The Geography of Job Search and Mismatch Unemployment[†]

Ioana Marinescu[‡]
Roland Rathelot[§]

November 8, 2013

Job Market Paper

Abstract

Can we reduce unemployment by moving job seekers to areas with better job opportunities? To answer this question, we need data on the distance between job seekers and the jobs they apply to. Using novel data from the popular website CareerBuilder.com, we quantify how application probability declines with distance from the job seekers' zip code of residence. 82% of applications are sent to jobs within the same city (Core-Based Statistical Area, CBSA), but only 46% are sent to jobs within the same county. We build a simple search-and-matching framework in which job seekers have a distaste for distance and use our data to estimate its parameters. Using our model, we find that US unemployment could be reduced by up to 3% by reallocating job seekers across zip codes. This magnitude of mismatch is similar to what we find using data aggregated at the CBSA level. Our evidence suggests that the CBSA is an acceptable definition of a local labor market.

Keywords: local labor markets, job search, misallocation.
JEL: J21, J61, J62, J64.

[†]We would like to thank Pierre Cahuc, Xavier D'Haultfoeuille, François Fontaine, Florence Goffette-Nagot, Thomas Le Barbanchon, Etienne Lehmann, Francis Kramarz, Manasa Patnam, Julien Prat as well as participants to seminars in CREST, Cachan/Paris-Sud, GATE (Lyon), Mannheim, Uppsala and Louvain-la-Neuve 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
[§]CREST. 18 Bld G. Péri 92245 Malakoff. +33141176036. roland.rathelot@ensae.fr. Corresponding author.

1 Introduction

Are vacancies and job seekers distributed across space in a way that minimizes aggregate unemployment? How can we measure the impact of localized labor market policies on surrounding areas? Investigating these questions requires data on the geographic distribution of job seekers and the vacancies they apply to. However, until now, systematic data on the geography of job search has been missing. In this paper, we provide novel evidence on the geography of job search in the United States. We discuss how this evidence can inform the evaluation of localized labor market shocks. Finally, we demonstrate how the geography of job search affects the measurement of mismatch unemployment.

Understanding the geography of job search is important in many contexts. For instance, the literature on the impact of localized labor market shocks typically assumes that these shocks are best measured at some specific level of geographic aggregation, such as the state (e.g. for the impact of immigration, Borjas, Freeman, and Katz, 1997), the city (e.g., for the impact of immigration, Card, 2001), or the county (e.g. for the impact of a local economic development program, Kline and Moretti, 2013). The choice of the level of geographic aggregation matters for the estimation. For example, if the units of analysis are small compared to job seekers' search radius (e.g. zip codes), estimates are likely to be affected by spillover effects. If observation units are large (e.g. states), spillover effects may be limited. However, the estimated treatment effect will be diluted as the shock may be too far away from some of the job seekers considered as treated. Evidence on the geography of job search can therefore strengthen our ability to choose the geographic unit of analysis when evaluating the impact of many types of localized labor market shocks.

Understanding the geography of job search also allows us to better measure the degree of geographic mismatch, i.e. the degree to which we could reduce unemployment by reallocating job seekers across space. The basic idea of geographic mismatch is that there are too many job seekers relative to vacancies in some places and too few job seekers in some other places. For example, we may be able to lower the aggregate unemployment rate by moving some unemployed job seekers from states heavily hit by the recent housing market crash to other less affected states. In the existing literature, unemployment is minimized when the job seekers to vacancy ratio is equal across states. Unfortunately, the definition of the unit of analysis mechanically affects measured geographic mismatch: higher levels of aggregation, such as the state, yield lower levels of measured mismatch than lower levels of aggregation, such as the county.

To justify the choice of a level of aggregation for the data used in measuring mismatch with existing indices, one needs to make some stark assumptions about the geography of job search. These assumptions are as follows: (i) job seekers are equally likely to apply to all jobs in their area of residence (e.g. the county), (ii) job seekers do not apply to jobs located outside of their area of residence. In order to understand how these stark assumptions affect measured mismatch, we need to provide evidence about the degree to which they are satisfied for different definitions of the job seekers' area of residence, such as the state or the county.

In this paper, we use data from the largest online job search website in the US, CareerBuilder.com. The location of job seekers and vacancies is provided at the zip code level.[1] The data pertains to 2012. We start by documenting how distance affects the application behavior of job seekers. We find that 46% of job applications are made to jobs within the same county as the job seeker and 82% to jobs within the same Core-Based Statistical Area (CBSA). In terms of geographic distance, 90% of the applications are made to jobs less than 100km away. The application probability strongly declines with distance even after accounting for composition effects at both origin and destination. A vacancy located 5km away is 20% less likely to be applied to than a vacancy located in the same zip code as the job seeker. For the purpose of understanding the impact of localized labor-market shocks, this suggests limited spillovers across CBSAs but large spillovers across counties in the same CBSA. Furthermore, treatment intensity likely decreases with distance even within the confines of a CBSA.

Using our data, we compute the degree of geographic mismatch based on the same mismatch index as Şahin, Song, Topa, and Violante (2012), hereafter SSTV. Like them, we find that the degree of mismatch depends on the level of aggregation of the data. For example, we find that 5% of hires are lost to mismatch when using county-level data, while this number is nearly half as large when using CBSA-level data, with 2.7% of hires lost to mismatch. We then build a new mismatch index that takes into account the geography of job search at the zip code level. The index is based on a theoretical search-and-matching model in which, for job seekers, the utility associated with hiring is a decreasing function of the geographic distance between the job and their home. We use our data to estimate the structural parameters of the model. We find that 3.2 % of hires are lost due to mismatch. This result implies that geographic mismatch accounts for 3% of US unemployment. This amount of geographic mismatch is about the same as what we find using the mismatch index from SSTV and our data aggregated at the CBSA level. Our analysis of mismatch yields two main conclusions. First, the mismatch index used by SSTV gives the most accurate measurement when used with data aggregated at the CBSA level rather than at the state or county level. Second, using a measure of geographic mismatch that takes into account the geography of job search and is robust to issues of data aggregation, we conclude that geographic mismatch alone is not a key factor in US unemployment in 2012.

Our paper is most closely related to two recent working papers: Manning and Petrongolo (2011) and SSTV. Manning and Petrongolo (2011) use UK data on local stocks of job seekers and vacancies to estimate a search-and-matching model. Using their model, they estimate job seekers' preference for proximity and simulate the impact of several localized labor-market policies. In contrast to them, we directly observe the location of the vacancies that job seekers apply to, which allows us to document the geography of job search without having to rely on the assumptions of a structural model. Our results indicate that American job seekers are far more willing to apply to distant jobs

[1] We document that zip codes are small compared to the typical range of job search, so that, in this context, zip codes are almost as good as geographic coordinates.

than their British counterparts.

As already mentioned, our paper also builds on the contributions by Şahin, Song, Topa, and Violante (2011) and SSTV. Given that the specific magnitude of mismatch is sensitive to the level of aggregation of the data, SSTV focus their work on *changes* in mismatch during the Great Recession. In this paper, we build a new mismatch index that takes into account the geography of job search at the zip code level and is therefore arguably robust to issues of geographic aggregation. This allows us to pin down the *level* of mismatch in 2012: geographic mismatch accounts for 3% of US unemployment. Since SSTV show that there is little change in geographic mismatch during the Great Recession, the 2012 level is likely to give us the right order of magnitude for mismatch in 2005-2012. Furthermore, we show that analyzing mismatch with data aggregated at the CBSA level gives a measurement that is close to what we find using our more sophisticated mismatch index. This suggests that the CBSA is an acceptable definition of the local labor market for the purpose of computing geographic mismatch.

Since we document how likely job seekers are to apply to jobs far away from home, our paper is also related to the literature on geographic mobility in the US. This literature typically measures moves across states (Molloy, Smith, and Wozniak, 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, Hunt, and McDowell, 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, Neumark, and McInerney (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, as mentioned above, 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.[2]

The next section presents the data and new empirical facts about the geography of job search. In the third section, we introduce the notion of mismatch unemployment and provide simple insights using our data. In the fourth section, we present our theoretical framework and provide new results about geographic mismatch. Section five concludes.

2 Describing the Geography of Job Search

2.1 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 dataset 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 vacancies 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

[2]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, Freeman, and Katz, 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, Dorn, and Hanson, 2013a,b), the equilibrium effects of active labor market policies (Davidson and Woodbury, 1993; Blundell, Dias, Meghir, and Reenen, 2004; Ferracci, Jolivet, and van den Berg, 2010; Gautier, Muller, van der Klaauw, Rosholm, and Svarer, 2012; Crépon, Duflo, Gurgand, Rathelot, and Zamora, 2013), the heterogeneity of the negative duration dependence with local conditions (Kroft, Lange, and Notowidigdo, 2013), spatial mismatch (Patacchini and Zenou, 2005; Hellerstein, Neumark, and McInerney, 2008; Boustan and Margo, 2009; Åslund, Østh, and Zenou, 2010), or the impact of place-based policies (Neumark and Kolko, 2010; Ham, Swenson, Imrohoroglu, and Song, 2011; Gobillon, Magnac, and Selod, 2012; Givord, Rathelot, and Sillard, 2013; Busso, Gregory, and Kline, 2013).

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 education requirement of those vacancies, CareerBuilder.com vacancies are not disproportionately directed at college educated workers (Marinescu, 2013). Overall, our data is therefore representative of the US distribution of vacancies and job seekers at the geographic level, while it is likely to be reasonably representative of the distribution of vacancies in terms of industry and education requirements.

2.2 Evidence

To assess the impact of a plant opening on local unemployment, one would typically use a difference-in-differences estimation strategy. One would select treatment and control areas defined as administrative entities (e.g. counties, CBSAs, or states). The same strategy can be used to estimate the impact of a other localized shocks to labor demand or labor supply, such as place-based policies, immigration or specific training and assistance provided to job seekers.

There are at least two potential sources of bias in difference-in-differences estimates of a plant opening on local unemployment. First, a widely recognized source of bias is spillover effects. If the control area is close enough to the treated area, unemployment in the control area may also decrease, leading to a downward bias in estimates. Second, there may be variation in treatment intensity: areas closer to the new plant may experience a larger decrease in unemployment. Assuming away spillovers to the control group, such variation in treatment intensity will lead to smaller estimates the larger the size of the treated area. The extent of the bias in the estimates arising from variation in treatment intensity depends on which population we really want to measure the average effect on. However, there is clearly a downward bias in the estimate when the treated area includes unaffected sub-areas. Documenting the geography of job search allows us to better understand the extent of the biases due to spillover effects and geographic variation in treatment intensity.

