How Local Are Labor Markets?
Evidence from a Spatial Job Search Model

Alan Manning

London School of Economics and CEP (LSE)

Barbara Petrongolo*

Queen Mary University and CEP (LSE)

March 2017

Abstract

This paper models the optimal search strategies of the unemployed across space to characterize local labor markets. Our methodology allows for linkages between numerous areas, while preserving tractability. We estimate that labor markets are quite local, as the attractiveness of jobs to applicants sharply decays with distance. Also, workers are discouraged from searching in areas with strong job competition from other jobseekers. However, as labor markets overlap, a local stimulus or transport improvements have modest effects on local outcomes, because

ripple effects in job applications dilute their impact across a series of overlapping markets.

In recent years there has been a resurgence of interest in the consequences of localization of economic activity for workers' welfare (Moretti, 2011) and in policies aimed to improve labor market outcomes in disadvantaged areas (see Glaeser and Gottlieb, 2008, for a survey, and recent work by Busso, Gregory and Kline, 2013). In the US, federal, state and local governments combined spend nearly \$50 billion per year on local development policies. The returns to these policies depend crucially on the effective size of local labor markets. If labor markets are very local, an effective intervention needs to be targeted to the disadvantaged areas themselves and more distant interventions will not benefit the target group. But if labor markets are not as local, targeted intervention is ineffective as it may simply benefit workers from other, more advantaged areas. A broader question concerns the incidence of local shocks to labor demand and their impact on labor mobility and labor market equilibrium (see among others Blanchard and Katz, 1992, Bound and Holzer, 2000, and Notowidigdo, 2013).[1] In addressing these and related questions, the size of local labor markets is important in so far as it helps to define appropriate treatment and control areas for the evaluation of local policies and shocks.

Most research on the topic relies on divisions of geographical space into a relatively small number of non-overlapping areas, which are either purely administrative units (e.g. states or counties in the US), or are intended to be self-contained labor markets (examples would be the BEA's 179 Economic Areas and the 720 Commuting Zones for the US, or the 320 Travel to Work Areas for the UK). These classifications are very valuable for understanding

---

[1] Another related issue is the spatial mismatch hypothesis (Kain, 1968), suggesting that the unemployment rate of blacks in the inner city was so high because many jobs had moved to the suburbs and these jobs were no longer in the local labor market of those living in the city (see also Hellerstein, Neumark and McInerney, 2008, and Boustan and Margo, 2009, and Zenou, 2013, for recent studies).

2

spatial differences in economic outcomes, but suffer from important limitations when it comes to research into local labor markets. First, the cost of distance within such areas is implicitly assumed to be zero. Because people commute large distances to work in the centre of big cities, large metropolitan areas are generally classified as single labor markets. But those who live in the northern suburbs might not think of the southern suburbs as part of their labor market. Second, the non-overlapping nature of labor markets constructed in this way causes inevitable discontinuities around the boundaries. Someone living just inside a large metropolitan area would be classified as living in a large labor market, while someone living just across the border would be classified as living in a modestly-sized labor market. However, these individuals live in essentially the same labor market.

In reality, the economy cannot be divided into non-overlapping segments, and there is always some commuting across borders (see, for example, Monte, Redding and Rossi-Hansberg, 2015; Amior and Manning, 2015). Typically, the labor market for one individual at one location overlaps with that for a second individual at a different but not too distant location, whose labor market then overlaps with that for a third individual, whose labor market may not overlap at all with that of the first individual. The aim of this paper is to develop a model of local labor markets that takes into account overlaps and associated interdependencies, and to derive implications for the evaluation of local policies.

We argue that a structural model is necessary to understand the anatomy of local spillovers and to evaluate the impact of local intervention, and propose a job search model that allows for linkages across a large number of small areas. Jobseekers decide whether to apply to job vacancies at different locations, based on the probability of an application being successful, in turn depending on how many other jobseekers are applying to these jobs, and on the utility enjoyed on the job, which depends on the distance to it and the wage paid. Linkages across areas arise because the number of applicants to jobs in a given area is likely to be influenced, even if only very slightly, by unemployment and vacancies in all other areas, as they are ultimately linked

3

through a series of overlapping markets. We relate job matches in a ward to the number of applications received by local vacancies, as predicted by our job search model, taking as given the distribution of unemployment and vacancies. We estimate the model parameters by minimizing the distance between the model prediction and the observed vacancy filling rate.

In our empirical analysis we use unemployment and vacancy data on 8850 Census wards in England and Wales, and combine these with micro data on wages and the use of transport modes, which allow us to model commuting costs between any two wards in the economy. Our estimates show evidence of high costs of distance. For example, the probability of a random job 5km distant being preferred to random local job is only 19 percent. We find that workers are discouraged from applying to jobs in areas where they expect relatively strong competition from other jobseekers, and the hypothesis of constant returns in search markets is not rejected, implying that larger scale markets do not systematically offer more efficient matching of workers to jobs. The estimated model predicts commuting patterns across UK wards which replicate very accurately actual commuting patterns, even though commuting data are not used for estimation.

The paper makes a number of contributions. First, on a methodological level, our proposed model allows interdependencies across a very large number of areas, while preserving tractability in analysis and estimation. Despite the fact that workers in any of 8850 wards may apply to jobs in any of 8850 locations – with over 78 million combinations of origin and destination wards – and the decision of each worker is influenced by the search strategies of other workers, we show that the equilibrium allocation of applications can be solved for using an efficient contraction mapping.

Second, we provide microfoundations for the effective size of local labor markets. We characterize the size of local labor markets based on our estimate for the cost of distance, as this is the main determinant of the set of jobs that an unemployed worker, currently in a particular location, is willing to apply for. The cost of distance is identified as the rate of decline of job applications with distance from the applicant's location, and its estimate embodies both the

4

actual time and cost of travelling and the extent to which the attractiveness of jobs varies within areas. Intuitively, workers are prepared to travel longer distances when they expect wider variation in jobs' attractiveness within areas, as this makes it easier to find a distant job that provides higher utility than a local job.

Third, we provide a deeper understanding of the likely impact of place-based policies and how they should be evaluated. The observation that commutes are generally short or, equivalently, that our estimated cost of distance is relatively high, may suggest that local intervention would have heaviliy concentrated effects on target areas. We show that this argument is deceptive if labor markets are overlapping, an insight for which structural modelling is critical. Even though the labor market for an individual worker may be quite local, a local shock sends a ripple effect through surrounding areas, diffusing its impact over a much wider area than the typical commute. The extent of the ripple effect may be limited by "firebreaks", i.e. natural or institutional borders across which few workers commute.[2] Ripple effects have important implications for the evaluation of place-based policies, as defining small treatment areas around the shock location risks missing most of the impact and possibly contaminating control areas. We suggest the treatment area should be considerably larger than the median commute, although a test based on a large treatment area, across which impact is highly diluted, may lack statistical power, unless the scale of the shock is commensurate to its (relatively large) size. In summary, ripple effects imply that typical evaluations of place-based policies may either yield biased estimates of the effects of interest or be under-powered.

