Mismatch Unemployment and the Geography of Job Search[†]

Ioana Marinescu[‡]
Roland Rathelot[§]

December 1, 2014

Abstract

To what extent can we reduce unemployment by relocating job seekers and jobs closer to each other? It depends on the degree to which job seekers are stuck in their current location. Using data from the leading job board CareerBuilder.com, we estimate a job seeker's willingness to apply to a job as a function of the job's distance to the job seeker's zip code of residence. Plugging this estimate into a new theoretical model, we find that optimally relocating job seekers would decrease unemployment by at most 5%, implying that geographic mismatch is a negligible driver of US unemployment.

Keywords: local labor markets, job search, misallocation, mismatch, applications, vacancies, unemployment.
JEL: J21, J61, J62, J64.

[†]We would like to thank Pierre Cahuc, Raj Chetty, Bruno Crépon, Xavier D'Haultfoeuille, François Fontaine, Florence Goffette-Nagot, Larry Katz, Philip Kircher, Francis Kramarz, Thomas Le Barbanchon, Nathan Hendren, Etienne Lehmann, Manasa Patnam, Barbara Petrongolo, Julien Prat, Craig Riddell, Antoinette Schoar, Etienne Wasmer and participants to seminars at CREST, Cachan/Paris-Sud, GATE (Lyon), Mannheim, Uppsala, Louvain-la-Neuve, Bonn, EIEF (Rome), FGV (Rio), PUC (Rio), Sciences Po, Stockholm, Vienna, Warwick, Aarhus, Copenhagen, Montreal, Pompeu Fabra, UIC, Harvard, Harris School (Chicago), Booth (Chicago), Queen Mary as well as the SaM Meeting at Edinburgh, the ESEM at Toulouse, and the workshop Topics in Labor Economics for useful comments. We are grateful to CareerBuilder.com for letting us use their data. Any opinions expressed here are those of the authors and not of any institution.

[‡]Harris School of Public Policy, University of Chicago. ioanamarinescu@uchicago.edu
[§]University of Warwick. r.rathelot@warwick.ac.uk. Corresponding author.

1 Introduction

In the aftermath of the Great Recession, the recovery of the US labor market has been sluggish. While the unemployment rate increased from about 5% to 10% during the Great Recession, it was still as high as 8% in 2012, three years after the official end of the recession as defined by the NBER Business Cycle Dating Committee. One potential explanation for this sluggish recovery is geographic mismatch: job seekers in depressed areas may not be able or willing to relocate to areas with better job prospects. Two reasons make this explanation plausible. First, there has been a sustained decline in internal migration since the 1980's (Molloy et al., 2011). Second, because of the collapse in house prices during the crisis, home owners may be unwilling to incur large capital losses in order to move ("house lock").

Despite the plausibility of the mismatch hypothesis, geographic mismatch did not increase in the aftermath of the Great Recession (Şahin et al., 2014). However, the mismatch index used in prior literature cannot precisely determine the *level* of mismatch because it assumes that job seekers do not search for jobs across geographic boundaries. Furthermore, as demonstrated by Şahin et al. (2014), the level of mismatch is sensitive to the assumption made about the geography of job search: if we assume that job seekers search within their state (but not across states), measured mismatch is lower than if we assume that job seekers search within their MSA (but not across MSAs). Therefore, while the macro literature shows that mismatch did not increase in the aftermath of the Great Recession, it cannot determine the exact level of geographic mismatch because it does not account for job search across geographic units.

To determine with greater accuracy the level of geographic mismatch, we need a model that takes into account the fact that job seekers apply across geographic units, such as the model in Manning and Petrongolo (2011). Exploiting fine-grained geographic data on the location of job seekers and vacancies in the U.K., Manning and Petrongolo (2011) infer where job seekers apply for jobs. Since job seekers do not seem to apply very far from home, geographic mismatch in the U.K. may be large. However, while Manning and Petrongolo (2011) take into account the geography of job search, their paper focuses on the impact of place-based policies rather than mismatch unemployment.

In this paper, we use zip code level data on the geography of job search in order

to determine the level of geographic mismatch in the US in 2012. Our data comes from CareerBuilder.com, a leading job board in the US. The data is broadly representative of the US labor market. It contains information of the zip code of vacancies and job seekers, and on the applications from job seekers to vacancies. The application data allows us to directly estimate a job seeker's willingness to apply to a vacancy as a function of the distance between the vacancy and the job seeker's zip code of residence. We develop a new model of geographic mismatch that allows for job applications across geographic units (e.g. zip codes). Using our model and our empirical estimates of the probability of applications across zip codes, we find that US aggregate unemployment would decrease by at most 5% if job seekers were reallocated to maximize hires.

To compute mismatch, we use a directed search model where workers choose where to send their applications, and vacancies closer to home yield higher utility. Our theoretical model is close to the one by Manning and Petrongolo (2011), but we also borrow elements from Albrecht et al. (2006), Chade and Smith (2006), and Galenianos and Kircher (2009). The model yields the number of hires with the existing allocation of job seekers and vacancies. We can also compute the number of hires under the allocation of job seekers that maximizes hires. Geographic mismatch is measured as the difference between the latter and the former.

Documenting the geography of job search, we find that job seekers are more likely to apply to jobs closer to home: for example, a job seeker is 50% as likely to apply to a vacancy that is 25 miles away relative to a vacancy that is in the job seekers' own zip code. Still, we find that 16% of applications are to out-of-state vacancies. Overall, the American job seekers' willingness to apply to far away jobs is much greater than that of British job seekers (Manning and Petrongolo, 2011).

How sensitive is the level of geographic mismatch to our estimates of job seekers' willingness to apply to far away jobs? We compute geographic mismatch for a range of distastes for distance and find that it varies very little, even for moderate increases in the distaste for distance. We show that the robustness of our results to changes in the distaste for distance can be explained by the fact that job seekers are already allocated fairly optimally given the location of vacancies. Therefore, the low level of geographic mismatch in the US can be explained by a combination of a high enough willingness to apply to far away jobs and a fairly balanced distribution of job seekers and vacancies across zip codes.

Since job seekers do not only differ in location but also in occupation, one may wonder to what degree our estimate of geographic mismatch is sensitive to heterogeneity across occupations. Indeed, Şahin et al. (2014) show that mismatch across 2-digit SOC is much greater than mismatch across counties. However, the estimate from Şahin et al. (2014) assumes that job seekers only apply to jobs within their own occupation. We re-estimate geographic mismatch by taking into account job seekers' applications across commuting zones and across 2-digit SOC. We find that our estimates of geographic mismatch are remarkably robust to the occupation heterogeneity. Thus, even after taking account heterogeneity by occupation, we still find that geographic mismatch can account for about 5% of US unemployment.

Since we document how likely job seekers are to apply to jobs far away from home, our paper is related to the literature on geographic mobility in the US. This literature typically measures moves across states (Molloy et al., 2011). We complement this work by showing which locations job seekers consider during their job search, before making an actual move. Furthermore, we are able to analyze the distribution of locations considered by job seekers both within and across states. A strand of the literature on geographic mobility investigates whether people move to places with better economic conditions (Greenwood et al., 1986; Bound and Holzer, 2000; Wozniak, 2010). Our work is complementary to this literature since we can analyze whether workers direct their applications to areas with better economic conditions. We show that geographic mismatch based on applications within job seekers' area of residence is higher than geographic mismatch that takes into account job seekers' applications across areas. This implies that, when applying to vacancies outside of their area of residence, job seekers tend to choose vacancies in areas with lower job seeker to vacancy ratios.

Our work is also related to the literature that investigates the distance between the place of residence and the place of employment. This literature uses either matched employer-employee datasets or commuting surveys. Hellerstein et al. (2008) use a matched employer-employee dataset and find that 14% of workers work in the zip code where they live and 92% work in the CBSA where they live. Using the American Community Survey, McKenzie (2013) finds that 27% of workers travel for work outside of their county of residence in a typical week. We complement this research with evidence on the job search process. We find that vacancies job seekers apply to tend to be further away from their place of residence than jobs are to employed workers' place of residence.

