Worker Sorting and Agglomeration Economies

Theodore Papageorgiou*

Penn State University, Department of Economics†

November 2010

Abstract

This paper argues that larger cities allow workers to find better occupational matches. We introduce a framework where workers are initially uncertain about the quality of their match with each occupation. They can switch occupations within cities at no cost, whereas moving across cities is costly. Larger cities offer workers more options, who in turn become more selective, and, in equilibrium, earn higher wages. Using data from the SIPP, we find support for the setup's implications regarding worker mobility: conditional on wages, workers in metropolitan areas are more likely to switch occupations; they are also less likely to move; the negative effect of metro areas on the moving probability is significantly larger for workers who recently moved there; workers who move and switch occupations experience significant wage gains, whereas moving without occupational switching does not affect wages; workers who move from a metro area continue to earn higher wages in their new location, but only if they do not switch occupations.

Keywords: Agglomeration Economies, Occupations, Urban Wage Premium, Moving Probability

JEL Classification: J24, J31, R23

---

*I am grateful to Joe Altonji, Apostolos Burnetas, Ed Coulson, Ed Glaeser, Ed Green, Boyan Jovanovic, Fabian Lange, as well as seminar participants at the Search and Matching Workshop at the University of Pennsylvania and Fall 2010 Cornell-Penn State Macro Workshop for useful comments.

†418 Kern Building, University Park, PA 16802, USA. E-mail: *tpapageorgiou@psu.edu*. Web: *http://sites.google.com/site/theodorepapageorgiou/*

1 Introduction

Workers in more densely populated areas are paid higher wages and produce more output. Since concentrating a large number of workers and firms in one region can be costly, several economists have speculated that agglomeration economies exist. In a recent survey, Glaeser and Gottlieb (2009) note that "there remains a robust consensus among urban economists that [agglomeration] economies exist." Agglomeration economies generally refer to any mechanism that makes economic agents more productive as the level of economic activity in that area increases. Over the years, economists have proposed several mechanisms such as human capital externalities and reduced transportation costs.[1] Nevertheless, Moretti (2010) notes that "the channels that generate these economies remain more elusive."

This paper argues that workers in larger cities are more productive because they form better occupational matches. We construct a model of occupational choice and examine its empirical implications regarding the effect of population density on wages, moving probabilities and occupational mobility within cities. The model's implications regarding mobility, both within as well as across cities, differentiate our proposed mechanism from other potential explanations such as human capital externalities.

In our setup, workers can switch occupations within a city at no cost; on the other hand, moving across cities entails a moving cost. Furthermore, information frictions induce workers to switch across occupations. In particular, the quality of a match between a worker and an occupation is initially uncertain.[2] Larger cities have more occupations and thus offer workers more opportunities to experiment. As a result workers become more selective and are willing to abandon an unpromising match more easily. We show that in equilibrium, workers in larger cities are better matched, they earn higher wages and they are less likely to move.

Using data from the 1996 panel of the Survey of Income and Program Participation (SIPP), we compare workers residing in metropolitan and non-metropolitan areas and we find support for the model's implications.

In our setup, it's not clear whether mobility should be higher or lower in metropolitan areas: on one hand workers in larger cities are better matched, on the other they have better outside options. Indeed empirically the impact of metro on internal mobility is

[1] See for instance Jacobs (1969), Lucas (1988), Jovanovic and Rob (1989), Krugman (1991), Glaeser et al. (1992) and Eaton and Eckstein (1997).

[2] Alternatively one could reinterpret mobility to be the result of shocks to a worker's productivity in a given occupation and the implications would still hold.

very close to zero. Conditional on wages however, which capture match quality, workers in metropolitan areas are in fact more likely to switch, consistent with them having better alternatives.

Furthermore, our framework predicts that workers in metropolitan areas should be less likely to move, since larger cities offer more options. In the data we find indeed the coefficient on metro to be negative and significant. Moreover, the effect of metro is particularly strong in the data for recent movers, who are now significantly less likely to move from a metropolitan area compared to a worker who recently moved to a non-metro area. This is consistent with our framework which predicts that these workers are the most unsure of their occupational match and haven't had the opportunity to sample the occupations offered by the city. On the other hand, a theory of human capital externalities would have difficulty explaining why workers who recently moved to a non-metro area are more likely to move again.

Our paper is not the first to argue that workers form better matches in larger cities. A number of papers have made this point either by arguing that there may be increasing returns to matching between workers and firms and therefore matching is faster in cities (Petrongolo and Pissarides (2006), Gautier and Teulings (2009)) or, as in our case, by arguing that heterogeneous workers have more options in larger cities (Helsley and Strange (1990), Kim (1991), Costa and Kahn (2000)). The present paper differs from the literature in two respects: our setup has testable implications regarding mobility (both within and across cities), which we examine in the data. Furthermore it focuses on occupational, rather than firm matching, which as we argue is more important for workers.[3]

Our paper is also related to earlier work by Glaeser and Maré (2001) who document a number of empirical regularities that are consistent with workers becoming *permanently* more productive in larger cities; Glaeser and Maré do not however specify *how* workers become more productive. We revisit some of their results and argue that they are consistent with larger cities allowing workers to find better occupational matches. For instance, Glaeser and Maré (2001) find that a significant fraction of the wage premium stays with workers when they leave a large city; we show that this holds only for workers who keep their old occupation in their new location, consistent with cities allowing workers to find better matches.

