Cities, Matching and the Productivity Gains of Agglomeration

Fredrik Andersson
Urban Institute

Simon Burgess
University of Bristol, CMPO and CEPR

Julia I. Lane
National Science Foundation

November 2005

* This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications, and is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The data used are a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging (R01-AG18854-01), and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau is preparing to support external researchers' use of these data; please contact U.S. Census Bureau, LEHD Program, FB 2138-3, 4700 Silver Hill Rd., Suitland, MD 20233, USA. Thanks are due to Harry Holzer, and seminar participants at the University of Bristol for helpful comments.

Abstract

The striking geographical concentration of economic activities suggests that there are substantial benefits to agglomeration. However, the nature of those benefits remains unclear. In this paper we take advantage of a new dataset to quantify the role of one of the main contenders – the matching of workers and jobs. Using individual level data for two large US states we show that thicker urban labour markets are associated with more assortative matching between workers and firms. Another critical condition is required for this to generate higher productivity: complementarity of worker and firm quality in the production function. Using establishment level productivity regressions, we show that such complementarity is found in our data. Putting together the production and matching relationships, we show that production complementarity and assortative matching is an important source of the urban productivity premium.

Keywords: Urban Productivity, Matching, Agglomeration
JEL-Code: R23, R12, J24

1. Introduction

Cities are home to 75% of Americans, yet these occupy less than 2% of the land area of the lower 48 states[1]. Indeed, urbanisation has increased dramatically over the last century, not only in the United States but also in other developed countries. This suggests that the benefits of agglomeration are very substantial. However, the nature of those benefits remains unclear. The main possibilities include lower transport costs deriving from high density and increasing returns to scale in production (Krugman, 1991; Ciccone and Hall, 1996), knowledge spill-overs between firms (Lucas, 1988; Rauch, 1993), accelerated human capital acquisition (Marshall, 1890; Glaeser, 1999), and improved labour market matching (Becker and Murphy, 1992; Wheeler, 2001). It has proved difficult to evaluate the relative importance of these theories principally because of the lack of suitable microdata.

The contribution of this paper is to use a unique dataset to quantify the role of one of the main contenders – the matching of workers and jobs. Using linked panels of workers and firms for two US states, we show that thicker urban labour markets are associated with more assortative matching in the labour market. Another condition is required for this to generate higher productivity: complementarity of worker and firm quality in the production function (see for example, Kremer and Maskin, 1996). This of course also provides the incentive for workers and firms to match assortatively. Our data are uniquely suited to address the issue: we have universe longitudinal data on workers and firms, high quality data on their place of work and their place of residence, and also new measures of worker and firm quality. We use the linked panels of workers and firms to estimate worker quality (value of human capital) and firm quality (wage mark-up), using the methodology of Abowd, Kramarz and Margolis (1999) and Abowd, Lengermann and McKinney (2003). We match on to that the spatial coordinates of each worker and each firm. With explicit, market-based measures of worker and firm quality, we can directly investigate their joint distribution, and characterise how this varies over space and over

[1] These facts are from Rosenthal and Strange (2003).

varying labour market thickness in particular. We address the issue of differences in industrial structure by repeating the analysis within industry.

Our data show that there is a significant urban productivity premium. The raw average productivity differential between firms located in counties with a population per square mile in the upper decile and those located in counties with population per square mile below the median is between 0.15 and 0.3 log points across the states in our sample, in favour of the urban firms[2]. These raw productivity differentials cannot be accounted for by differences in industry structure between urban and rural areas – in fact the urban productivity premium is larger within industry. We show that the two conditions for matching to matter are met: there is complementarity in production, and workers and firms are more assortatively matched in dense labour markets. Putting these together, we calibrate the effect on productivity and show that labor market matching is an important source of the urban productivity premium.

The rest of the paper is organised as follows. The next section briefly reviews the literature on this topic, and the following section sets out a modelling framework for our analysis. Section 4 describes the data and section 5 presents the results. The final section offers some conclusions.

2. Literature

Rosenthal and Strange's (2003) recent survey of the evidence on agglomeration economies proposes three main categories of effects, largely following Marshall (1890), though they also add two others. These are knowledge spillovers, input sharing and labour market pooling. Our focus is on the last of these.

In labour market pooling, the matching or pairing of agents becomes an issue, in particular the degree of positive assortative matching (PAM). Becker (1973) first discusses PAM in the context of marriage in a frictionless world. He shows that

[2] Similar productivity differentials are found if we use employment per square mile in the firm's census tract as a density measure.

complementarity in the household production function generates PAM. Shimer and Smith (2000) provide a general analysis and proof of a similar result in a model with search frictions[3]. They provide examples to show that the result does not carry over straightforwardly, but they are able to establish restrictions on the production function that ensure PAM[4]. Burdett and Coles (1999) provide a very useful overview of the issues and some simple models. They set out a model with discrete *ex ante* heterogeneity, Nash bargained utilities, and an exogenous arrival rate of offers. They show that five types of pure strategy equilibria[5] will occur for different specifications of the joint production function. In particular sufficient complementarity in production yields PAM (the 'elite' equilibrium in their description). A particularly important result is that as the offer rate increases (as search frictions decline), the market equilibrium tends to the elite outcome (pp. F325, F326). In our context, assortative matching plus complementarity in production lead to an urban productivity premium. In equilibrium, once workers and jobs are allowed to relocate across areas, this relocation will accentuate the productivity premium by inducing a difference in mean match quality between city and rural areas.

Delacroix (2003) provides a very useful model that we will use in the subsequent section. He assumes heterogeneity on both sides of the market and a random meeting technology. Agents can transfer utility between themselves via the wage setting process. Once created, matches are subject to exogenous match breakdowns. The model equilibrium is defined by decision rules on who each agent will accept matches with. A number of equilibria with different sorting properties are possible. A key assumption on the nature of the production function rules out others: that it is supermodular. This implies that high quality agents are better off matched with other high quality agents. For example, if there are two low quality agents, labeled L, and two high quality agents, labeled H, then a pairing of {LL, HH} produces more than {LH, LH}. This is a crucial property of the production function; without it there is no particular reason to expect an equilibrium with PAM. With this assumption, Delacroix is essentially left with an assortatively matched

[3] See also Collins and McNamara (1990) for a related analysis of assortative matching.
[4] These are supermodularity of the production function, but also of its log first- and cross-derivatives (see p. 344).
[5] They note that mixed strategy equilibria can occur but they ignore them.

equilibrium and a pooled equilibrium. Kremer and Maskin (1996) discuss this class of production technologies in some detail, along with the implications for matching. Using simulations, Delacroix shows that the PAM equilibrium is more likely as the exogenous offer arrival rate increases. This is the theoretical basis for saying that dense urban labor markets lead to more sorted matching and therefore, given the nature of the production function, higher productivity.

However, as Rosenthal and Strange make clear, there is little direct econometric evidence on the importance of the different sources, and of labour market issues in particular. Moretti (2003) surveys the evidence on human capital externalities and productivity spill-overs in cities. Ellison and Glaeser (1999) and Rosenthal and Strange (2001) show that some proxies for labour market pooling explain the regional degree of spatial correlation quite well. Baumgartner (1988) shows that the division of labour is finer in big cities, which suggests a more efficient labour market If labour market pooling is interpreted as providing a more efficient search environment, with lower search frictions, these results then support the pooling hypothesis.

In other, related, work, Wheeler (2001) considers the issue of differences in urban-rural wage distributions, and in particular looks at the degree of wage inequality and the return to human capital. His model is also based on a search and matching approach, but adopts a somewhat unusual search and matching technology. Becker and Murphy (1992) study productivity and the division of labour. They formalise Smith's idea that the division of labour is constrained by the extent of the market, but add the idea that it is also constrained by coordination costs, broadly construed – principal-agent and hold-up problems, communication and coordination costs. The relationship to the productivity premium in cities is twofold – the division of labour will be finer (and hence productivity higher) because of a bigger market, and because of lower coordination costs. Benabou (1993) discusses residential segregation and productivity within cities based on complementarity between high and low skill workers. The segregation arises from spill-overs in the acquisition of education. He does not consider worker-firm matching.

Most of the empirical research has been based on surveys of workers, and matched employer-employee data are necessary to fully investigate the possibility of PAM. Very few studies using such data have had a spatial dimension. Abowd and Kramarz (2000) shows that there is an important element of positive sorting between workers and firms once data are aggregated with respect to firm characteristics such as industry or size. In addition, the study of Burgess, Lane and McKinney (2001) presents results showing that the role of assortative matching has increased over time and that this has contributed to the changes in wage inequality[6]. Andersson (2003) uses matched Swedish employer-employee data to estimate a spatial labour demand model and shows that high-wage workers sort into urban areas, and that there is an important element of positive assortative matching within urban environments. The paper closest to ours is Combes, Duranton and Gobillon (2003) who use a French panel dataset. They are interested in spatial wage disparities and can control for worker fixed-effects, industry fixed-effects, and the characteristics of the local labour market.

However, it is possible that one reason that urban labour markets have an urban wage premium is that there is more sorted matching in such markets. If there are complementarities between firm and worker quality in the production function, then this should lead to higher productivity, and hence higher wages in urban areas. No previous work has directly examined this potential contribution using linked worker-firm data. In this paper we examine the extent to matching between different types of firms and workers varies with labour market density. We follow this by examining the evidence of complementarity in production, and estimating the contribution of such factors to the urban wage premium.