Finally, the evidence we provide about the geography of job search is relevant to the literature on the impact evaluation of many types of local labor market shocks: shocks to labor demand such as a plant opening/closure, place-based policies, etc., or shocks to labor supply such as immigration, training, job search assistance programs, etc.[1]

The next section presents the data. In the third section, we present our theoretical framework. In the fourth section, we provide results about the geography of job search and the level of geographic mismatch. Section five provides robustness tests and extensions. Section six concludes.

2 Data

We use proprietary data provided by CareerBuilder.com, the largest US employment website. We merge three data sets extracted from CareerBuilder's database. The first one is a random sample of registered users whose accounts were active between April and June 2012. For each job seeker, we have the residence location at the zip code level and whether the job seeker is currently unemployed. In order for our results to be comparable with prior literature on job search, we restrict the data to unemployed users. After dropping those who do not reside in the US, who live in Alaska and Puerto Rico, and those whose location is unknown, we end up with a data set of 451,783 users.

The second data set is a sample of vacancies published on the website between April and June 2012, and therefore available to the job seekers to apply to. For each job, we know its location at the zip code level. The raw data set includes 1,116,314 vacancies. Removing non-consistent observations, duplicates and vacan-

[1]This issue is relevant to measure the impact of immigrants on natives' wages or employment rates (Card, 1990; Altonji and Card, 1991; Friedberg and Hunt, 1995; Borjas et al., 1996, 1997; DiNardo and Card, 2000; Card, 2001, 2005; Borjas, 2003; Ottaviano and Peri, 2006), the impact of local shocks on labor demand (Blanchard and Katz, 1992; Bound and Holzer, 2000; Notowidigdo, 2011), the impact of trade and FDI on labor market outcomes (Autor et al., 2013a,b), the equilibrium effects of active labor market policies (Davidson and Woodbury, 1993; Blundell et al., 2004; Gautier et al., 2012; Crépon et al., 2013; Ferracci et al., 2014), the heterogeneity of the negative duration dependence with local conditions (Kroft et al., 2013), spatial mismatch (Patacchini and Zenou, 2005; Hellerstein et al., 2008; Boustan and Margo, 2009; Åslund et al., 2010), or the impact of place-based policies (Neumark and Kolko, 2010; Ham et al., 2011; Gobillon et al., 2012; Givord et al., 2013; Busso et al., 2013).

cies not located in the US (or located in Alaska or Puerto Rico) leaves 1,111,353 vacancies. Removing observations without zip code information leaves 696,975 observations, which means that 37% of the sample is lost due to this restriction. Finally, the third data set connects the two previous data sets by showing which jobs each job seeker applied to: each observation corresponds to an application and is characterized by a job seeker ID and a vacancy ID. On average, job seekers sent around 12.8 applications and vacancies receive 15.8 applications.

We verified that the location of vacancies and job seekers in this data is representative of the location of vacancies and job seekers in the US in general. Across US regions, vacancies in our dataset are distributed very similarly to vacancies in the nationally representative Job Openings and Labor Turnover Survey (JOLTS) in April-June 2012: the correlation between the two distributions is 96%. Across US states, job seekers in this data are also distributed very similarly to the unemployed in the Current Population Survey in April-June 2012, with a correlation of 88%.

Furthermore, some background work (Marinescu and Wolthoff, 2012) was done to compare the industry distribution of job vacancies in CareerBuilder.com with the distribution in JOLTS. Compared to the distribution of vacancies across industries in JOLTS, some industries are overrepresented in CareerBuilder data, in particular information technology, finance and insurance, and real estate, rental and leasing. The most underrepresented industries are state and local government, accommodation and food services, other services, and construction. While the vacancies on CareerBuilder are not perfectly representative of the ones in the US economy as a whole, they form a substantial fraction of the market. Indeed, the number of vacancies on CareerBuilder.com represented 35% of the total number of vacancies in the US in January 2011 as counted in JOLTS (Marinescu and Wolthoff, 2012).

In terms of occupation (2-digit SOC codes), the distribution of job seekers' occupations in CareerBuilder data is very similar to the CPS (correlation of 0.71), and the distribution of vacancies' occupations in the CareerBuilder data is essentially identical to the distribution of vacancies in all online jobs (correlation of 0.95 with Help Wanted Online data). Overall, our data is therefore broadly representative of the US distribution of vacancies and job seekers.

3 Theory

We use a directed search model where workers choose where to send their applications based on the *location* of the vacancies. Specifically, vacancies closer to the job seekers' home yield higher utility. Our theoretical model is close to the one by Manning and Petrongolo (2011), where agents must choose a set of places to apply to, but we also borrow elements from Albrecht et al. (2006), Chade and Smith (2006), and Galenianos and Kircher (2009).

3.1 Environment and strategies

Each firm has one vacancy. All workers and all firms are identical, risk neutral, and they produce one unit of output when matched and zero otherwise. The utility of an employed worker is defined below, and an unmatched worker has a utility of zero. Workers observe all vacancies. Workers and vacancies are spread across S locations. Each location j has V_j vacancies and U_j unemployed workers. $i(u)$ and $j(v)$ denote the geographic units where unemployed worker u and vacancy v are located. A strategy of a worker is a set of $\bar{a}$ vacancies that he applies to. The timing of the game is the following.

1. Job seekers apply to vacancies: each job seeker sends $\bar{a}$ applications.
2. Firms gather the applications they receive: each application has a probability q to be valid in the sense that the applicant will produce positive output if he is hired. q is a scale parameter: it helps us calibrate the model by capturing the fact that the matching rate in the labor market is lower than what can be predicted on the basis of the number of applications that firms receive.
3. Firms can only make one offer. If a vacancy has more than one valid application, the firm picks randomly the job seeker to whom it makes an offer.
4. Offers are sent to job seekers.
5. Job seekers can only accept one job offer. If a job seeker has received more than one offer, he accepts the offer that generates the highest utility.
6. Matches are realized. If a firm's chosen applicant rejects the job offer then the firm remains unmatched.

3.2 Outcomes and equilibrium

The application of a worker u to a vacancy v provides the worker utility $w_{uv} = f(d_{i(u)j(v)})\varepsilon_{uv}$, the product of a deterministic decreasing function of the geographic

distance $d_{i(u)j(v)}$ between the job seeker and the vacancy, and an idiosyncratic term that is job-worker pair specific. ε is assumed to be uncorrelated across job seekers within vacancies or across vacancies within job seekers. This idiosyncratic term allows for workers in a given location i to have different preferences over vacancy locations j, and even across jobs within a given location.

We assume that the probability π_{uv} that a worker u gets an offer for vacancy v conditional on applying only depends on the location of the vacancy: $\pi_{uv} = \pi_{j(v)}$. This assumption, which is crucial for the tractability of the model, is also present in Manning and Petrongolo (2011). It is not a straightforward assumption as it entails in particular that employers do not treat candidates differently based on where these candidates live. Because we do not observe matches in our data, we are not able to assess the extent to which this assumption holds.

Let $\mathbf{v} = \{v_1, \dots v_{\bar{a}}\}$ be the $\bar{a}$-tuple of vacancies worker u applies to. We use the convention that $w_{uv_1} \geq w_{uv_2} \geq \dots w_{uv_{\bar{a}}}$. The expected utility associated with strategy $\mathbf{v}$ is:

$$\mathcal{U}(\mathbf{v}) = \pi_{j(v_1)} w_{uv_1} + \sum_{k=2}^{\bar{a}} \left[\prod_{\ell=1}^{k-1} (1 - \pi_{j(v_\ell)}) \right] \pi_{j(v_k)} w_{uv_k} \tag{1}$$

With probability π_{v_1}, the job seeker gets an offer from the highest utility vacancy. Whatever other offers he might get, he takes v_1 and his utility is w_{uv_1}. He only takes an offer from vacancy v_k if he does not get any offer from higher utility vacancies $v_{k'}, k' < k$, which happens with probability $\prod_{\ell=1}^{k-1}(1 - \pi_{j(v_\ell)})$.

Determining which strategy maximizes expected utility is complex: an algorithm such as the one described in Chade and Smith (2006) should be used. In the general case, it is *not* an optimal strategy to apply to the $\bar{a}$ highest expected utility jobs. Instead, workers should first apply to the highest expected utility job, and then gamble upwards by applying to jobs that have lower probability of yielding an offer but higher utility.