The extent of spillover effects depends on how far away job seekers apply for jobs. We therefore start by investigating the share of applications where the vacancy is within the residence area of the job seeker (Figure 1). We vary the size of the geographic unit we consider. For 86% of applications, the job seeker's state of residence is the same as the state where the vacancy is located. This share

is only slightly smaller (82%) when considering CBSAs. Even though states and CBSAs concentrate a large number of applications, there are still many applications that are out of CBSA or even out of state. Therefore, some caution should be exercised when examining states and CBSAs that are close to each other: spillovers may not be negligible in such cases. When considering the county level, a much smaller share of applications have the job seeker and vacancy located in the same area (46%). Finally, this share becomes minuscule (2%) when considering zip codes. We conclude that the extent of spillovers due to job search across CBSAs and states is likely to be small, while spillover effects between neighboring counties and zip codes are likely to be large.

If job seekers are more likely to apply to vacancies close to their zip code of residence, the treatment intensity for a local shock like a plant opening is likely to vary with the size of the area considered as treated. Indeed, the smaller the geographic unit considered and the larger the share of jobs in the geographic unit a job seeker applies to: for example, as we go from the CBSA of residence to the zip code of residence, the share of available jobs an average job seeker applies to almost doubles[3] (Figure 2). As we move from the CBSA level to the state level, there is a large drop off in the share of jobs in a given state that an average job seeker residing in the same state applies to: the number is almost divided by four. This pattern implies that job seekers find the highest share of jobs that they deem worth applying to in their own zip codes. As we consider larger areas, job seekers become more and more selective when choosing among available jobs. We can also investigate this issue from the point of view of the vacancy: for an average vacancy, what is the share of job seekers residing in its area that apply to it? As we move from the CBSA to the zip code, the share of job seekers in the area who apply to an average vacancy in the same area doubles (Figure 3). On the other hand, as we move from the CBSA level to the state level, there is a large drop off in the share of job seekers in a given state applying to an average vacancy in that state. This pattern implies that vacancies have the highest recruitment success in their own zip codes and they are able to attract a smaller and smaller share of job seekers as we consider areas larger than the zip code. Therefore, treatment intensity for a local labor market shock is likely to decline quickly as we consider areas larger than the zip code, and the drop off is especially large as we move from the CBSA to the state level. This implies that it is important to consider the size of the treated area for the purpose of estimating the impact of a plant opening: the geography of job search is such that treatment effects are likely to become much smaller as the size of the treated area expands.

So far, we have examined the geography of job search by considering geographic aggregates that are commonly used when evaluating the impact of labor market shocks. However, since we have individual data at the zip code level, we can say more about the impact of geographic distance on job search behavior. This is interesting in itself, but it is also helpful to gauge variation in treatment effects across geographic units depending on the distance between these units.

We first illustrate the geography of job search at the zip code level by an example from a specific

[3]The absolute numbers are very small, but this is not readily interpretable because we only see applications over a limited time period. Furthermore, not all jobs are relevant to all job seekers because of the variation in job and job seeker characteristics due to factors other than location.

CBSA. Consider job seekers living in zip codes 60639 and 60561, both in the CBSA of Chicago. These two zip codes are less than 30km from each other. Figure 4 presents two maps displaying the geographical patterns of the applications of job seekers living in these zip codes. The applications made by these job seekers are concentrated within areas smaller than the Chicago CBSA. This is consistent with the general pattern that most applications are contained within the CBSA (Figure 1). Applications are still quite spread out within the Chicago CBSA. This gives a concrete illustration to the fact that applications are not well contained in areas smaller than the CBSA, such as counties. Finally, we can see that applicants in each zip code concentrate their applications in neighboring zip codes. This is consistent with the general fact documented above: as distance from their zip code of residence increases, job seekers become more and more selective when choosing among available jobs (Figure 2).

Examining the role of continuous geographic distance in application behavior allows us to understand potential spillover effects as a function of distance. Using our individual data up to the zip code level, we can examine this issue quite accurately. We start by plotting, among applications, the distribution (cdf and pdf) of distance between the job seeker and the job they applied to. More than 80% of applications are concentrated within less than 100km of the the job seeker's zip code of residence (Figure 5), a share similar to the share of applications within CBSA. 100 km likely corresponds to the longest distance that a job seeker may be willing to commute to a job. The mode of the distance between the zip code of residence and the zip code of a job applied to is 15 km. Thus, most applications go to areas within commuting distance of the zip code of residence, but the typical distance between zip code of residence and zip code of the job is quite large and job seekers do not send the bulk of their applications to jobs very close to their zip code of residence. These results imply that spillover effects arising from job search behavior are likely to be concentrated within 100km of a localized labor market shock.

The distribution of distance between the job seeker and the job they applied to results from both the job seeker's preference for closer jobs and the spatial distribution of jobs. It is interesting to separate these two by estimating how distance influences the probability that a job seeker will apply to a given job. To do so, we consider the application behavior of every job seeker in Illinois to any vacancy located in Illinois, Indiana or Missouri. We include Indiana and Missouri because users from Illinois frequently apply outside of their state and these two neighboring states are the ones in which they apply the most. The number of observations is 1,355,172 pairs of zip codes (949 zip codes of users times 1428 zip codes of jobs). Among these pairs, 94% have 0 application. The number of applications by pair varies from 0 to 1396, with a mean at .62.

Let $A_{z,z'}$ denote the number of applications from job seekers located in z to vacancies located in z'. U_z and $V_{z'}$ are respectively the numbers of job seekers and job vacancies in zip codes z and z'. The maximum number of applications by job seekers located in z to jobs located in z' is thus $U_z V_{z'}$, so that $A_{z,z'} \in \{0, 1, \ldots, U_z V_{z'}\}$. Denote $p_{z,z'}$ as the probability that there exists an application between a given pair of job seeker and vacancy, the former being located in z and the latter in

z'. The number of applications is assumed to be distributed as a binomial of parameters $U_z V_{z'}$ and $p_{z,z'}$. Because, in our case, $U_z V_{z'}$ is large and $p_{z,z'}$ is small, the binomial is well proxied by a Poisson: $A_{z,z'} \sim Poisson(\mu_{z,z'})$, with $\mu_{z,z'} = U_z V_{z'} p_{z,z'}$. Now, we assume that the probability $p_{z,z'}$ is a function of a polynomial of the geographic distance between z and z', $d_{z,z'}$. The distance between two zip codes is defined as the distance between the centroids. The application probability may also depend on the (potentially unobserved) characteristics of the vacancies' zip codes and the job seekers' zip codes; we include the fixed effects δ_z and $\lambda_{z'}$ to account for this unobserved heterogeneity. For a cubic polynomial, the probability can be written as:

$$p_{z,z'} = \exp(\delta_z + \lambda_{z'} + \beta_1 d_{z,z'} + \beta_2 d_{z,z'}^2 + \beta_3 d_{z,z'}^3)$$

So that

$$\mu_{z,z'} = U_z V_{z'} \exp(\delta_z + \lambda_{z'} + \beta_1 d_{z,z'} + \beta_2 d_{z,z'}^2 + \beta_3 d_{z,z'}^3)$$

To estimate this model, we use Poisson regressions. Cubic polynomials are chosen as the fourth term is never significantly different from zero. We test several specifications, with or without fixed effects.[4] Fixed effects for destination zip codes are included to take into account that some areas may have better amenities and attract more job seekers, whatever the distance. Fixed effects for origin zip codes account for the heterogeneity of the job seekers living in different zip codes.
Distance has a large negative impact on the probability of applying to any given vacancy (Table 1). Interestingly, introducing fixed effects in the specification has little effect on the estimates. Job zip code fixed effects change almost nothing to the estimates, while the results for job seeker zip codes fixed effects and two-way fixed effects are also very similar. We also use a negative binomial specification (col. 2) to explore whether a more general specification than Poisson is needed. In this alternative specification, the coefficients estimates are very similar and the overdispersion parameter is very close to one, which confirms that the Poisson model is satisfactory for our data.

Our preferred model is model (5), which is the most flexible one. The probability of application decreases by 21%, when the distance between job seekers and vacancies goes from 0 to 5 km, by 19% from 20 to 25km, by 13% between 100 and 105km and by only 7% between 200 and 205km. Figure 6 shows the shape of the relative probability of application as a function of distance for Models 1 and 3-5. This graph strikingly confirms that controlling for the zip code of destination has little impact on the estimates. It is difficult to distinguish the curve related to the simple Poisson model from the one with jobs zip codes fixed effects, and the one with job seekers zip codes fixed effects from the one with two-way fixed effects. The relative influence of distance on applications tends to decrease with distance: job seekers make strong distinctions between jobs within a few kilometers of their zip code of residence but are less sensitive to differences in distance to job once jobs are further away. This may be rationalized by fixed costs of commuting or moving: for example, once taking a job necessitates to move place of residence, it is relatively less important where exactly the job is located. This feature is consistent with the empirical distribution displayed in Figure 5.

[4] To estimate the two-way fixed effect Poisson model, we use a zig-zag algorithm similar to the one in Guimarães and Portugal (2010). The estimation is performed with Stata.

Finally, let us remind the reader that the estimation is based, for the sake of computation feasibility, on job seekers located in Illinois and jobs in Illinois, Indiana and Missouri. There might be two issues with this choice. First, internal validity can be questioned: as we drop other (more distant) destination states, the distaste for distance is likely to be upward biased. In other terms, spillover effects deduced from these estimates could be downward biased. Second, there is the issue of external validity: are job seekers from Illinois representative of job seekers from the US? To assess this second question, we run the same regressions on other states: California vs. California, Nevada, Arizona, and Oregon; Florida vs. Florida, Georgia and Alabama; Massachusetts vs. Massachusetts, New-Hampshire, Vermont, Rhode Island, New York and Connecticut. We estimate the two-way fixed-effect Poisson model for all these different samples and find very similar results to the Illinois case. The one that differs the most in Massachusetts, in which job seekers seem to exhibit a slightly larger distaste for distance.

The estimated impact of distance on applications we find is much smaller than what Manning and Petrongolo (2011) find for the UK using local counts of jobs and job seekers to estimate the coefficients of a structural model. They found that the probability of a random job 5km distant being preferred to a random job in the worker's residential location is 11%. This figure is several orders of magnitude smaller than the figure of 79% we find. There might two plausible reasons for the discrepancy between the two figures. First, our figure is directly based on application data and on the estimation of application counts between two zip codes, while theirs rely on the estimation of a structural model. Second, distate for jobs far away may indeed be higher in the UK than in the US.

