type-L individuals in each city, v_H^i, v_L^i, are therefore

$$v_H^i = q_{HH}\mu^i + q_{HL}(1-\mu^i), \qquad v_L^i = q_{HL}\mu^i + q_{LL}(1-\mu^i). \tag{2}$$

---

5 See Milgrom and Roberts (1990) for discussion of supermodularity and forms of complementarities.

6 See Combes, Duranton and Gobillon (2008), Andersson, Burgess and Lane (2007).

Locating in city 1 has a cost attached to it, denoted *c*. This might be the rent or commuting cost differential of city 1 compared to city 2 and, for the moment, we treat it as exogenous. Workers of each type locate in the city offering the higher return, so city 1 is chosen by type-H-individuals if $v_H^1 - v_H^2 \geq c$ and type-L individuals if $v_L^1 - v_L^2 \geq c$. Using (2), these conditions become,

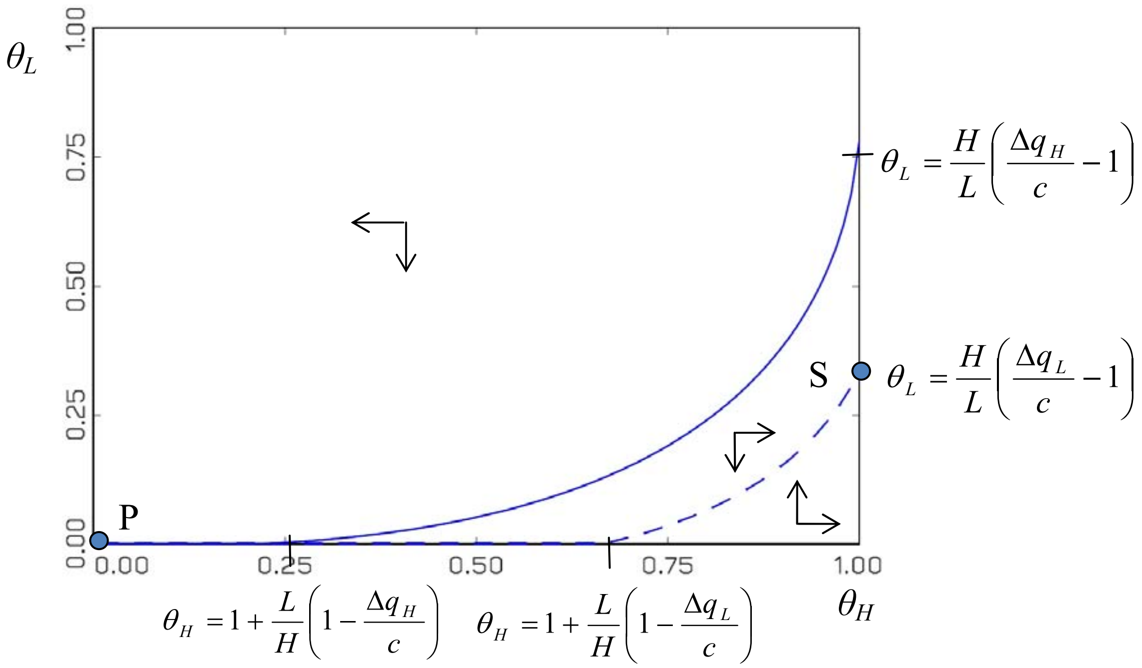
$$
\begin{aligned}
v_H^1 - v_H^2 &= (q_{HH} - q_{HL})(\mu^1 - \mu^2) \geq c, \\
v_L^1 - v_L^2 &= (q_{HL} - q_{LL})(\mu^1 - \mu^2) \geq c.
\end{aligned}
\tag{3}
$$

The equilibrium location of workers is where values of θ_H and θ_L have adjusted to make workers indifferent between cities, or at a corner solution where all workers of a particular type are in their preferred location.