We assume that, from the point of view of the job seeker, the probability π_v and the reward w_{uv} associated with a vacancy v are not negatively correlated. If the correlation is negative but small enough or if the correlation is positive, there is little to no opportunity for gambling upwards and applying to the $\bar{a}$ vacancies with the highest expected utility $\pi_v w_{uv}$ yields an expected utility that is very close to following the optimal strategy as defined by the algorithm in Chade and Smith

(2006).[2]

We argue that the probability of getting an offer is not highly correlated with utility. Utility is the product of two terms: $f(d)$ is strictly decreasing with geographic distance and ε is an idiosyncratic shock. By assumption, ε is a random draw across vacancies, and, thus will not generate any correlation between π and w. What about the correlation between π and $f(d)$? If π exhibits no spatial autocorrelation, it will not be correlated with a function that is monotone with distance. Indeed, if π_d (d stands for distance) is a spatially determined random variable such that $Cov(\pi_d, \pi_{d-1}) = 0$ (no autocorrelation). Then, if we take a distance metric x_d (or a monotonic function of a distance metric), we have that $Cov(x_d, \pi_d) = 0$.

If π exhibits large spatial autocorrelation, it means that there are large areas with high π and others with low π. In this case, take a job seeker in an area with a high π, the closest vacancies, which are the ones that provide, on average, higher utility, are also the ones in which the probability to get an offer is higher. Therefore, for these job seekers living in areas with high π, it is definitely optimal to apply to the highest expected utility job. Conversely, in locations that have below average labor market tightness, job seekers do face a trade-off between applying to low offer rate (low π_j) but nearby jobs versus higher offer rate but far away jobs. For these job seekers, applying to the highest expected utility vacancies is not optimal: they should also "gamble upwards" by applying to vacancies close to home even though the probability of getting an offer is low. This issue will depend on the degree of autocorrelation of π. However, these job seekers are likely to be a small share of the population. Therefore, applying to the highest expected utility vacancies will be close to optimal for most job seekers, thus justifying our assumption.

We now simplify the problem by assuming that workers' optimal strategy is to the $\bar{a}$ vacancies with the highest expected utility. How can we represent this strategy in a tractable way? Let b be the lowest expected utility vacancy a worker is willing to apply to. The optimal strategy is to apply to all vacancies whose expected utility exceeds b. We allow the expected utility threshold to differ across job seekers' locations. Job seekers should apply to vacancies that are above the

[2]Manning and Petrongolo (2011) solve this issue by assuming that π is so enough that the probability to receive two offers or more is negligible. In this case, $\mathcal{U}(\mathbf{v}) = \sum_k \pi_{j(v_k)} w_{uv_k}$ so that the optimal strategy is too apply to the vacancies with the highest expected utility.

expected utility cutoff $b_{i(u)}$ that varies with the job seekers' location, i.e. apply if:

$$\pi_{j(v)} f(d_{i(u),j(v)}) \varepsilon_{uv} > b_{i(u)} \tag{2}$$

Assuming that ε has a Pareto distribution of parameter α, equation (2) above implies that the probability of application for a job seeker in i to a vacancy in j is given by:

$$p_{ij} = \left(\frac{\pi_j f(d_{ij})}{b_i} \right)^{\alpha} \tag{3}$$

The average number of applications sent by a job seeker in location i is equal to $\bar{a}$.

$$\bar{a} = \sum_{\ell} p_{i\ell} V_{\ell} \tag{4}$$

Combining equations (3) and (4), we can obtain an expression for b_i^{α}.

$$b_i^{\alpha} = \frac{\sum_{\ell} \pi_{\ell}^{\alpha} f^{\alpha}(d_{i\ell}) V_{\ell}}{\bar{a}}, \forall i$$

Replacing b_i^{α} by its expression in equation (3), we obtain the following expression for p_{ij}.

$$p_{ij} = \bar{a} \frac{\pi_j^{\alpha} f^{\alpha}(d_{ij})}{\sum_{\ell} \pi_{\ell}^{\alpha} f^{\alpha}(d_{i\ell}) V_{\ell}}, \forall i, j \tag{5}$$

The number of applications received by a vacancy located in j from job seekers located in i is distributed as a Poisson$(p_{ij}U_i)$. Summing applications coming from all origins, the distribution of the number of applications received by a vacancy in j is a Poisson(r_j), where $r_j = \sum_k p_{kj} U_k$ denotes the expected number of applications received by a vacancy in j. Remember that an application is assumed to be valid with probability q. The number of valid applications is thus distributed as $X_j \sim \text{Poisson}\,(qr_j)$.

From the point of view of a job seeker, the probability π_j that an application she is about to make is going to be successful is equal to q multiplied by the expectation of $1/(X_j + 1)$, where X_j is the expected number of applications made by other job seekers to the job. Since $X_j \sim \text{Poisson}\,(qr_j)$, π_j is given by:

$$\pi_j = q\mathcal{R}\,(qr_j) \tag{6}$$

where $\mathcal{R}(x) = [1 - \exp(-x)]/x$.

Replacing p_{kj} by its expression in (5), we obtain, for all j:

$$\pi_j = q\mathcal{R}\left(q\bar{a}\pi_j^{\alpha} \sum_k \frac{f^{\alpha}(d_{kj}) U_k}{\sum_{\ell} \pi_{\ell}^{\alpha} f^{\alpha}(d_{k\ell}) V_{\ell}} \right) \tag{7}$$

The number of offers received by a job seeker in i from vacancies in j is distributed as Poisson $(\pi_j p_{ij} V_j)$. The total number of offers received by this job seeker in i from all locations is thus distributed as $Y_i \sim$ Poisson $(\sum_\ell \pi_\ell p_{i\ell} V_\ell)$.

A job seeker will match if and only if he receives at least one offer. For a job seeker living in k, this occurs with probability $1 - \exp(\sum_\ell p_{k\ell}\pi_\ell V_\ell)$. Therefore, the number of matches is equal to:

$$M = \sum_k U_k \left[1 - \exp\left(\sum_\ell p_{k\ell}\pi_\ell V_\ell\right)\right]$$

This expression of the total number of hires involves $p_{ij}\pi_j$, the probability that a given job seeker in i applies to a given vacancy in j and that he gets an offer from this vacancy. The part between square brackets is the probability to have at least one offer.
Finally, using expression (5) to substitute p by its expression:

$$M = \sum_k U_k \left[1 - \exp\left(-\bar{a}\frac{\sum_\ell \pi_\ell^{1+\alpha} f^\alpha(d_{k\ell}) V_\ell}{\sum_\ell \pi_\ell^\alpha f^\alpha(d_{k\ell}) V_\ell}\right)\right] \tag{8}$$

We can check that this matching function exhibits constant returns to scale: if the number of job seekers and vacancies is multiplied by the same scalar factor in every location, the number of matches is also multiplied by this factor. Interestingly, we note that the probability of getting an offer π_j and the probability that a job seeker matches are unaffected by such a change in scale. We also check that the number of matches is equal to zero when either the number of vacancies or job seekers is equal to zero.

3.3 Solving the model

Knowing parameters $\bar{a}$, α, q, $f(.)$, as well as vectors U and V, how can we determine the total number of hires M? First, equation (7) defines a system of S equations where the S π_j are the unknowns, and $\bar{a}$, α, q, $f(.)$, U and V are the parameters. We do not have a general proof for the existence or unicity of a solution vector π for equation (7). However, we note that, in practice, the application $x \mapsto q\mathcal{R}\left(q x_j^\alpha \bar{a} \sum_k \frac{f^\alpha(d_{kj}) U_k}{\sum_\ell x_\ell^\alpha f^\alpha(d_{k\ell}) V_\ell}\right)$ is a clear contraction mapping that reaches an equilibrium very fast, for a large range of parameters. Trying several starting points always leads to the same solution, which argues in favor of a unique equilibrium. Manning and Petrongolo (2011) are able to prove formally the existence and unicity of the solution in a similar case, at the cost of imposing an exponential function form on the relationship between π and $\sum_k p_{kj} U_k$. Numerically, we

iterate equation (7) to obtain its fixed-point solution.

Second, once π is known, it is straightforward to find the total number of hires, using equation (8).

3.4 Particular case: distinct markets

So far, we have assumed that markets were interconnected: job seekers could apply to vacancies in their own unit as well as in others. Most of the existing literature make the simplifying assumption that markets are distinct, that is: (i) job seekers can only apply to vacancies within their own unit, (ii) job seekers are equally likely to apply to all vacancies within their own unit (there is no geographic distance within units).