The rest of the paper is organized as follows: Section 2 describes the environment and worker behavior. Section 3 derives model implications regarding the urban wage

[3]Kambourov and Manovskii (2009) for instance argue that controlling for occupational tenure, firm and industry tenure are not important in explaining wages. Gautier and Teulings (2009) also emphasize matching between workers and tasks.

premium, the internal and external mobility. It also investigates their empirical relevance using data from the SIPP. Section 4 concludes.

2 Theory

2.1 Economy

Time is continuous. There is a population of workers who are risk neutral and have discount rate r.

There is also a measure of cities. Each city is characterized by the number of the occupation available in that city, $m \in \{1, 2..., M\}$. The distribution of occupations across cities is exogenous and let q_m denote the fraction of cities with m occupations. Within each city, there is a large mass of firms for each occupation.

While in a city a worker can work in only one occupation any time. Moreover, a worker can switch occupations within a city at no cost. Flow output for worker i, in occupation k, in city l at time t is given by:

$$dY_{tl}^{ik} = \alpha_l^{ik} dt + \sigma dW_{tl}^{ik}$$

where dW_{tl}^{ik} is the increment of a Wiener process and $\alpha_l^{ik} \in \{\alpha_G, \alpha_B\}$ is mean output per unit of time and $\sigma > 0$.

Assume, without loss of generality, that $\alpha_G > \alpha_B$ and that productivities, α_l^{ik}, are independently distributed across occupations, cities and workers. Furthermore, assume α_l^{ik} is unknown, and let $p_{0l}^{ik} \in (0,1)$ be the worker's prior belief that $\alpha_l^{ik} = \alpha_G$. When he enters a city, the worker draws his prior, p_{0l}^{ik}, for all occupations in that city. Each prior, p_{0l}^{ik}, is drawn independently from a known distribution with support $[0,1]$ and density $g(\cdot)$.

Workers can move from one city to another and search for cities is undirected. A worker leaves his current city either endogenously, or exogenously according to a Poisson process with parameter $\delta > 0$. Moving from one city to another entails a fixed cost $c > 0$.

The sequence of actions is the following: a worker moves to a new city. He observes the number of occupations there, m, and draws his prior p_{0l}^{ik} for each occupation. He then chooses one of the occupations and begins working there, or alternatively he may decide pay c and move to another city.

2.2 Behavior

Workers observe their output and obtain information regarding the quality of their match in that specific occupation. Let p_{tl}^{ik} denote the posterior probability that the match of worker i with occupation k is good, i.e. $\alpha_l^{ik} = \alpha_G$. In particular, a worker observes his flow output, dY_{tl}^{ik}, and updates p_{tl}^{ik}, according to (Liptser and Shyryaev (1977)):

$$dp_{tl}^{ik} = p_{tl}^{ik}\left(1 - p_{tl}^{ik}\right)\frac{\alpha_G - \alpha_B}{\sigma}\frac{dY_{tl}^{ik} - \left(p_{tl}^{ik}\alpha_G + \left(1 - p_{tl}^{ik}\right)\alpha_B\right)dt}{\sigma} \tag{1}$$

The last term on the right hand side is a standard Wiener process with respect to the unconditional probability measure used by the agents. p_{tl}^{ik} is a sufficient statistic of the worker's beliefs regarding α_l^{ik}. To minimize notation, from now on, we drop the t subscript, as well as the i and l superscripts.

Firm competition for the services of workers, ensures that a worker's compensation equals his expected output in the occupation n, he's employed:[4]

$$w(p^n) = p^n\alpha_G + (1 - p^n)\alpha_B$$

Each posterior evolves independently of each other and each one evolves only when the worker is employed in the corresponding occupation. Therefore the worker's problem is a multi-armed bandit one. The solution to these type of problems consists of working in the occupation (arm) with the highest Gittins index (see Gittins and Jones (1974)). In the appendix we follow the work of Whittle (1980, 1982) and Karatzas (1984) and show that a worker's optimal strategy is to work at occupation n, such that:

$$n \in \arg\max_{k\in\{1,..m\}}\left\{p^k\right\}$$

We also need to specify when a worker chooses to leave his current city and pay the fixed cost and move (optimal stopping). Consider a worker in a city with m occupations, whose posterior probability of a good match is the same for all occupations and equal to p. Denote his value function as:

$$V(p, m)$$

where $p^k = p$, for all k. Then, as shown in the appendix, a worker optimally pays the fixed cost c and moves to a new city, when the posterior of all his occupations reaches the

[4] Alternatively, one can assume that workers sell their output every period to the firms. In that case, the value function of the worker remains the same, since it depends on the expectation of next instant's output. All the implications derived later continue to hold.

same value, $\underline{p}$, such that:

$$V\left(\underline{p}, m\right) = J$$

where J is the value of a worker about to move to a new city:

$$J = -c + \sum_{m=1}^{M} q_m EV\left(m\right)$$

where, with slight abuse of notation, $EV\left(m\right)$ is the expected value of a worker who moves into a city with m occupations available for him to work in and q_m denotes the fraction of cities with m occupations.

In other words, consider a worker who has just moved to a new city. He immediately draws a prior, p_0^k, for each of the m occupations that are available for him to work in. If all m draws are below $\underline{p}$, then he immediately pays the fixed moving cost c and starts over in another city. Otherwise, he picks the occupation with the greatest value of the prior and begins work there. If the value of his posterior in that occupation falls below the value of the second best occupation, he immediately switches and begins work there. A worker leaves his current city endogenously, only when value of the posteriors of all his occupations reach $\underline{p}$.[5] Some workers however may find that one of the occupations they try out is a good match for them, in which case their posterior drifts towards one and their wage increases. These workers leave their match and city only exogenously, when they are hit by a δ shock.