The plan of the paper is as follows. The next Section describes the data sets used. Section 3 proposes a model of job search across space. The model incorporates interdependencies across a large number of small areas and illus-

_____

[2]Statistical authorities attempt to find such firebreaks when defining local labor markets, as the criterion typically used is based on limited cross-border commutes, relative to within-border commutes. Indeed our model provides a framework justifying the criterion used to define local units. Nonetheless, it is important to recognize that there is considerable cross-border commuting across available measures of commuting zones.

trates that the size of labor markets can be inferred from the optimal search strategies of the unemployed. Section 4 reports our main estimates and compares the model's predictions to actual commuting patterns. Section 5 uses the model estimates to map the simulated impact of place-based policies on the spatial distribution of the unemployment outflow. Section 6 concludes.

I   Data and descriptive statistics

The geographic units used in this paper are Census Area Statistics 2003 Wards for England and Wales. There are 8850 CAS wards in England and Wales, with an average population of about 5900.

We use data on several labor market indicators at the ward level, drawn from various sources. Data on unemployment (the "claimant count") and vacancies are from the UK Public Employment Service (PES). The series used are produced by the Department for Work and Pensions and available through NOMIS (nomisweb.co.uk). The UK PES is structured as a network of government funded employment agencies (Jobcentres), whereby each town or neighborhood within a city has at least one Jobcentre. Jobcentre services are free of charge both to jobseekers and employers. To be entitled to receive welfare payments, unemployed benefit claimants are required to register at a Jobcentre, and sign-on every two weeks. The UK PES is much more widely used than its US equivalent (see Manning, 2003, Table 10.5, and OECD, 2000).

Employers wishing to advertise job vacancies submit a form with detailed job specifications to a centralized service called Employer Direct. Each vacancy is assigned to the employer's local Jobcentre, and has a dedicated recruitment adviser, who can assist with the recruitment process. Regardless of the Jobcentre in charge, the ward for each vacancy is defined using the full postcode of the job location. Vacancies are advertised on the centralized employment website https://jobsearch.direct.gov.uk; through the Jobcentre phone service; and on the Jobcentre network, which can be accessed at Jobcentre offices around the country. Jobseekers can sample job openings via one or more of these methods, using various search criteria (sector, occupation, working hours, distance from

6

a given postcode etc.). The detailed geographic information on both claimant unemployment and job vacancies recorded at Jobcentres makes them a unique data source for studying job matching patterns at the very local level. The series run monthly from April 2004 onwards. However, we restrict our sample period to April 2004-April 2006, because changes introduced in May 2006 to vacancy handling procedures make later vacancy series less suitable for our purposes.[3]

As not all jobseekers are claimant unemployed and not all vacancies are advertised through the PES, our data cannot cover the entire job matching activity in the economy. Appendix A provides a more detailed discussion of data coverage, and shows that our data capture an important section of both supply and demand of the job search process in the UK, especially for low-skilled workers and jobs. But imperfect coverage would be problematic if the portion of the job search process covered by our data varies systematically across areas, something on which unfortunately we have no information. As a check against the possibility of biases, we assess the power of our job search model in explaining commuting patterns across wards using a representative sample of UK employees, independently of how they searched for jobs.

In the data presented below and all estimated specifications, we obtain the vacancy and unemployment outflows as differences between the corresponding inflows and the monthly variations in the stocks. Due to measurement error, for about 0.5 percent of observations the vacancy outflow implied by the stock-flow accounting identity is negative, and we drop the corresponding observations. Table A1 in the Appendix presents descriptive statistics on unemployment and vacancy stocks and flows from May 2004-April 2006, a period of historically low and stable unemployment, and Table A2 shows pair-

---

[3]Prior to May 2006, vacancies notified to Jobcentres were followed up with the employer until they were filled, and the number of vacancies filled at a Jobcentre was used as one of the main indicators of its performance. From May 2006, the Jobcentre performance evaluation is no longer based on vacancies being filled, thus vacancies notified to Jobcentres are not followed-up, and have an ex-ante closure date agreed with the employer, upon which they are automatically withdrawn. This systematically understates the stock of unfilled vacancies from May 2006 onwards.

wise correlations.[4] Wards have on average 106 unemployed and 92 vacancies. The average vacancy outflow rate is about one third. Most of the variation in unemployment and vacancies is across, rather than within, wards, reflecting wide variation in wards' size, while the unemployment-to-vacancy (U-V) ratio varies widely across both space and time. Figure A1 shows spatial variation in the U-V ratio for a representative month in our sample, February 2005. Different wards are shaded according to the quartile of the corresponding U-V ratios, with darker shades representing higher quartiles. There is no obvious pattern emerging from this map – rather we observe a patchwork of very different labor market outcomes across quite small areas, and there are no large regions in which, say, all high-unemployment wards are clustered together.

Additional data used are ward-level wages and commuting flows across wards. Wage data are drawn from two sources. First, we construct a proxy of ward-level wages by combining data on the industry structure of employment at the ward level with industry-level earnings, exploiting the local variation in the composition of employment and systematic inter-industry wage differentials to predict ward-level wages. The industry composition of wards is obtained from the Business Register and Employment Survey (BRES) for the period 2004-2006, and is measured at the 4-digit level, and hourly earnings are obtained from the Annual Survey of Hours and Earnings (ASHE) for the same period.[5] The resulting ward-level wage measure is given by

$$\widehat{W_b} = \sum_j s_{jb}\omega_j, \tag{1}$$

where s_{jb} denotes the share of industry j in employment of ward b, and ω_j denotes industry-specific median earnings. The ward-level average of $\widehat{W_b}$ is about 9£ per hour, in 2004 £. In our estimated model, we include this wage

---

[4]Unemployment and vacancy measures are "time aggregated", obtained as the sum of the stock measured at the end of the previous month, plus the monthly inflow (see Berman, 1997). By using lagged stocks, we lose the initial month in the sample period.

[5]The BRES is the official, plant-level, source of employee estimates by detailed geography and industry. The ASHE is an employer-based survey, covering a 1% random sample of employees in the UK.

measure among the components of the utility of working in a given ward.

Second, we use data on actual earnings available in ASHE. However, the number of wage observations per ward is small. The median cell size is 40, but 30 percent of wards have 20 wage observations or less, giving a very noisy measure of ward-level wages. In order to avoid measurement error biases due to small cell size, we do not use actual wages on the righ-hand side of our estimated models. However, we use actual wages to test the model prediction that wages should, ceteris paribus, be negatively affected by the expected number of applicants.

Data on commuting flows are obtained from the restricted access version of ASHE, which contains information on postcode of work and postcode of residence of employees, and we assign postcodes to wards to characterize commuting patterns across wards. Nearly half of employees in England and Wales have commutes shorter than 6 km, and two thirds have commutes shorter than 10 km. We will compare the distribution of actual commutes to the distribution of commutes predicted by our model estimates.