We can adapt the previous to the subcase of distinct markets. Just suppose that $f(d_{ii}) = 1$ within the unit and $f(d_{ij}) = 0$ if $i \neq j$. The probability of applying to a vacancy within one's unit is $p_{jj} = \bar{a}/V_j$, the expected number of application received by a vacancy is $r_j = \bar{a}U_j/V_j$ and $\pi_j = q\mathcal{R}(qr_j)$. The total number of matches is then

$$M = \sum_k U_k \left[1 - \exp(-q\bar{a}\mathcal{R}(q\bar{a}U_k/V_k))\right] \tag{9}$$

3.5 Estimation

In this subsection, we turn to the issue of estimating the structural parameters of the model.

We take $\bar{a}$ as the average number of applications by job seekers as observed in the data. For the number of vacancies by unit V_j, we have to take into account the fact that we only observe a sample of job seekers and vacancies: the distribution of vacancies across locations should be correct, on average, but the global tightness may be wrong. Adjusting for tightness is important because the number of matches depends on labor market tightness. To obtain the total number of unemployed people in the US, we use the numbers reported by the Bureau of Labor Statistics based on the Current Population Survey. To obtain the total number of vacancies in the US, we use JOLTS. For each month of April to June 2012, we compute national tightness by dividing the total number of vacancies by the total number of unemployed job seekers. We take the average of these three measures of tightness as our measure of national labor market tightness. Keeping the geographic distribution as in our data, we then inflate the number of vacancies such

that the global tightness is equal to the national labor market tightness.

For function f, we start from equation (3). Another way to write the expression for the probability that a given individual in i applies to a given vacancy in j, p_{ij}, is:

$$p_{ij} = \tilde{\lambda}_i \tilde{\mu}_j f^{\alpha}(d_{ij})$$

We assume that $f(.)$ has an exponential form $f(d) = \exp(s(d))$, where $s(d)$ is a spline function. Therefore,

$$p_{ij} = \exp(\log \tilde{\lambda}_i + \log \tilde{\mu}_j + \alpha s(d_{ij})) = \exp(\lambda_i + \mu_j + \alpha s(d_{ij})) \tag{10}$$

Using the number of applications sent by job seekers in i to vacancies in j, we can identify the spline $\alpha s(.)$, not α and $s(.)$, which is enough to identify $f(.)^{\alpha}$. We estimate the coefficients of the spline function by a Poisson regression with two-way fixed-effects to account for the presence of job seeker location fixed effects λ_i and vacancy location fixed effects μ_j.

Now, parameters α and q still have be determined. In this version of the paper, we assume that $\alpha = 1$ and calibrate q based on the national job finding rate (using the CPS, this is the number of unemployment to employment transitions in a given month divided by the number of unemployed workers in the previous month). However, it is possible to find α in the following way. We can use the fact that, in our data, we observe the average number of applications received by vacancies in all units, and that we know the national job finding rate. For two values of α and q, and given the other parameters, we can solve the model and the average number of applications received by vacancies in j, $\hat{r}_j$, should be close to r_j:

$$r_j = \pi_j^{\alpha} \bar{a} \sum_k \frac{f^{\alpha}(d_{kj}) U_k}{\sum_{\ell} \pi_{\ell}^{\alpha} f^{\alpha}(d_{k\ell}) V_{\ell}}$$

Also, the target in terms of job finding rate provides us with a target $\hat{M}$ in terms of the total number of hires by multiplying by the number of job seekers: M should be close to $\hat{M}$. We can estimate α and q such that we minimize $\sum_j (r_j - \hat{r}_j)^2 / S$ and $\sum_j (M - \hat{M})^2$.

3.6 Mismatch index

3.6.1 No distaste for distance

Starting from equation (5), we examine the case in which job seekers have no distaste for distance, i.e. $f(d_{ij}) = 1, \forall i, j$.

$$p_{ij} = \bar{a}\frac{\pi_j^\alpha}{\sum_\ell \pi_\ell^\alpha V_\ell}, \forall i, j \tag{11}$$

In this case, p_{ij} does not depend on i. Let $\bar{U}$ and $\bar{V}$ be the total number of job seekers and vacancies in the economy. As for π, we have, for all j:

$$\begin{aligned}\pi_j &= q\mathcal{R}\left(q\bar{a}\pi_j^\alpha \sum_k \frac{U_k}{\sum_\ell \pi_\ell^\alpha V_\ell}\right) \\ &= q\mathcal{R}\left(\bar{U}\frac{q\bar{a}\pi_j^\alpha}{\sum_\ell \pi_\ell^\alpha V_\ell}\right)\end{aligned}$$

The only term that depends on j on the right-hand side is π_j itself. Therefore, solving for π_j is the same for any zip code j. Hence π is equal across zip codes in the case of no distaste for distance. Since π is equal across zip codes, we can rewrite π as a function of parameters, i.e.:

$$\pi = q\mathcal{R}\left(q\bar{a}\frac{\bar{U}}{\bar{V}}\right) \tag{12}$$

If $f(d_{ij}) = 1$, the total number of matches is:

$$M = \sum_k U_k \left[1 - \exp\left(-\bar{a}\frac{\sum_\ell \pi_\ell^{1+\alpha} V_\ell}{\sum_\ell \pi_\ell^\alpha V_\ell}\right)\right] \tag{13}$$

Since π is equal across zip codes, the total number of matches when there is no distaste for distance is:

$$M = U\left[1 - \exp\left(-\bar{a}\pi\right)\right]$$

Replacing π by its expression in equation (12),

$$M = \bar{U}\left[1 - \exp\left(-q\bar{a}\mathcal{R}\left(q\bar{a}\frac{\bar{U}}{\bar{V}}\right)\right)\right] \tag{14}$$

3.6.2 Optimal allocation

As seen above, in the general case, the total number of matches is given by:

$$M = \sum_k U_k \left[1 - \exp\left(-\bar{a}\frac{\sum_\ell \pi_\ell^{1+\alpha} f^\alpha(d_{k\ell}) V_\ell}{\sum_\ell \pi_\ell^\alpha f^\alpha(d_{k\ell}) V_\ell}\right)\right]$$

We want to find the allocation of job seekers U_k that maximizes M. In the expression above, only π_j (and of course U_k itself) depends on the distribution of U_k across zip codes k.

If the allocation of job seekers is such that π_j is equal to the same value π across zip codes, then the number of matches M is the same as in the case when there is no distaste for distance. We conjecture therefore that the allocation of job seekers at equalizes π_j across zip codes maximizes the number of hires. We detail this reasoning below.

If the allocation of job seekers equalizes π_j across zip codes, then we can write equation (7) as:

$$\pi = q\mathcal{R}\left(q\bar{a}\sum_k \frac{f^\alpha(d_{kj})U_k}{\sum_\ell f^\alpha(d_{k\ell})V_\ell}\right)$$

As there is bijective relationship between π_j and r_j, the average number of applications received by a vacancy, in this case, does not depend on its location. Thus, r_j will be equal to the total number of applications divided by the number of vacancies: $\sum_k p_{kj}U_k = \bar{a}\bar{U}/\bar{V}$. Using the expression of p_{ij} in equation (5), we can rewrite the previous equation as:

$$\pi = q\mathcal{R}\left(q\bar{a}\frac{\bar{U}}{\bar{V}}\right)$$

which is exactly the expression of π in the case in which there is no distaste for distance. If the allocation of job seekers is such that π_j is equal to the same value π across zip codes, the total number of matches is:

$$M^* = \bar{U}\left[1 - \exp\left(-q\bar{a}\mathcal{R}\left(q\bar{a}\frac{\bar{U}}{\bar{V}}\right)\right)\right]$$

Therefore, even with a positive distaste for distaste, allocating job seekers so that π (or r, the number of applications received by each vacancy) is equal across zip codes leads to the same number of matches than when there is no distaste for distance.