2.3 Discussion

Before investigating the model's implications, we discuss some of the model's assumptions.

In our setup, residing in a large city is preferred to being in a smaller one since, as we'll see in the next section, workers are able to find better occupational matches and enjoy higher wages. Therefore if workers could choose where to locate, they would all pick the most densely populated regions. In our setup they cannot choose by assumption, since search across cities is undirected.

One could have extended the setup to allow for directed search as in the standard spatial equilibrium model (see, for instance, the framework in Glaeser and Gottlieb (2009), which is based on work by Rosen (1979) and Roback (1982)). In that setup, wages in larger cities are higher due to agglomeration economies, but the price level (cost of living)

[5]For some tasks, the drawn prior may be below $\underline{p}$. The optimal strategy for the worker involves ignoring those tasks and never working there.

adjusts to ensure workers are indifferent across regions in equilibrium. Indeed in the data, wages adjusted for the cost of living, do not vary with population density, while nominal wages are clearly increasing in population density.[6] Thus our assumption of undirected search is not crucial for our results, but allows to focus on the mechanism that generates agglomeration economies without unnecessarily complicating the setup.

Furthermore in our setup, city size, here captured as the number of occupations, is assumed to be exogenous. Since occupations are generally associated with some industries, one can rephrase the question as what determines the number (and composition) of industries across cities. As in the standard spatial equilibrium, one would expect a combination of the following three factors to explain why some cities are larger and why some are smaller: fixed city productivity, which captures natural advantages that certain cities enjoy that contributed to their growth, such as New York's natural harbor; housing supply, in particular land scarcity; amenities, such as warm climate.

Moreover one may imagine an additional reason that would lead to larger cities having more occupations: certain occupations require relatively larger markets in order to exist. For instance opera singers exist only in larger cities, since an opera would not be profitable in a small town.

3 Model Implications

The next subsection discusses the data used in our investigation of the model's implications. We begin by looking at implications relating to the urban wage premium, then examine occupational mobility within cities and finally mobility across cities or regions. The predictions of our setup concern cities with more occupations compared to cities with fewer occupations. In the data however we do not have information on the number of occupations per city. As we show however, in the external mobility section, our setup predicts that cities with more occupations have higher populations, so we can use that in our empirical investigation instead.

3.1 Data

We use data from the 1996 panel of the Survey of Income and Program Participation (SIPP). In the 1996 SIPP, interviews were conducted every four months for four years and included approximately 30,000 households. It includes information on the worker's

[6] See for instance Figure 2 in Glaeser and Maré (2001) and compare with their Figure 1.

wage, occupation, current employer, industry, employer size, as well as the usual demographics, such as gender, age, race, education and marital status.[7] The 1996 panel of the SIPP uses dependent interviewing, which is found to reduce occupational coding error (Hill (1994)). This makes it preferable to use when investigating occupational switching, compared to other panel datasets, such as the National Longitudinal Survey of Youth 1979. Furthermore, the SIPP follows original respondents when they move to a new address, unlike, for instance, the Current Population Survey which is an address-based survey.

The SIPP includes three variables that provide information regarding the geographical location of the respondents. The first identifies the worker's state. The second variable, called TMETRO, records whether the respondent is located in a metropolitan area or not. The third variable, called TMSA, identifies one of 93 MSAs (Metropolitan Statistical Areas) and CMSAs (Consolidated Metropolitan Statistical Areas), as defined by the Office of Management and Budget.

Our theory predicts that workers in metropolitan areas should benefit from the increased variety of occupational choice and eventually become better matched. Our empirical work involves looking the behavior of workers in metropolitan areas comparing it to that of those in non-metro areas.[8] We also use the state and the TMSA and TMETRO variables, to identify whether a worker has moved.[9]

3.2 Urban Wage Premium

A worker's posterior in his current occupation, $p^n \equiv \max_{k \in \{1,..m\}} \left\{p^k\right\}$, is increasing in the total number of occupations, m: workers with more potential matches are better matched on average. Thus, workers in cities with higher m, are more likely to have higher output. Since firm competition ensures workers are paid their marginal product, this leads to the following result:

Implication *Workers in cities with more occupations, m, earn on average higher wages.*

[7] In our investigation, we exclude workers in the armed forces. Wages are deflated to real 1996 dollars using the Consumer Price Index.

[8] Accoring to the Survey of Income and Program Participation Users' Guide (Third Edition, 2001), "[t]o protect respondent confidentiality, the Census Bureau recorded and identified a small random sample of metropolitan households in the public use files as nonmetropolitan." This implies that the results that follow may be underestimating the true differences between metropolitan and non-metropolitan workers with respect to the economic variables considered (wages, moving probabilities etc.)

[9] In our specification, a worker moves when (at least) one of the three location variables change from one wave to the next.

	Initial	Moved <4 years	Full Sample
	ln(wage)	ln(wage)	ln(wage)
Metro	0.033	0.05	0.105
	(0.018)*	(0.01)***	(0.002)***

Table 1: Wage Premium Evolution. Source: 1996 Panel of Survey of Income and Program Participation. 4-month intervals. Controls include gender, race, education, marital status, firm size, quadratic in age, 11 industry dummies.

Furthermore, since the highest draw from $g(\cdot)$ is increasing on average in the number of draws, the above proposition holds for initial wages as well:

Implication *Initial wages are higher in cities with more occupations m.*

Table 1 presents the evolution of the urban wage premium. The last column of Table 1, confirms the well-known empirical regularity that workers in metropolitan areas are paid significantly more than workers in non-metropolitan areas. In the first column of Table 1, we see that initial wages, i.e. the wages of workers who just moved to the area, are also higher in metropolitan areas. This is consistent with second implication above. However it is interesting to note that the coefficient is three times smaller. Expanding the set to include workers who moved within the past four years leads to an increase of the urban wage premium to almost half of that of the full sample.