II   The job search model

This Section builds an estimable model of job search behavior across space. Absent data on job applications, a structural model is necessary to understand the anatomy of local spillovers and to evaluate the impact of local intervention.

Our approach develops in four steps. We first model the job search behavior of individuals, directing their applications to vacancies in any area that offer the highest expected utility. Second, we specify expected utility as the product of the utility the worker would enjoy on the job — which depends on the wage earned and commuting costs — and the probability that her application is successful — which depends on the number of applicants and, hence, the job search behavior of others. Third, we characterize a fixed point in the mapping across job applications, so that the spatial distribution of job applications is the best response to itself. Finally we relate job matches in a ward to the number of applications received by local vacancies, as predicted by the model.

9

This is the relationship that we bring to the data in Section III.

A   The job application process

There are U_a unemployed workers and V_a vacancies in each area a of the economy. Each worker decides which of the existing vacancies to apply for, and applications are simultaneous, in the non-sequential search tradition of Burdett and Judd (1983). Assume that the expected utility for an unemployed worker in area a applying to a vacancy in area b can be written as $\Omega_{ab}\varepsilon_i$, where Ω_{ab} depends on the probability that the application is successful and the attractiveness of working in b for a resident of a (specified in detail later), and ε_i is an idiosyncratic utility component, which is assumed to be i.i.d. across workers and vacancies. Such idiosyncratic component may capture wage variation around the area mean, non-monetary components to utility, or any individual-specific preferences about living in a and working in b.

Workers are assumed to have a cost function for sending N applications of the form:

$$C(N) = \frac{c}{1+\eta}N^{1+\eta}.$$

Under the assumption that the probability of more than one application being successful is infinitesimal,[6] the net expected utility from job search can be

---

[6] The assumption that the probability of more than one application being successful is infinitesimal plays an important role. If this assumption is not met, vacancies cannot be ranked by the expected utility offered, and this substantially complicates the worker decision problem. To see this more formally, let's denote the expected utility from a vacancy as pu where p is the probability of an application being successful and u is the utility enjoyed if successful. Suppose that jobs can be ordered in terms of utility, i.e. $u_1 > u_2 < ...$ Furthermore, assume that jobs that offer a higher utility have a lower probability of success, so that $p_1 < p_2 < ...$ In this case a worker only accepts job i if no job applications to lower jobs have been successful. The expected utility from applying to a set of jobs is thus $\sum_i D_i p_i u_i \Pi_{j=1}^{i-1}(1 - p_j)^{D_j}$. This leads to a decision rule in which the marginal benefit of applying to vacancy i can be written as $p_i(u_i - E_i)(1 - Q_i)$, where Q_i is the probability of getting a better job than i and E_i is the expected utility from jobs worse than i, conditional on a better job not being obtained. The effect of other applications on the decision to apply to vacancy i is no longer limited to their effect on marginal costs. But the difference between this specifciation of marginal benefit and the one we use is small if Q_i and E_i are small. Chade and Smith (2006) provide a more complete analysis of optimal decision rules in this case.

expressed as:

$$\sum_{b,i} D_{bi}\Omega_{ab}\varepsilon_i - \frac{c}{1+\eta}\left(\sum_{b,i} D_{bi}\right)^{1+\eta},$$

where D_{bi} is a binary variable taking the value 1 if a worker applies to vacancy i in area b and zero otherwise. Mazimization of expected utility implies that a worker applies to all vacancies for which the expected utility of doing so is higher or equal to the marginal cost $C'(N)$:[7]

$$\Omega_{ab}\varepsilon_i \geq c\left(\sum_i D_i\right)^{\eta} = cN^{\eta}. \tag{2}$$

If we assume that ε is Pareto distributed with exponent k, the probability that a worker in a applies to a vacancy in b can be written as:

$$\Pr\left(\Omega_{ab}\varepsilon_i \geq cN_a^{\eta}\right) = \left(\frac{\Omega_{ab}}{cN_a^{\eta}}\right)^k,$$

where N_a denotes the total number of applications sent by a worker in a. The number of expected applications sent from each worker in a to vacancies in b, denoted by N_{ab}, is thus given by:

$$N_{ab} = V_b\left(\frac{\Omega_{ab}}{cN_a^{\eta}}\right)^k. \tag{3}$$

---

[7]In the interests of simplicity we do not impose the number of applications to be an integer. Imposing this would yield a first-order condition for the optimal number of applications in the form of an inequality, which would be much harder to treat analytically. In addition, there may be mixed strategy equilibria, in which case the average number of applications could be a non-integer. One could reinterpret the number of applications in this model as a decision about their (continuous) search effort, combined with a decision about the distribution of such effort across vacancies at different locations. This mechanism would deliver a specification relating the vacancy outflow rate in an area to the number of applicants, where the number of applicants is re-interpreted as the rate at which job seekers apply to vacancies. This formulation is more similar to that used in the sequential search literature, e.g. Pissarides (2000).

Summing the N_{ab} terms across all possible destination areas b yields:

$$N_a = \sum_b V_b \left(\frac{\Omega_{ab}}{cN_a^\eta} \right)^k ,$$

which can be solved for N_a :

$$N_a = \left(c^{-k} \sum_b V_b \Omega_{ab}^k \right)^\gamma , \tag{4}$$

where $\gamma = 1/(1 + \eta k)$. The parameter η (or - equivalently - its transformation γ) is related to the returns to scale in the matching function. The issue of constant versus increasing returns is a recurrent question in the matching literature, as increasing returns lead to the possibility of multiple equilibria (Diamond, 1982). If $\eta = 0$ ($\gamma = 1$), the marginal cost of an application is constant and a worker applies to a vacancy if the expected utility of doing so is above this marginal cost. In this case doubling the number of vacancies leads to a doubling in the number of applications each worker makes. The average number of applicants per vacancy remains unchanged, and so does the probability of filling each vacancy. The total number of matches also doubles, thus there are constant returns to scale to vacancies alone. But if one doubles both vacancies and the number of unemployed workers, the number of applications will rise four-fold, as both the applications per worker and the number of workers double. This implies increasing returns to scale.

At the other extreme, consider $\eta = \infty$ ($\gamma = 0$). In this scenario each worker has a fixed number of applications to make, and will direct them to vacancies that offer the highest expected utility. A doubling of vacancies and unemployment leads to a doubling of applications, as applications per worker are unchanged and the number of workers has doubled. Hence applications per vacancy are unaltered, the probability of filling a vacancy is unaltered, and the total number of matches will double. This implies constant returns to scale.

Our set-up makes it harder to rationalize the possibility of decreasing returns, which would require some extra form of congestion in the model. How-

ever, our model is consistent with decreasing returns to vacancies and unemployment in individual areas, as doubling vacancies and unemployment in a particular area would result in a lower probability of filling jobs in that area due to spillovers across neighboring areas.