How can we find this allocation? Denote $\tilde{V}_k = \sum_\ell f^\alpha(d_{k\ell})V_\ell$, X the matrix of term $[f^\alpha(d_{ij}/\tilde{V}_i]_{ij}$ and b a vector of ones (of dimension the number of zip codes). The allocation of job seekers such that π is constant across zip codes is equal to:

$$U^* = \frac{\bar{U}}{b'X^{-1}b}X^{-1}b$$

We conjecture that M^*, obtained with the allocation U^* in the presence of distaste for distance and with any allocation when there is no distaste for distance, is the highest number of matches that one can obtain. This is consistent with the view that, starting from any allocation, decreasing the utility cost to apply far away will lead to an increase of the number of hires.

Our interconnected-markets mismatch index is then defined as one minus the ratio between the number of matches with the actual allocation of job seekers and the maximum number of matches:

$$\mathcal{M}_i = 1 - \sum_k \frac{U_k}{\bar{U}} \frac{1 - \exp\left(-\bar{a}\frac{\sum_\ell \pi_\ell^{1+\alpha} f^\alpha(d_{k\ell})V_\ell}{\sum_\ell \pi_\ell^{\alpha} f^\alpha(d_{k\ell})V_\ell}\right)}{1 - \exp\left(-q\bar{a}\mathcal{R}\left(q\bar{a}\frac{\bar{U}}{\bar{V}}\right)\right)} \tag{15}$$

4 Results

4.1 The Geography of Job Search

To understand the geography of job search, we must understand how important distance is in job seekers' application behavior. We use a Poisson regression to estimate the probability p_{ij} that a job seeker in zip code i applies to a vacancy in zip code j (see equation 10)above). Estimating p_{ij} helps us pin down the distaste for distance function f. In turn, f will be used to calculate the degree of geographic mismatch.

Specifically, the model we estimate is the following. The number of applications from job seekers in zip i to vacancies in j is modeled as a Poisson with parameter μ_{ij}:

$$\mu_{ij} = U_i V_j \exp[\alpha_i + \lambda_j + s(d_{ij})] \tag{16}$$

where α_i and λ_j are fixed effects for job seekers' and vacancies' zip codes respectively. U_i and V_j the number of job seekers in i and vacancies in j, and $s(.)$ a spline function whose parameters are estimated. The estimated spline function that captures how far away job seekers apply is displayed graphically in Figure 1.

Overall, applications decrease with distance, as shown in Figure 1. One potential concern is that job seekers in different locations are quite different in their distaste for distance. Similarly, vacancies in some locations may be more attractive to all job seekers, which could bias our estimates of the distaste for distance. Reassuringly, the estimate of the spline is not sensitive to job seeker zip code and vacancy

zip code fixed effects (Figure 1). This suggests that our estimate of the distaste for distance is robust to unobserved heterogeneity across job seekers in different zip codes, and across vacancies in different zip codes.

[Figure 1 about here.]

For very short distances, the probability of application increases with distance (Figure 1). This is likely due to the fact that small geographic areas are specialized in either residential or commercial/industrial activities,. This specialization could arise organically, or may be due to zoning laws. If job seekers are more likely to live in specialized residential areas, then they are less likely to find suitable jobs in their home zip codes. Instead, they will find more suitable jobs in nearby areas that specialize in commercial/industrial activities. Area specialization may thus explain why job seekers are slightly more likely to apply to jobs in nearby zip codes than in their zip code of residence.

The probability of application decreases drastically up to 50 miles of distance (Figure 1). This is probably due to the fact that job seekers would have to move in order to take jobs beyond 50 miles. While the distaste for distance at short distances probably reflects distaste for longer commutes, distaste for distance at longer distances is more and more reflective of a distaste for moving.

The magnitude of the distaste for distance we estimate here implies that American job seekers are much more willing to apply to vacancies far away from home than the British job seekers studied by Manning and Petrongolo (2011). Some of the difference in these estimates may be explained by the fact that the job seekers studied Manning and Petrongolo (2011) have somewhat lower skill compared to the average job seeker in the UK, while our sample is fairly representative of the average US job seeker in terms of occupation. Furthermore, Manning and Petrongolo (2011) estimate the distaste for distance based on a matching model rather than on direct observation of job seekers' application behavior, which could introduce some noise in the estimates. However, the difference in the two estimates is large enough that there may also be a difference between the two countries in the distaste for distance for a given skill level.

Overall, we find that job seekers are less likely to apply to vacancies further away from their zip code of residence, and these results are robust to controls for job seeker and vacancy zip code fixed effects. What is yet to be determined is whether

job seekers' preference for jobs close to home is high enough to generate substantial geographic mismatch. This is the topic of the next section.

4.2 Mismatch Unemployment

4.2.1 Mismatch with distinct units

Existing mismatch indices implicitly assume that job seekers are equally likely to match with any job within their location, and will never match with a job outside their home location. These mismatch indices are then used to measure differences between the geographic distribution of jobs and job seekers. One of the most commonly used measures of mismatch is the following:

$$\mathcal{M} = 1 - \sum_i \left(\frac{v_i}{\sum_i v_i} \right)^{\gamma} \left(\frac{u_i}{\sum_i u_i} \right)^{1-\gamma} \tag{17}$$

where v_i, u_i are the number of vacancies and unemployed workers in geographic area (labor market) i respectively.

$\mathcal{M}$ has been introduced by Nickell (1982) and Jackman and Roper (1987) to measure the share of unemployment that is due to mismatch. [3] The idea is to rely on a simple search-and-matching model of unemployment (Pissarides, 2000): in each market, the number of hires is assumed to be a constant-return-to-scale Cobb-Douglas function of the number of vacancies and job seekers, of parameter γ, e.g. proportional to $v^{\gamma}u^{1-\gamma}$. In what follows, we take $\gamma = .5$, following Şahin et al. (2014). Following Nickell (1982), Şahin et al. (2014) show that $1 - \mathcal{M}$ is the ratio of the actual number of hires and the optimal number of hires that a planner can obtain by allocating job seekers across markets. Therefore, the Cobb-Douglas mismatch index $\mathcal{M}$ represents the percentage shortfall in hires obtained with the actual allocation of job seekers relative to the optimal allocation of job seekers.

In order to calculate this mismatch index, one must choose a geographic unit for the location of job seekers, such as the MSA. The choice of the geographic unit of observation is subject to a trade-off. Working with too broad areas is likely to create a downward bias on the index. If one considers there is only one area (say the whole United States), all applications from job seekers residing in this area are obviously sent within the same area. In this case, the index will obviously

[3] See Jackman et al. (1989); Dickens (2010); Lazear and Spletzer (2012) for a dissimilarity index, which provides a measure of the proportion of the unemployed who are in the "wrong" market. Using this other measure yields very similar qualitative results for Figure 2.

be equal zero but will understate the actual geographic mismatch. Conversely, if we use zip codes as the unit of observation, we have the opposite problem. Many applications are directed to vacancies that are not located in the area where the job seeker resides, and we run the risk of overestimating geographic mismatch. As demonstrated by Şahin et al. (2014), choosing a larger area to define the location of job seekers mechanically yields higher mismatch according to the Cobb-Douglas index $\mathcal{M}$.

Figure 2 shows how mismatch varies with the size of the geographic area where job seekers are assumed to look for jobs. Under the assumptions underlying the index $\mathcal{M}$, we find that 1.6% of hires are lost due to the misallocation of job seekers when job seekers' search area is defined as the state. If we define the job seekers' search area as the MSA or the commuting zone (CZ), mismatch is about 2.5%. When search areas are counties, this figure doubles, to 4.9%. At the zip-code level, the fraction of hires lost due to misallocation of job seekers is 22.9%, a very large figure. The magnitude of geographic mismatch therefore strongly depends on the size of the geographic area where job seekers are assumed to look for jobs, with smaller areas yielding much larger values for the Cobb-Douglas mismatch index $\mathcal{M}$.

[Figure 2 about here.]

4.2.2 Mismatch with distinct markets using our model

Our mismatch index $\mathcal{M}_i$ differs from the Cobb-Douglas mismatch index $\mathcal{M}$ in two ways. First, $\mathcal{M}_i$ uses a different matching function. Second, $\mathcal{M}_i$ allows for applications across geographic units. In order to understand differences in measured mismatch when using $\mathcal{M}_i$ versus $\mathcal{M}$, we first report mismatch results for a version of our mismatch index that prevents job seekers from applying across geographic areas. For this version of the mismatch index, we use the mismatch index $\mathcal{M}_i$, and take the distaste for distance f to be an identity matrix, so that job seekers get no utility from applying to jobs that are not in their home area. We call this index $\mathcal{M}_d$ the "distinct markets mismatch index", referring to the fact that labor markets are assumed not to interact across geographic areas.