3. Modelling Framework

We now set up a simple empirical model that allow us to take the Delacroix approach to the data. We set up the production side and then labor market matching. We discuss the properties of the model, and the implications for productivity in equilibrium both with

[6] Note, however, that Abowd, Kramarz and Margolis, 1999; and Abowd, Lengermann and McKinney, 2003 find that individual measures of the quality of workers and firms are more or less uncorrelated – correlations only appear when aggregating workers to firms or industries.

short-run fixed locations, and the long-run when workers and firms can relocate between urban and rural areas.

3.1 Production

We denote the quality of worker i, θ_i, the quality of firm j as ψ_j. We assume that the firm's productivity (y_j) depends on both its own quality, the quality of its workers and the interaction between the two. This formulation assumes constant returns to scale, but in the empirical work below we show that the results are robust to controlling for firm size. We parameterize this in a very straightforward way:

$$\ln y_j = a_0 + a_1\psi_j + a_2\sum_{i\in j}\theta_i + \beta\left(\psi_j . \sum_{i\in j}\theta_i\right) \qquad (1)$$

In the empirical section we estimate versions of this production function. If the parameter β is positive, a high ψ firm will be more productive with high θ workers and the production function will be supermodular. This is one of two necessary assumptions in the model to generate an urban productivity premium.

3.2 Matching

Matching equilibrium is described by the joint density function of firm and worker qualities. We write this as $f(\theta, \psi; \delta)$, giving the probability that a firm with quality $\psi = \Psi$ is found matched with a worker of quality $\theta = \Theta$ in a market with labour market density δ. In the empirical results below we characterise $f(\theta, \gamma; \delta)$ over space (δ) and calibrate the implications for the distribution of productivity. Positive assortative matching (PAM) implies a correlation between θ and γ that is increasing in density δ. The second key assumption is that dense urban labour markets generate higher offer arrival rates and hence are more likely to lead to a PAM equilibrium.

3.3 Model properties

Using (1) and the matching function, expected productivity for a firm with quality Ψ_j is given by:

$$\mathrm{E}\left(\ln y_j \mid \Psi_j\right) = a_0 + a_1\Psi_j + a_2\sum_{i\in j}\mathrm{E}(\theta \mid \Psi;\delta) + \beta\left(\Psi_j . \sum_{i\in j}\mathrm{E}(\theta \mid \Psi;\delta)\right) \qquad (2)$$

For illustrative purposes, we summarise the joint distribution of θ and ψ as a simple linear regression relationship, Eθ = *const* + *slope*.ψ, or

$$\mathrm{E}\theta = const + [\rho(\delta).v].\psi \qquad (3)$$

where ρ is the correlation coefficient of θ and ψ from *f(.),* and we explicitly include its dependence on labour market density, and $v = (\mathrm{var}(\theta)/\mathrm{var}(\psi))^{1/2}$. Substituting this into (2) and assuming for simplicity just one worker per firm yields:

$$\mathrm{E}\left(\ln y_j \mid \Psi_j\right) = a_0 + \Psi_j\left(a_1 + \beta.const + a_2.v.\rho(\delta)\right) + \Psi_j^2\left(\beta.v.\rho(\delta)\right) \qquad (4)$$

Note that if there is no complementarity in the production function ($\beta = 0$), and no PAM process ($\rho = 0$), then productivity is simply equal to a constant plus $a_1\Psi$. If β is positive, the matching equilibrium is influenced by density, and in particular an equilibrium displaying PAM is more likely in dense markets, and a pooled equilibrium more likely in sparse labour markets.

The properties of this are apparent from Figure 1. Panel A shows that for low ψ firms, there is little difference in productivity between high and low density locations, whereas for high ψ firms there is a substantial difference. Equivalently, there is much greater dispersion of productivity levels across space among high ψ firms than for low ψ firms. Panel B also shows that high ψ firms face a strong gradient of productivity in density, while low ψ firms do not. It also shows a high dispersion of productivity levels across firms in denser locations. These are testable in our data.

We detail the nature of equilibrium in the appendix, which demonstrates that under reasonable assumptions the productivity gap between urban locations and rural locations is increasing in the importance of the complementarity, as captured by β; the difference in sorting, as measured by the correlation coefficients ρ in each market; and the scope for reallocation, as measured by the variance of θ and γ.

We also show that the equilibrium impact of assortative matching on productivity exceeds the short-run impact, because agents have incentives to relocate. In steady state, depending on the importance of relocating costs, the relocation mechanism accentuates the impact of the matching effect by raising mean θ and ψ in the city and reducing them in the rural area.

Thus in equilibrium, assortative matching has two effects on the urban productivity premium – the direct effect from sorting plus complementarity in production, and the consequent relocation which accentuates this by inducing a difference in mean match quality. We can see this also referring back to panel B of Figure 1, and thinking of the mass of firms distributed vertically between the high ψ and low ψ lines. As we show in appendix, once relocation is allowed, this distribution will be concentrated near the lower line in a rural (low ρ) area and concentrated near the higher line in an urban area (high ρ). Thus the overall gradient of productivity with respect to density will be steeper than a line simply bisecting the two curves shown.

The implications of this for the econometric work below are as follows. We can recover the short run impact of assortative matching by controlling for the distribution of θ and ψ in each area. In the context of this simple model, the unconditional relationship between productivity and density would reflect the long run relationship with relocation. However, it seems unlikely that in the world generating our data, no other factor produces a difference in mean quality between areas, in which case it is harder to estimate the long run impact of assortative matching. The results above show that we should expect a greater difference in mean quality across areas with bigger differences in density, and this provides one channel for gauging the importance of the relocation story.

4. Data

As indicated in the Introduction, our data provide a unique opportunity to directly examine the spatial interaction between firms and workers. In particular, we have universe longitudinal data on workers and firms, and high quality data on their place of work and their place of residence. This enables us to construct very detailed measures of density. We also make use of new market-based measures of worker and firm wage mark-ups.

The new database that enables us to match workers with past and present employers has been assembled at the Longitudinal Employer-Household Dynamics Program at the U.S. Census Bureau (Abowd, Haltiwanger and Lane, 2004) . This database consists of quarterly records of the employment and earnings of almost all individuals from the unemployment insurance systems of a number of US states in the 1990s – these provide the key link between workers and firms. These type of data have been extensively described elsewhere (Haltiwanger, Lane and Spletzer, 2005), but it is worth noting that there are several advantages over household based, survey data. In particular, the earnings are quite accurately reported, since there are financial penalties for misreporting. The data are current, and the dataset is extremely large. The Unemployment Insurance records have also been matched to internal administrative records at the Census Bureau that contain information on date of birth, place of birth, race, and sex for all workers, thus providing limited demographic information[7]. One limitation of the data is that there are no direct data links between workers and *establishments*, but only between workers and *firms*. Thus, for about 30% of the workforce – whose employing firm consists of more than one establishment -- we cannot tell with certainty in which particular establishment a worker is employed, if the employing firm consists of more than one establishment. Thus,

[7] Given the sensitive nature of the dataset, it is worth discussing the confidentiality protection in some detail. All data that are brought in to the LEHD system have been anonymized in the sense that standard identifiers and names are stripped off and replaced by a unique "Protected Identification Key" or PIK. Only Census Bureau employees or individuals who have Special Sworn Status are permitted to work with the data, and they have not only been subject to an FBI check but also are subject to a $250,000 fine and/or five years in jail if the identity of an individual or business is disclosed. All projects have to be reviewed by the Census Bureau and other data custodians, and any tables or regression results that are released are subject to full disclosure review

probabilistic links are used to impute a place of work for workers who work for multi-unit businesses.[8]

The geographic information that exists on the dataset is extremely detailed. The physical location of each establishment is geocoded to the latitude and longitude level, as is the place of residence of each worker (from 1999 on). This information is available on a longitudinal, annual basis (geocoded businesses are available all years and residences have been geocoded in 1999, 2000 and 2001). This allows us to describe the geographical distribution of workers and employers as well as commuting and mobility patterns. In this study we use data on workers and their employers in 2001 for two large states – California and Florida - covering about 47 million workers employed in about 7 million firms.

We use two different density measures – population and employment per square mile – for two different geographical units – Census Tracts and Counties.[9] There are advantages and disadvantages associated with each of these measures. Aggregated measures of density in a county could be somewhat misleading, to the extent that counties often cover large areas containing both urban and rural parts. Also, counties do not respect the boundaries of local labour markets. Census tracts, on the other hand, are relatively small areas of between 1,500 and 8,000 individuals, and while they are not designed to be a local labour market, they are chosen to be relatively homogeneous in terms of population characteristics, economic status, and living conditions. Thus, tract-based density measures will pick up some of the within-county variation in density. However, the small size of tracts is not unproblematic either. While population and employment density measures are very highly correlated at the County level, this is not always true in tracts. In urban areas tracts cover a small area by construction, which in many cases means that

[8] These probabilistic links have been estimated in the LEHD data using multiple imputation techniques, based on a model that takes into account of the relative location of workers and establishments, the employment distribution across establishments and dynamic employment restrictions imposed by worker and job flow dynamics. We have verified that the statistical properties of the probabilistic links do not affect our results, by comparing our results with those obtained from subsamples in which direct links between workers and firms are available (i.e., in the Minnesota data and for workers employed in single-establishment firms in the two states).

either the population per square mile is high – if it is in the residential areas of the city – or the employment per square mile is high – if it is in the commercial districts of the city – but the two measures are not necessarily highly correlated. To check whether our results are sensitive to the level of geographical aggregation, we estimate our results using all four measures of density. Gautier and Teulings (2000) propose a different empirical measure of labour market density based on revealed preference on commuting patterns which they implement on data from PUMAs, which have around 100,000 people in. However, since these are far more aggregated than our areas we decided to keep with the standard jobs per unit area as our measure of employment density.