Combining (3) and (4) gives a solution for the number of applications made by workers in a to vacancies in b:

$$N_{ab} = c^{-k\gamma} V_b \Omega_{ab}^k \left(\sum_{b'} V_{b'} \Omega_{ab'}^k \right)^{\gamma - 1}. \tag{5}$$

The intuition behind (5) is that the number of applications sent from area a to area b depends on job opportunities in area b (V_b) and the expected utility from those jobs (Ω_{ab}). The term in brackets can be interpreted as a weighted average of vacancies across the whole economy, where weights are given by the expected utility offered to residents of a. This term captures the "effective" size of the economy, and would simply work as a normalization in the case of constant returns ($\gamma = 0$). When $\gamma = 0$, (5) has clear similarities with a logit model.

The number of applications received by vacancies in b is given by the sum of applications that workers in any area a send to area b. Thus the ratio of applications per vacancy in b, denoted by A_b, is given by

$$A_b = \frac{\sum_a N_{ab} U_a}{V_b} = c^{-k\gamma} \sum_a U_a \Omega_{ab}^k \left(\sum_{b'} V_{b'} \Omega_{ab'}^k \right)^{\gamma - 1}. \tag{6}$$

This is an expression for market tightness in area b and states that the expected number of applications per job in b depends on the distribution of the unemployed across all possible origin areas a (U_a) and the expected utility they would enjoy from jobs in b (Ω_{ab}^k).

To make equation (6) operational, we assume that expected utility can be expressed as:

$$\Omega_{ab} = p(A_b) W_b f_{ab}, \tag{7}$$

13

where $p\left(A_b\right)$ is the probability of an application being successful, assumed to depend negatively on the number of applicants, W_b is the wage offered by jobs in area b, and f_{ab} represents the intrinsic attractiveness of a job in area b for a resident of a, which is primarily a function of distance.

Substituting (7) into (8) leads to the key result of our spatial job search model:

$$
A_b = \frac{\sum_a N_{ab} U_a}{V_b} = c^{-k\gamma} \sum_a U_a \left(p\left(A_b\right) f_{ab} W_b\right)^k \left[\sum_{b'} V_{b'} \left(p\left(A_{b'}\right) f_{ab'} W_{b'}\right)^k\right]^{\gamma-1}.
\tag{8}
$$

Expression (8) captures all interdependencies across areas. According to (8), the number of applications to jobs in b is likely to be influenced (even if only very slightly) by unemployment and vacancies in all other areas, because they are ultimately linked through a series of overlapping labor markets. With 8,850 wards, this expression represents a system of 8,850 equations in 8,850 unknowns. But Appendix B shows that, under reasonable conditions, (8) is a contraction mapping, which can be solved iteratively and efficiently to obtain A_b. This is the feature that allows us to estimate a model with a very large number of areas.

It is helpful to highlight the relationship between our model of job search, in which vacancies receive a number of applications and then, possibly, choose one of them, and the more common modelling strategy based on a flow arrival rate of job applicants, in which the first acceptable candidate is chosen (e.g. Pissarides, 2000). In our model one could reinterpret the number of applications as a decision about the rate at which one applies for jobs, combined with a decision about the distribution of these applications over vacancies at different locations. This mechanism would also lead to a specification relating the vacancy outflow rate to the number of applicants, in which the number of applicants is re-interpreted as the rate at which job applicants apply to the firm.

Our approach also has similarities to the way in which markets for differentiated goods are represented in industrial organization models, using contrac-

tion mappings to characterize equilibrium (see Berry, Levinsohn and Pakes, 1995). One can think of a product as being a job in a particular area. Compared to most IO applications, we have strong a priori information on which of these products are the closest substitutes – those closer in space – which allows us to reduce the dimensionality of product heterogeneity. Consumers (workers in our applications) are also differentiated by space, and one can think of information on unemployment and vacancies as being information on the distribution of different consumer and product types. Job applications play the role of prices, in the sense that more applications discourage consumers from purchasing a product of a particular type and encourage them to divert their demand to other products. Our outcome variable, the number of matches, represents the market outcome in a quantity space. The equation we estimate is essentially a reduced-form equation for the quantity traded as a function of the demand and supply fundamentals, and the demand function can be identified under the assumption of exogenously fixed supply of vacancies.

B   Endogenous Wages

The discussion so far has treated wages in an area as exogenous, but in reality wages are likely to respond, among other factors, to the ease of filling vacancies. We next endogenize wages, by assuming that they are set to maximize the expected profits from a vacancy, as given by:

$$\Pi = (Y_b - W) \Psi (A_b (W)),$$

where Y_b denotes productivity, potentially varying by area, $\Psi (A)$ denotes the probability of filling a vacancy that receives A applications, and $A_b (W)$ is the number of applications to a vacancy in area b, paying wage W. We assume $\Psi (A)$ is increasing and concave in A, and bounded between zero and one, with $p(A) = \Psi (A) /A$.

The first-order condition for wage setting can be written as:

$$W = (Y_b - W) \varepsilon_{\Psi A} \varepsilon_{AW}, \tag{9}$$

where $\varepsilon_{\Psi A}$ is the elasticity of the probability of filling a vacancy with respect to the number of applications and ε_{AW} is the elasticity of applications with respect to the offered wage. To derive ε_{AW}, note that, using (3), applications per job in area b must solve:

$$A_b = \sum_a U_a \left(\frac{p(A_b) W f_{ab}}{c N_a^\eta} \right)^k,$$ (10)

where N_a is exogenous to the individual firm in b. Using (10):

$$\varepsilon_{AW} = \frac{k}{1 - k\varepsilon_{pA}} = \frac{k}{1 - k(\varepsilon_{\Psi A} - 1)},$$ (11)

where the second equality follows from $\varepsilon_{pA} = (\varepsilon_{\Psi A} - 1)$. Substituting (11) into (9) and re-arranging yields:

$$W_b = \frac{k}{1 + k} \varepsilon_{\Psi A} (A_b) Y_b.$$ (12)

This result implies that the wage offered in an area is increasing in local productivity, and is influenced by the number of applicants whenever this affects the elasticity of the probability of filling a job $\varepsilon_{\Psi A} (A_b)$. It is reasonable to expect that $\varepsilon_{\Psi A} (A_b)$ declines with A_b – and the functional form used below does have this feature – implying that offered wages decline with the expected number of applicants. This can be thought of as the conventional relationship between wages and labor market tightness. Substituting (12) into (8) gives applications per job:

$$A_b = c^{-k\gamma} \sum_a U_a \left(\frac{k}{1+k} \Psi' (A_b) Y_b f_{ab} \right)^k \left[\sum_{b'} V_{b'} \left(\frac{k}{1+k} \Psi' (A_{b'}) Y_{b'} f_{ab'} \right)^k \right]^{\gamma-1}.$$ (13)