To sum up this section, we present novel evidence on the geography of job search. This evidence is highly relevant when evaluating the impact of a localized labor demand or labor supply shock, such as a plant opening. We show that job seekers prefer jobs closer to their zip code of residence but the preference for proximity is much smaller than prior evidence based on UK data suggests. Given the overall pattern of job search as a function of distance, spillovers from a localized shock should be mostly contained within 100km of the shock. Furthermore, the impact of any shock is likely to significantly decline with distance, even within 100 km. Choosing the CBSA where the local shock occurred as the treated zone will do a decent job at minimizing spillover effects to other areas. At the same time, such a choice will tend to weaken the treatment effect as treatment intensity decreases with distance from the shock within the CBSA. A workable strategy may be to choose a smaller area than the CBSA for the treatment group, such as a county. In this case, however, one should avoid using other counties in the same CBSA as controls because of spillover effects: instead, it is preferable to choose observationally similar counties further than 100km away.

Evidence on the geography of job search is important when evaluating the impact of localized labor market shocks. Indeed, as we just argued, estimates will be affected by the size of the areas chosen as treatment and control. The case of mismatch is quite similar. Intuitively, geographic mismatch arises when there are some areas with too many job seekers relative to vacancies and some areas with too few job seekers. However, the size of the area chosen to assess whether there are too few

or too many job seekers is going to affect our estimates of the degree of geographic mismatch. In the next section, we document how such a choice affects estimated mismatch in our data, and we discuss how we can include data on job applications to improve our understanding of geographic mismatch.

3 Measuring mismatch unemployment

3.1 Standard measures

If geographic mobility is costly or if mobility is determined by amenities other than job availability, CBSAs with many vacancies and few job seekers can coexist with CBSAs with few vacancies and many job seekers. More generally, high levels of unemployment can occur when limited mobility is accompanied by large discrepancies between the distribution of job seekers and jobs across locations, occupations, industries, etc. Mismatch indices are used to measure differences between the distribution of jobs and job seekers. One of the most commonly used measure of mismatch is the following:

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

where v_i, u_i are the number of vacancies and unemployed workers in 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.[5] 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, Song, Topa, and Violante (2012). Following Nickell (1982), Şahin, Song, Topa, and Violante (2012) 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 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.

3.2 Which unit of observation should be used?

The choice of the geographic unit of observation is subject to a trade-off. In the impact-evaluation case, the trade-off is between the presence of spillover effects and treatment intensity. For geographic mismatch, the same kind of issues arise. 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 but not all vacancies in the area are equally relevant to all job seekers. In this case, the index will obviously 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

[5] See Jackman, Layard, and Pissarides (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 7.

to vacancies that are not located in the area where the job seeker resides, and we run the risk of overestimating geographic mismatch. As the size of the area decreases, the index will mechanically tend to one. In the end, there is no perfect choice: no definition will both grant that (i) vacancies are equally relevant to all job seekers in the area and (ii) job seekers do not apply outside of the area. However, there might exist a level for which the two biases have about the same size, so that the resulting index is an acceptable approximation of geographic mismatch.
Figure 7 shows how the index varies with the unit of observation. Under the assumptions underlying the index, we find that 2% of hires are lost due to the misallocation of job seekers when units of observation are defined as states or CBSAs. When units of observation are counties, this figure more than doubles, to 5%. At the zip-code level, the fraction of hires lost due to misallocation of job seekers is 20%, a very large figure. The magnitude of geographic mismatch therefore strongly depends on the size of the unit of observation, with smaller units yields much larger values for the mismatch index.

3.3 Mismatch measures using applications

The measure of mismatch we just discussed may be very misleading. Imagine that, independently of their places of residence, job seekers apply to vacancies such that the number of applicants by vacancy is roughly constant. In this case, the previously-defined mismatch indices may indicate high mismatch while the relevant mismatch is actually much smaller. In other terms, the mismatch between job seekers' and vacancies' locations may be corrected by the application pattern of job seekers. If job seekers located in areas with lower tightness apply to jobs in areas with higher tightness, their behavior will correct the location mismatch.

To provide a first insight about this issue, we consider indices based on applications. Specifically, we define two indices as:

$$\mathcal{M}_a = 1 - \sum_i \left(\frac{v_i}{\sum_i v_i}\right)^{\gamma} \left(\frac{a_i}{\sum_i a_i}\right)^{1-\gamma}$$
$$\mathcal{M}_{\tilde{a}} = 1 - \sum_i \left(\frac{v_i}{\sum_i v_i}\right)^{\gamma} \left(\frac{\tilde{a}_i}{\sum_i \tilde{a}_i}\right)^{1-\gamma}$$

where a_i is the number of applications sent to vacancies in market i and $\tilde{a}_i$ is the number of applications sent to market i, weighted by the inverse of the number of applications sent by the job seeker (so that $\sum_i a_i = \sum_i u_i$). The reason for using applications instead of job seekers is to investigate whether application behavior mitigates the imbalance in the geographic distribution of job seekers and vacancies. The difference between $\mathcal{M}_a$ and $\mathcal{M}_{\tilde{a}}$ is that, in the latter, job seekers who send more applications do not contribute more to the index than those sending less applications.

Figure 8 compares the indices $\mathcal{M}$, $\mathcal{M}_a$, $\mathcal{M}_{\tilde{a}}$ for several geographical units. The index based on weighted applications (which corrects for the number of applications by job seeker) is equal, for large units, to the index based on job seekers. For small units, the index based on weighted applications is lower than the index based on job seekers. This indicates that job seekers tend to

choose the destinations of their applications in such a way that mismatch is reduced. However, the phenomenon is quite marginal across CBSAs and states and is only substabtial across zip codes. This last result makes sense if zip codes are specialized, some being more residential, other more commercial or industrial. The index based on applications, conversely, is always higher than the other two (expect at the zip-code level). This reflects another phenomenon: job seekers located in areas with lower tightness on the labor market tend to send more applications. Because the majority of their applications are still sent to surrounding low-tightness areas, this has the consequence to increase the index.

Another way to understand how across-market applications affect mismatch is to compare mismatch indices based on internal applications with mismatch indices based on all applications. Following the previous definitions, the mismatch index based on internal applications is defined as:

$$\mathcal{M}_n = 1 - \sum_i \left(\frac{v_i}{\sum_i v_i} \right)^{\gamma} \left(\frac{n_i}{\sum_i n_i} \right)^{1-\gamma}$$

where n_i are internal applications in market i respectively. Internal applications in market i are defined as applications to jobs in market i that come from job seekers residing in market i. If applications across markets tend to correct mismatch, we expect mismatch based on all applications to be lower than mismatch based on internal applications. Indeed, regardless of the level of geography considered, we find that mismatch using all applications is lower than mismatch using only internal applications (Figure 9). For example, if we use CBSAs as a definition of the labor market, mismatch is reduced by 30% when we go from considering only internal applications to considering all applications received by jobs within a CBSA. We conclude that applications across labor markets can substantially reduce the measured level of geographic mismatch.

Job seekers' application behavior tends to correct imbalances in the labor market. However, the extent of this correction remains difficult to assess as it depends on the level of aggregation chosen. In the next section, we offer a solution to this problem by using a new mismatch index, which directly accounts for job seekers' distaste for distance.

4 A measure of mismatch that accounts for the geography of job search

The goal of this section is to build a measure of mismatch unemployment compatible with our findings about the geography of job search. The mismatch measures we presented in the preceding section assume that job seekers apply within their CBSA or their state, and that all vacancies in this area are equally relevant. By contrast, we consider a framework in which geographic *distance* determines the extent to which a vacancy is relevant to a job seeker. As explained in the previous section, measuring mismatch requires us to map an allocation of job seekers to the number of hires in the economy. In order to define such a mapping, we have to build a simple theoretical model that accounts for the fact that job seekers may send applications to several places, depending on the distance between the vacancies and themselves.

4.1 Theoretical framework

The model presented here is a static urn-ball search-and-matching model. In the canonical model, in which job seekers apply only once, unemployment comes from the fact that job seekers cannot observe each other's strategies (Butters, 1977; Hall, 1979; Peters, 1984). Some vacancies receive several applications while others do not receive any. In this model, the probability for a job seeker to receive several offers is often assumed to be sufficiently small so that job seekers receive either zero or one offer. In the multiple application case, this assumption is less likely to hold and another friction is introduced. Because employers do not observe each other's offers: some workers receive several offers while others do not receive any.[6]

In the simplest version of our model, job seekers have an exogenous number of applications and decide to send them to vacancies in different places. Their decision is based on a trade-off between the probability of being hired (which depends endogenously on the number of applicants by vacancy) and the distance between their place of residence and the location of the vacancy. Firms receive applications and decide randomly whom to hire among applicants who turn out to be qualified. The matching function we obtain is an elaborated version of the classical urn-ball matching function that incorporates the feature that all job seekers may apply to a vacancy, although those further away have lower probability to do so.

Manning and Petrongolo (2011) present a model that has inspired ours in several dimensions. Like us, they use a static job search model with multiple applications, they consider that the application decision is a product between a distance-specific term and the destination-specific probability to be hired. Unlike us, they endogenize the average number of applications sent by job seekers (we also do this in an extended version of the model, see appendix). To keep their model tractable, they make an assumption about the functional form of the probability to be hired and assume that the coordination failure on the firm side can be neglected. However, Wolthoff (2010) finds that search frictions and recruitment frictions are each responsible for around half of the loss in social surplus compared to a Walrasian world. We therefore decided to include the coordination failure on the firm side in our model.

Many models in the literature stress the importance of taking into account the heterogeneity of job seekers and jobs on the labor market.[7] Introducing geographical space in the modelling of labor markets is however considered as a difficult exercise (Zenou, 2009). In most contributions in the urban economics literature, geographic distance is introduced in search-and-matching to point out either the role of commuting and the existence of informational imperfections (Wasmer and Zenou, 2002; Brueckner, Thisse, and Zenou, 2002; Zenou, 2009; Rupert and Wasmer, 2012).

[6] Albrecht, Gautier, and Vroman (2003, 2004, 2006); Galenianos and Kircher (2009); Kircher (2009); Wolthoff (2010) are examples of theoretical frameworks in which job seekers send multiple applications. Both Albrecht, Gautier, and Vroman (2006); Decreuse (2008) allow for an endogenous number of applications and show that, under certain conditions, job seekers prospect too many segments.

[7] Barnichon and Figura (2013) argue that the relationship between the aggregate job finding rate and labor market tightness does not hold after 2007, as a result of changes in the segmentation of the labor market.

Job seekers and vacancies are spread across S spots. In spot s, there are U_s job seekers and V_s vacancies. The timing of the game is the following.