This implies that the mechanism that generates wage difference among workers in metro and non-metro areas, appears to be relevant mostly *after* a worker arrives in a metro area. Furthermore, as Glaeser and Maré (2001) point out, the large difference between the coefficients also implies that selection is not the main driving force behind agglomeration economies.[10]

The gradual increase of the urban wage premium is also consistent with our setup. Even though workers arriving in areas with more options form better matches initially, over time one expects the difference to become larger: workers in metro areas are more likely to sort in an occupation that is a good match for them and see their wages increase, as their expected output goes up.

3.3 Internal Mobility

This section analyzes occupational switching behavior of residents. We look at their probability of switching occupations and whether residing in a metro area affects it.

[10]If the wage premium was the result of selection, then recent migrants would receive higher wages since they would be the most able.

If workers in cities enjoy a better selection of occupational choices, then we would expect their internal mobility decisions to differ from workers in less populated areas.

Following Karlin and Taylor (1981), we can show that, ignoring δ shocks, the probability of an occupation switch for a worker whose posterior in his current occupation is equal to p, is given by:

$$\text{Occ Switch Prob} = \Pr\left(p \text{ reaches } p^{\text{sec}} \text{ before } 1\right) = \frac{1-p}{1-p^{\text{sec}}}$$

where p^{sec} is value of the worker's second highest posterior. Clearly the above probability is decreasing in p and increasing in p^{sec}.

One might expect the setup to predict that internal mobility to be higher in larger cities. That is not however, necessarily the case: as we saw above, workers in larger cities have higher posteriors in their current occupations p. Their second highest posterior, p^{sec}, is also increasing in m, the total number of occupations. Therefore, without additional assumptions, the number of occupations has an ambiguous effect on internal mobility.[11] Put differently, workers in larger cities are both better matched, which tends to decrease their switching probability, but also have better outside options, which increases the probability they switch.

If however we condition on quality of the current match by controlling for the wage, then our setup predicts that mobility is higher in larger cities. Comparing two workers that earn similar wages, the one located in a larger city is more likely to switch occupations because he has better outside options:

Implication *Conditional on p (wage), internal mobility is higher in cities with more occupations m.*

Indeed in the first column of Table 2 we note that occupational mobility is not higher in metropolitan areas, indicating that the two effects, better matching and better outside options, roughly offset each other.[12] On the other hand, when we control for the wage in the second column, the impact of metro is positive and very significant, consistent with our framework.[13]

Another way of examining whether in fact the occupational options of metro area

[11] For instance, with more restrictions on the shape of $g(\cdot)$, the initial beliefs distribution, it could be that one effect unambiguously dominates.

[12] Bleakely and Lin (2007) find that occupational and industrial mobility is higher in more densily populated areas for younger workers. In our theoretical setup, even newly arrived workers do not necessarily switch more often in larger cities, since they are better matched, as shown in the first column of Table 1. Therefore this is not an unambiguous implication of our setup.

[13] Workers in metro areas switch occupations *at* higher wages; they also switch *to* higher wages (both results are available upon request). Both facts are consistent with our framework.

	Occ. Switching	Occ. Switching
	Prob. (Probit)	Prob. (Probit)
	Residents	Residents
Metro	0.0003	0.007
	(0.001)	$(0.002)^{***}$
ln(wage)		-0.067
		$(0.002)^{***}$

Table 2: Metro Impact on Internal Mobility. Source: Source: 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, marital status, firm size, quadratic in age, 11 industry dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equals 0.0917 (no wage control) and 0.109 (with wage control).

	Prob. of Keep Old Occ.
	& Not Move (Probit)
Metro	0.159
	$(0.079)^{**}$

Table 3: Metro Impact in case of Plant Closings. Source: 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, marital status, firm size, quadratic in age, 11 industry dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equals 0.3391.

residents are better is to consider the following: imagine that there is a set of workers who have discovered their preferred occupation. If these workers are "dropped" into a large city we would expect them to have a higher probability of being able to work in their preferred occupation. Empirically we can look at workers who have been laid off due to plant closings. We consider one of three alternatives for these workers: keep their old occupation and not move, switch occupations and not move or move to another region. As shown in Table 3, workers in metro areas are more likely to keep their old occupation and not move, consistent with our hypothesis.[14]

[14] Our result is consistent with the finding of Bleakley and Lin (2007), who use the Current Population Survey's Displaced Worker Supplement and find that displaced workers who live in metro areas are less likely to switch occupations (see their Table 7).

3.4 External Mobility

In this section we investigate external mobility, namely the movement of workers across cities. We analyze the interaction between the moving probability, the size of a city and also whether a worker has recently moved there. We also investigate the impact of moving on wages and how it relates to occupation switching. Finally we revisit a finding by Glaeser and Maré (2001) regarding the urban wage premium when a worker moves from a city.

3.4.1 Moving Probability

We next consider the probability that a worker leaves a city. As we saw, a worker leaves a city when his posterior for all occupations is less or equal to $\underline{p}$. Consider a worker who has moved to a new city with m occupations. Assume that $d \leq m$ of his draws are above $\underline{p}$. Then the probability he leaves the city endogenously, conditional on d, is given by:

$$\Pr\left(p^1 \text{ reaches } \underline{p}\right) \times \Pr\left(p^2 \text{ reaches } \underline{p}\right) \times ... \times \Pr\left(p^d \text{ reaches } \underline{p}\right)$$

Since $\Pr\left(p^k \text{ reaches } \underline{p}\right) < 1$ for all k with $p_0^k > \underline{p}$, then the probability that a worker moves endogenously is decreasing in d.