Both the exogenous and endogenous wage models will be estimated in Section III. To make these models estimable we need to impose further structure on their components, as discussed next.[8]

---

[8] As is typical in a search environment, decentralized equilibrium in our model is in

C   The Cost of Distance Function

The main determinant of the attractiveness of a job in b for someone living in a, f_{ab}, is the cost of commuting between a and b. We model the cost of distance between any two wards based on a framework of transport choice that originated in McFadden (1973) and is widely used in the transportation literature (see, among others, Small, Winston and Yan, 2005). Consider an individual living in a and working in b, at distance d_{ab}, with a choice of transport mode $m = 1, .., M$ (walking, bike, car or public transport in our application) for the journey. The utility of using mode m can be written as:

$$u_m = \widetilde{\delta}_{0m} - \widetilde{\delta}_{1m} d_{ab} + \xi_m,$$

where $\widetilde{\delta}_{1m}$ denotes the cost of distance for mode m, reflecting a combination of time and monetary costs. For example, walking involves negligible monetary costs but considerable time costs, while driving has higher monetary costs but lower time costs. $\widetilde{\delta}_{0m}$ measures the attractiveness of mode m for a journey of distance zero, and ξ_m measures uncertainty about the time and monetary costs involved. Under the assumption that ξ_m has an extreme value distribution, the probability of choosing mode m is given by the logit model:

$$\Pr\left(Y = m \,|\, a, b\right) = \frac{e^{\widetilde{\delta}_{0m} - \widetilde{\delta}_{1m} d_{ab}}}{\sum_{m'} e^{\widetilde{\delta}_{0m'} - \widetilde{\delta}_{1m'} d_{ab}}}.$$

It can be shown that the expected utility for the choice of alternatives modes of transport is represented by the inclusive value. This is given by:

$$IV_{ab} = \ln \sum_{m'} e^{\widetilde{\delta}_{0m'} - \widetilde{\delta}_{1m'} d_{ab}}.$$

---

general inefficient. When a firm creates a vacancy in a certain area, she makes it harder for firms in the same area and other surrounding areas to fill their vacancies, as the number of applications per vacancy falls, but she makes it easier for unemployed workers to receive a job offer. Therefore vacancy creation generates both negative and positive externalities, whose relative size cannot be a priori determined, which are not internalized in wage setting. In a spatial job search model, there are thus issues about the efficiency of the geographic distribution of vacancies, as well as about their overall number.

The parameters of the logit model can only be identified relative to some base category, which we assume to be driving. This is the most common commuting mode, used by a non-trivial proportion of people for journeys of all distances. Having denoted driving by $m = 1$, the logit model can be rewritten as:

$$\Pr\left(Y = m \mid a, b\right) = \frac{e^{\delta_{0m} - \delta_{1m} d_{ab}}}{\sum_{m'} e^{\delta_{0m'} - \delta_{1m'} d_{ab}}}, \tag{14}$$

where $\delta_{0m} \equiv \widetilde{\delta}_{0m} - \widetilde{\delta}_{01}$ and $\delta_{1m} \equiv \widetilde{\delta}_{1m} - \widetilde{\delta}_{11}$.

This model is estimated on commuting data from the 2001 Census Special Workplace Statistics, containing information on the number of people commuting between any two wards and the mode of transport used. The regression results, reported in Table A3, show that the marginal cost of distance is highest for walking, followed by cycling and, last, public transport and car. We use these estimates to predict the inclusive value relative to driving, and take the exponential to ensure non-negativity:

$$\exp(IV_{ab}) \equiv \widehat{IV}_{ab} = \sum_{m'} e^{\delta_{0m'} - \delta_{1m'} d_{ab}}. \tag{15}$$

We include $\widehat{IV}_{ab}$ in the determination of the utility to apply for jobs at different locations:

$$f_{ab} = \widehat{IV}_{ab} e^{\delta_0 - \delta d_{ab}}, \tag{16}$$

where e^{δ_0} is a constant term, d_{ab} is the distance between a and b, and δ measures the exponential rate of decay of the attractiveness of a given job with distance, when the transport mode is driving.

In the empirical specifications (8) and (13), all f_{ab} terms are raised to power of k. In fact, k is not well-identified in this model, and we simply aim to identify $\delta^* \equiv \delta k$. The reason why δ and k play similar roles is that an increase in the variance of the idiosyncratic component of utility – represented by a rise in k – has observationally equivalent consequences on the probability of applying for a job as a fall in the cost of distance. Intuitively, if idiosyncratic characteristics of jobs vary widely within areas, workers are more likely to find

a distant job that offers higher utility than a local job, and thus are willing to travel longer distances. A lower cost of distance would produce a similar outcome. δ^* is thus identified as the rate of decline of job applications with distance, and its estimate embodies both the actual time and monetary cost of travelling, and the extent to which the attractiveness of jobs varies within areas.

For estimation purposes, in both the logit model (14) and the job application models (8) and (13) we measure distance between any two different wards a and b as the geographic distance between their centroids, implicitly assuming that commutes start and end in the centroid of each ward. For within-ward commutes, we use instead the average distance between two random points in a circle:

$$d_{aa} = \frac{128\pi^{-1.5}}{45} area^{0.5},$$

where $area$ measures the size of the ward in square km. Thus we are approximating each ward with a circle of identical area, and each commute with the average within-ward commute. To assess the goodness of these approximations we compute commuting distances across postcodes, which identify one street or part of it, using information on postcode of residence and postcode of work from ASHE. We then assign postcodes to wards and compute commuting distances as described. The correlation between the two distance measures is 0.9995 overall and 0.9035 for commutes within 10 km. Figure A2 shows the binned scatterplot (by 100-metre bins) of the postcode-based distance against the ward-level distance, which virtually coincides with the 45 degree line except at distances below 200 metres, which represent 1 in 10,000 observations in ASHE.

D   The Probability of Filling a Vacancy

We use an urn-ball framework to model the probability of filling a vacancy, in which firms play the role of urns and applications the role of balls. Because of a coordination failure, a random placing of the balls in the urns implies that some urns will end up with more than one ball and some with none,

19

with overcrowding in some jobs and no applications in others. Conditional on receiving an application, a vacancy may still remain unfilled if one allows for worker heterogeneity and the possibility that an applicant may not be suitable for the job. Having denoted by q the probability of an individual applicant being suitable for the job, the probability that a given vacancy is not filled by any applicant is $(1-q)^A$ and the vacancy outflow rate is $1-(1-q)^A$. For small enough q, $(1-q)^A \simeq \exp(-qA)$. We further introduce a scale parameter λ, that affects the level of job matches at given applications. This can be thought of the probability that a match is made, conditional on receiving at least one acceptable application, and is related to overall matching effectiveness. Combining these elements, the probability of filling a vacancy can be written as:

$$\Psi(\tilde{A}) = \lambda \left(1 - e^{-\tilde{A}} \right), \tag{17}$$

where $\tilde{A} \equiv qA$ denotes the expected number of acceptable applicants.