1. Job seekers apply to vacancies.
2. Firms gather the applications they receive: each application has a probability q to be valid in the sense that the applicant can do the job.
3. Firms can only make one offer. If a vacancy has more than one valid application, a tie breaking procedure determines who will be offered the job.
4. Offers are sent to job seekers.
5. Job seekers can only accept one job. If a job seeker has received more than one offer, he accepts the offer that generates the highest utility.
6. Matches are realized.

To simplify exposition, we introduce first a model in which local labor markets are assumed to be distinct. In a second step, we present the full-fledged model in which job seekers can apply everywhere but are sensitive to distance.

4.1.1 Distinct markets

Let us start by assuming that local markets are distinct. Job seekers can apply to all vacancies in their own spot, but not elsewhere. In this case, all markets are independent and the solution is found independently for each market.

Assume that the number of applications by job seekers is exogenous and constant, equal to $a_s = \bar{a}$. Job seekers are equally likely to apply to any job in s. Therefore, the probability p_s that a job seeker in s applies to a vacancy located in s is equal to:

$$p_s = \frac{\bar{a}}{V_s}$$

Because job seekers do not coordinate their applications, the number of applications received by a vacancy located in s from job seekers located in s is distributed as a Binomial(U_s, p_s). As U_s is large and p_s is small, this distribution is well proxied by a Poisson$(p_s U_s)$.

An application has a probability q to be valid for the firm. The number of valid applications is distributed as a Binomial parameterized by the number of received applications and the probability q that the application is valid. As the composition of a Binomial and a Poisson, the number of valid applications is distributed as a Poisson$(qp_s U_s)$.

Employers choose randomly among the valid applications which job seeker they will make an offer to. From the point of view of a job seeker, what is the probability π_s that an application he is about to make is going to generate an offer? He will get an offer with a probability q if no-one else

applies (associated with the event "number of applications equal to zero"), with a probability $q/2$ if one other job seeker applies, and so on. All in all, π_s is equal to q multiplied by the expectation of $1/(X+1)$ where X is a random variable distributed as a Poisson(qp_sU_s).[8]

$$\pi_s = \frac{1}{p_s U_s}\left[1 - \exp\left(-q p_s U_s\right)\right]$$

Replacing p_s the probability of applying to a job in s by its value, this may be written as:

$$\pi_s = \frac{V_s}{\bar{a} U_s}\left[1 - \exp\left(-q\frac{\bar{a} U_s}{V_s}\right)\right] \tag{2}$$

Following a similar reasoning, the number of offers received by a job seeker in s is distributed as a Poisson$(\pi_s p_s V_s)$, or, equivalently, a Poisson$(\pi_s \bar{a})$. The job seeker will choose among the offers she received the one which provides her the maximum amount of utility. From our point of view or the point of view of the firm, this is equivalent to choosing randomly (the utility brought by a job is private information). Because a job seeker with h offers has a probability $1/(h+1)$ to accept a marginal offer, the probability that a job seeker will accept a given offer is:

$$\xi_s = \frac{1 - \exp\left(-\pi_s \bar{a}\right)}{\pi_s \bar{a}} \tag{3}$$

A vacancy is filled if at least one valid application is received and if the job seeker who is offered the job accepts it. The probability for a vacancy to be filled is:

$$\xi_s\left[1 - \exp\left(-q\frac{\bar{a} U_s}{V_s}\right)\right]$$

The expected number of matches in s is equal to:

$$M_s = V_s \xi_s\left[1 - \exp\left(-q\frac{\bar{a} U_s}{V_s}\right)\right] \tag{4}$$

and the total number of matches is:

$$M = \sum_s V_s \xi_s\left[1 - \exp\left(-q\frac{\bar{a} U_s}{V_s}\right)\right]$$

This matching function is a generalized version of the well-known expression that arises from the canonical urn-ball model. There are two modifications compared to the canonical matching function. First, job seekers are allowed to send multiple applications, which is reflected in the $\bar{a}$ term in the expression above. Second, job seekers may get more than one offer. Therefore, the number of matches depends on job seekers' propensity to accept offers ξ_s.

This matching function exhibits two welcome properties. First, it exhibits constant returns to scale (Pissarides and Petrongolo, 2001). Second, the number of matches is equal to zero when either the number of vacancies or job seekers is equal to zero.

[8] If X is distributed as a Poisson(λ), $E[1/(X+1)] = (1 - \exp(-\lambda))/\lambda$.

4.1.2 Overlapping markets

Now we turn to our case of interest: job seekers are allowed to apply to all labor markets, not only their own. We assume that the only parameter that matters is the geographical distance between their place of residence and the vacancy. In this model, all local markets are tied to each other, so the whole model must be solved simultaneously. In our application, locations are defined at the smallest unit available in our data: zip codes. The results from the previous section showed that zip codes are very small compared to the scale relevant for job search. For this reason, we consider that zip codes are almost as good as the exact geographic location of job seekers.

Assume that the number of applications by job seekers is exogenous and constant, equal to $\bar{a}$. However, by contrast with the case of distinct markets, the allocation of these applications among the different destinations has to be decided. The optimal allocation is such that the marginal gain from applying to each destination is equalized. Let b_{ij} denote the share of applications sent to vacancies located in s_j by a job seeker located in s_i. Let F_{ij} denote the utility to work in s_j for a job seeker initially living in s_i. π_j still denotes the probability to get an offer from a given vacancy in s_j (conditional on applying). Note that π_j is assumed not to depend on the job seekers' places of residence: the probability of receiving an offer does not depend on the location of the job seeker, only on the location of the job. Assume that $b_{i\ell}\bar{a}$ applications have already been sent to each destination $\{s_\ell\}_{\ell=1...S}$. For all j and j', the marginal gain to send an extra application to s_j should be equal to the one for sending an application to $s_{j'}$.

$$\left[\Pi_\ell (1-\pi_\ell)^{b_{i\ell}\bar{a}}\right] \pi_j F_{ij} = \left[\Pi_\ell (1-\pi_\ell)^{b_{i\ell}\bar{a}}\right] \pi_{j'} F_{ij'}$$

which is equivalent to $\pi_j F_{ij} = \pi_{j'} F_{ij'}$.

Now, let us assume that, besides the idiosyncratic preferences of each job seeker for specific jobs, the only relevant variable for the utility F_{ij} is the geographical distance between s_i and s_j, d_{ij}. The idea is that, for short distances, the worker will incur a commuting cost, which grows with the distance, and that, for longer distances, she may have to face moving costs. We assume that F_{ij} decreases with the number of applications sent to s_j by job seekers in s_i. This assumption reflects the idea that, for each job seeker, some jobs are better than others, and the job seeker applies first to his favorite jobs. Given this behavior, any new job a job seeker applies to must be worse than the jobs he has already applied to. Therefore, for each spot s_j, the utility of being hired in any given job declines with the number of applications a job seeker has already sent. For simplicity, we choose $F_{ij} = f(d_{ij})/a_{ij}$, where $f(d_{ij})$ is a function of the distance d_{ij}. Introducing this expression in the former equation and using that b_{ij} sum to one (over j), we find that:

$$b_{ij} = \frac{\pi_j f(d_{ij})}{\sum_\ell \pi_\ell f(d_{i\ell})}$$

The probability p_{ij} that a job seeker in s_i applies to a vacancy located in s_j is equal to:

$$p_{ij} = \frac{\bar{a}}{V_j} b_{ij} = \frac{\bar{a}\pi_j f(d_{ij})}{V_j \sum_\ell \pi_\ell f(d_{i\ell})} \tag{5}$$

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($\sum_k p_{kj}U_k$). The number of valid applications is distributed as a Poisson($q\sum_k p_{kj}U_k$).

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 $1/(X+1)$ if X is a random variable distributed as a Poisson($q\sum_k p_{kj}U_k$).

$$\pi_j = \frac{1}{\sum_k p_{kj}U_k}\left[1-\exp\left(-q\sum_k p_{kj}U_k\right)\right] \tag{6}$$

Similarly, the number of offers received by a job seeker is s_i corresponding to vacancies in s_j is distributed as a Poisson($\pi_j p_{ij}V_j$). The total number of offers received by this job seeker is thus distributed as a Poisson($\sum_\ell \pi_\ell p_{i\ell}V_\ell$). Because we assume, as in the distinct market case, that a job seeker with h offers has a probability $1/(h+1)$ to accept a marginal offer, the probability that a job seeker s_i will accept a given offer is:

$$\frac{1}{\sum_\ell \pi_\ell V_\ell p_{i\ell}}\left[1-\exp(-\sum_\ell \pi_\ell V_\ell p_{i\ell})\right]$$

From the point of view of a firm in s_j, the average probability that an applicant will accept an offer for its job is:

$$\xi_j = \sum_k \frac{p_{kj}U_k}{\sum_{k'} p_{k'j}U_{k'}}\frac{1-\exp(-\sum_\ell \pi_\ell V_\ell p_{k\ell})}{\sum_\ell \pi_\ell V_\ell p_{k\ell}} \tag{7}$$

The first term in the sum reflects the share of applications coming from each spot k while the second term is the probability that an applicant from k will accept the offer.
A vacancy is filled if at least one valid application is received and if the job seeker who is offered the job accepts it. The probability for a vacancy to be filled is:

$$\xi_j\left[1-\exp\left(-q\sum_k p_{kj}U_k\right)\right]$$

The total number of matches is:

$$M = \sum_j V_j\xi_j\left[1-\exp\left(-q\sum_k p_{kj}U_k\right)\right] \tag{8}$$

This matching function is similar to the one obtained for distinct markets. We can check that the function exhibits constant returns to scale, in the sense that if the number of job seekers and vacancies is multiplied by the same scalar factor in every spot, the number of matches is also multiplied by this factor. Interestingly, we note that the probability to get an offer π_j, the number of applications a_i, the share b_{ij}, the probability of acceptance ξ_j and the probability that a vacancy is filled are all 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.

4.1.3 Model resolution and calibration

In the case of distinct local labor markets, the resolution of the model is simplified by the fact that the problem for each spot can be solved independently from the others. The number of vacancies V_s in each location s, and the number of applications are considered exogenous. In each location s, the probability of being hired π_s is first obtained as a function of the number of job seekers U_s, using equation (2). Then, the probability of acceptance ξ_s is determined by equation (3), as a function of π_s. Finally, the number of matches is determined by equation (4), as a function of ξ_s and U_s. Numerically, the resolution is fast because endogenous variables depend on each other in the form of a chain, with no loop. Moreover, because the resolution can be done independently in each spot, the computational cost should be linear in S, the number of spots.