If $c = 0$ there is evidently an equilibrium with $\mu^1 = \mu^2$, so the two cities are identical. What happens if $c > 0$? The equilibrium composition of the cities can be seen most easily by plotting values of θ_H and θ_L at which workers of each type are indifferent between the cities, and this is done in fig. 1 (using equations (1) with (3) set equal to zero, as given in the appendix). The figure is drawn for the case in which a good match is more valuable for type-H individual than for a type-L (super-modularity) i.e. $q_{HH} - q_{HL} > q_{HL} - q_{LL}$. We denote $\Delta q_H \equiv q_{HH} - q_{HL}$, $\Delta q_L \equiv q_{HL} - q_{LL}$, giving intercepts as indicated. Along the solid and dashed curves workers of types H and L respectively are indifferent between the two cities. These curves are upwards sloping as an increase in θ_L (the number of low ability workers in city 1) reduces the quality of the match and the return to locating in city 1; for indifference, this must be compensated by an increase in θ_H. Above each of the curves workers prefer to be in city 2 and below in city 1, this represented on the figure by arrows giving directions of change towards preferred locations. The relative positions of the curves are for the case $q_{HH} - q_{HL} > q_{HL} - q_{LL}$ and are reversed otherwise.

The case illustrated has two equilibria. The pooling equilibrium is at point P, with no workers in the expensive city. More interesting is point S, the separating equilibrium, in which all type-H workers are in city 1, ($\theta_H = 1$), these workers being strictly better off than they would be in city 2. Type-L workers are indifferent between cities and fraction $\theta_L = (\Delta q_L / c - 1) H / L$ of them are in city 1, the remainder being in city 2.

Figure 1: Self-selection into cities



Solid line: type-H indifferent; dashed line, type-L indifferent.

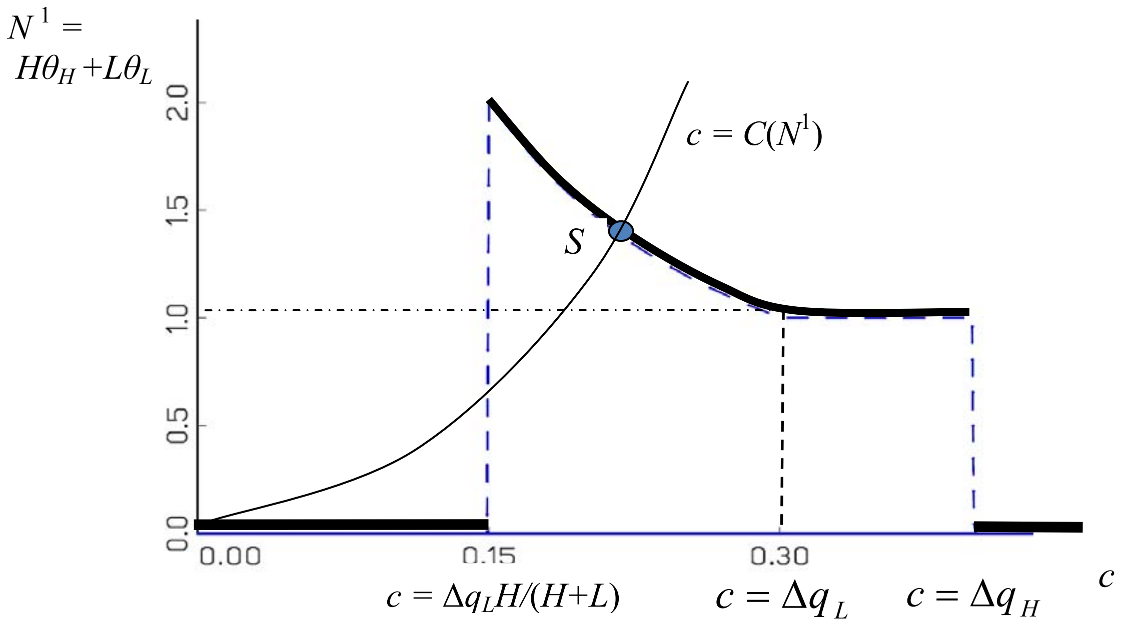
Parameters: H = L = 1, $q_{HH} = 0.8$, $q_{HL} = 0.4$, $q_{LL} = 0.1$, $c = 0.225$.