E   Combining Model Elements

Expressions (17) and (16) can be combined with (8) to derive an estimable job application model for the exogenous wage case. As a wage proxy, we use the predicted ward-level wage based on the local industry structure, $\widehat{W}$, defined in (1):

$$\tilde{A}_b^{1+k} = \widetilde{c} \widehat{W}_b^{\rho k} \left(1 - e^{-\tilde{A}_b} \right)^k \sum_a U_a e^{-\delta^* d_{ab}} \widehat{IV}_{ab}^k \left[\sum_{b'} V_{b'} \left(\frac{1 - e^{-\tilde{A}_{b'}}}{\tilde{A}_{b'}} \right)^k \widehat{W}_{b'}^{\rho k} e^{-\delta^* d_{ab'}} \widehat{IV}_{ab'}^k \right]^{\gamma-1}, \tag{18}$$

where $\widetilde{c} \equiv c^{-k\gamma} \lambda^{k\gamma} e^{k\delta_0\gamma} q^{k\gamma-1}$ bundles all constant terms that cannot be separately identified. We introduce a further parameter ρ as an elasticity on the wage term, allowing us to test for the explanatory power of wages on the allocation of job applications rather than simply assume it as in (8), in which the null hypothesis is $\rho = 1$.

For the endogenous wage model, we need a measure of ward-level productivity Y_b. As there are no available data on productivity at such local

level, we again use the local industry structure as a proxy, based on evidence that high-productivity areas tend to specialize in high-productivity industries. Specifically, we assume that Y_b is isoelastic in $\widehat{W}_b$:

$$Y_b = \rho_0 \widehat{W}_b^{\rho_1}. \tag{19}$$

By combining (13), (17), (16) and (19) we obtain:

$$\tilde{A}_b = \breve{c} \widehat{W}_b^{k\rho_1} e^{-k\tilde{A}_b} \sum_a U_a e^{-\delta^* d_{ab}} \widehat{IV}_{ab}^k \left[\sum_{b'} V_{b'} e^{-k\tilde{A}_{b'}} \widehat{W}_{b'}^{k\rho_1} e^{-\delta^* d_{ab'}} \widehat{IV}_{ab'}^k \right]^{\gamma-1}, \tag{20}$$

where $\breve{c} \equiv c^{-k\gamma} \lambda^{k\gamma} e^{k\delta_0\gamma} q^{k\gamma-1} \rho_0^{k\gamma-1}$.

Models (18) and (20) give mappings from unemployment and vacancies – via job applications – to the vacancy outflow rate $\Psi(\tilde{A}) = \lambda(1 - e^{-\tilde{A}})$. The mapping depends on all model parameters, whose interpretation is summarized as follows:

- δ^* captures the rate of decline of job applications with distance and encompasses both time and monetary costs of distance, as well as heterogeneity in job characteristics. For simplicity, we will refer to this as "cost of distance".

- γ measures the returns to scale in the matching function.

- ρ (or ρ_1) measures the impact of wages on applications.

- λ controls job matches for given applicants, and thus captures matching effectiveness.

- $\tilde{c}$ (or $\breve{c}$) controls the general level of suitable applications.

- k measures the extent of heterogeneity in utility across jobs in an area.

The model is then estimated by non-linear least squares, choosing parameters to minimize the difference between the observed vacancy outflow rate and

21

the predicted rate:

$$\sum_b \left[\frac{M_b}{V_b} - \Psi(\widetilde{A}_b) \right]^2, \tag{21}$$

where M_b denotes the number of vacancies filled in b.[9]

III   Results

A   Main Estimates

This Section provides estimates of our structural model. For reasons of computing capacity, we are unable to estimate the main regression equation (21) on the whole sample period, and thus this is separately estimated for each month from May 2004-April 2006. This, however, has the advantage that each month's estimate can be thought of as a draw from the data, providing a measure of the standard error of our estimates based on their variation across different months, which we can then compare with standard errors produced by the non-linear least squares procedure.

Table 1 reports time averages of the parameters of interest. The parameter k was never well-identified. The estimates reported fix it at 1, and the model fit was not sensitive to this normalizing assumption. Poor identification of k can be traced to the fact that, unlike for other parameters, k is not linked to a readily identifiable feature of the data, and is only identified by a functional form assumption on $\Psi(\widetilde{A})$.[10] The Table also reports standard errors, obtained

---

[9]In estimation, we treat the spatial distribution of unemployment and vacancies as exogenous, as in most of the literature on the estimation of matching functions (see Petrongolo and Pissarides, 2001, for a survey). This is consistent with evidence from the policy simulation exercises described in Section IV, showing that the returns to job search – as measured by the probability of leaving unemployment or filling a vacancy – are not very responsive to local shocks. The direction of the implied biases from assuming exogenous unemployment and vacancies is not ex-ante clear. Suppose that one location has a more efficient matching function. This leads to higher vacancy outflows and a lower stock of unemployment and vacancies, with ambiguous effects on the U-V ratio and the number of applicants per job. On the other hand, if the unemployed and firms are mobile, such an area will attract workers and firms, leading to both higher unemployment and vacancies, with again ambiguous effects on the U-V ratio.

[10]To see this, note that, if $\Psi(A)$ were iso-elastic, k would not be identified at all.

with two alternative methods. The first measure (s.e.$_1$) is obtained as the standard deviation of the monthly parameter estimates. These are valid if the underlying parameters are constant over time and there is no serial correlation in the estimates, and Appendix C shows that the extent of serial correlation is small and never significant. The second measure (s.e.$_2$) reports the "sandwich" standard error obtained from the non-linear least squares estimator, allowing for possible heteroskedasticity and spatial correlation in the residuals, but assuming that the covariance between residuals from areas more than 100km apart is zero. Appendix C gives full detail on how sandwich standard errors are obtained.[11]

Estimates reported in columns 1-3 of Table 1 refer to the exogenous wage model. The estimate for δ^*, representing the cost of distance, is positive and highly significant, based on either standard error measure. A point estimate of 0.22 implies a relatively strong decay of job applications with distance. To see this, consider choice between two random jobs paying the same wage. Job 1 is local while job 2 is located further afield ($d_1 < d_2$), and ε_1 and ε_2 denote the respective idiosyncratic utility components. Despite the longer commute, a worker may still prefer job 2 over job 1 if the associated idiosyncratic utility component is high enough. From (7) and (16), job 2 is preferred to job 1 if $\varepsilon_2 e^{-\delta^*(d_2-d_1)}\widehat{IV}_2/\widehat{IV}_1 > \varepsilon_1$. Under the assumption that ε_1 and ε_2 are Pareto distributed, simple algebra shows that this happens with probability $\frac{1}{2}e^{-\delta^*(d_2-d_1)}\widehat{IV}_2/\widehat{IV}_1$. Assume, for instance, $d_1 = 1km$ and $d_2 = 5km$. With $\delta^* = 0.22$, a worker would prefer the more distant job over the local one in only 19 percent of cases. Our findings are in line with relatively high costs of commuting detected by Bonhomme and Jolivet (2009), who exploit information on job satisfaction from the ECHPS, and find that workers in Europe are typically willing to forgo large fractions of their salaries to become satisfied with their commuting distances or costs, ranging from 40 percent in France to 14 percent in Austria (though they do not report estimates for the UK). This

---

[11]See also Driscoll and Kraay (1998) for a discussion on how to obtain standard errors corrected for heteroskedasticity, serial and/or spatial correlation in panel data with a large number of geographic units.

is also consistent with the finding of Krueger at al. (2008) that commuting has a more detrimental impact on subjective well being than work itself.