However d, the number of draws above $\underline{p}$ is increasing in the total draws, m. Thus the probability that a worker moves endogenously is decreasing in m:

Implication *The rate at which workers move out of a city, is lower in cities with more occupations m.*

This implies that workers stay longer in cities with more occupations, m. Since the flow into a city is the same regardless of the number of occupations, the above result immediately leads to the following corollary:[15]

Corollary *Cities with more occupations, m, have large populations.*

We now turn to the data. In our framework workers move either endogenously, because their current occupational match is sufficiently unpromising, or exogenously, for reasons beyond our modelling assumptions. As Table 4 indicates, moving is generally associated with occupational switching.

Furthermore, as seen in the first column of Table 5, the moving probability is lower for residents of metro areas. Since larger cities offer worker more occupational options, workers are less likely to move.

[15] In fact the flow into larger cities is slightly larger, since the probability that all prior draws are less than $\underline{p}$, is decreasing in m. This reinforces the following corollary.

	Moving
	Prob. (Probit)
Switch Occupation dummy	0.018
	$(0.001)^{***}$

Table 4: Moving Probability and Occupation Switching. Source: 1996 Panel of Survey of Income and Program Participation. 4-month probabilities. Controls include gender, race, education, marital status, quadratic in age, 11 industry dummies. Coefficient represents marginal effect evaluated at the average value of the 4-month probability, which equals 0.0231.

	Moving	Moving
	Prob. (Probit)	Prob. (Probit)
Metro	-0.012	-0.009
	$(0.001)^{***}$	$(0.001)^{***}$
Recent		0.03
		$(0.002)^{***}$
Metro × Recent		-0.029
		$(0.002)^{***}$

Table 5: Impact of Metro and Recent on Moving Probability. Source: 1996 Panel of Survey of Income and Program Participation. Recent refers to workers who moved to the area less than 4 years ago. 4-month probabilities. Controls include gender, race, education, marital status, firm size, quadratic in age, 11 industry dummies. Coefficients represent marginal effects evaluated at the average value of the 4-month probability, which equals 0.0242 (only metro control) and 0.0226(with recent controls).

The better options offered by larger cities are particularly important for recent movers, since they are the ones who are unsure of their occupational match (which is why, all else equal, they are more likely to move) and haven't had the opportunity to sample the occupations offered by the city. One would expect these workers to benefit the most from being in a metro area. As shown in the second column of Table 5, for recent movers, the effect of being in a metro area is indeed three times higher.

On the contrary, models of human capital externalities cannot easily explain why workers who recently moved have a very high probability of moving out of a non-metro area. One may argue that there are exogenous reasons that force people to move (like the δ shocks in our setup). However it's not clear under the human capital hypothesis, why these shocks should depend on the area's population density, the time a worker has spent in a region or the interaction of the two.

3.4.2 Moving, Wages and Occupation Switching

In this section we look at the impact of moving on wages and also how it relates to occupation switching. In our framework, workers pay fixed cost and move because they expect a better match in their new location. Thus, by revealed preference:

Implication *Workers who move endogenously experience wages increases.*

On the other hand, if a worker is hit by a δ shock and is transferred to a new location he may or may not see his wage increase. In our model, occupations are city specific and when a worker moves, he switches occupations. Although this assumption simplifies the analysis considerably, it does not allow for moving without switching occupations. One would be interested in considering a more general model in which there is some overlap of occupations across cities.[16] Let's therefore consider the following extension of our setup:

When hit by a δ shock, then a worker's old occupation may be available in the new location. This probability is increasing in the number of occupations, m, of the new location. If the old occupation is in fact available, then the worker stays there if:

$$\max\left\{p^{old}, p^{n}\right\} = p^{old}$$

where p^{old} is the posterior associated with old occupation and $p^{n} = \max_{k\in\{1,..m-1\}}\left\{p^{k}\right\}$.[17] In that case the wage doesn't change and equals $w\left(p^{old}\right)$.

Implication *No wage change when moving without switching occupations.*

We examine empirically these implications. The first column of Table 6 shows that moving is associated with a modest wage increase. In the second column we include controls for employer and occupation switching, along with their interaction with moving. There are a couple of things to notice here.

First, given that one could easily modify the setup to focus on employer matching instead, it's interesting to note that employer switching, either within an area, or combined with moving, appears to have either no effect or a negative effect on wages. On the other

[16] In a more general model, there is a fixed set of occupations and each occupation may be available in more than one cities. Workers die at some rate and new workers enter at the same rate. In that case, when a worker moves to a new city, some of the occupations may already be familiar to him. One expects our setup's implications to hold in this version of the model as well, as workers in larger cities will tend to be better matched. Such a model however would require keeping track of the worker's beliefs regarding every occupation he has ever sampled, a complicated object that is analytically intractable.

[17] This implies that the probability of switching occupations when moving depends negatively on the pre-move wage. This is true in the data as well (results available upon request). One however, cannot derive predictions regarding the probability of switching occupations when moving to a larger city, since larger cities increase the probability that the old occupation is available, but decrease the probability of keeping it since outside options are better. In the data it turns out that workers who move to metro areas are less likely to switch occupations, so the first effect dominates.