The orders of magnitude of the distinct markets mismatch index $\mathcal{M}_d$ are similar to those obtained with the Cobb-Douglas mismatch index $\mathcal{M}$ (compare Figure 3 and Figure 2). Just like $\mathcal{M}$, $\mathcal{M}_d$ is sensitive to the size of the area where job seekers are assumed to look for jobs, with larger areas yielding smaller levels of

mismatch. Therefore, while our matching function is not a Cobb-Douglas, the distinct markets mismatch index $\mathcal{M}_d$ yields very similar results to those arising from the Cobb-Douglas mismatch index $\mathcal{M}$.

[Figure 3 about here.]

4.2.3 Comparing interconnected and distinct markets using our model

Mismatch is most accurately captured by our mismatch index $\mathcal{M}_i$ (equation 15 above) at the zip code level, because it allows for detailed geography and for interconnected labor markets. Using our preferred measure of mismatch, we find that geographic mismatch is very small: 4.97% of hires are lost due to the misallocation of job seekers when job seekers' location is defined at the zip code level (Figure 4).

Allowing for applications across boundaries is important to determine the magnitude of mismatch. Indeed, the interconnected markets mismatch index $\mathcal{M}_i$ is a lot less sensitive to the definition of the labor market than the distinct markets mismatch index $\mathcal{M}_d$. $\mathcal{M}_d$ varies between 31.1% if the labor market is assumed to be the zip code of residence of the job seeker and 3.66% if the labor market is a whole state. By contrast, $\mathcal{M}_i$ goes from 4.97% to 1.29% when going from the zip code level to the state level.

Regardless of the definition of a labor market, the interconnected markets mismatch index $\mathcal{M}_i$ yields lower mismatch that the distinct markets mismatch index $\mathcal{M}_d$ (Figure 4). The percent difference between $\mathcal{M}_d$ and $\mathcal{M}_i$ is largest when labor markets are defined by zip codes, but it is also quite large when labor markets are defined by counties or states. Because $\mathcal{M}_i$ allows for applications across labor markets, $\mathcal{M}_i$ is systematically lower than $\mathcal{M}_d$.

[Figure 4 about here.]

5 Robustness and Extensions

5.1 Mismatch for various distastes for distance

To what extent can low mismatch be attributed to our specific estimate of the distaste for distance? Would mismatch increase a lot if distaste for distance were greater? There are at least two reasons why this is an interesting question. First,

people who do not use Internet for job search may have higher distaste for distance. Second, British job seekers studied in Manning and Petrongolo (2011) do in fact exhibit higher distaste for distance. To test the sensitivity of our results, we re-calculate the number of matches M (see equation 8) at the zip code level and for different distastes for distance f, ranging from infinity (isolated zip codes) to zero (single market for the whole country). Mismatch $\mathcal{M}_i$ is equal to the number of matches with no distaste for distance minus the number of matches M, which varies with distaste for distance. Therefore, variations in M are directly informative about changes in the geographic mismatch $\mathcal{M}_i$.

Geographic mismatch increases with distaste for distance, but this increase is very small, even for moderate changes in the distaste for distance. To demonstrate this, we plot the number of matches M realized with various distastes for distance given the actual allocation of job seekers. We normalize the results by the number of matches M obtained with the distaste for distance for American job seekers that we have estimated above. The blue bars in Figure 5 show the results from this exercise. We can see that a decrease in the distaste for distance barely increases the number of matches, which is another way of saying that there is little geographic mismatch. Conversely, if we increase the distaste for distance, we also don't see a large drop in the number of matches, at least for moderate increases in the distaste for distance. For example, if we move from the American distaste for distance to the British distaste for distance (as estimated in Manning and Petrongolo (2011), we find that matches would decrease by less than 5%.

On the other hand, if job seekers were allocated randomly across zip codes, greater distaste for distance would have a larger impact on geographic mismatch. The red bars in Figure 5 show the results from this exercise. In this case, going from the American distaste for distance to to the British distaste for distance, matches decline by more than 20%. Going to isolated zip codes decreases the number of matches by almost 60%. These results suggest that distaste for distance did not matter much when using the actual allocation of job seekers because job seekers already live pretty close to vacancies on average.

Based on this analysis, geographic mismatch is low because distaste for distance is low enough, and job seekers are already fairly close to vacancies. In a dynamic framework, low distaste for distance can explain why job seekers are relatively well allocated across space. Indeed, if distaste for distance is low enough, then, over time, job seekers relocate to follow vacancies, so that, at any given point in

time, job seekers live close to vacancies on average.

[Figure 5 about here.]

5.2 Geography interacted with occupations

Mismatch unemployment may be the result of a different geographic distribution of job seekers and job vacancies, but it might also result from a different distribution of job seekers and job vacancies across occupations. Moreover, the occupation and spatial dimensions may interact to further increase mismatch.

In this subsection, we carry a simple exercise to assess whether occupations play an important role in mismatch. Instead of defining geography-based labor markets as we have done so far, we define a labor market as a location *and* an occupation. For vacancies, the occupation is directly obtained from the listing. For job seekers, we use the occupation of their last job. In order to keep computations tractable, we build labor markets as the intersection of SOC-2 occupations and Commuting Zones, obtaining around 10,000 CZ*SOC-2 labor markets. For the same reasons that we do not assume that job seekers only apply in their home location, we do not assume that job seekers whose last job was in a given occupation will restrict their applications to the same occupation. Therefore, we estimate a model similar to the one described in equation (16), and we add a dummy for applying to an occupation that is different from the one held in the last job. The expected probability for job seeker in labor market i to apply for a job in labor market j is:

$$\mu_{ij} = U_i V_j \exp[\alpha_i + \lambda_j + s(d_{ij}) + \alpha_o \mathbb{1}_{\{o_i \neq o_j\}}]$$

where d_{ij} is the geographic distance between the centroids of the CZs corresponding to i and j, and o_i is the SOC-2 occupation of job seekers in i. Note that we include, as before, job seeker labor market fixed effects α_i, and therefore we also control for fixed effects by occupation.

The estimation of this model pins down a new function f to be used in the mismatch index $\mathcal{M}_i$. The third column in Figure 6 presents the result of this exercise. We find that, indeed, the mismatch index corresponding to the CZ*SOC-2 labor markets is higher than the one corresponding to the CZ labor markets, but by very little. Qualitatively, our results are unchanged: mismatch unemployment remains around 5%. If the data and computation capacities allowed to refine the exercise, building for instance zip*SOC-2 labor markets and estimating more flexible models, we conjecture that the resulting index would be higher than 5%, but

only modestly so. Overall, we conclude that our measure of geographic mismatch is robust to accounting for heterogeneity in SOC-2 occupations.

[Figure 6 about here.]

6 Conclusion

In this paper, we have documented how far job seekers are willing to apply to jobs and, based on this evidence, we have measured the degree of geographic mismatch. Our measure of geographic mismatch is based on a search and matching model of the labor market in which job seekers choose where to send their applications. Quantitatively, we find that US aggregate unemployment would be reduced by at most 5% if job seekers were reallocated so as to maximize hires. Therefore, geographic mismatch is a negligible driver of US unemployment.

We further seek to understand why geographic mismatch is so low. First, we show that our results are not driven by our neglecting heterogeneity by occupation: indeed, after accounting for heterogeneity by occupation, geographic mismatch remains essentially the same. Second, we find that low mismatch can be explained by a job seekers' high enough willingness to apply far away from home combined with a geographic distribution of job seekers that mimics relatively closely the distribution of vacancies.

Our results suggest that policies that attempt to combat geographic mismatch will have a limited effect on aggregate unemployment. There are two reasons why this is the case. First, we show that mismatch is just not a big problem at the macro level. While the prior macro literature could not pin down the exact level of geographic mismatch, the results were suggestive of relatively low levels of mismatch (Şahin et al., 2014). Based on better evidence on the geography of job search, we conclude that geographic mismatch is indeed not a major cause of unemployment at the macro level. Second, place-based policies will tend to be ineffective even at the micro level, i.e. in locations that are especially job poor. Prior literature investigating the impact of the geography of job search on place-based policies in the UK (Manning and Petrongolo, 2011) already suggested that such policies are rather ineffective. Indeed, job seekers from other areas will apply to newly created jobs in target areas, so the positive employment effect for target area residents will be muted. Because we find that, in the US, job seekers apply much further away from their home area than in the UK, our results suggest that the impact

of place-based policies in contemporary US will be even more muted.