In the case of overlapping labor markets, the resolution is more complicated. We consider $\{V_s\}_s$, $\bar{a}$ and $f(.)$ as exogenous parameters. Equation (5) shows that the probability p_{ij} for a job seeker in s_i to apply to a given vacancy in s_j depends on the probability to be hired π_j. But π_j depends both on the whole allocation of job seekers $\{U_i\}_i$ and the vector of probabilities of application $\{p_{ij}\}_i$, as shown by equation (6). These two equations define a system of S equations (or a mapping) of $\{\pi_j\}_j$ as a function of itself. For all j,

$$\pi_j = \left(\sum_k \frac{\bar{a}\pi_j f(d_{ij}) U_k}{V_j \sum_\ell \pi_\ell f(d_{i\ell})} \right)^{-1} \left[1 - \exp\left(-q \sum_k \frac{\bar{a}\pi_j f(d_{ij}) U_k}{V_j \sum_\ell \pi_\ell f(d_{i\ell})} \right) \right]$$

Does the equilibrium exist? Is it unique? To answer this question, it would be sufficient to prove that this mapping is contracting. We have run many simulations for various values of the parameters: the mapping was contracting in all the cases we tested. However, we could not arrive to a formal proof so far, using, for instance, the Blackwell sufficient conditions. We solve this system numerically by iterating this mapping until the value of π does not change any longer.

Once π_j (and thus p_{ij}) have been found, equation (7) provides a way to compute the acceptance probability ξ_j. Then, we can use equation (8) to obtain the number of matches, as function of $\{p_{ij}\}$, $\{\xi_j\}$ and $\{U_i\}$. Here, two hurdles make the computational cost much higher than in the distinct-market case. First, the number of job seekers in a given location impacts the number of matches in another one. This feature is consistent with our finding that job seekers' application patterns tend to correct mismatch, but makes things more complicated, as the whole model must be solved simultaneously. Second, because p and π depend on each other, the iterating process to solve the system is the computational bottleneck. With a number of units of around 23,000, the system can be solved in around 8 seconds.[9]

There are three set of parameters that we need to estimate. First, the number of vacancies in each location V_s is given by the value we observe in the data. Second, the average number of applications $\bar{a}$ is also taken as its empirical counterpart, equal to 12.8. Third, we have to choose the function $f(.)$, which determines the distaste for distance of the job seekers. To simplify, we impose a functional

[9] The code is written in Matlab and is available upon request.

form, the exponential of a polynomial in distance, and we denote β as the vector of coefficients of the polynomial. We have two possibilities to estimate β, which have their pros and cons. The first one is to rely on a reduced-form estimation of f, similar to the ones presented in the first section of the paper. The advantage of this method is to account for the nature of count data (through the Poisson specification), and to allow for fixed effects in s_i and s_j, which makes the estimation less dependent on the exact structure of the model. The second estimation strategy is to rely on the structural form of the model. The idea is to combine equations (5) and (6) to form a system of equations of the following form. For all j,

$$a_{.j} = g(a_{.j}, \{a_{.\ell}\}_\ell, \beta)$$

where $a_{.j} = \sum_k a_{kj}$. We observe in the data the number of applications sent to a given location, which are consistent estimates of $a_{.j}$. Thus, we can obtain a structural estimate of β by generalized minimum distance estimation. The advantage of this method is to rely on the structural form of the model and to incorporate the fact that the allocation of applications results from strategic decisions of job seekers. The drawaback of this method is that the estimate will depend on whether the model includes all first-order determinants of the application decision. For simplicity, we use the first strategy but we plan to assess the sensitivity of the results when the second strategy is used instead. The β chosen are the ones reported in the column 5 in Table 1.

4.2 Computing the mismatch index

Now that we have obtained a matching function, which maps an allocation of job seekers to a number of hires, we can define the mismatch index as:

$$\tilde{\mathcal{M}} = 1 - \frac{M}{M^*}$$

where M is the number of hires given the observed allocation of job seekers and M^* is the maximum number of hires achievable by a social planner able to freely allocate job seekers across markets.

We denote by $M(u)$ the number of matches associated with the allocation $u = \{u_{s_1}, \dots u_{s_S}\}$ of job seekers (keeping other variables and parameters fixed). $M(u)$ is computed by solving the model for the allocation u (instead of the current allocation U). M^* is the result of the optimization:

$$M^* = \max_u M(u) \text{ s.t. } \sum_s u_s = \sum_s U_s$$

The solution allocation is denoted as U^*, so that $M(U^*) = h^*$.

In the case in which markets are distinct, the solution can be analytically found because our matching function exhibits constant returns to scale. As shown in Şahin, Song, Topa, and Violante (2012) for instance, the optimal allocation is such that the tightness U_s/V_s is constant over locations s. In what follows, we name this allocation the *proportional* allocation, as the number of job seekers should be made exactly proportional to the number of vacancies.

In the case of overlapping markets, conversely, there is no easy way to find the analytical solution. The intuition is that, in this case, it is not U_s that should be made proportional to V_s but rather the amount of applications converging to s from all origins. The optimal allocation must, in this case, be found numerically. The optimization problem is a tricky one as S can be very large. At the zip-code level, S is higher than 23,000 and standard optimization tools will not work in limited time, even with high-performance computers. We design an ad-hoc algorithm relying on approximated (but faster to compute) versions of the gradient (see Appendix 2 for details on the algorithm used). Depending on the starting point, about 10 days are necessary to reach the optimum.

As stated before, we use the data at the finest geographic level available: the zip codes. There are 23,585 zip codes in our data set and we compute the number of job seekers and vacancies at this level. In the data, U_s and V_s are positively correlated: a naive regression of U_s on V_s exhibits a coefficient of .15 (with a standard error of .003).

4.3 Results

Now, we turn back to the measurement of mismatch unemployment. Using the model with overlapping markets and considering the zip code as the unit of analysis, we compute the index $\tilde{\mathcal{M}}$ and find a value of 0.032, i.e. 3.2% of hires are lost due to the misallocation of job seekers. This is our preferred measure of geographic mismatch.

How does this value compare with the values obtained using other approaches? First, we can compare it to the mismatch index $\mathcal{M}$ obtained under the assumption of distinct markets and using a Cobb-Douglas matching function. We find that our preferred measure of geographic mismatch is slightly higher than the value found at the CBSA level (2%) and smaller than the one found at the county level (see Figure 7). This conclusion is consistent with the intuition that emerged from our empirical analysis: the local labor market relevant to a job seeker seems to lie between the county and the CBSA.

Given the value we found for geographic mismatch, how much would the aggregate unemployment rate in the U.S. fall if we eliminated mismatch? Let us consider that the unemployment rate is not far from its steady-state value, which depends on the exit rate from employment μ and the hiring rate λ as follows:

$$u_{ss} = \frac{\mu}{\lambda + \mu}$$

We take $\lambda = 36\%$ (as in the calibration above) and $\mu = .031$, so that the initial steady-state unemployment rate is equal to the actual value of the unemployment rate in the US at the time our data was collected, i.e. 8%. The planner value λ^* for the hiring rate is equal to $\lambda^* = \lambda/(1 - \tilde{\mathcal{M}})$. Given these expressions, a mismatch index of 0.032 means that, if job seekers were allocated according to the optimal allocation (instead of the one observed in the data), the unemployment rate would be decreased by .2 percentage point (3%). This figure is subject to the usual disclaimer that applies to any attempt to compute mismatch unemployment. That a planner can move job seekers without

costs is an extreme assumption: this exercise should be taken as a way to compute the higher bound of the contribution of mismatch to the unemployment rate.

We check the sensitivity of our results to the functional form of the matching function by re-computing our mismatch index on data aggregated at the CBSA level and comparing it to the mismatch index using a Cobb-Douglas matching function. To mimic the results obtained with the Cobb-Douglas matching function (Figure 7), we assume that CBSAs are isolated and use the distinct-market version of our model. In this case, the optimal allocation is the proportional one. At the CBSA level, we find a mismatch index $\tilde{\mathcal{M}}$ of 2.8%. This value is very similar to the 2.7% for $\mathcal{M}$ on the same sample. This shows that the functional form of the matching function (Cobb Douglas vs. urn-ball) has only a second-order impact on the value of the mismatch index. Table 2 summarizes our results and the comparison with alternative measures.

Finally, is it much ado about nothing? If geographic mismatch accounts for a maximum of 3% of total unemployment, is it still worthwhile looking for policies that would reduce mismatch? An alternative way to assess whether 3% is low in absolute terms is to run a counterfactual exercise. Imagine that an economic policy managed to increase the number of vacancies while keeping the geographic distribution of jobs and job seekers fixed as in the data. How much does the number of vacancies need to grow to achieve a similar increase (3%) in the number of matches as eradicating geographic mismatch? Using our model with overlapping markets at the zip code level, we find that the number of vacancies would have to be increased by 18%. To understand how such an impressive figure can arise from a small degree of mismatch, let us have a closer look at the matching function. Around the values observed in the data for U and V, our matching function exhibits an elasticity on unemployment equal to 80% and an elasticity on vacancies equal to 20%. Even if estimated elasticities vary substantially across empirical studies, these values are clearly in the range of plausible elasticities. They are equal to those obtained by Layard, Nickell, and Jackman (1991) and close to the 70/30% obtained by Pissarides (1986). Eradicating geographic mismatch would decrease unemployment by 3%, yielding the same effect as an 18% increase in the number of vacancies. To the extent that generating such a large increase in vacancies is no trivial task for a policy maker, this suggests that attempting to alleviate geographic mismatch may be worthwhile even if the expected improvement in the unemployment rate is small.

5 Conclusion

The geography of job search matters. Using data from CareerBuilder.com, this paper shows that job seekers are more likely to apply to jobs closer from home: the probability of applying to a job strongly decreases with distance, at an initial rate of 20% for 5 km. However, American job seekers are far more likely to apply to jobs far away from home than their British counterparts. We also investigate how well administrative units can capture the geography of job search. We find that most job applications are contained within the CBSA of residence of the job seekers. However, job seekers very often apply across counties. This implies that the CBSA is a good choice for the unit of treatment if one wishes to evaluate localized labor market shocks while avoiding spillover effects.

We then demonstrate how the geography of job search allows us to better understand geographic mismatch. We measure geographic mismatch using a new index that incorporates the geography of job search at the zip code level. We find that geographic mismatch accounts for 3% of US unemployment, which is similar to what we find using an existing mismatch index and data aggregated at the CBSA level. While existing mismatch indices are sensitive to the degree of geographic aggregation of the data, our index is robust to this issue and allows us to conclude that geographic mismatch is not a major driver of US unemployment. Furthermore, we find that the CBSA is a good level of analysis for the purpose of computing geographic mismatch. Overall, our analysis suggests that the CBSA is a good definition of the labor market for a number of applications.