	$\ln(\text{wage})_t$	$\ln(\text{wage})_t$
$\ln(\text{wage})_{t-1}$	0.859	0.868
	$(0.001)^{***}$	$(0.001)^{***}$
Move	0.008	0.006
	$(0.003)^{***}$	$(0.003)^{*}$
Move × Occupation Switch		0.020
		$(0.011)^{*}$
Move × Employer Switch		-0.011
		(0.011)
Occupation Switch		0.011
		$(0.002)^{***}$
Employer Switch		0.002
		(0.002)

Table 6: Impact of Occupation Switching and Moving on Wage. Source: 1996 Panel of Survey of Income and Program Participation. 4-month intervals. Controls include gender, race, education, marital status, quadratic in age, 11 industry dummies.

hand, workers who move and switch occupations make significant wage gains, consistent with the implications of our theory.

Second, moving without occupation switching is not associated with wage increases, consistent with the second implication above: a worker who moves but remains in the same occupation should keep his old wage.

Finally, we should note that even switching occupations within cities entails some wage gain, albeit smaller than in the case of moving and switching. Our setup predicts that workers who switch within cities should not experience wage gains. This is an immediate consequence of our assumption that the cost of switching occupations within cities is zero. One may imagine nonetheless that there are some costs associated with occupation switching, which should be compensated by corresponding wage gains when switching.[18] The results of our setup depend on the assumption that the costs of switching occupations are larger across cities, rather than within. Indeed, the corresponding gains of workers who switch occupations and move are larger than those who switch without moving.

It's also worth considering the path of wages before moving. In Table 7 we see that if a worker is going to move in period t, then his wage evolves negatively from period $t-2$ to $t-1$, indicating a declining wage path. This indicates that for at least some of the moves we observe, labor market considerations are important in the decision to move.

[18] Introducing costs to switching occupations within cities complicates the analytical solution to the worker's occupational choice problem substantially. In particular, Gittins indeces cannot be used in the presence of even $\varepsilon > 0$ cost to switching.

	$\ln(\text{wage})_{t-1}$
Move_t	-0.008
	$(0.003)^{**}$
$\ln(\text{wage})_{t-2}$	0.86
	$(0.001)^{***}$

Table 7: Wage Path Before Moving. Source: 1996 Panel of Survey of Income and Program Participation. 4-month intervals. Controls include gender, race, education, marital status, quadratic in age, 11 industry dummies.

	$\ln(\text{wage})_{t-1}$
$\text{Move}_t \times \text{Occupation Switch}_t$	-0.021
	$(0.007)^{***}$
$\text{Move}_t \times \text{No Occupation Switch}_t$	-0.0003
	(0.004)
$\ln(\text{wage})_{t-2}$	0.864
	$(0.001)^{***}$

Table 8: Wage Path Before Moving and Switching Occupations. Source: 1996 Panel of Survey of Income and Program Participation. 4-month intervals. Controls include gender, race, education, marital status, quadratic in age, 11 industry dummies.

In our setup, workers move endogenously following a downward revision of their beliefs, which is also reflected in their wages:

Implication *Workers experience wage decreases before moving endogenously.*

Indeed as we see in Table 8, wages are declining beforehand only in the case of workers who move *and* switch occupations, consistent with the predictions of our theory. Workers who move and keep the same occupation do not experience decreasing wages before moving, consistent with them experiencing an exogenous δ shock, which does not depend on their previous wage path. Thus the results of Table 7 are driven by workers who move and switch occupations.[19]

3.4.3 Urban Wage Premium when Moving

Finally, we revisit a result by Glaeser and Maré (2001) and attempt to interpret it in light of our framework. Glaeser and Maré (2001) look at the wages of workers who move

[19] If we control for occupation switching separately in Table 10, only occupation switching is statistically significant, whereas moving is not. This is consistent with the predictions of our setup that workers who switch occupations experience wage declining paths beforehand, regardless of whether they move or not; moving does not imply a steeper wage decline, compared to switching occupations without moving.

	ln(wage)	ln(wage)	ln(wage)
	After Move	After Move	After Move
		No Occup Switch	w/ Occup Switch
Previously in Metro	0.106	0.127	0.037
	(0.012)***	(0.014)***	(0.027)

Table 9: Urban Wage Premium when Moving. Source: 1996 Panel of Survey of Income and Program Participation. 4-month intervals. Controls include currently in metro, gender, race, education, marital status, current and last period's firm size, quadratic in age, 11 industry dummies.

and find that those who moved away from an urban area earn higher wages in their new location compared to those who move from a rural area. Indeed, as shown in the first column of Table 9, workers who moved from a metro area earn significantly higher wages in their new location. They interpret this as evidence that workers become *permanently* more productive while in an urban area, but do not provide further evidence on the underlying mechanism.

In our setup this mechanism is explicit: cities allow workers to find better occupational matches. Moreover, if a worker leaves a city, but is able to keep his occupational match, then our setup predicts that he'll continue to reap the benefits: as we saw above, his wage doesn't change. In order to investigate our explanation further, we split the set of workers who moved between those who switched occupations and those who did not (second and third column). Indeed, workers who stayed in the same occupation, received significantly higher wages if they moved from a metropolitan area. On the other hand, among workers who moved and switched occupations, the impact of metro was much smaller and statistically not significant, consistent with the predictions of our setup.

4 Conclusion

This paper argues that larger cities allow workers to find better occupational matches. Its predictions are consistent with the observed patterns in the data regarding the joint implications of moving, occupation switching and wage changes. Furthermore, it is consistent with the observed internal mobility patterns: conditional on wages, workers in metropolitan areas are more likely to switch occupations.