In addition to this information, the LEHD program staff have constructed measures of individual worker fixed effects, θ, and of firm fixed effects, ψ - which have been attached to the records of each worker and firm in the dataset. While these are straightforward to describe, they were computationally impossible with a dataset of this size until new econometric techniques were recently developed (see Abowd, Lengermann and McKinney, 2003)[10].

Briefly, the model that is estimated is:

$$\ln w_{ijt} = c + x_{it}\beta + \theta_i + \psi_j + \varepsilon_{ijt} \qquad (5)$$

where θ_i is the individual effect, $x_{it}\beta$ is the effect of time-varying personal characteristics – a quartic in labor force experience and work history dummies all interacted with sex , ψ_j is the firm effect, c is the intercept, and ε_{ijt} is the statistical residual. The left hand side is the log of annualized earnings. The estimates are generated from universe data consisting of 287,241,891 observations on 68,329,212 individuals,

---

[9] Population estimates by County and Census Tract are available to us in Decennial Census data. Employment estimates are based on LEHD data.

[10] There are literally millions of individual fixed effects and millions of firm fixed effects to estimate. Thought of in the conventional way to estimate fixed effect models, the X'X matrix is an extremely large and sparse matrix, which is simply infeasible to compute.

working at a total of 3,662,974 firms during the period 1992 to 1999. The model, which separately identifies person and firm effects, explains some 85% of earnings variation.

The individual fixed effect is a wage markup that summarizes the individual wage premium (or discount) that an individual carries with him/her as s/he moves from firm to firm. This human capital measure can be thought of as the market value of the portable component of an individual's skill set - and includes some factors that are often observable to the statistician, such as years of education and gender; and some factors that are often not, such as innate ability, "people skills," "problem solving skills," perseverance, family background, and educational quality.

Symmetrically, the firm fixed effect is a summary measure of the wage premium (or discount) that each firm pays to observationally equivalent workers. This firm wage markup can reflect a variety of different factors such as the organisational structure, the degree of rent-sharing, the capital intensity, or the degree of unionisation at a firm (see Andersson, Holzer and Lane, 2004, for a non-technical description). Although it is impossible to separate out how much of the firm wage markup is due to each of these factors, it does capture the key elements of the firm's production and personnel decisions.

These new measures enable the effect of worker and firm characteristics on earnings outcomes to be separated for the first time, and give insights into the sources of wage differences. For example, as shown in Abowd et al (2002), the average wage premium for firms in the security, commodity brokers and services industry is about 79%. However, about half of this premium is due to workers possessing above average skill sets, and about half is due to firms paying a premium to observationally equivalent workers. In the engineering industry, the markup is about 29%, but one third is due to high quality workers, and two thirds to a firm wage premium. By contrast, the food store industry is a low-wage industry in which workers are paid about 25% below average. Yet workers in this industry, on average, have the same skill set as the average worker in the rest of the economy: almost all of the wage discount is due to firms in that industry paying observationally equivalent workers less. In the hotel and lodging services industry workers are paid 30% less, and the reason for the discount is about even between lower skilled workers and low mark-up firms.

Table 1 shows the correlation between different wage components in California, Florida, Illinois, Maryland, Minnesota, North Carolina, and Texas over the period 1985 to 2000.[11] The first thing to note is the explanatory power of this decomposition. The correlation between the residual and the wage measure is 0.402, which translates into an R^2 of about 85%. The second thing to note is the importance of firm effects. The simple pairwise correlation of the estimated firm effect and earnings is 0.484. This number is substantially higher than the correlation between the effects of observable personal characteristics and earnings and comparable to the correlation between the effects of unobservable person characteristics and earnings. Finally, note that firm and worker effects are virtually uncorrelated, which is true for the two individual states we study here. We show below that there is an important element of positive assortative matching once the spatial dimension of data is incorporated.

Our final key measure is the productivity of the establishment. The data from the Economic Census in 1997 provide measures of sales at the establishment level, which, together with employment, is used to create a proxy for productivity – sales (or revenue) per worker. This is similar to the measure used by Haltiwanger, Lane and Spletzer (1999; 2001). Although clearly the preferred productivity measure would be value-added per hour, Haltiwanger, Lane and Spletzer point out that there is a close correspondence both conceptually and in terms of measurement between this measure of gross output at the establishment level and the industry-level measures published by BLS[12] . The standard BLS measure of labour productivity at the detailed industry level is output per hour.

---

[11] The information in this table is extracted from Table 6 in Abowd, Lengermann and McKinney (2003).

[12] As they point out "It is worth noting that for most sectors there are not highly reliable measures of value added per hour even at the industry level that differ from the measures of gross output per hour. The reason is that materials usage data is poor in most sectors other than manufacturing. As Triplett and Bosworth (2001) note, for most service sector industries the correlation between gross output per hour measures from

5. Results

We first present results on establishment level productivity analysis. We repeat this analysis within industry. We then model the matching outcome, and finally turn to calibrate the impact of density on productivity outcomes.

5.1 Productivity

We confirm the urban productivity premium by calculating simple correlations of density with productivity and productivity dispersion, at both the tract and the county levels. As is clear from an examination of the results in Table 2, geographic areas that have higher employment density are indeed more productive – and Table 3 confirms that workers in dense geographic areas are also higher paid. Productivity dispersion should also be positively correlated with density, and the data confirm this.

In order for assortative matching to contribute to higher productivity, there must be complementarities in production between high quality workers and high wage markup firms. In order to investigate this, we examine the relationship between the worker premium θ and the firm wage premium, ψ, and productivity (at the establishment level), while controlling for firm size and industry[13]. The results of this are reported in table 4. Clearly, the statistical strength and magnitude of the coefficients on worker and firm quality show the importance of these inputs. However, the critical part for our purposes is the degree of complementarity between those inputs, which is captured by the interaction term. In both California and Florida this is statistically significant and positive. This provides the incentive for firms and workers to match assortatively, and the mechanism that yields higher productivity if they are successful in doing so. We return to its quantitative significance in section (c) below.

We also check alternative less restrictive specifications. Since productivity at the establishment level might depend less on average worker quality than on the distribution of workforce quality, we calculate three different measures – the quality of workers in the

---

BLS and value-added per hour measures from BEA is extremely high for many service sector industries because the measurement of materials usage is poor" – a finding reinforced by Foster et al, 2001.

25th, 50th, and 75th percentile of the human capital distribution at the establishment – and include these as independent variables in the regression. The results in Table 5 also support the finding of complementarity in production, and hence that the estimates of the productivity relationship using mean quality do not do any great violence to the data. In addition, we check whether differences in industrial structure are driving these findings of a significant complementarity between worker and firm quality, and report the results in Table 6 of running the analysis separately by industry. The California results indicate that the interaction is significantly positive in 5 industries, and significantly negative in none. In Florida, there is a positive effect in all but wholesale trade.

These results support the existence of widespread and strong positive quality interactions (complementarity) in productivity and set the scene for the possibility of finding PAM equilibria in high arrival rate labour markets.

5.2 Matching

The second empirical question is whether high markup firms and high quality workers are more likely to be found together in urban than in rural areas. In this section we characterise the joint density function of worker and firm wage markups, $f(\theta, \psi)$, and the way in which this varies with the density of the labour market. Clearly, although we can only define a matching correlation over an aggregate of individuals and firms, there is no unique way to define a market over which to compute this correlation. We therefore adopt two different straightforward scales – a census tract, which is smaller than a labour market, and a County. Whilst theory suggests a clear cut-off between pooled matching and assortative matching, the fact that our empirical areas will not perfectly correspond to true labour markets means that we should expect to see smoother changes. We characterise the joint density using a number of techniques: first graphically and with maps; second non-parametrically, estimating the joint density with kernel estimators in different labour markets; and third using regression.

---

[13] We did not experiment in terms of different functional forms in these regressions.

5.2.1 Graphs and maps

Simple cross-plots of employment density and the matching correlation are presented in Figures 2 and 3 for the two states. These show a clear positive relationship between the two, regardless of whether the detail is at the tract or the county level[14]. The underlying spatial structure underlying these maps is illustrated by the maps for California which are shown in Figures 4 and 5. These show considerable variation in density and correlation, particularly the tract level maps in Figure 5.

5.2.2 Non-parametric characterisation

While the simple correlation coefficient and the regressions reported below provide useful summaries of the bivariate density function, $f(\theta, \psi)$, it is helpful to get some impression of the overall distribution. In order to illustrate the difference in the distributions between areas, we compute the density over (θ, ψ) space separately in urban and rural tracts[15], and subtract the latter from the former.

In Figure 6 we plot the contours of this difference in distribution for tracts in California, focussing on manufacturing and retail industries. To reduce other sources of heterogeneity, we focus on a particular group of workers – males, aged 35 – 55 years old. Panel A (manufacturing) of the figure shows a clear north-east – south-west axis, with higher density in urban tracts in the top right quadrant. Whilst part of this reflects both higher firm and worker fixed effects in urban areas , there is also an impression that more of the mass of the density is close to the 45^{o} line in the urban area, and that the distribution is more diffuse in the rural area. Panel B looks at the retail sector. Again there is a clear north-east – south-west orientation, but the pattern is otherwise less clear because of the compression of firm effects relative to worker effects.

5.2.3 Parametric characterisation

Table 7 provides county and tract level regressions of the matching correlation on density. The dependent variable is the matching correlation over the areal units shown –

[14] These graphs use population density, but the results are the same using employment density.

county and tract. We control for the mean level of worker and firm markup. The regressions show a positive and statistically significant effect of (employment) density on the matching correlation. This is true for both spatial scales in both California and Florida. Of the other variables, worker quality seems to matter in California, but not in Florida.