Finally, geography is only one dimension that distinguishes different labor markets. Labor economists have long recognized that labor markets are stratified by skills. When it comes to mismatch in the labor market, "skills gap" or occupation mismatch is quite important, as documented by Şahin, Song, Topa, and Violante (2012). In future work, we plan to address the interaction of the geographic and occupational dimensions in differentiating labor markets.

References

Albrecht, J. W., P. A. Gautier, and S. B. Vroman (2003): "Matching with Multiple Applications," *Economics Letters*, 78(1), 67–70.

——— (2004): "Matching with Multiple Applications Revisited," *Economics Letters*, 84(3), 311–314.

——— (2006): "Equilibrium Directed Search with Multiple Applications," *Review of Economic Studies*, 73(4), 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*, NBER Chapters, pp. 201–234. National Bureau of Economic Research, Inc.

Åslund, O., J. Østh, and Y. Zenou (2010): "How Important Is Access to Jobs? Old Question - Improved Answer," *Journal of Economic Geography*, 10(3), 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(3), 220–25.

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

Barnichon, R., and A. Figura (2013): "Labor Market Heterogeneity and the Aggregate Matching Function," mimeo CREi.

Blanchard, O. J., and L. F. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 23(1), 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(4), 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(4), 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(2), 246–51.

——— (1997): "How Much Do Immigration and Trade Affect Labor Market Outcomes?," *Brookings Papers on Economic Activity*, 28(1), 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(1), 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(1), 11–23.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (2002): "Local Labor Markets, Job Matching, and Urban Location," *International Economic Review*, 43(1), 155–171.

Busso, M., J. Gregory, and P. Kline (2013): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103(2), 897–947.

Butters, G. R. (1977): "Equilibrium Distributions of Sales and Advertising Prices," *Review of Economic Studies*, 44(3), 465–91.

Card, D. (1990): "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 43(2), 245–257.

——— (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19(1), 22–64.

——— (2005): "Is the New Immigration Really so Bad?," *Economic Journal*, 115(507), F300–F323.

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(2), 531–580.

Davidson, C., and S. A. Woodbury (1993): "The Displacement Effect of Reemployment Bonus Programs," *Journal of Labor Economics*, 11(4), 575–605.

Decreuse, B. (2008): "Choosy Search And The Mismatch Of Talents," *International Economic Review*, 49(3), 1067–1089.

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(2), 360–367.

Ferracci, M., G. Jolivet, and G. J. van den Berg (2010): "Treatment evaluation in the case of interactions within markets," Working Paper Series 2010:1, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Friedberg, R. M., and J. Hunt (1995): "The Impact of Immigrants on Host Country Wages, Employment and Growth," *Journal of Economic Perspectives*, 9(2), 23–44.

Galenianos, M., and P. Kircher (2009): "Directed Search with Multiple Job Applications," *Journal of Economic Theory*, 144(2), 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(1), 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(9-10), 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(2), 223–234.

Guimarães, P., and P. Portugal (2010): "A simple feasible procedure to fit models with high-dimensional fixed effects," *Stata Journal*, 10(4), 628–649(22).

Hall, R. E. (1979): "A Theory of the Natural Unemployment Rate and the Duration of Employment," *Journal of Monetary Economics*, 5(2), 153–169.

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(7-8), 779 – 797.

Hellerstein, J. K., D. Neumark, and M. McInerney (2008): "Spatial mismatch or racial mismatch?," *Journal of Urban Economics*, 64(2), 464–479.

Jackman, R., R. Layard, and C. Pissarides (1989): "On Vacancies," *Oxford Bulletin of Economics and Statistics*, 51(4), 377–394.

Jackman, R., and S. Roper (1987): "Structural Unemployment," *Oxford Bulletin of Economics and Statistics*, 49(1), 9–36.

Kircher, P. (2009): "Efficiency of Simultaneous Search," *Journal of Political Economy*, 117(5), 861–913.

Kline, P. M., and E. Moretti (2013): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," Working Paper 19293, National Bureau of Economic Research.

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(3), 1123–1167.

Layard, R., S. J. Nickell, and R. A. Jackman (1991): *Unemployment, Macroeconomic Performance and the Labour Market*. Oxford University Press, Oxford.

Lazear, E. P., and J. R. Spletzer (2012): "Hiring, Churn and the Business Cycle," *American Economic Review: Papers and Proceedings*, 102(3), 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. (2013): "Online Job Search and Unemployment Insurance during the Great Recession," .

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(3), 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(367), 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(1), 9–44.

Patacchini, E., and Y. Zenou (2005): "Spatial mismatch, transport mode and search decisions in England," *Journal of Urban Economics*, 58(1), 62–90.

Peters, M. (1984): "Bertrand Equilibrium with Capacity Constraints and Restricted Mobility," *Econometrica*, 52(5), 1117–27.

Pissarides, C. (1986): "Unemployment and Vacancies in Britain," *Economic Policy*, 1(3), 499–559.

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

Pissarides, C. A., and B. Petrongolo (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39(2), 390–431.

Rupert, P., and E. Wasmer (2012): "Housing and the labor market: Time to move and aggregate unemployment," *Journal of Monetary Economics*, 59(1), 24–36.

Şahin, A., J. Song, G. Topa, and G. L. Violante (2011): "Measuring Mismatch in the U.S. Labor Market," mimeo.

——— (2012): "Mismatch Unemployment," NBER Working Papers 18265, National Bureau of Economic Research, Inc.

Wasmer, E., and Y. Zenou (2002): "Does City Structure Affect Job Search and Welfare?," *Journal of Urban Economics*, 51(3), 515–541.

Wolthoff, R. P. (2010): "Applications and Interviews: A Structural Analysis of Two-Sided Simultaneous Search," IZA Discussion Papers 5416, Institute for the Study of Labor (IZA).

Wozniak, A. (2010): "Are College Graduates More Responsive to Distant Labor Market Opportunities?," *Journal of Human Resources*, 45(4), 944–970.

Zenou, Y. (2009): *Urban Labor Economics*. Cambridge University Press.

Table 1: Probability for an application as a function of distance

	Number of applications				
	Poisson	Neg. Bin	FE Poisson		
	(1)	(2)	(3)	(4)	(5)
Distance (× 100km)	−3.64 (.05)	−4.64 (.03)	−3.55 (.15)	−4.63 (.09)	−4.64 (.09)
Distance (× 100km), squared	.70 (.02)	1.11 (.01)	.51 (.02)	.96 (.04)	.96 (.04)
Distance (× 100km), cubic	−.05 (.00)	−.09 (.00)	−.02 (.01)	−.07 (.01)	−.06 (.01)
Overdispersion parameter	-	.98 (.01)	-	-	-
Fixed effects	-	-	jobs zip	job seekers zip	jobs and job seekers zip
Nobs	1,355,172	1,355,172	1,038,206	1,355,172	1,038,206

Source: CareerBuilder database.

Notes: 1 star means 99%-significant. Column 1 is a Poisson model. Column 2 is a Negative Binomial model. Columns 3 to 5 are Poisson models with fixed effects. All specifications include $\log U_z V_{z'}$ with a coefficient constrained to one. All models are estimated by maximum likelihood.

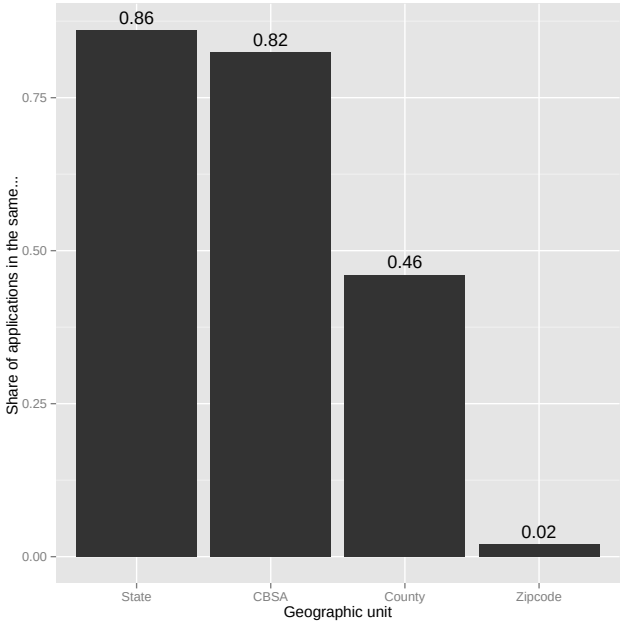
Table 2: Geographic mismatch indices

Matching function	Urn-ball		Cobb-Douglas	
Local markets	overlapping	distinct	distinct	distinct
Unit of analysis	zip code	CBSA	CBSA	County
Index value	.032	.028	.027	.049

Source: CareerBuilder database and authors' calculations.

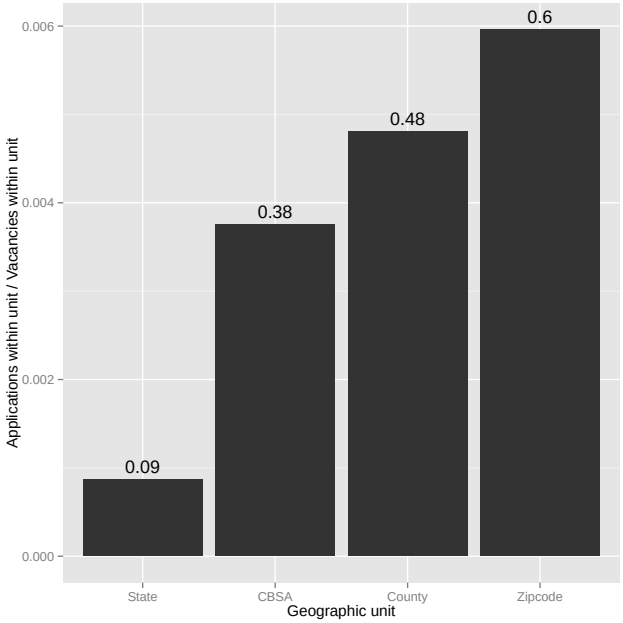
Notes: The two first columns correspond to the index $\tilde{\mathcal{M}}$ computed with our matching function: equation (4) for the distinct version and equation (8) for the overlapping version. The last two columns correspond to the index $\mathcal{M}$ computed under the distinct-market assumption using a Cobb-Douglas with parameter .5; see equation (1).

Figure 1: Share of applications where the vacancy and the job seeker are located within the same geographic area



Source: CareerBuilder database.