Unlike theories that emphasize human capital externalities in cities, the setup can jointly account for observed differences in both wages and moving probabilities across metro and non-metro areas. For instance, it is consistent with recent movers to non-metro areas being more likely to move again, compared to recent movers in metropolitan areas.

Appendix

A Worker's Optimal Strategy

Following Whittle (1980, 1982) and Karatzas (1984), the solution to this problem consists of finding a retirement value for each occupation and then work in the occupation with the highest retirement value. This retirement value serves as an index for each occupation, which corresponds to that occupation's Gittins index (see Gittins and Jones (1974) and Bergemann and Valimaki (2006)).

For every occupation, k, with probability p^k of being good and retirement value W^k, we want to compute the optimal retirement policy. In other words, at every instant, workers can work in this occupation or retire and obtain value W^k.

In that case, the value function of a worker with posterior p^k and the option of retiring and obtaining value W^k, $V\left(p^k, W^k\right)$ is given by:

$$rV\left(p^k, W^k\right) = w\left(p^k\right) + \frac{1}{2}\left(\frac{\alpha_G - \alpha_B}{\sigma}\right)^2 \left(p^k\right)^2 \left(1-p^k\right)^2 V_{pp}\left(p^k, W^k\right) - \delta\left(V\left(p^k, W^k\right) - J\right)$$

where $V_{pp}(\cdot)$ is the second derivative of $V(\cdot)$ with respect to p. The flow benefit of the worker consists of his flow output, plus a term capturing the option value of learning, which allows him to make informed decisions in the future. Finally, the worker leaves his current city exogenously at rate δ and pays the fixed c and moves to a new one.

Guessing that $V(\cdot)$ is increasing in p^k, the optimal stopping rule is retire when p^k reaches $\underline{p}\left(W^k\right)$ such that:

$$V\left(\underline{p}\left(W^k\right), W^k\right) = W^k$$

$$V_p\left(\underline{p}\left(W^k\right), W^k\right) = 0$$

The solution to the above differential equation is given by:

$$\begin{aligned} V\left(p^k, W^k\right) &= \frac{w\left(p^k\right) + \delta J}{r} \\ &+ \frac{\alpha_G - \alpha_B}{r}\left(\underline{p}\left(W^k\right) + \frac{1}{2}\theta - \frac{1}{2}\right)^{-1} \underline{p}\left(W^k\right)^{\frac{1}{2}+\frac{1}{2}\theta}\left(1-\underline{p}\left(W^k\right)\right)^{\frac{1}{2}-\frac{1}{2}\theta} \\ &\times \left(p^k\right)^{\frac{1}{2}-\frac{1}{2}\theta}\left(1-p^k\right)^{\frac{1}{2}+\frac{1}{2}\theta} \end{aligned}$$

where:

$$\underline{p}\left(W^k\right) = \frac{(\theta-1)\left((r+\delta)W^k - \alpha_B - \delta J\right)}{\alpha_G + \alpha_B + (\alpha_G - \alpha_B)\theta - 2(r+\delta)W^k + 2\delta J} \tag{2}$$

and $\theta = \sqrt{\frac{8(r+\delta)}{\left(\frac{\alpha_G - \alpha_B}{\sigma}\right)^2} + 1}$. Note that $\underline{p}\left(W^k\right)$ is strictly increasing in W^k.

The index of occupation k is the highest retirement value at which the worker is indifferent between working at occupation k or retiring with $W^k = W\left(p^k\right)$, i.e.:

$$W\left(p^k\right) = V\left(p^k, W^k\right) \tag{3}$$

For eq. (3) to hold, it must be the case that:

$$p^k = \underline{p}\left(W^k\right) \tag{4}$$

Substituting conditions (3) and (4) into the threshold condition, equation (2), we obtain:

$$\begin{aligned} p^k &= \frac{(\theta-1)\left((r+\delta)W\left(p^k\right) - \alpha_B - \delta J\right)}{\alpha_G + \alpha_B + (\alpha_G - \alpha_B)\theta - 2(r+\delta)W\left(p^k\right) + 2\delta J} \Rightarrow \\ W\left(p^k\right) &= \frac{1}{r+\delta}\frac{(\alpha_G + \alpha_B + (\alpha_G - \alpha_B)\theta + 2\delta J)p^k + (\theta-1)(\alpha_B + \delta J)}{2p^k + \theta - 1} \end{aligned}$$

which is strictly increasing in p^k. Therefore the optimal strategy of the worker is to work in the occupation with the highest posterior.

In what follows, we show that all workers leave their current city when the posterior of all their occupations reaches the same value, $\underline{p}$, regardless of the total number of occupations, m.

Consider the value of a worker employed in a firm with m occupations, whose posterior probability of a good match is equal for all occupations. Denote his value functions as:

$$V(p, m)$$

where $p^k = p$, for all k. $V(p,m)$ is non-decreasing in m, as a worker can ignore any extra occupations he wishes. Moreover $V(p,m)$ must be increasing in p. To see this, note that as we showed above, p corresponds to the Gittins index of all occupations. Because the optimal strategy of the worker is to work in the occupation with the highest Gittins index, then by revealed preference, his value must be increasing in p^n and therefore p as well.