We then exploit all the data we have and run individual level regressions for both of our states. These are not to be interpreted causally, but rather as summaries of the joint density of worker and firm quality, and its dependence on density. An individual worker's fixed effect is regressed on her matched firm fixed effect, local density, and an interaction of quality and density. As before, this latter interaction term is the focus of interest as it shows how the matching equilibrium varies over labour markets of different densities. The results in Table 8 show very clearly that the interaction term is positive and significant, regardless of the spatial scale used, and for both states. Again we interpret the size of the effect through the impact on productivity.

In summary, the results of this section have established that the degree of assortative matching does increase significantly with the thickness of the labour market.

5.3 Calibration of productivity effect

The results in the previous sections have established the necessary preconditions for assortative matching to contribute to productivity. In this section we calibrate the effects across areas with different densities by examining differences across firms in two different employment density levels – low density and high density. The first panel of Table 9 demonstrates that the mean productivity of firms is about .19 log points higher in high density areas than low density areas; firm fixed effects are about .15 log points higher; and worker fixed effects about .14 log points higher.

We use the estimates above in equation (2) (repeated here) to examine the impact of location on a firm's expected productivity. For example, for a firm with markup Ψ_j in

---

[15] This uses a bivariate kernel estimation technique. See Press et al (1988).

California, we compute its expected productivity at two locations with different employment densities.

$$\mathrm{E}\left(\ln y_j \mid \Psi_j\right) = a_0 + a_1\Psi_j + a_2\sum_{i\in j}\mathrm{E}f\left(\theta \mid \Psi;\delta\right) + \beta\left(\Psi_j.\sum_{i\in j}\mathrm{E}f\left(\theta \mid \Psi;\delta\right)\right) \qquad (2)$$

In the second panel of Table 9, we see that the fitted differences for the average firm in a high density area for all firm values is around 0.1 log points. This can be compared to the overall urban premium of 0.19 log points shown in the first panel – and makes it clear that this approach predicts a substantial part of the urban productivity premium. Interestingly, the importance varies depending on which part of the firm fixed effect distribution is fitted to equation (2) – ranging from .107 for firms one standard deviation above the mean to .083 for firms one standard deviation below mean firm fixed effect.

However, this exercise, while controlling for spatial differences in firm fixed effects, does not control for difference in mean worker fixed effects between urban and rural areas. Of course, in the model set out above, the sorting between rural and urban areas occurs in response to the matching. To isolate the pure direct role of matching, separately from the sorting effect, we need to control for both the worker and firm fixed effects in each area. We construct the following counter-factuals of matched worker-firm pairs and compare the resulting productivity with actual productivity. First, we count the number of job slots in each area – Census Tract or County - and keep this constant to maintain area differences in density. We then generate a random allocation of firms and workers, with a firm and worker fixed effects randomly assigned to each job slot. This simulates the short-run before relocation of firms and workers takes place. The next step is to randomly match a worker (a quality value) to each firm. This yields a random *sorting* and *matching* situation with no expected differences in productivity across areas. Since we now want to isolate the short-run or pure effect of non-random matching, we keep the resources (worker and firm qualities) in each area constant and simply reshuffle them according to three different matching regimes: perfect positive assortative matching, perfect negative assortative matching, and actual matching, where the latter is based on the actual correlations between worker and firm effects in each area. Finally, having created new worker-firm pairs, we use the estimated productivity function to calculate productivity for each pair. In each case, when we probabilistically match workers and firms, we do this 100 times and present the mean outcome.

The results are presented in Table 10, holding the resources constant across tracts, and in Table 11, holding the resources constant across Counties. The results show that the effects of matching patterns on productivity are sizeable. The first block of each table refers to outcomes with random sorting of workers and firms across rural and urban areas. Taking the top three rows, it is clear that with complementarity in the production function positive assortative matching produces considerably higher productivity than the other matching regimes. Holding worker and firm fixed effects constant, the difference between positive and negative assortative matching is 0.056 in urban tracts and 0.062 in urban Counties. Thus patterns of worker-firm matching matter. Because of the random sorting in this part of the table, rural and urban areas have (in expectation) the same resources, and so this difference between matching regimes is the same for both areas. In row 4 we allow for the difference in actual matching patterns across rural/urban areas, reflecting the greater degree of positive assortative matching in cities. Productivity in rural areas is about the same as with random matching (0.002 at tract scale and –0.001 for Counties). In urban areas, it is about half way between random and perfect positive assortative matching, 0.016 at tract scale and 0.015 at County scale. Thus, the direct short-run productivity effect of differences in matching between urban and rural areas is 0.013 at tract scale and 0.017 at County scale. It is a feature of all the results in these two tables that the effects are stronger at County scale. This makes sense – Counties hold considerably more firms and workers and so the scope for productivity gains from relocation is greater.

The second block of tables 10 and 11 relate to the long run influence of matching, and are based on the actual distribution of worker and firm fixed effects between urban and rural areas. First, we see that the sorting of higher fixed effect workers and firms into urban areas is important for productivity differences and dominates the direct short-run effect of matching. Second, at the County scale (table 11), the actual sorting of firms matters considerably more than the actual sorting of workers, and explains more of the urban productivity difference. Third, the role of matching can be seen by comparing row 7 (random matching) and row 8 (actual matching). At the County scale, productivity increases 0.047 log points (from 0.147 to 0.194) in urban areas once we apply the actual matching patterns in the data, and 0.031 (from –0.142 to –0.109) in rural areas; the differences are smaller at tract scale. The difference in these differences, the direct contribution of matching to the urban productivity premium given the actual sorting of workers and firms, is 0.016. This is essentially the same as the 0.017 figure based on random sorting. At tract scale, the difference is in fact marginally greater in rural tracts giving a small negative contribution of –0.005 to the premium – compare 0.243 (actual matching) and 0.248 (random matching). We attribute this anomaly to the much smaller scale of tracts making them susceptible to outliers in the re-matching process. Overall, these results show that pure differences in matching patterns are quantitatively important for productivity. They also show that the long-run allocation of high fixed effect workers and firms to urban areas is more important, and we finally turn to explore the implications of our model for that.

Equation A12 in the Appendix shows that the difference in the long-run and short-run contributions of matching to the urban-rural productivity differential is $(2+\beta)(\bar{\theta}_c - \bar{\theta}_r)$. We can use our estimates to quantify this for California: β is estimated at 0.028 (Table 4), $\bar{\theta}_c$ is 0.10, and $\bar{\theta}_r$ is –0.04 (Table 9). This gives a value of 0.284, and combining this with the direct effect of matching of 0.017 (County scale) yields a long-run productivity differential of 0.301. Given the simple nature of the assumptions in the model, the fact this is so close to the actual value of 0.305 across Counties is no doubt coincidence. But the fact that it is of the right order of magnitude is interesting and does suggest that matching and the consequent relocation is an important component of productivity differentials.

6. Conclusions

In this paper we address the puzzle of the urban productivity premium. While it is clear that this is substantial, the literature is unclear what it derives from. We take one of the main contenders and test it using a new micro dataset. Our results suggest that assortative matching in thick urban labour markets plus complementarities in production play an important role in generating high productivity in cities. Using non-parametric techniques and simple regression analysis we show that the degree of matching of firm and worker quality does vary with labour market density, and we establish that there is evidence of complementarity in production. Putting these together, we show that this contributes to the urban premium.

The paper also illustrates the insights of the search and matching approach to labour markets, and the power that the new emerging datasets offer in addressing long-standing questions. There are other related issues that we can tackle: for example, segregation and networks in cities, earnings and local labour markets, residential and commuting patterns. Complementarity in production plus assortative matching also imply greater wage inequality in denser labour markets. We leave all these to future work.

References

Abowd, John, Francis Kramarz, and David Margolis (1999) "High-Wage Workers and High-Wage Firms", *Econometrica*, Vol. 67, No. 2, (March 1999): 251-333.

Abowd, John, and Francis Kramarz, (2000) "Inter-Industry and Firm-Size Wage Differentials: New Evidence from Linked Employer-Employee Data", LEHD Technical Working Paper

Abowd, John M., Francis Kramarz, Paul Lengermann and Sébastien Roux (2002), "Inter-industry and Firm-size Wage Differentials: New Evidence from Linked Employer-Employee Data", working paper.

Abowd, John, Lengermann, Paul, and McKinney, Kevin (2003) "The Measurement of Human Capital in the U.S. Economy" March, LEHD Technical Working Paper

Acemoglu, D. (1997) Matching, Heterogeneity and the Evolution of Income Distribution. March, *Journal of Economic Growth*, volume 2, pp. 61-92.

Andersson, Fredrik (2003) "The Spatial Wage Distribution, Sorting of Workers and Urban Agglomeration" in *Causes and Labor Market Consequences of Producer Heterogeneity*. *Economic Studies* 73, Uppsala.

Andersson, Fredrik, Harry Holzer, and Julia Lane. (2002) "The Interaction of Workers and Firms in the Low-Wage Labor Market." LEHD working paper (http://lehd.dsd.census.gov).

Andersson, Fredrik, Harry Holzer and Julia Lane, (2003) "*Moving Up Or Moving On Workers, Firms and Advancement in the Low-Wage Labor Market*" Russell Sage Foundation draft manuscript

Becker, G. S. and Murphy, K. M. (1992) 'The Division of Labor, Coordination costs and Knowledge' *Quarterly Journal of Economics* vol. 107, pp. 1137 – 1160.