The mechanism that supports this separation is self-selection of the simplest form, dating back to Spence (1973). The assumption that payoffs are super-modular (which in this context is simply $q_{HH} - q_{HL} > q_{HL} - q_{LL}$) means that good matches are worth more to type-H individuals than to type-L. Differential city costs then provide a mechanism which induces sorting by serving to keep some low ability workers out of the high cost city. Reversing this inequality (removing supermodularity) reverses the positions of the curves on fig. 1 and destroys the argument. The implication of the separating equilibrium for productivity differentials between cities is clear. The high cost city has higher productivity than the low cost one for two reasons. The first is that it has a higher proportion of type-H individuals who (regardless of population mix) expect to do better than type-L. In addition to this, all workers in the expensive city make a better average quality match. Even type-L individuals are more productive than they would be in the low cost city because they face a higher probability of matching with a type-H.[7]

[7] If super-modularity is dropped, so $q_{HH} - q_{HL} < q_{HL} - q_{LL}$, then the position of the curves is reversed, and there is no separating equilibrium as can be readily checked by inspection of a figure analogous to fig.1.

What level of costs supports the separating equilibrium? Fig. 2 has the population of city 1 on the vertical axis and the cost differential c on the horizontal. Point S corresponds exactly to S on fig.1, and the bold discontinuous line through S gives equilibrium city size as c varies.[8] At S city 1 contains all type-H workers and around 1/3 of type-L workers. Increasing c from this point reduces θ_L and the number of type-L workers, because they need better matches to compensate for higher city costs; this variation gives the downwards slope of the line through S. At $c > \Delta q_L$ there are no type-L workers left in city 1 (S has $\theta_H = 1$, $\theta_L = 0$, see fig. 1); all type-H workers match with other type-H workers. There is no incentive for a type-H worker to deviate to city 2 as long as $c < \Delta q_H$; however, at levels of c greater than this type-H workers prefer to pool with type L in city 2 than to pay c, so no workers choose to locate in city 1. What happens at low c? At $c \le \Delta q_L H/(H+L)$ all type-L workers would choose to locate in city 1 ($\theta_H = 1$, $\theta_L = 1$, see fig. 1). There is then no-self selection effect operating, so all workers are better off pooling in a city that does not incur cost c. Separating equilibria therefore exist only in the interval $c \in (\Delta q_L H/(H+L), \Delta q_H)$.

Figure 2: City costs and population.



[8] Fig. 2 has the same parameter values as fig.1, except that c now varies, moving S on the vertical axis of fig. 1.

We can compare the real incomes of workers of each type in the pooling and separating equilibria, by noting that in the pooling equilibrium type-H and type-L workers get respectively v_H^P, v_L^P,

$$v_H^P = \frac{q_{HH} H + q_{HL} L}{H + L}, \qquad v_L^P = \frac{q_{HL} H + q_{LL} L}{H + L}$$

In the separating equilibrium type-L workers get $v_H^2 = q_{LL}$, and type-H workers get $v_H^1 - c = c(q_{HH} / \Delta q_L - 1)$, found by using $\theta_H = 1$, and $\theta_L = (\Delta q_L / c - 1)H / L$ in equations (1) and (2). This is *increasing* in *c*, so that direct city costs are more than offset by better quality matches. Comparison of these values gives the following results. Type-L workers are certainly worse off with separation than with pooling. If *c* takes the lowest value that supports separation, $c = \Delta q_L H /(H + L)$, so too are type-H workers; however, at higher values of *c* the comparison is ambiguous, depending on the size of the two groups in overall population. The real income difference between type-H and type-L is greater in the separating equilibrium than the pooling equilibrium for all values of *c* that support pooling, if it is the case that $q_{LH} > 2q_{LL}$. As is usual in models of this type, separation may therefore both be Pareto inferior to pooling and increase inequality.

To this point we have taken the cost of locating in city 1, *c*, as exogenous, and fig. 2 illustrates how it may be endogenised. A standard urban model relates the cost of urban living to city size (commuting costs and rent, Alonso 1964). The relationship is generally increasing and concave (as a function of N), as illustrated by the curve $C(N^1)$, drawn to go through point S. Thus, it is possible that the very fact that an expensive city attracts high ability workers gives the city its size and hence a high cost of living. The figure suggests how a full general equilibrium model could be constructed, although such a model would of course contain a number of cities, all with endogenous populations, city costs c^i, and with workers having employment opportunities in multiple sectors, some tradable and others non-tradable.

Finally, we have worked with a single profession or occupation and exogenous cost differential *c*. If there are multiple professions with the characteristics outlined here, then the model predicts that these professions will cluster in the high-cost city. There is no inherent linkage that would cause these professions to co-locate, but high ability individuals in these professions will all be drawn to the city with costs high enough to support separation of types in their profession. Urban diversity arises simply as different professions share the screening benefits of high costs.