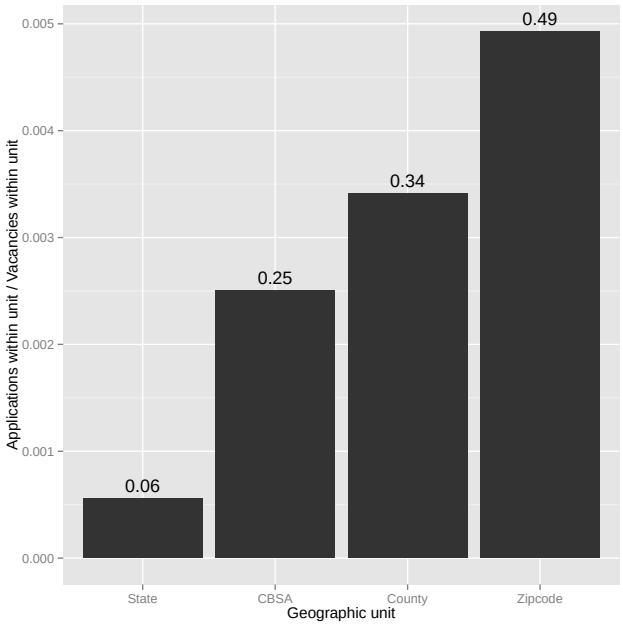
Figure 2: Ratio between the number of applications that a job seeker send in her residence area and the number of available vacancies in this area



Source: CareerBuilder database.

Note: The numbers at the top of the bars are multiplied by 100.

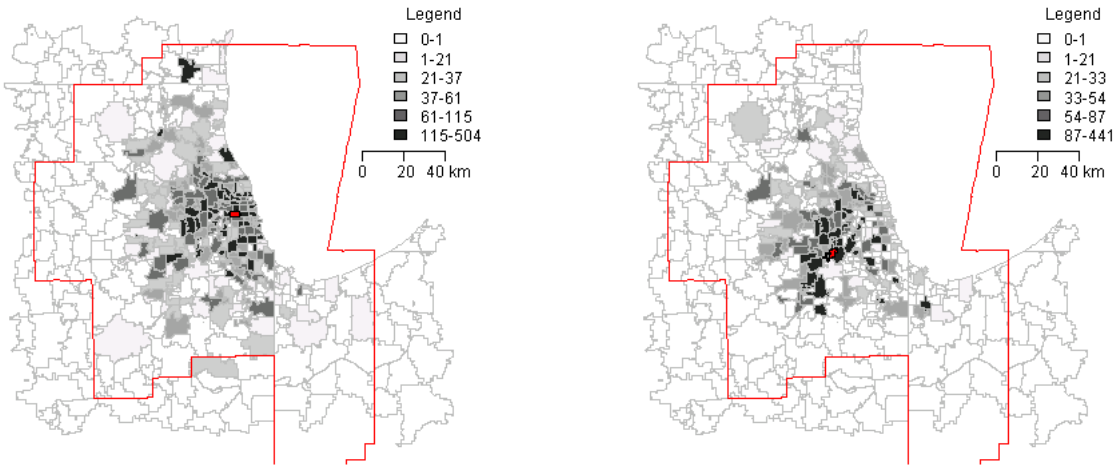
Figure 3: Ratio between the number of applications that a vacancy received from job seekers living in the same area and the number of job seekers within the area



Source: CareerBuilder database.

Note: The numbers at the top of the bars are multiplied by 100.

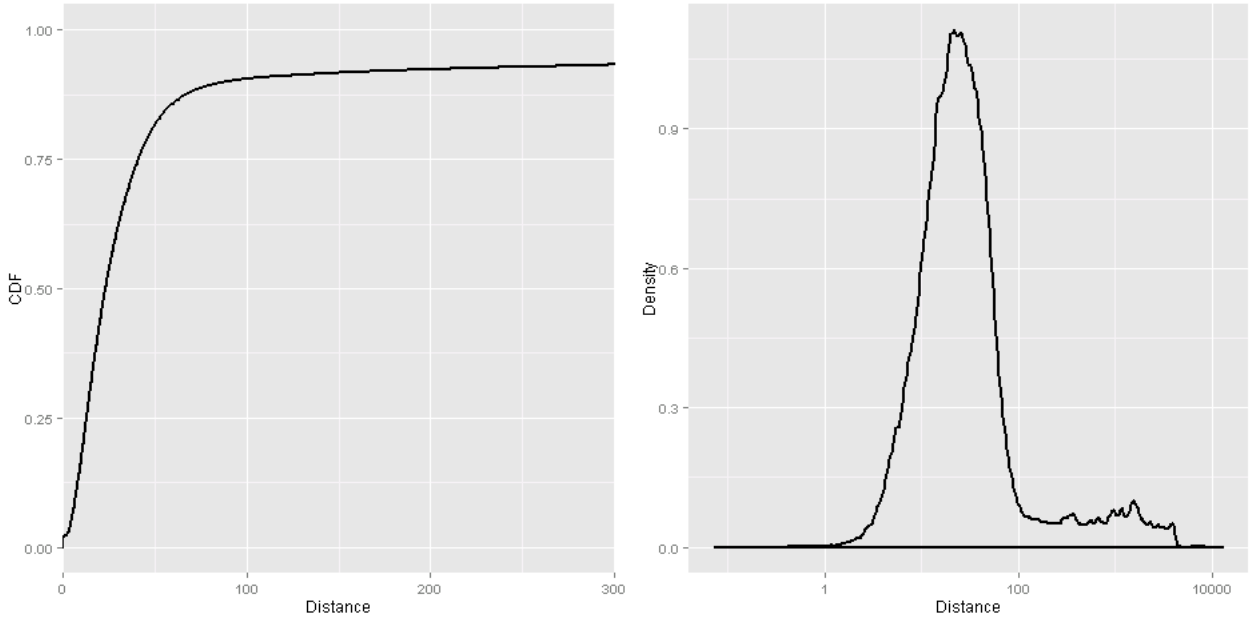
Figure 4: Geographic pattern of the applications sent by job seekers living in Chicago CBSA zip codes 60639 (left panel) and 60561 (right panel)



Source: CareerBuilder database.

Note: The area delimited in red is the Chicago CBSA.

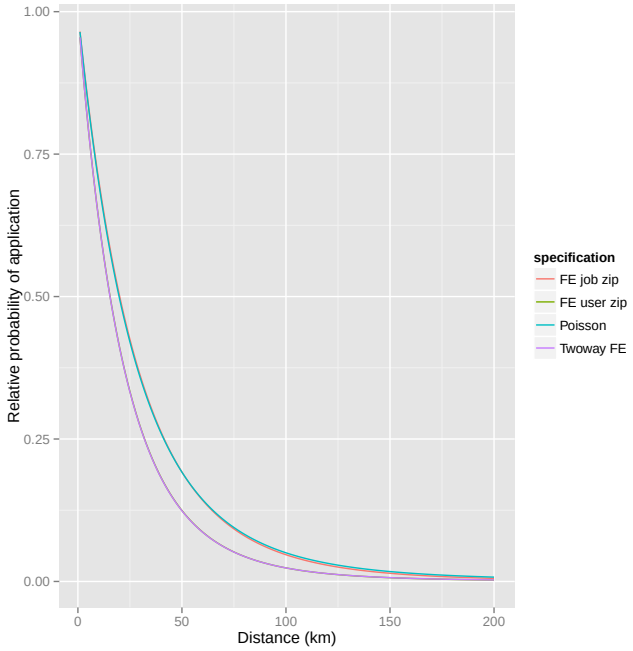
Figure 5: Distribution of the distance between job seekers and vacancies in observed applications: cumulative distribution function (left panel) and probability distribution function (right panel)



Source: CareerBuilder database.

Note: The right panel uses a log scale for the x axis.

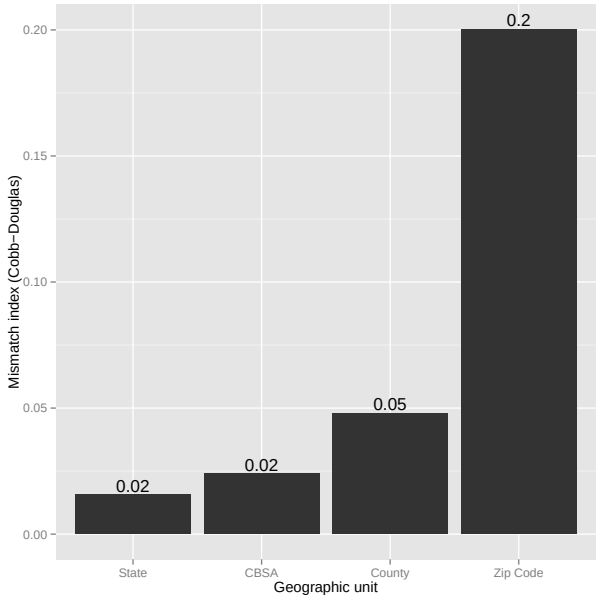
Figure 6: Relative probability of application as a function of geographic distance: predictions from Poisson model with or without fixed effects



Source: CareerBuilder database.

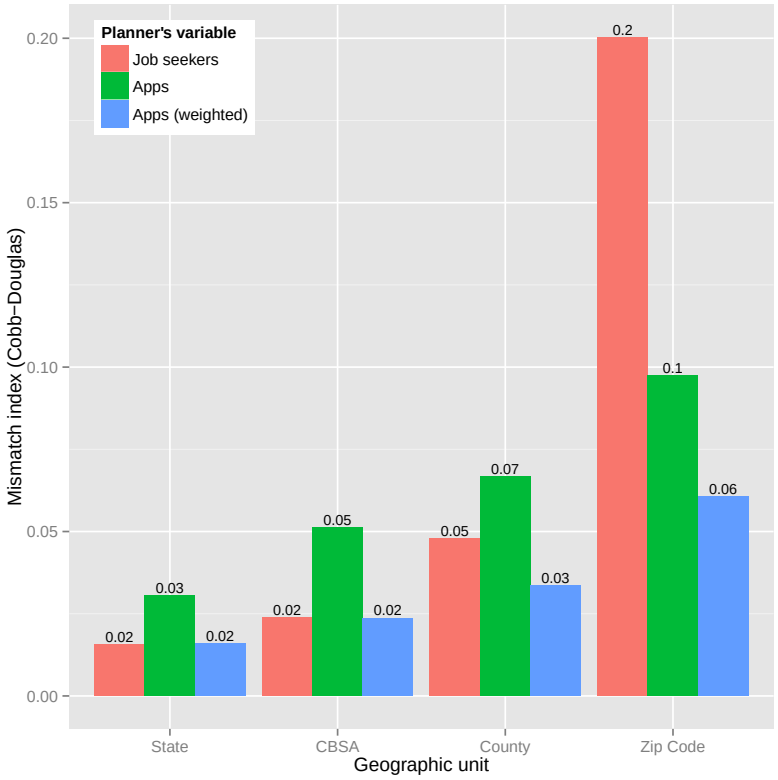
Note: The values of coefficients correspond to the ones reported in the columns 1, 3, 4, and 5 of Table 1.