Becker, G. S. (1973) 'A Theory of marriage: Part I' *Journal of Political Economy* vol. 81 pp. 813 – 846.

Benabou, R. (1993) Workings of a City: Location, Education and Production. *Quarterly Journal of Economics*, vol. 108, pp. 619 – 652.

Burdett, K. and Coles, M. G. (1999) 'Long-term Partnership Formation: Marriage and Employment' *Economic Journal* vol. 109, pp. F307 – F334.

Burdett, K., and M. Coles (1997) 'Marriage and Class' *Quarterly Journal of Economics* vol. 112, pp. 141 – 168

Collins, R. J. and McNamara, J. M. (1990) 'The job-search problem with competition: an evolutionary stable dynamic strategy' *Advances in Applied Probability* vol. 25, pp. 314 – 333.

Delacroix, A. (2003) 'Heterogeneous Matching with Transferable Utility: Two Labor Market Applications' *International Economic Review.* Vol. 44 pp. 313 - 342

Diamond, D.B. and G.S. Tolley. (1982) "The Economic Role of Urban Amenities" in *The Economics of Urban Amenities,* edited by Diamond and Tolley, New York: Academic Press, 3-54.

Duranton, G. and Puga, D. (2003) 'Micro-foundations of Urban Agglomeration Economies' forthcoming in Henderson, J.V. and Thisse, J-F. (eds) *The Handbook of Regional and Urban Economics*. North Holland, Amsterdam.

Foster, Lucia, John Haltiwanger, and C.J. Krizan (2001). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." In (Dean, Hulten, and Harper, eds.) *New Developments in Productivity Analysis*, NBER/University of Chicago Press.

Gautier, P. A. and Teulings, C. N. (2000) 'An Empirical Measure for Labor market Density' Tinbergen Institute Discussion Paper TI 2000-036/3.

Haltiwanger, J. C, J.I. Lane and J. R. Spletzer "Productivity Differences Across Employers: The Role of Employer Size, Age, and Human Capital" *American Economic Review*, May 1999. Pg. 94-98

Haltiwanger, J. C, J.I. Lane and J. R. Spletzer "Wage, Productivity and the Dynamic Interaction of Businesses and Workers" *Labour Economics*, forthcoming.

Kremer, M., and E. Maskin (1996) 'Wage Inequality and Segregation by Skill' NBER Working paper 5718.

Marshall, A. (1890) *Principles of Economics*. MacMillan, London.

Mood, A. M., Graybill, F. A., and Boes, D. C. (1974) *Introduction to the Theory of Statistics*. 3rd Edition. McGraw-Hill, London.

Moretti, E. (2003) Human Capital Externalities in Cities forthcoming in Henderson, J.V. and Thisse, J-F. (eds) *The Handbook of Regional and Urban Economics*. North Holland, Amsterdam.

Petrongolo, B. and Pissardies, C.(2001) Looking into the black-box: A survey of the matching function *Journal of Economic Literature* 39, 390-431, 2001.

Petrongolo, B. and Pissardies, C. (2003) ‘Scale Effects in Markets with Search’ CEP mimeo, LSE.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1988), *Numerical Recipes: The Art of Scientific Computing,* Cambridge: Cambridge University Press.

Rosenthal, S. and Strange, W. (2003) ‘Evidence on the Nature and Sources of’Agglomeration Economies’ forthcoming in Henderson, J.V. and Thisse, J-F. (eds) *The Handbook of Regional and Urban Economics*. North Holland, Amsterdam.

Shimer, R. and L. Smith (2000) ‘Assortative Matching and Search’ *Econometrica* vol. 68 pp. 371 – 398.

van den Berg, G. and van Vuuren, A. (2003) The Effects of Search Frictions on Wages. mimeo Free University, Amsterdam.

Wheeler, C. (2001) ‘Search, Sorting and Urban Agglomeration’ *Journal of Labor Economics* vol. 19 no. 4 pp. 879 – 899
.