3. Cities and reputation

Spatial cost differentials are one way of keeping low ability people out of matches. Are there others? The previous model is based on inability to directly observe an individual's type but in reality aspects of peoples' performances – such as the outcomes of projects they have undertaken – are observable, so that reputations can be built. But how does a reputation get disseminated and stored within a group of potential project partners? Can people with a track record of failing projects become anonymous, so their record is lost? One way to think about this is as professions developing a shared knowledge of who is 'in-the-loop' or in the 'in-group'. This is not knowledge of the track-record of every individual, nor a formal assessment process, but instead an evolving group view of who is 'in' and who is 'out'. An informal group of insiders develops and members of the group implicitly put themselves up to judge, be judged, and share their judgements with others.

Is there a geographical basis for such shared group knowledge? In some professions not: in academia (or economics at least) the information network is near global. In other professions group formation is more likely to have an urban basis. London based lawyers, architects or designers are likely to know who is in the London in-group, but not the Manchester one. The need to observe project outcomes and to know and recall individuals' names puts a natural geographical limit on the formation and membership of a group.

To explore these ideas we modify the model of the preceding section. We focus for the moment on a single city. Partnerships get made between members of an 'in group', and the proportion of the group who are type-H is μ; the group is not a perfect screening device ($\mu \leq 1$) so contains some type-L individuals, and this composition is determined below. Since matching takes place within the group the probability of an individual matching with a type-H is μ. As before, partnerships undertake projects, and we now interpret q_{ij} as the probability of success of a project undertaken by a pair of type i, j. If the return to a successful project is unity and the return to a failure zero, then expected returns to type-H and L group members from undertaking a project are,

$$v_H = q_{HH}\mu + q_{HL}(1-\mu), \qquad v_L = q_{HL}\mu + q_{LL}(1-\mu). \tag{4}$$

We move to a dynamic setting, with repeated rounds of projects. Participation in projects requires continued membership of the group, and this depends on reputation. We model this in the simplest possible way, supposing that if a project succeeds membership of the group is maintained. If a project fails, then membership of the group is terminated with probability γ. This termination process is not formal; it is simply that word gets around the group that you are associated with a failure. We model how γ is determined in what follows.

Partnerships are formed within the group and the earnings for non-group members are normalised at zero. With discount rate δ the expected present values of group membership for type-H and type-L individuals, V_H, V_L, are therefore

$$V_H = v_H + \frac{V_H}{1+\delta}\left[1-\gamma(1-v_H)\right], \qquad V_L = v_L + \frac{V_L}{1+\delta}\left[1-\gamma(1-v_L)\right]. \tag{5}$$

Thus, a type H individual expects v_H from the current project, and has probability $1-\gamma(1-v_H)$ of still being in the group for the next project round; $(1-v_H$ is the probability of a project failing and γ the probability that you are then ejected). Rearranging,

$$V_H = \frac{v_H}{\delta+\gamma(1-v_H)}, \qquad V_L = \frac{v_L}{\delta+\gamma(1-v_L)}. \tag{6}$$

Initial entry to the group is done on the basis of comparison with an outside option. Since this is normalised at zero, we will suppose that there is entry fee c, so type-L individuals will enter the group until $V_L = c$. When this holds all type-H individuals will enter the group, since $V_H > V_L = c$. Entry of type-L individuals reduces the average quality of the group, lowering μ, so will take place until μ is low enough that it is not worthwhile for further type-L to enter. Setting in $V_L = c$ in (6) and using (4) this means that μ is determined from

$$v_L = q_{LL} + \mu(q_{HL} - q_{LL}) = \frac{c(\gamma+\delta)}{\gamma c+1}, \quad \text{so} \quad \mu = \frac{1}{\Delta q_L}\left\{\frac{c(\gamma+\delta)}{1+\gamma c} - q_{LL}\right\}. \tag{7}$$

The value of μ from this equation can be used to give the return to type-H individuals of group membership. Using (7) in (4) the returns on a particular project are $v_H = \mathrm{v}(\gamma, c)$, with the function increasing in both arguments (see appendix).[9] Using this in (5) gives the expected present value of group membership, V_H, as a function v(γ, c) and γ, $V_H = V_H\left(\mathrm{v}(\gamma,c),\gamma\right)$. This function is plotted on fig. 3, which has γ, the probability a participant in a failing project is ejected from the group, on the horizontal axis, and higher curves drawn for higher entry costs, c.[10]