Figure 7: Mismatch index between job seekers and vacancies, for several geographic definitions of the markets



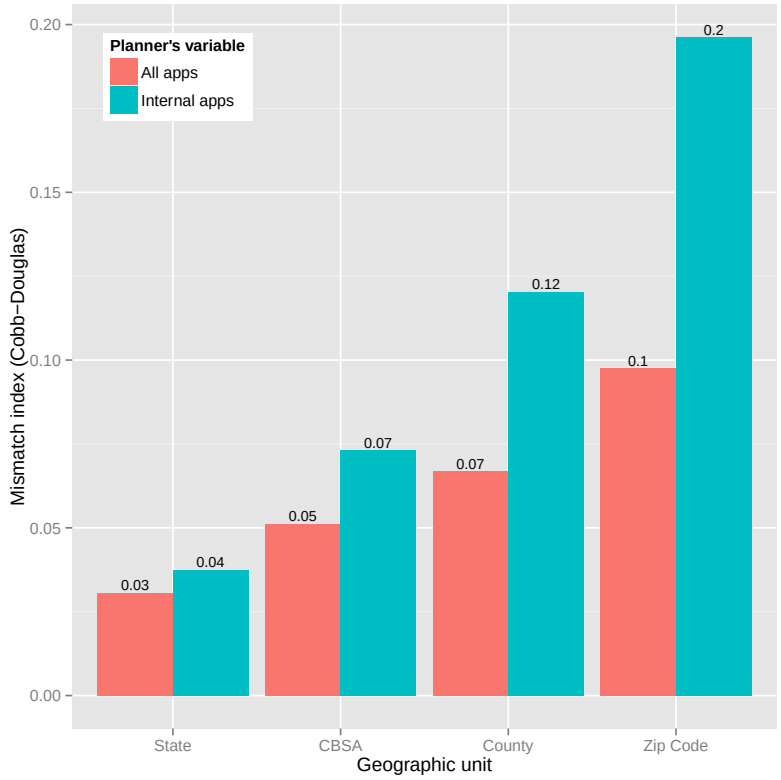
Source: CareerBuilder database.

Figure 8: Mismatch index between jobs seekers (or applications or weighted applications) and vacancies, at several geographic levels



Source: CareerBuilder database.

Figure 9: Mismatch index between applications (or internal applications) and vacancies at several geographic levels



Source: CareerBuilder database.

Appendix 1: A search model with endogenous number of applications

5.1 Distinct markets

Let us now relax the assumption about the exogeneity of the number of applications. Instead, the number of applications is determined endogenously such that the marginal gain of an application is equal to its marginal cost in each spot s. In this case, the probability p_s that a job seeker in s applies to a vacancy located in s is:

$$p_s = \frac{a_s}{V_s}$$

a_s is the total number of applications that a job seeker in s sends to jobs in s. We now determine the optimal value of a_s. To simplify the analysis, we assume that the number of applications is continuous and not discrete. Let π_s denote the probability to be hired associated with an application to a job in s. F_s is the utility that a job seeker experiences if he is hired in a job in s. Before any application is sent, the marginal gain is $\pi_s F_s$ while the marginal cost is denoted by m. When a_s applications have already been sent, a marginal application only brings utility if the previous applications were unsuccessful, i.e. with probability $(1-\pi_s)^{a_s}$. Job seekers stop applying when the marginal gain equals the marginal cost:

$$(1-\pi_s)^{a_s}\pi_s F_s = m$$

π_s will be endogenously determined as it depends on other job seekers' decisions to apply. The more job seekers apply to a job, the higher the competition and the lower the probability to be hired for each individual job seeker. $F_s(a_s)$ is the utility that a job seeker experiences if he is hired in a job in s, conditional on having already sent a_s applications. We assume that F_s decreases with a_s. For simplicity, we assume that $F_s(a_s) = a_s^{-1/\alpha}$. Note that we assume that the utility is idiosyncratic so that workers are not necessarily attracted by the same jobs (there are no good jobs and bad jobs or, more precisely, the definition of a good job depends on each individual). The marginal cost m may also depend on the number of applications. The literature frequently assumes convex costs, so that the marginal cost increases with a_s. We assume that the cost is written ca_s^{η}: if $\eta > 0$, the cost is convex; if $\eta = 0$ the cost is linear; if $\eta < 0$, the cost is concave.

Denoting $\bar{\eta} = 1/\alpha + \eta$, the total number of applications a_s that a job seeker in s sends to jobs in s is determined by equating marginal gains with marginal costs:

$$(1-\pi_s)^{a_s}\pi_s = ca_s^{\bar{\eta}}$$

A closed-form solution for a_s can be found using the Lambert W function:

$$a_s = -\frac{\bar{\eta}}{\log(1-\pi_s)} W\left(-\frac{\log(1-\pi_s)}{\bar{\eta}}\left[\frac{\pi_s}{c}\right]^{\frac{1}{\bar{\eta}}}\right)$$

The Lambert W is increasing in its argument for positive values: therefore, the expression above shows that the number of applications decreases with the cost parameter c, as is intuitive. One can also check that the number of applications increases with the probability of a successful application

π_s. This is an interesting property because it suggests that, when labor market tightness is higher, so that π_s is higher, job seekers send more applications, i.e. job search effort is higher. Therefore, allowing for an endogenous number of applications may have important consequences for the cyclical properties of search and matching models.

5.2 Overlapping markets

We could also assume that the number of applications is determined endogenously such that the marginal gain of an application is equal to its marginal cost. a_i is the total number of applications for a job seeker living in s_i. The marginal gain should be equal to the marginal cost; for all j:

$$\Pi_\ell(1-\pi_\ell)^{b_{i\ell}a_i}\pi_j F_{ij} = m$$

We keep the assumptions we used before for the expressions of the cost term. The utility term F_{ij} is assumed to take the functional form: $F_{ij} = [f(d_{ij})/a_{ij}]^\alpha$. Using these assumptions, the previous equation can be rewritten, for all j:

$$(1-\bar{\pi}_i)^{a_i}\pi_j\left(\frac{f(d_{ij})}{a_{ij}}\right)^{1/\alpha} = ca_i^\eta$$

with $(1-\bar{\pi}_i) = \Pi_\ell(1-\pi_\ell)^{b_{i\ell}}$.

Multiplying both terms by $a_{ij}^{1/\alpha}$ and summing over the j, we obtain an equation that determines a_i.

$$(1-\bar{\pi}_i)^{a_i}\left[\sum_j \pi_j^\alpha f(d_{ij})\right]^{1/\alpha} = ca_i^{1/\alpha+\eta}$$

Denoting $\bar{\eta} = 1/\alpha + \eta$, a closed-form solution can be found using the Lambert W function:

$$a_i = -\frac{\bar{\eta}}{\log(1-\bar{\pi}_i)}W\left(-\frac{\log(1-\bar{\pi}_i)}{\bar{\eta}}\left[\frac{\sum_j \pi_j^\alpha f(d_{ij})}{c^\alpha}\right]^{\frac{1}{\alpha\bar{\eta}}}\right)$$

By examining the expression above, we can see how endogeneizing the number of applications changes the matching process compared to a situation with a fixed number of applications. In particular, the term $\sum_j \pi_j f(d_{ij})$ reflects the fact that job seekers in spot s_i send more applications when there are more vacancies nearby ($f(d_{ij})$ is small for many spots s_j) and when π_j, the probability of getting an offer from these jobs in s_j, is higher. As in the case of distinct markets, this means that job seekers send more applications where labor market tightness is higher so that there is low competition for jobs. What is different from the case of distinct markets is that labor market tightness is considered not only for the job seekers' own place of residence but also for other markets. The share of applications send by job seekers in i to jobs in j, b_{ij}, is determined as previously:

$$b_{ij} = \frac{\pi_j^\alpha f(d_{ij})}{\sum_\ell \pi_\ell^\alpha f(d_{i\ell})}$$

The probability p_{ij} that a job seeker in s_i applies to a vacancy located in s_j is equal to:

$$p_{ij} = \frac{a_i}{V_j}b_{ij} = \frac{a_i\pi_j^\alpha f(d_{ij})}{V_j\sum_\ell \pi_\ell^\alpha f(d_{i\ell})}$$

Appendix 2: The algorithm used to find the planner's optimal allocation

The planner's problem is to find the allocation of job seekers which maximizes the number of matches in the economy:

$$M^* = \max_u M(u) \text{ s.t. } \sum_s u_s = \sum_s U_s$$

When we solve this program at the zip-code level, the dimension of the vector u is more than N =25,000. Moreover, the problem is constrained by the fact that using classical (Newton-Raphson inspired) tools for optimization under constraints proves not to be manageable in periods in days or weeks, especially because each evaluation of the function takes a few seconds (around 4). Computing the numerical Hessian once would, for instance, take roughly 4 N^2 seconds, that is 500 million seconds, or more than 15 years. Even parallelizing the problem would not help much.

In order to find the optimal solution, we use the following algorithm that allow us to find the solution in one or two days. The whole optimization routine is coded under Matlab. First, we try to find the best possible starting value. We use the solution of the distinct-market model as the starting value for the fuzzy-market case. The solution of the distinct-market model is obtained using a similar algorithm as the one we describe now, except that the framework is much simpler and the solution is reached in a couple of minutes.

The general idea of the algorithm is to compute the vector of discrete gradients, $g_s(u) = M(u + \delta_s) - M(u), \forall s$, where δ_s is a vector with 1 at position s and zeros elsewhere. At the equilibrium allocation, the planner should be indifferent between adding a job seeker in the different units. In other terms, he should be unwilling to remove a job seeker from a unit to allocate her to another unit. Therefore, the goal is to find the allocation that equalize the $g_s(u), \forall s$. Note that $M(u)$ are increasing and concave in u, so that $g_s(u)$ are positive but decreasing in u.

We proceed in two stages. The first stage is heuristic and only approximately converging to the result but is faster. First, we compute the discrete gradients $g_s(u)$ taking advantage of the matrix structure (enough RAM should be available so that several matrices $N \times N$ can be stored and manipulated). Second, we take all the units such that the gradient is below a given share of the median of $g_s(u)$ and have at least one job seeker. We construct a try allocation by removing one job seeker to all these units and allocating them to the units with the highest $g_s(u)$. We only keep the try if the number of matches is increased, and we iterate. At some point, this stage does not improve the result and we switch to the slower second stage.

The second stage also starts by computing the discrete gradients. The minimum and the maximum are isolated and one job seeker is transfered from the minimum to the maximum. Several thousands of iterations are necessary to reach the maximum.