Figure 1: Productivity, Density and Firm Quality

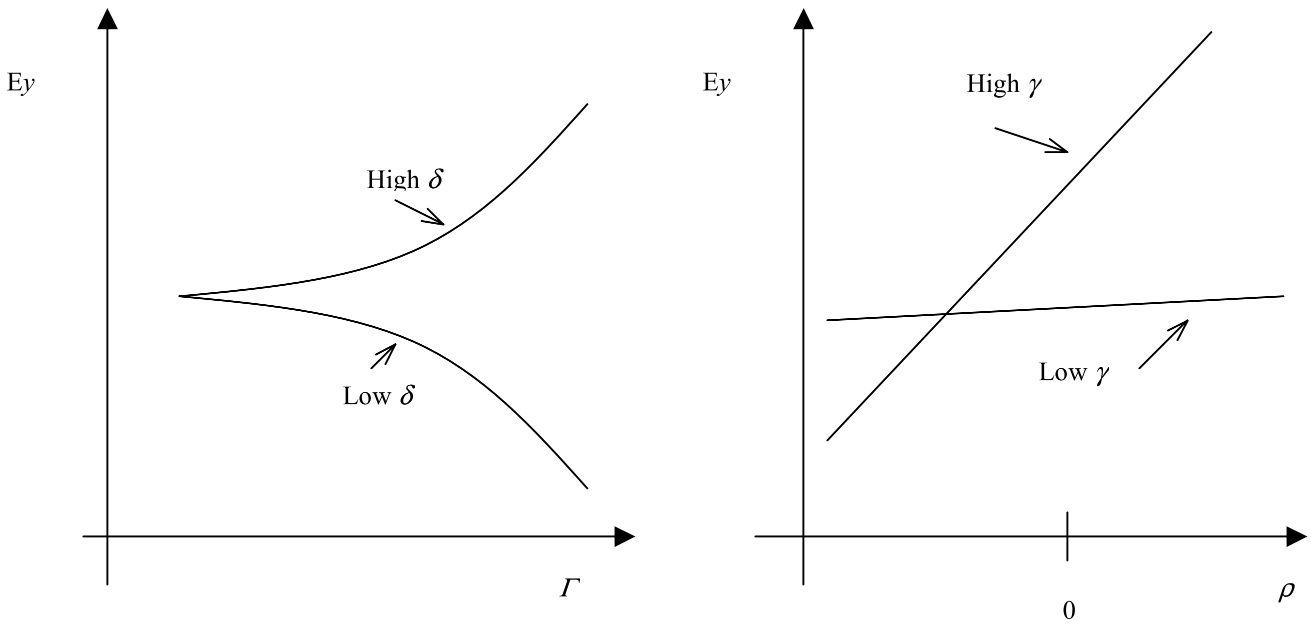


Figure 2: Matching and Density in California

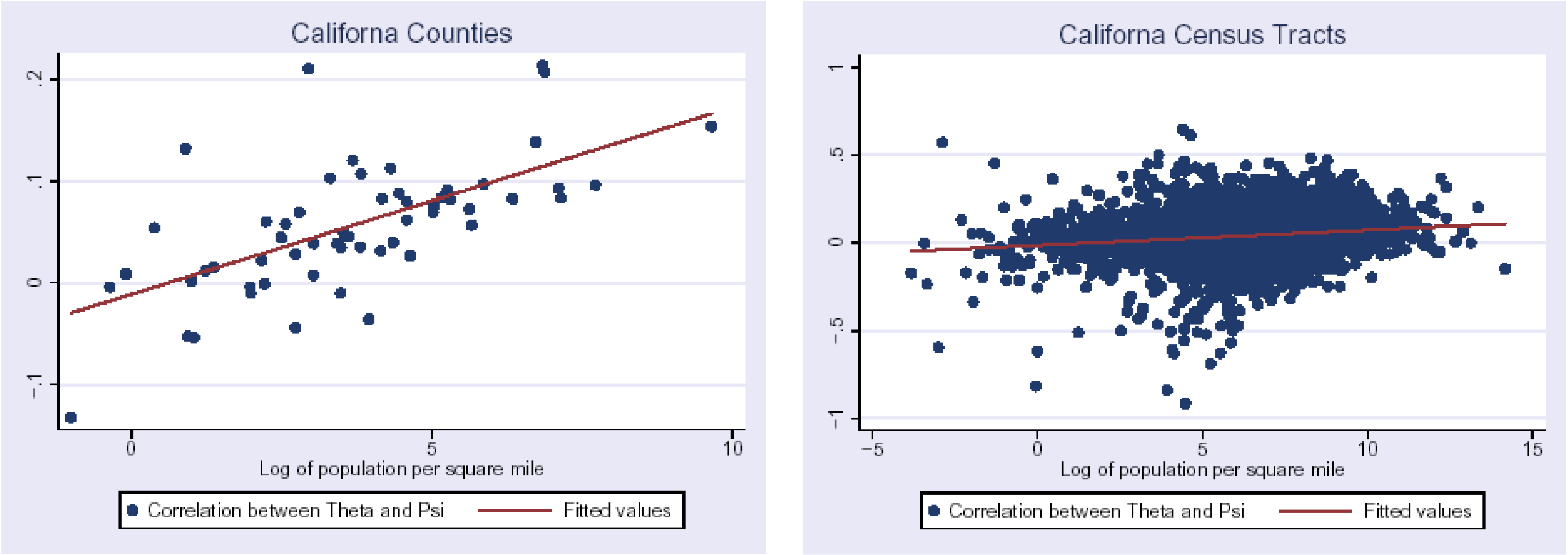


Figure 3: Matching and Density in Florida

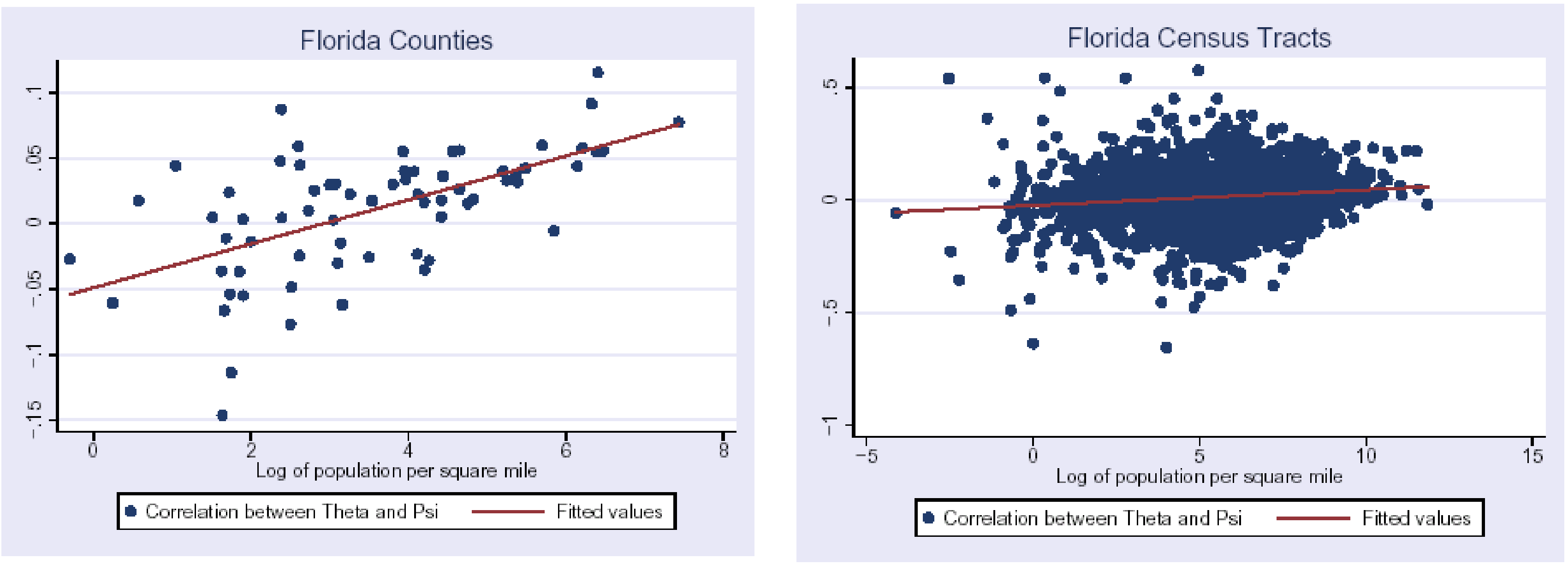
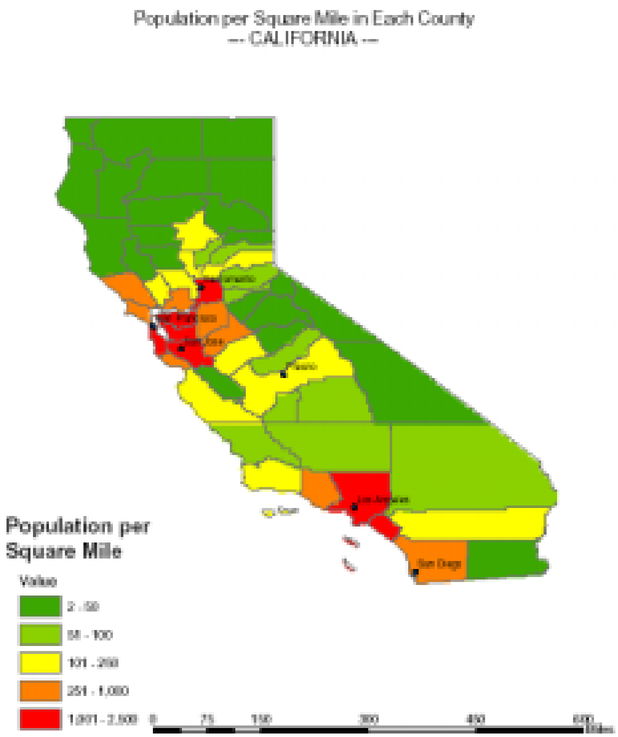


Figure 4: Matching and Density in California, County Level Maps



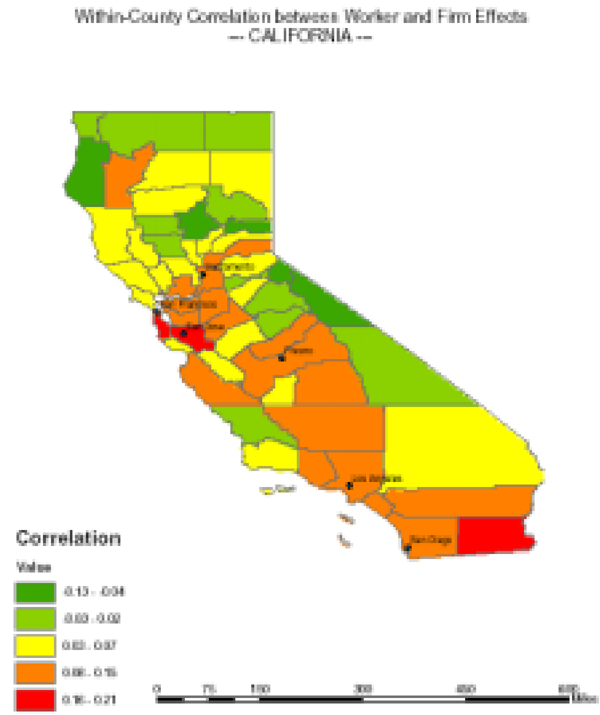


Figure 5: Matching and Density in California, Tract Level Maps

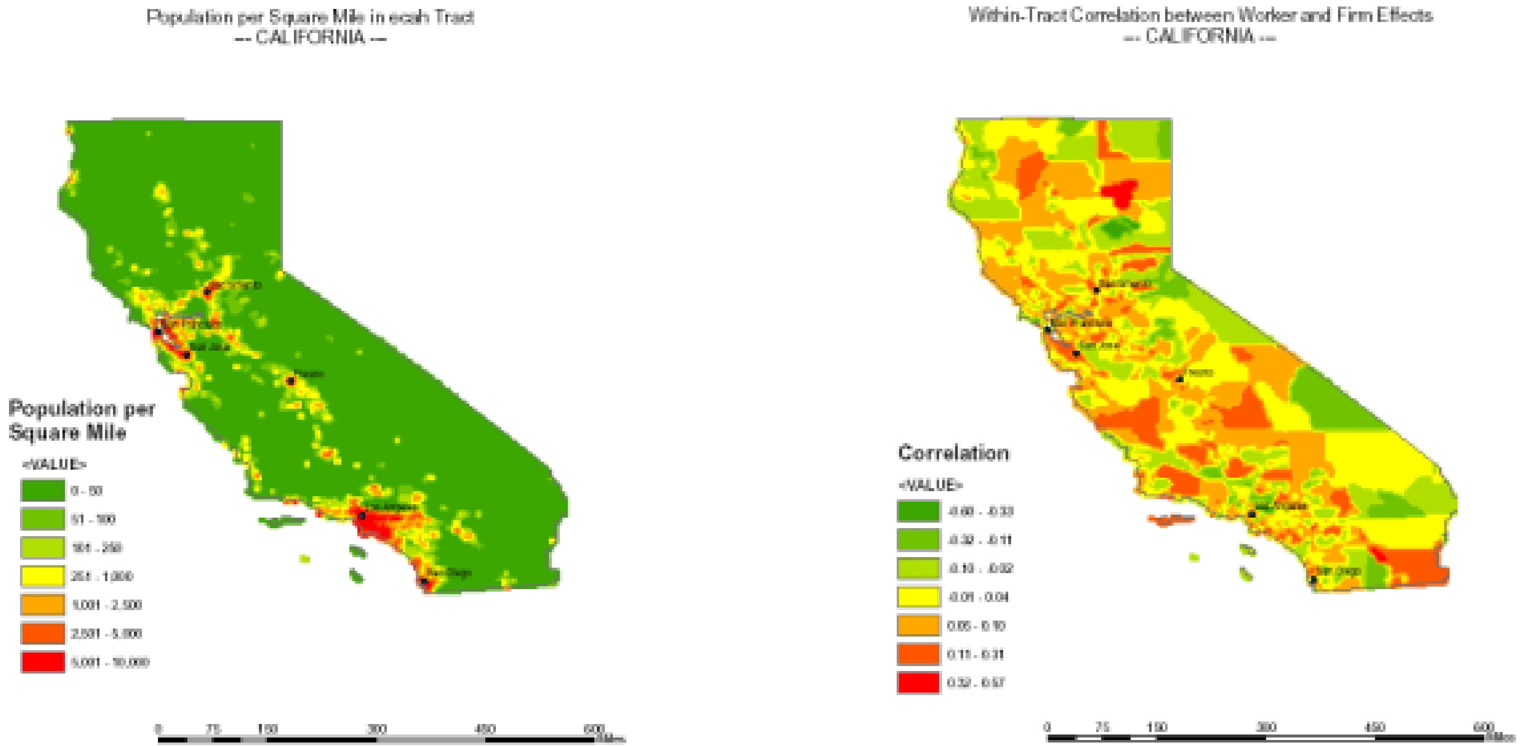


Figure 6: Kernel estimates of *(θ, ψ)* density in California, Men aged 35-55: urban vs rural Census Tracts