Overall, our results show that geographic mismatch cannot account for the high level of unemployment in 2012 relative to the beginning of the Great Recession. Instead, other causes must have been at play: to manage aggregate unemployment effectively, policy makers should therefore use policies that are not geography dependent, such as monetary policy.

References

Albrecht, J. W., P. A. Gautier, and S. B. Vroman (2006): "Equilibrium Directed Search with Multiple Applications," *Review of Economic Studies*, 73, 869–891.

Altonji, J. G. and D. Card (1991): "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in *Immigration, Trade and the Labor Market*, National Bureau of Economic Research, Inc, NBER Chapters, 201–234.

Åslund, O., J. Østh, and Y. Zenou (2010): "How Important Is Access to Jobs? Old Question - Improved Answer," *Journal of Economic Geography*, 10, 389–422.

Autor, D. H., D. Dorn, and G. H. Hanson (2013a): "The Geography of Trade and Technology Shocks in the United States," *American Economic Review*, 103, 220–25.

——— (2013b): "Untangling Trade and Technology: Evidence from Local Labor Markets," NBER Working Papers 18938, National Bureau of Economic Research, Inc.

Blanchard, O. J. and L. F. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 23, 1–76.

Blundell, R., M. C. Dias, C. Meghir, and J. V. Reenen (2004): "Evaluating the Employment Impact of a Mandatory Job Search Program," *Journal of the European Economic Association*, 2, 569–606.

Borjas, G. J. (2003): "The Labor Demand Curve Is Downward Sloping: Reexamining The Impact Of Immigration On The Labor Market," *The Quarterly Journal of Economics*, 118, 1335–1374.

Borjas, G. J., R. B. Freeman, and L. F. Katz (1996): "Searching for the Effect of Immigration on the Labor Market," *American Economic Review*, 86, 246–51.

——— (1997): "How Much Do Immigration and Trade Affect Labor Market Outcomes?" *Brookings Papers on Economic Activity*, 28, 1–90.

Bound, J. and H. J. Holzer (2000): "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s," *Journal of Labor Economics*, 18, 20–54.

Boustan, L. P. and R. A. Margo (2009): "Race, Segregation, and Postal Employment: New Evidence on Spatial Mismatch," *Journal of Urban Economics*, 65, 11–23.

Busso, M., J. Gregory, and P. Kline (2013): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103, 897–947.

Card, D. (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43, 245–257.

——— (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19, 22–64.

——— (2005): "Is the New Immigration Really so Bad?" *Economic Journal*, 115, F300–F323.

Chade, H. and L. Smith (2006): "Simultaneous Search," *Econometrica*, 74, 1293–1307.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013): "Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment," *The Quarterly Journal of Economics*, 128, 531–580.

Davidson, C. and S. A. Woodbury (1993): "The Displacement Effect of Reemployment Bonus Programs," *Journal of Labor Economics*, 11, 575–605.

Dickens, W. T. (2010): "Has the Recession Increased the NAIRU?" Mimeo Northwestern and the Brookings Institution.

DiNardo, J. and D. Card (2000): "Do Immigrant Inflows Lead to Native Outflows?" *American Economic Review*, 90, 360–367.

Ferracci, M., G. Jolivet, and G. J. van den Berg (2014): "Evidence of Treatment Spillovers Within Markets," Forthcoming in the *Review of Economics and Statistics*.

Friedberg, R. M. and J. Hunt (1995): "The Impact of Immigrants on Host Country Wages, Employment and Growth," *Journal of Economic Perspectives*, 9, 23–44.

Galenianos, M. and P. Kircher (2009): "Directed Search with Multiple Job Applications," *Journal of Economic Theory*, 144, 445–471.

Gautier, P., P. Muller, B. van der Klaauw, M. Rosholm, and M. Svarer (2012): "Estimating Equilibrium Effects of Job Search Assistance," Tinbergen Institute Discussion Papers 12-071/3, Tinbergen Institute.

Givord, P., R. Rathelot, and P. Sillard (2013): "Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines program," *Regional Science and Urban Economics*, 43, 151–163.

Gobillon, L., T. Magnac, and H. Selod (2012): "Do Unemployed Workers Benefit from Enterprise Zones? The French Experience," *Journal of Public Economics*, 96, 881–892.

Greenwood, M. J., G. L. Hunt, and J. M. McDowell (1986): "Migration and Employment Change: Empirical Evidence on the Spatial and Temporal Dimensions of the Linkage," *Journal of Regional Science*, 26, 223–234.

Ham, J. C., C. Swenson, A. Imrohoroglu, and H. Song (2011): "Government Programs can Improve Local Labor Markets: Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Community," *Journal of Public Economics*, 95, 779 – 797.

Hellerstein, J. K., D. Neumark, and M. McInerney (2008): "Spatial mismatch or racial mismatch?" *Journal of Urban Economics*, 64, 464–479.

Jackman, R., R. Layard, and C. Pissarides (1989): "On Vacancies," *Oxford Bulletin of Economics and Statistics*, 51, 377–394.

Jackman, R. and S. Roper (1987): "Structural Unemployment," *Oxford Bulletin of Economics and Statistics*, 49, 9–36.

Kroft, K., F. Lange, and M. J. Notowidigdo (2013): "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*," *The Quarterly Journal of Economics*, 128, 1123–1167.

Lazear, E. P. and J. R. Spletzer (2012): "Hiring, Churn and the Business Cycle," *American Economic Review: Papers and Proceedings*, 102, 575–579.

Manning, A. and B. Petrongolo (2011): "How Local are Labor Markets? Evidence from a Spatial Job Search Model," CEPR Discussion Papers 8686, C.E.P.R. Discussion Papers.

Marinescu, I. E. and R. P. Wolthoff (2012): "Wages, Applications, and Skills," .

McKenzie, B. (2013): "County-to-County Commuting Flows: 2006-10," U.S. Census Bureau.

Molloy, R., C. L. Smith, and A. Wozniak (2011): "Internal Migration in the United States," *Journal of Economic Perspectives*, 25, 173–196.

Neumark, D. and J. Kolko (2010): "Do Enterprise Zones Create Jobs? Evidence from California's Enterprise Zone Program," *Journal of Urban Economics*.

Nickell, S. (1982): "The Determinants of Equilibrium Unemployment in Britain," *The Economic Journal*, 92, 555–575.

Notowidigdo, M. J. (2011): "The Incidence of Local Labor Demand Shocks," NBER Working Papers 17167, National Bureau of Economic Research, Inc.

Ottaviano, G. P. and G. Peri (2006): "The economic value of cultural diversity: evidence from US cities," *Journal of Economic Geography*, 6, 9–44.

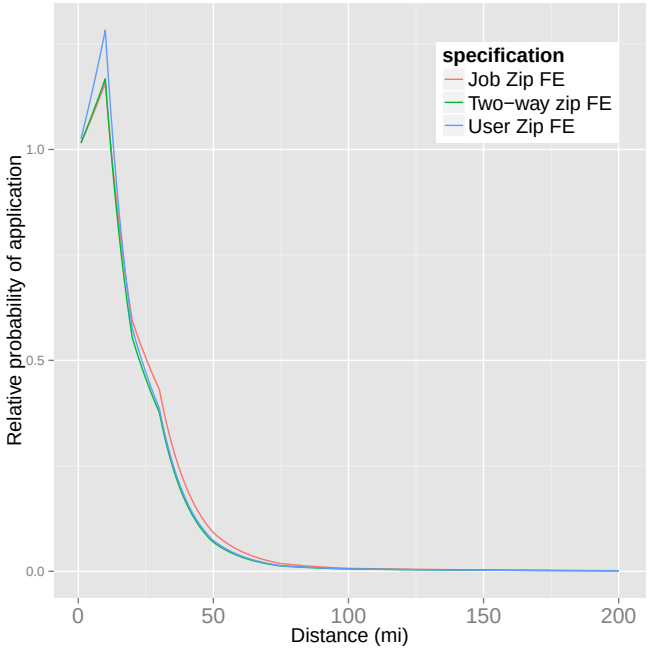
Patacchini, E. and Y. Zenou (2005): "Spatial mismatch, transport mode and search decisions in England," *Journal of Urban Economics*, 58, 62–90.