With this in mind, what determines the value of γ, the probability of a failure causing ejection from the group? Geography and technology set an upper bound which we denote $\hat{\gamma}$. Thus, if there is any chance that individuals can re-enter the group (e.g. by becoming anonymous) then ejection is not absolute and γ is less than unity. The upper bound $\hat{\gamma}$ is likely to be sector and city specific, but its exact level is not our primary focus. More importantly, γ is chosen by members of the group and we assume that this is a type-H majority. Once again, we stress that this is not a formal or precise choice, but emerges from group interaction. The key point is that the group may choose to not use the maximum possible ejection probability $\hat{\gamma}$ because of the risk of making errors. It may not be the interests of the group to blacken the reputation of all of its members who find themselves engaged on a failing project, because this will include some type-H individuals.

Fig. 3 draws out implications. The function $V_H\left(\mathrm{v}(\gamma,c),\gamma\right)$ is convex and for relatively small c has negative slope at $\gamma = 0$ (see appendix). Type-H people will want a group to form (i.e. to set a positive value of γ which ejects some participants in failing projects) only if $V_H\left(\mathrm{v}(\gamma,c),\gamma\right) > V_H\left(\mathrm{v}(0,c),0\right)$. On the figure, this holds for the curves on or above $V_H\left(\mathrm{v}(\gamma,c_2),\gamma\right)$ but fails for curve $V_H\left(\mathrm{v}(\gamma,c_1),\gamma\right)$, $c_2 > c_1$. Thus, if the cost of entering the group is low it will not be worthwhile to introduce any ejection process. A large number of type-L enter the group, so type-H people have a high probability of making failing matches and thereby risking ejection; they will choose $\gamma = 0$, and no group will form. If the cost of entering the group is high enough – so the quality mix is good enough – then the highest feasible value, $\hat{\gamma}$, will be chosen.

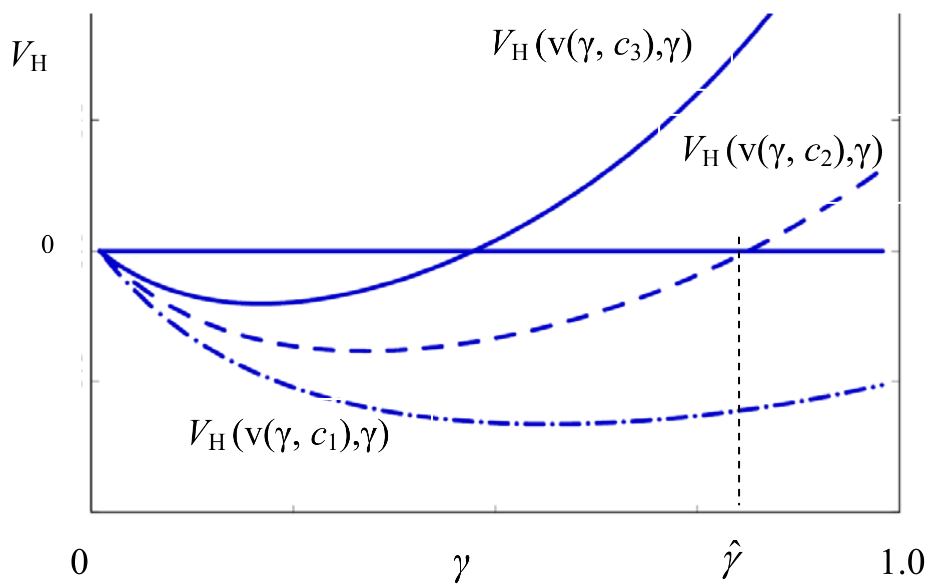
This result clearly depends on the U-shape of $V_H\left(\mathrm{v}(\gamma,c),\gamma\right)$. The intuition is as follows. Higher γ reduces V_H through its direct effect, increasing the risk of error; some type-H people engaged in failing matches may be ejected. However, it raises V_H through deterring entry of type-L individuals, creating better quality matches, raising μ and hence v_H.