If $m = 1$, i.e. the worker is located in a city with only one occupation, then a necessary condition for a worker to leave his current city is for his posterior, p, to reach a level $\underline{p}(1)$, such that:[20]

$$V\left(\underline{p}(1),1\right) = J$$

In other words, $\underline{p}(1)$ denotes the value of the posterior at which point a worker with one occupation pays the fixed cost c and moves to another city. We are looking for equilibrium in which when a worker chooses to move, his posterior for all occupations have the same value. We show later that this has to be true. Denote this value of the posterior by $\underline{p}(m)$. Then a necessary condition for a worker with m occupations to move is:

$$V\left(\underline{p}(m),m\right) = J$$

Since $V(\cdot)$ is weakly increasing in m, then it has to be that $\underline{p}(m) \leq \underline{p}(1)$ for $m > 1$. To see this, note that if $\underline{p}(m) > \underline{p}(1)$, then it could never be the case that $V\left(\underline{p}(m),m\right) = V\left(\underline{p}(1),1\right) = J$, since $V(\cdot)$ is increasing in p.

Consider the worker who has only one occupation with posterior $\underline{p}(1)$. Since he is indifferent between trying it out and moving to another city, then the option of trying out this occupation is clearly zero. If he has more occupations with posteriors equal to $\underline{p}(1)$, then their option value is also equal to zero and the worker will continue being indifferent between trying them out and moving. Thus he is no better off and therefore it cannot be that $\underline{p}(m) < \underline{p}(1)$. Therefore it must be the case that $\underline{p}(m) = \underline{p}(1) = \underline{p}$ for all m.

Finally we need to show that a worker moves only when all his occupations have posterior equal to $\underline{p}$ (or less, which happens if the initial draw from $g(\cdot)$ is less than $\underline{p}$). To see this, consider a worker with one occupation with posterior greater than $\underline{p}$. Since $V(\cdot)$ however is increasing in p^n, then it must be the case that this value is greater than $V\left(\underline{p},m\right) = J$ and it can never be optimal for him to move to another city.

References

[1] Bergemann, D. and J. Valimaki (2008): "Bandit Problems," in *The New Palgrave Dictionary of Economics,* ed. by S. Durlauf and L. Blume, New York: Macmillan.

[2] Bleakley, H. and J. Lin (2007): "Thick-Market Effects and Churning in the Labor Market: Evidence from U.S. Cities," Federal Reserve Bank of Philadelphia Working Paper 07-23, Philadelphia, PA.

[20] This is a necessary, but not sufficient condition.

[3] Costa, D. L. and M. E. Kahn (2000): "Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990," *Quarterly Journal of Economics*, 115(4), 1287-1315.

[4] Eaton, J. and Z. Eckstein (1997): "Cities and Growth: Theory and Evidence from France and Japan," *Regional Science and Urban Economics*, 27(4-5), 443-474.

[5] Gautier, P. A. and C. N. Teulings (2009): "Search and the City," *Regional Science and Urban Economics*, 39(3), 251-265.

[6] Gittins, J. C. and D. M. Jones (1974): "A Dynamic Allocation Index for the Sequential Design of Experiments," in *Progress in Statistics*, vol. 1, ed. by J. M. Gani, K. Sarkadi, and I. Vincze, 241-266, Amsterdam: North-Holland.

[7] Glaeser, E. L. and J. D. Gottlieb (2009): "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, 47(4), 983-1028.

[8] Glaeser, E. L., H. D. Kallal, J. A. Scheinkman and A. Shleifer (1992): "Growth in Cities," *The Journal of Political Economy*, 100(6), 1126-1152.

[9] Glaeser, E. L. and D. C. Maré (2001): "Cities and Skills," *Journal of Labor Economics*, 19(2), 316-342.

[10] Helsley, R. W. and W. C. Strange (1990): "Matching and Agglomeration Economies in a System of Cities," *Regional Science and Urban Economics*, 20(2), 189-212.

[11] Hill, D. H. (1994): "The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey," *Journal of Official Statistics*, 10(4), 359-380.

[12] Jacobs, J. (1969): *The Economy of Cities,* New York: Vintage Books.

[13] Jovanovic, B. and R. Rob (1989): "The Growth and Diffusion of Knowledge," *The Review of Economic Studies*, 56(4), 569-582.

[14] Kambourov, G. and I. Manovskii (2009): "Occupational Specificity of Human Capital," *International Economic Review*, 50(1), 63-115.

[15] Karatzas, I. (1984): "Gittins Indices in the Dynamic Allocation Problem for Diffusion Processes," *The Annals of Probability*, 12(1), 173-192.

[16] Karlin, S. and H. M. Taylor (1981): *A Second Course in Stochastic Processes*, New York: Academic Press.

[17] Kim, S. (1991): "Heterogeneity of Labor Markets and City Size in an Open Spatial Economy," *Regional Science and Urban Economics*, 21(1), 109-126.

[18] Krugman, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483-499.

[19] Liptser, R. and A. Shyryaev (1977): *Statistics of Random Processes,* Vol. 2 Berlin: Springer-Verlag.

[20] Lucas, R. E. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), 3-42.

[21] Moretti, E. (2010): "Local Labor Markets," *Handbook of Labor Economics*, forthcoming.

[22] Petrongolo, B. and C. Pissarides (2006): "Scale Effects in Markets with Search," *The Economic Journal*, 116(1), 21-44.

[23] Roback, J. (1982): "Wages, Rents and the Quality of Life," *Journal of Political Economy*, 90(6), 1257-1278.

[24] Rosen, S. (1979): "Wage-based indexes of urban quality of life," in *Current Issues in Urban Economics*, ed. by P. Mieszkowski and M. Straszheim, 74-104, Baltimore and London: Johns Hopkins University Press.

[25] Whittle, P. (1980): *Optimization over Time: Programming and Stochastic Control*, New York: Wiley.

[26] Whittle, P. (1982): "Multi-Armed Bandits and the Gittins Index," *Journal of the Royal Statistical Society*, Series B, 42(2), 143-149.