Panel A: Manufacturing

Panel B: Retail Trade

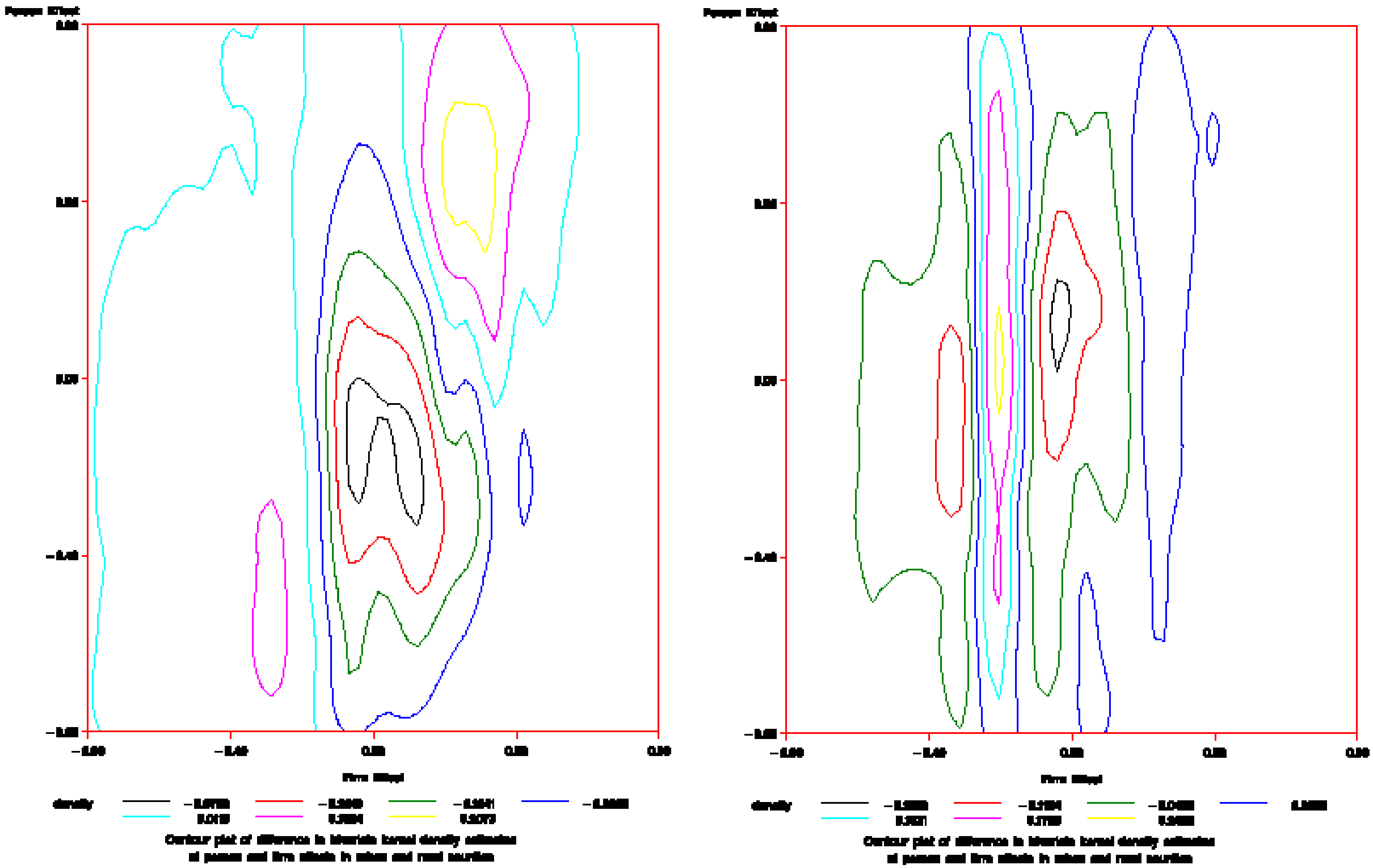


Table 1: Summary of Estimated Wage Components

Component	Standard Deviation	Correlation with y	$x\beta$	θ	ψ	ε
Log real annual wage rate (y)	0.881	1.000	0.224	0.468	0.484	0.402
Time-varying personal characteristics ($x\beta$)	0.691	0.224	1.000	-0.553	0.095	0.000
Person effect (θ)	0.835	0.468	-0.553	1.000	0.080	0.000
Firm effect (ψ)	0.362	0.484	0.095	0.080	1.000	0.000
Residual (ε)	0.354	0.402	0.000	0.000	0.000	1.000

Note: Based on 287,241,891 annual observations from 1986 to 2000 for 68,329,212 persons and 3,662,974 firms in California, Florida, Illinois, Maryland, Minnesota, North Carolina, and Texas. Source: Table 6 in Abowd, Lengermann and McKinney (2003) and based on data from the LEHD Program Employment Dynamics Estimates Database.

Table 2: Correlations between productivity, productivity dispersion and density

	California		Florida	
	Tract	**County**	**Tract**	**County**
Corr(P,E)	0.385**	0.489**	0.521**	0.726**
Corr(PD,E)	0.255**	0.709**	0.207**	0.275*
N	7049	58	3154	67

Note: Corr(P,E) is the correlation between the mean of log of labor productivity and log of employment per square mile across the geographical units (Tract or County) within each state. PD is the standard deviation of log of labor productivity across firms within the geographical unit. ** significant at 1%, * significant at 5%.

Table 3: Correlations between wages, wage dispersion and density

	California		Florida	
	Tract	**County**	**Tract**	**County**
Corr(W,E)	0.366**	0.525**	0.431**	0.748**
Corr(WD,E)	0.181**	0.559**	0.328**	0.659**
Corr(W9010,E)	0.095**	0.247	0.172**	0.504**
N	7049	58	3154	67

Note: Corr(W,E) is the correlation between the mean of log of annualized earnings and log of employment per square mile across the geographical units (Tract or County) within each state. PD is the standard deviation of log of annualized earnings across all workers within the geographical unit. ** significant at 1%, * significant at 5%.

Table 4: Firm-level productivity regressions as a function of mean human capital

	California	**Florida**
(1): Mean person effect	0.430	0.346
	(163.46)**	(82.46)**
(2): Firm effect	0.638	0.568
	(246.76)**	(129.79)**
Interaction term between (1) and (2)	0.028	0.059
	(14.92)**	(15.92)**
Constant	4.828	4.739
	(686.22)**	(468.25)**
Observations	400770	152367
R-squared	0.29	0.24

Note: The dependent variable is the log of labor productivity. In addition the specification includes controls for size and industry of firm (not reported in table). ** significant at 1%, * significant at 5%.

Table 5: Firm-level productivity regressions as a function of human capital distribution.

	California	**Florida**
(1): 25:th percentile of person effect	0.060	0.102
	(32.67)**	(6.04)**
(2): 50:th percentile of person effect	0.142	0.063
	(35.07)**	(16.14)**
(3) 75:th percentile of person effect	0.164	0.111
	(8.05)**	(16.53)**
(4): firm effect	0.631	0.571
	(188.57)**	(95.02)**
Interaction term between (1) and (4)	-0.016	0.017
	(2.79)**	(1.90)
Interaction term between (2) and (4)	0.035	0.026
	(3.29)**	(1.63)
Interaction term between (3) and (4)	0.024	0.024
	(3.88)**	(2.50)*
Constant	4.832	4.739
	(670.70)**	(457.08)**
Observations	396020	150756
R-squared	0.29	0.24

Note: The dependent variable is the log of labor productivity. In addition the specification includes controls for size and industry of firm (not reported in table). ** significant at 1%, * significant at 5%.

Table 6: Coefficients on the interaction term between mean person effect and firm effect from firm-level productivity regressions by industry

	California	Florida
Construction	-0.003	0.065**
Manufacturing	0.012	0.022**
Transportation & Utilities	0.065**	0.093**
Wholesale Trade	0.007	0.014
Retail Trade	0.148**	0.108**
FIRE	0.033**	0.046**
Business Services	-0.004	0.031**
Health Services	0.010*	0.034**
Educational Services	-0.033	0.175**
Other Services	0.013**	0.056**

Note: The dependent variable is the log of labor productivity. In addition the specification by industry includes controls for mean person effect, firm effect, size and a constant (not reported in table). ** significant at 1%, * significant at 5%.