[9] It must be the case that $c < 1/\delta$, i.e. the present value of an infinite stream of success exceeds the entry cost.
[10] Parameter values q_{HH}=1.0; q_{HL}=0.5; q_{LL}=0.0; δ=0.2; c_1=0.4; c_2=0.55; c_3=0.7.

The first of these effects is larger at low γ, since more type-H people then find themselves in failing matches. The rotation of the curve with c has the same intuition. Low c means that a lot of type-L people enter, and hence matches are of lower quality and there is a greater probability of ejecting type-H people. Hence the group 'chooses' to eject no-one (i.e. does not form). Only if c is high enough do type-H people choose to set a positive value of γ.

What we have established is then, that if the population mix is initially sufficiently high quality – a high value of c and associated self-selection process has kept out some type-L people and raised μ – then a group will form. The information network develops and people in the group put themselves up to judge and be judged. This further refines the group and raises productivity further. But in a city that initially has a low value of μ a group of this type will not form; the large number of type-L individuals means that the observation of failing projects does not provide a sufficiently accurate screen to be of use to type-H individuals. The information sharing mechanism of this section therefore is complementary with and amplifies the effects of the pure self-selection model of the preceding section.

Figure 3: Choosing the value of γ



4. Concluding comments

An extensive body of empirical research establishes that large cities have relatively high levels of earnings, a high cost of living, a high proportion of high skilled workers, and that the incremental returns to an urban environment are greater for high skilled workers than for those with lower skills (Glaeser and Resseger 2010). These effects cover a wide range of occupations, implying that cities are diverse. This paper shows how these features can be generated by a model with a very few basic ingredients. Output involves partnership of workers who cannot directly observe the ability of their partners. Working with good partners is more valuable for high ability workers than for low ability, and partnerships are formed within cities. The cost of living in a city then acts to induce self-selection, so high ability workers choose to live in a high cost city, and low ability workers are mainly located in a low cost city. At this separating equilibrium the high cost city exhibits all the characteristics listed above. Living in the city has no direct effect on technical efficiency, and raises costs rather than reducing them. However, willingness to incur these costs has the effect of signalling high ability, and therefore supports a separating equilibrium. These effects can be magnified if cities are also the basis for building reputational networks in which a shared (but not perfectly accurate) body of knowledge develops about individuals' abilities.

The model provides an example of the research frontier that is open for work on the micro-foundations of agglomeration. It combines an existing market failure (in this case, asymmetric information about ability) with 'micro-heterogeneity' i.e. the fact that economic agents are differentiated. While this paper looks at heterogeneity of workers, the case for bringing micro-heterogeneity of firms into the mainstream of economic geography is made by Ottaviano (2011). Building on the extensive literature on international trade with heterogeneous firms he suggests that firms will sort across locations, with lower-cost firms locating in the better (larger size or lower cost) locations. Both these contexts provide additional arguments for agglomeration and point to the qualitative aspects of agglomeration, as well as the quantitative aspects drawn out in the first wave of new economic geography models. The challenge for theorists is to develop models of heterogeneity which are tractable and insightful, and for econometricians, to distinguish between a yet wider set of determinants of urban agglomeration.

Appendix:

Section 2: Using (1) in (3), the indifference loci on fig. 1 are $\{\theta_H, \theta_L\}$ satisfying:

$$\frac{\Delta q_H}{c} = \frac{(H\theta_H + L\theta_L)(H(1-\theta_H) + L(1-\theta_L))}{HL(\theta_H - \theta_L)}, \quad \frac{\Delta q_L}{c} = \frac{(H\theta_H + L\theta_L)(H(1-\theta_H) + L(1-\theta_L))}{HL(\theta_H - \theta_L)}$$

Intercepts are illustrated on fig.1, evaluating at $\theta_L = 0$ and $\theta_H = 1$. The configuration of the curves is as shown if $\Delta q_H > \Delta q_L$.