The point estimate for γ is negative, implying decreasing returns in matching, although this is only significant based on s.e.$_1$ in column 2. Our model would require some extra form of congestion in order to deliver decreasing returns in the matching function.[12] The parameter ρ, representing the role of wages in the utility provided by jobs in an area, is significantly positive and not significantly different from one, in line with expression (7). The parameter λ, which represents the probability of filling a vacancy when there is at least one suitable applicant, is about 0.37 and highly significant. Finally the parameter $\widetilde{c}$, controlling the number of expected suitable applicants, $\tilde{A}$, is close to 1 on average and highly significant. At the sample means, the estimate of $\widetilde{c}$ implies that there is (almost exactly) 1 suitable applicant per vacancy.

The next three columns in Table 1 report estimates for the endogenous wage model, in which profit-maximizing wages are related to productivity via the parameter ρ_1. The main difference with respect to the exogenous wage model estimates are a higher cost of distance ($\delta^* = 0.36$), and a virtually zero estimate for γ, thus constant returns cannot be rejected on either standard error measure.

As typically with structural models, it is not straightforward to visualize the link between parameter estimates and specific data feature, and we devote Appendix D to highlight such links and discuss identification. Specifically, we estimate a reduced-form matching function in unemployment and vacancies at various distances to describe local matching patterns in a transparent way. As expected, the vacancy outflow rate in a ward rises with the unemployment stock and falls with the vacancy stock, and these effects gradually fade out with distance from the local ward, and the hypothesis of constant returns to

---

[12] A potential source of decreasing returns would be transport congestion: the higher the number of commuters from a to b, the lower the corresponding utility Ω_{ab} – other things equal – due to transport congestion. Alternatively, one could allow the probability to fill a vacancy to decrease with the size of the local vacancy pool, V_b, for any given level of job applications per vacancy, $\tilde{A}_b$, directly introducing decreasing returns to scale in the matching function. As in most specifications we obtain estimates for γ that are not significantly different from zero, we choose not to introduce further elements of search congestion into the model.

scale in matching is not rejected. Following the approach of Gentzkow and Shapiro (2014), we then provide a mapping from the data to the structural parameters, by relating monthly variation in structural parameter estimates to variation in reduced-form estimates. This exercize confirms the interpretation of structural parameters given above. For example, months in which distant unemployment and vacancies seem relatively less important in the reduced-form model are months in which our estimated cost of distance is relatively high. This exercize gives us confidence that the structural parameter estimates obtained do reflect relevant features of the data.

B   Alternative Specifications

In Table 2 we report estimates of alternative specifications of the job application model for February 2005, and Table A4 in the Appendix reports the corresponding average estimates for the whole sample period. The simple criterion used for picking a reference month in Table 2 is that it should not be December or a summer month, and that the parameter estimates for this month should be relatively close to the sample averages, so as to make the estimates of Table 2 well representative for the whole sample period. For reasons of computing capacity, we could not estimate the endogenous wage model with more than five parameters, and all specifications presented below refer to the exogenous wage model.

Column 1 in Table 2 reports estimates for the base model for February 2005, and indeed all estimates are very close to sample averages reported in columns 1-3 of Table 1. As constant returns to scale are not rejected, the specification of column 2 imposes constant returns ($\gamma = 0$) and shows estimates which are, unsurprisingly, very close to the estimates of the unrestricted model in column 1.

We next consider whether job applications are a sufficient statistic for describing job matches, as our model would predict. To do this we test whether the local U-V ratio has a significant explanatory power on local job matches, once one controls for applications per job, modifying the expected outflow rate

(21) to

$$E\left(\frac{M_b}{V_b}\right) = \lambda\left(1 - e^{-\tilde{A}_b - \alpha_1(U_{1b}/V_{1b})}\right),$$

where U_{1b}/V_{1b} denotes the U-V ratio within 1 km from the centroid of ward b, and job applications and the local U-V ratio are substitutes up to a parameter α_1. Column 3 reports estimates of this specification, and shows a small and non significant impact of U_{1b}/V_{1b}, confirming the prediction that applications should be a sufficient statistics for variation in the vacancy outflow. Importantly, we show in Appendix D.1 that, in a reduced-form vacancy outflow equation that does not control for applications, the local U-V ratio has a positive and highly significant impact on the vacancy outflow, but the results of column 3 show that this variable has no residual explanatory power once the role of labor market tightness is captured by the number of job applications. The other parameter estimates in column 3 remain close to those obtained on the main specification of column 1, although the return to scale parameter is now significantly negative.

We have so far assumed that labor markets only differ in wages offered and distance to potential applicants, and failure to recognize job and worker heterogeneity along other relevant dimensions could induce us to overstate the cost of distance. For example, if very few workers in a apply to jobs in b, our model would interpret that a and b are located too far apart to belong to the same local labor market, but another reason may be that workers in a do not have the right skills to perform jobs in b. To control for worker and job heterogeneity, we construct an index of mismatch between the skill composition of each origin labor market and that of each destination labor market, based on the occupational composition of claimants and job vacancies.[13] We extract data on claimants and job vacancies by ward and 1-digit occupation, and construct the following index of occupation dissimilarity between origin area

---

[13]For the unemployed the occupation refers to the type of job sought.

26

a and destination area b :

$$m_{ab} = \sum_{h=1}^{8} \left| \frac{U_{ha}}{U_a} - \frac{V_{hb}}{V_b} \right|,$$

where the occupation categories considered are: (1) managers and professionals; (2) associate professionals and technical occupations; (3) administrative and secretarial occupations; (4) skilled trades occupations; (5) personal service occupations; (6) sales and customer service occupations; (7) process, plant and machine operatives; (8) elementary occupations. We then introduce the mismatch indicator as a penalty in the utility of commuting from a to b:

$$f_{ab} = e^{(\delta_0 - \delta d_{ab} - \alpha_2 m_{ab})} \widehat{IV}_{ab},$$

where α_2 is an extra parameter to be identified. The results are reported in column 4, where the estimate for α_2 is positive, as expected, but very imprecise. While the February 2005 point estimate, together with the sandwich standard error, does not deliver a significant impact of mismatch, according to the estimates based on the whole sample period and bootstrapped standard errors (column 4 in Table A4) the impact of mismatch becomes significant at the 10 percent level. In either case, the estimate for the cost of distance is hardly affected by the inclusion of the mismatch term.