Pissarides, C. A. (2000): *Equilibrium Unemployment Theory, 2nd Edition*, MIT Press Books, The MIT Press.

Şahin, A., J. Song, G. Topa, and G. L. Violante (2014): "Mismatch Unemployment," *American Economic Review*, 104, 3529–64.

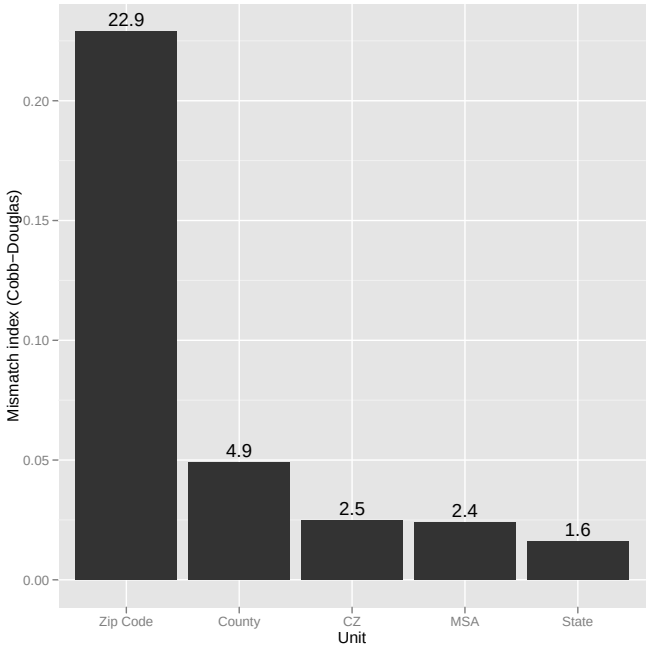
Wozniak, A. (2010): "Are College Graduates More Responsive to Distant Labor Market Opportunities?" *Journal of Human Resources*, 45, 944–970.

Figure 1: Relative probability of application as a function of geographic distance: predictions from Poisson model with or without fixed effects



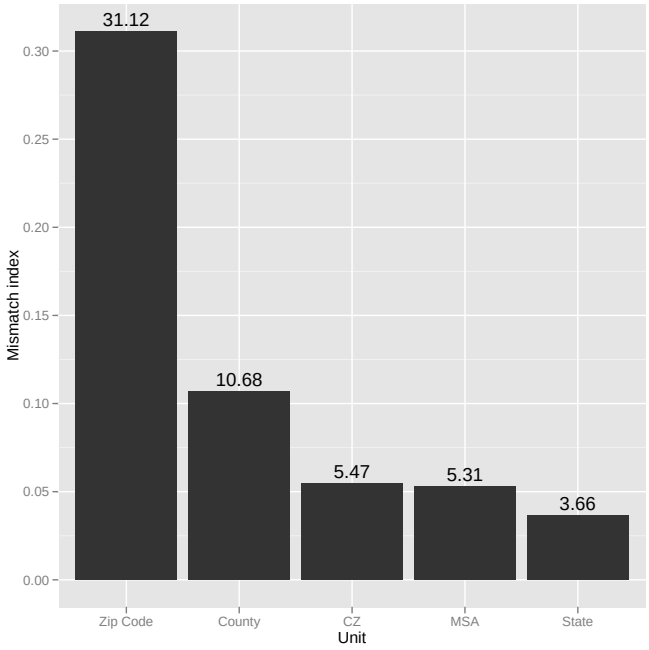
Source: CareerBuilder database.

Figure 2: Mismatch unemployment: distinct markets and Cobb-Douglas matching function



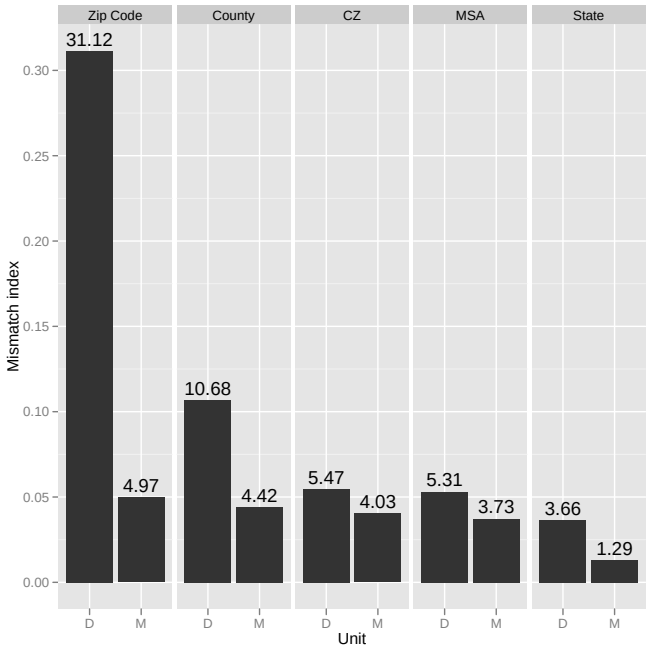
Source: CareerBuilder database.

Figure 3: Mismatch unemployment: distinct markets with the matching function from our model



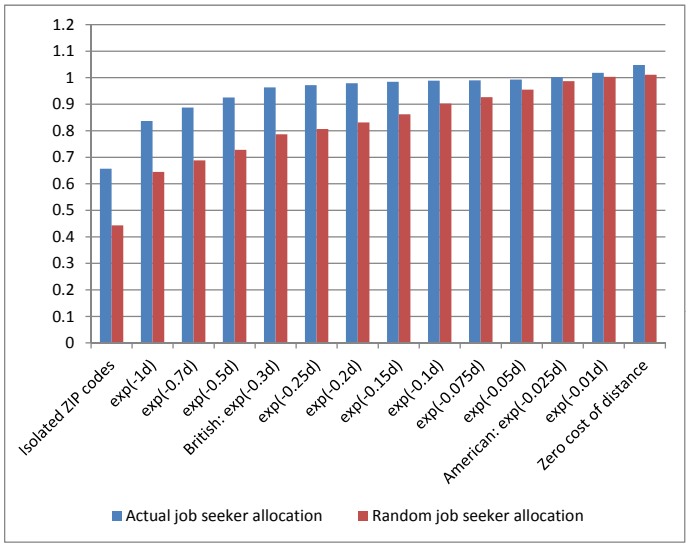
Source: CareerBuilder database.

Figure 4: Mismatch unemployment: interconnected and distinct markets



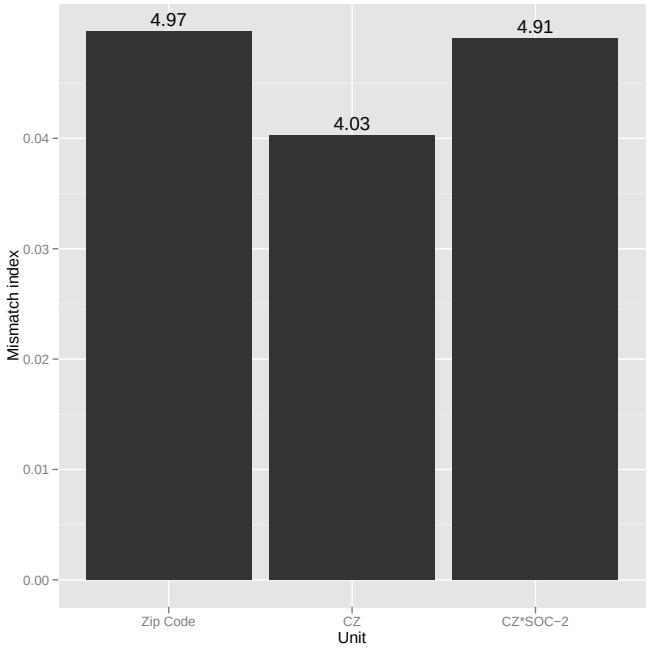
Source: CareerBuilder database.

Figure 5: Robustness to various distaste-for-distance parameters



Source: CareerBuilder database.

Figure 6: Mismatch unemployment: interconnected geographic units and occupations



Source: CareerBuilder database.