Section 3: Using (7) in (4) to eliminate μ gives

$$\mathrm{v}(\gamma, c) = v_H = q_{HL} + \frac{\Delta q_H}{\Delta q_L}\left\{\frac{c(\gamma+\delta)}{1+\gamma c} - q_{LL}\right\} = \frac{1+3\gamma c+2\delta c}{2(1+\gamma c)} \quad \text{(A1)}$$

The final equation in each of A1-A4 is for the case in which $q_{\mathrm{HH}} = 1$, $q_{\mathrm{HL}} = 0.5$, $q_{\mathrm{LL}} = 0$ (supermodularity is not required in section 3). Using this in equation (6),

$$V_H(\mathrm{v}(\gamma, c), \gamma) = \frac{\mathrm{v}(\gamma, c)}{\delta + \gamma(1-\mathrm{v}(\gamma, c))} = \frac{1+3\gamma c+2\delta c}{2\delta + \gamma - c\gamma^2}. \quad \text{(A2)}$$

So, denoting $D \equiv \delta + \gamma(1-\mathrm{v}(\gamma, c))$,

$$\frac{dV_H(\mathrm{v}(\gamma, c), \gamma)}{d\gamma} = \frac{\partial V_H}{\partial \gamma} + \frac{\partial V_H}{\partial \mathrm{v}}\frac{\partial \mathrm{v}}{\partial \gamma} = \frac{1}{D^2}\left[-\mathrm{q}(1-\mathrm{q}) + (\delta+\gamma)\left\{\frac{c(1-c\delta)}{(1+c\gamma)^2}\right\}\right]. \quad \text{(A3)}$$

At γ = 0, $$\frac{dV_H(\mathrm{v}(0, c), 0)}{d\gamma} = \frac{4\delta c - 1}{4\delta^2} \quad \text{(A4)}$$

which is negative for small *c,* and increasing in *c*.

References:

Alonso, W., 1964, '*Location and land use*', Harvard University Press, Cambridge MA.

Amiti, M. and C. Pissarides, 2005, 'Trade and industrial location with heterogeneous labour', *Journal of International Economics*, 67: 392-412

Andersson, F., S. Burgess and J. Lane, 2007, 'Cities, matching and the productivity gains of agglomeration,' *Journal of Urban Economics*, 61: 112-128.

Combes, P., G. Duranton and L. Gobillon, 2008. 'Spatial wage disparities: Sorting matters!' *Journal of Urban Economics*, vol. 63(2), 723-742, Duranton, G. and D. Puga, 2004 'Microfoundations of urban agglomeration economies' in J. V. Henderson and J. F. Thisse (eds.) *Handbook of Regional and Urban Economics*, Vol. 4, 2004. Amsterdam: NorthHolland,

Farrell, J. and S. Scotchmer 1988 'Partnerships', *Quarterly Journal of Economics* 103.2 279-297.

Fujita, M., P. Krugman and A.J. Venables, 1999 *'The spatial economy; cities, regions and international trade*', MIT press Cambridge MA.

Glaeser, E.L. 1999, 'Learning in Cities', *Journal of Urban Economics*, 46: 254-277

Glaeser, E. L. and M.G. Resseger 2010, 'The complementarity between cities and skills' *Journal of Regional Science*, 50: 221-244.

Helsley, R.W. and W.C. Strange 1990, 'Matching and agglomeration economies in a system of cities', *Regional Science and Urban Economics*, 20, 189-212.

Krugman, P.R. 1991, 'Geography and trade', MIT press, Cambridge MA

Milgrom, P., and J. Roberts. 1990. 'The Economics of Modern Manufacturing: Technology, Strategy, and Organization.' *American Economic Review* 80: 511–28.

Ottaviano, G. 2011 ''New' new economic geography; firm heterogeneity and agglomeration economies', forthcoming, *Journal of Economic Geography*.

Puga, D. 2010 'The magnitude and causes of agglomeration economies' *Journal of Regional Science*, 50: 203-219.

Putnam, R. 2000 *Bowling Alone; the Collapse and Revival of American Community.* New York: Touchstone.

Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language.* New York: Cambridge University Press.

Shimer, R. and L. Smith 2000 'Assortative matching and search' *Econometrica*, 68, 371-98.

Spence, A.M. 1974, *Market signalling*, Harvard University Press, Cambridge MA

Storper, M. and A.J. Venables 2004 'Buzz: face-to-face contact and the urban economy' *Journal of Economic Geography*, 4, 351-370.

Valley, K. and L. Thompson, R. Gibbons, M, Bazerman, 2002 'How Communication Improves Efficiency in Bargaining Games', *Games and Economic Behavior*, 38.1: 127-155.

Venables, A.J. 2002 'Labour sorting by cities; partnerships, self-selection and agglomeration', processed London School of Economics.