Table 7: Regressions of matching correlation, corr(θ,ψ) on density and average human capital estimates

	California		Florida	
	Tract	**County**	**Tract**	**County**
Mean of person effect	0.122 (9.61)**	-0.322 (2.66)*	0.002 (0.10)	-0.099 (0.79)
Mean of firm effect	0.053 (7.27)**	0.168 (1.79)	0.009 (0.71)	-0.012 (0.19)
Log of emp./sq. mile	0.005 (8.63)**	0.020 (4.55)**	0.007 (7.64)**	0.019 (5.35)**
Constant	0.015 (3.11)**	0.008 (0.27)	-0.024 (4.09)**	-0.054 (3.51)**
Observations	7013	58	3134	67
R-squared	0.06	0.47	0.02	0.38

Absolute value of t statistics in parentheses; * significant at 5%; ** significant at 1%

Table 8: Matching regressions. Dependent variable: estimated person effect (θ)

	California		Florida	
Geographical unit of density measure	**Census Tract**	**County**	**Census Tract**	**County**
(1): Firm effect	0.171 (53.60)**	0.180 (55.63)**	-0.044 (8.87)**	-0.039 (5.78)**
(2): Log of employment/sq. mile	0.019 (129.06)**	0.021 (112.48)**	0.016 (73.83)**	0.031 (77.12)**
Interaction term between (1) and (2)	0.006 (15.41)**	0.007 (13.37)**	0.023 (35.90)**	0.030 (25.66)**
Constant	-0.058 (45.00)**	-0.030 (24.51)**	0.003 (1.99)*	-0.049 (20.88)**
Observations	9,000,959	9,000,959	5376886	5376886
R-squared	0.01	0.01	0.00	0.00

Table 9: Calibrating Productivity Differences – California

	Employment Density		Difference
	Low Density	High Density	
Number of tracts	1387	106	
Mean density	3.74	11.34	7.60
Mean actual productivity	4.35	4.55	0.19
Mean Ψ	-0.04	0.11	0.15
Mean θ	-0.04	0.10	0.14
Fitted Marginal Productivity			
At Ψ = 0 (mean)	0.046	0.141	0.095
At Ψ = 0.4 (+ 1 SE)	0.128	0.235	0.107
At Ψ = -0.4 (- 1 SE)	-0.035	0.048	0.083

Notes: Employment density is log employment per sq. mile. We define a low density tract as one where this measure falls below 5.5; high density as one where it falls above 10.5

Table 10: Decomposition of productivity effects across California Tracts

Mean Log Productivity in:	Rural Tracts	Urban Tracts	Urban – Rural
Conditional on			
- random sorting and random matching	0.000	0.000	0.000
- random sorting and perfect positive assortative matching	0.027	0.029	0.002
- random sorting and perfect negative assortative matching	-0.028	-0.027	0.001
- random sorting and actual matching	0.002	0.016	0.013
- actual sorting of workers, random sorting of firms and random matching	-0.052	0.067	0.119
- actual sorting of firms, random sorting of workers and random matching	-0.084	0.035	0.119
- actual sorting of workers and firms, random matching	-0.154	0.094	0.248
- actual sorting of firms and workers and actual matching	-0.135	0.108	0.243

Note: The estimates are based on 100 boot-strapped samples. "Rural Tracts" are defined as Census Tracts with log of employment per square mile in the bottom 5th percentile. "Urban Tracts" are defined as Census Tracts with log of employment per square mile in the top 95th percentile.

Table 11: Decomposition of productivity effects across California Counties

Mean of Log Productivity in:	Rural Counties	Urban Counties	Urban – Rural
Conditional on			
- random sorting and random matching	0.000	0.000	0.000
- random sorting and perfect positive assortative matching	0.031	0.031	0.000
- random sorting and perfect negative assortative matching	-0.030	-0.031	0.000
- random sorting and actual matching	-0.001	0.015	0.017
- actual sorting of workers, random sorting of firms and random matching	-0.022	0.065	0.088
- actual sorting of firms, random sorting of workers and random matching	-0.088	0.124	0.213
- actual sorting of workers and firms, random matching	-0.142	0.147	0.293
- actual sorting of firms and workers and actual matching	-0.109	0.194	0.305

Note: The estimates are based on 100 boot-strapped samples. "Rural Counties" are defined as Counties with log of employment per square mile in the bottom 5th percentile. "Urban Counties" are defined as Counties with log of employment per square mile in the top 95th percentile.

Appendix

In this section we describe the equilibrium of the theoretical model in some detail. We consider here the equilibrium once re-location between markets is allowed[16].

Take two markets, a city (c) with high density and so high correlation between worker and firm quality (ρ_c), and a rural area (r) with low density (ρ_r). In each area there are *n* workers and *n* jobs, and θ and ψ in both areas are uniformly distributed between 0 and 1 [17], so $\bar{\theta}_j = \bar{\psi}_j = 1/2, j = c, r$. If moving between markets is impossible, workers and jobs would be matched according to ρ_c and ρ_r respectively. To keep notation simple, we assume that $\rho_r = 0$, and hence that $\Delta\rho \equiv \rho_c - \rho_r = \rho_c$. We continue to write $\Delta\rho$ to show that it is the correlation and density *differential* that matters.

Suppose that relocating costs *c* in either direction and for both workers and jobs. Clearly, the high value workers in the rural areas and the low value workers in the city will consider moving to the other market. The decision for jobs is identical. Specifically, a high θ rural worker will move if:

$$\theta + E_c(\psi \mid \theta) > \theta + E_r(\psi \mid \theta) - c \tag{A2}$$

where E_c (E_r) denotes the expectation under the city (rural) distribution. Substituting in from (A1), this yields a threshold value such that rural workers will move if:

$$\theta > \bar{\theta} + \frac{c.v}{\Delta\rho} \equiv \hat{\theta} \tag{A3}$$

Note that $\hat{\theta} \leq 1$, which implies the restriction that $c.v/\Delta\rho \leq 1/2$. Similarly, a low θ city worker will move if:

$$\theta < \bar{\theta} - \frac{c.v}{\Delta\rho} \equiv \breve{\theta} \tag{A4}$$

There are identical re-location thresholds for firms, $\hat{\psi}$ and $\breve{\psi}$. Thus the steady-state allocation will involve no workers (jobs) in the city below $\breve{\theta}$ $(\breve{\psi})$, and none in the rural market above $\hat{\theta}$ $(\hat{\psi})$. This implies two things. First, the population of workers in each market is selected through this relocation

[16] There are of course many models of urban-rural relocation (e.g. Diamond and Tolley, 1982). In this section we follow through the consequences of allowing such relocation for our productivity model.
[17] This very simple symmetric set up just keeps things simple and ensures that the relative variance term is unaffected by the relocation of workers and jobs. Generalizing to allow for different

process. Second, the relocation mechanism raises mean θ and ψ in the city and reducing them in the rural area. In this model, the rural mean is given by:

$$\bar{\theta}_r = \bar{\psi}_r = \frac{1}{4} + \left(\frac{c.v}{\Delta\rho} \right)^2 \tag{A5}$$

Note that this is increasing in *c* and decreasing in $\Delta\rho$. At *c* = 0, $\hat{\theta} = \bar{\theta}$ and the rural mean is 1/4. Similarly, the city mean is:

$$\bar{\theta}_c = \bar{\psi}_c = \frac{1}{2} + \frac{1}{4} - \left(\frac{c.v}{\Delta\rho} \right)^2 \tag{A6}$$

which is decreasing in *c* and increasing in $\Delta\rho$. The city/rural difference in mean θ (and mean ψ) is $1/2 - 2(c.v/\Delta\rho)^2 \geq 0$, depending negatively on ($c.v/\Delta\rho$), with $\Delta\rho$ in turn depending positively on $\Delta\delta$, the density differential. Areas with low relocation costs or big density differentials will see substantial differences in mean worker and job quality between cities and adjacent rural areas.

We consider the implications of relocation for productivity differences. We take the simplest model for productivity; take a firm with quality Ψ, matched with a worker of quality Θ:

$$y = \Theta + \Psi + \beta(\Theta.\Psi), \qquad \beta > 0 \tag{A7}$$

Mean productivity in a market depends on the joint distribution of θ and ψ. Noting that $\mathrm{E}(\theta\psi) = \bar{\theta} + \bar{\psi} + \mathrm{cov}(\theta, \psi)$ [18], we have for the short run (no relocation):

City: $$\mathrm{E}y_c = \bar{\theta} + \bar{\psi} + \beta\overline{\theta}\overline{\psi} + \beta\rho_c V \tag{A8}$$

Rural: $$\mathrm{E}y_r = \bar{\theta} + \bar{\psi} + \beta\overline{\theta}\overline{\psi} \tag{A9}$$

where $V = \sqrt{\mathrm{var}(\theta).\mathrm{var}(\psi)}$, and recalling that $\rho_r = 0$. The productivity gap is therefore $\beta\Delta\rho V$, increasing in β, the importance of the complementarity, $\Delta\rho$, the difference in sorting, and *V*, the scope for reallocation. Using the expressions for the city and rural quality means derived above, and recalling that $\bar{\theta}_j = \bar{\psi}_j$, $j = c, r$, we get for equilibrium:

---

distributions of job and worker quality, and different initial city and rural distributions would add little additional insight.

City: $\mathrm{E}y_c = 2\overline{\theta}_c + \beta\left(\overline{\theta}_c\right)^2 + \beta\rho_c V$ (A10)

Rural: $\mathrm{E}y_r = 2\overline{\theta}_r + \beta\left(\overline{\theta}_r\right)^2$ (A11)

In this case, the productivity differential is $2\left(\overline{\theta}_c - \overline{\theta}_r\right) + \beta\left(\overline{\theta}_c^2 - \overline{\theta}_r^2\right) + \beta\Delta\rho V > \beta\Delta\rho V$ since $\overline{\theta}_c > \overline{\theta}_r$. Using (A5) and (A6) above, this simplifies to:

$$\left(2+\beta\right)\left(\frac{1}{2} - 2\left(\frac{c.v}{\Delta\rho}\right)^2\right) + \beta\Delta\rho V \qquad \text{or}$$

$$\left(2+\beta\right)\left(\overline{\theta}_c - \overline{\theta}_r\right) + \beta\Delta\rho V \qquad \text{(A12)}$$

So the equilibrium impact of assortative matching on productivity exceeds the short-run impact by a factor depending positively on β, the importance of the complementarity, positively on $\Delta\rho$, the difference in sorting, and negatively on c, the relocation cost.

[18] See Mood, Graybill and Boes (1974) p. 180, though this essentially follows simply from the definition of a covariance.