To conclude, we compare the relative merits of the job application model proposed with a conventional matching function in vacancies and unemployment, by only including U_{1b}/V_{1b} as a regressor in place of job applicants $\tilde{A}$. We express the expected outflow rate as

$$E\left(\frac{M_b}{V_b}\right) = \lambda \left(1 - e^{-U_{1b}/V_{1b}}\right),$$

and report the corresponding λ estimate in column 5. The fit of the equation is substantially worse than that of the structural model, as shown by the residual sums of squares. Thus the job application model performs better at explaining the variation in job matching rates than a simple matching function in local

unemployment and vacancies.

C   Evidence from Wages

One of the predictions of endogenous wage setting is that wages offered should, ceteris paribus, decline with the expected number of job applicants, according to result (12). We show evidence on this prediction using data on hourly earnings from ASHE from 2003-2008. We estimate log wage regressions that control for individual characteristics and the number of applicants per ward, as predicted by the estimates of column 1 in Table 2, and ward-level productivity, as predicted by the industry structure. Not all wards are represented in this sample (8,752 out of 8,850), because there are no available wage observations for 98 wards.

The results from log wage regressions are reported in Table 3. Both applications and productivity are ward-level averages, and standard errors are clustered at the ward level. Column 1 regresses log hourly earnings on log productivity and the (log) predicted number of job applications from the estimated structural model, averaged across all months. The coefficient on the productivity variable is positive and significant, as expected. However the coefficient on applications is positive, while the model would predict a negative effect. The specification in column 1 omits individual characteristics that are correlated with wages, and the specification of column 2 controls for a complete set of age dummies, interacted with gender, and broad region dummies, to account for the fact that wages in London are much higher than elsewhere. The coefficient on the number of applications is now negative and significantly different from zero (although quantitatively very small, and indeed smaller than the structural model would predict). Column 3 also controls for 3-digit occupation, and the coefficient on the number of applications remains negative and significant, confirming that areas that are predicted to receive larger numbers of applications have, all else equal, lower wages.

D   Predicted Commuting Flows

One of the main results obtained, and namely a relatively high cost of distance, implies that job search is concentrated within fairly local areas, and the plausibility of our estimates ultimately relies on their ability to predict external evidence on actual job finding behavior. This subsection assesses the predictive power of our estimated model by exploiting its predictions for commuting patterns between any two wards.

Commutes from a to b are given by the number of applications that the unemployed in a send to jobs in b, times the probability that these are successful:

$$U_a N_{ab} p(\widetilde{A}_b). \tag{22}$$

The distribution of predicted commutes can be obtained as the share of workers who live in a and work in b, for all possible pairs (a, b). Given (22), this is equal to

$$\frac{N_{ab} p(\widetilde{A}_b)}{\sum_{b'} N_{ab'} p(\widetilde{A}_{b'})}. \tag{23}$$

Predictions are compared to actual commutes for all workers, as obtained from administrative data in ASHE. Predicted and actual commuting may not coincide if workers who accept a job in a ward tend to relocate, for example to shorten their commute, or if jobseekers filling Jobcentre vacancies have different commuting patterns from jobseekers who find jobs via other channels. However, external evidence on commuting behavior from the UK Labor Force Survey (LFS) suggests that this is not a major concern. The LFS contains information on commuting times for those in new jobs and those in continuing jobs and, for those in new jobs, on how they obtained the job. Table A5 in the Appendix presents evidence on the average length of commute for these groups. The average commute for the category of workers we model – those who have recently got a job through a Jobcentre – is the same as for the overall employed population. As the characteristics of workers in different cells may differ, and they may be related to commuting times, we also compare differences in commuting times controlling for the method used to find the

current job, age, gender, region and year (results not reported), and we find no significant difference between commuting times of those who found jobs via Jobcentres and those who are not on new jobs. This justifies comparisons between the commutes predicted by our model with the commuting data for the whole population.

To obtain actual commutes, we use we use information on postcode of residence and postcode of work of employees in ASHE for 2003-2008. We expand the sample period in ASHE relative to our original 2004-2006 sample to reduce noise due to small cell size, as each cell is a combination of any two origin and destination wards. We then group the obtained commuting flows by 1 km distance bins. Predicted commutes are computed using equation (23) and the estimates of column 1 in Table 2, and again aggregated by 1 km distance bins. Figure 1 plots the actual and predicted distributions of commutes, and shows that our estimated model indeed replicates very acturately the observed commuting behavior, and in particular the spike in the density of commutes at very short distances. This suggests that our high estimated cost of distance has clear empirical plausibility.

IV   Evaluating Place-Based Policies

There is a large and growing literature on the evaluation of place-based policies, increasingly using modern evaluation methods and better research designs (see Glaeser and Gottlieb, 2008, and Moretti, 2011, for recent surveys, and Kline and Moretti, 2013, for a model of place-based policies with frictional unemployment). A clear consensus on the impact of place-based policies is yet to emerge,[14] partly because of wide variation in the size of intervention, and because a typical place-based policy combines several elements.

One of the advantages of structural models is that they can be used for ex

---

[14]Some studies (e.g. Hanson, 2009, Neumark and Kolko, 2010, for the US; Einiö and Overman, 2016, for the UK; and Mayer, Meyneris and Py, 2015, for France) find little or no effects, or only displacement effects, of place-based policies, while others do detect benefits for local employment (e.g. Busso, Gregory and Kline, 2012, for the US; Duranton, Gobillon and Overman, 2011, Criscuolo et al, 2012, for the UK).

ante policy evaluation and this section illustrates our model's prediction about the impact of local policies. We consider two common types of intervention: an increase in vacancies in a targeted area (possibly induced by local subsidies to job creation); and an improvement in the transport infrastructure of targeted areas, designed to make jobs more accessible. In both cases the intervention considered is very "local", being targeted on neighborhoods, and we later discuss how predictions would differ if one considered interventions over a larger scale.

One limitation of our model for the evaluation of place-based policies comes from its assumption of given vacancies and unemployment. In other words the model cannot be used, for example, to assess the extent to which place-based policies lead to vacancy creation in targeted areas or displace vacancies in neighboring areas, or alter the residential decisions of workers (see Beaudry, Green and Sand, 2012, and Rupert and Wasmer, 2012, for recent attempts to combine search and residential mobility). However, our model can shed light on the likely size of these responses. In particular, we show that the overlapping structure of local labor markets largely dilutes the impact of local policies across space, with limited impact on the local vacancy and unemployment outflows. This result suggests that the returns to locating in a certain area are not greatly affected by the policy interventions considered.

A   Local Labor Demand Stimulus

A recurrent question in the design of local policy is whether unemployment may be alleviated in a depressed area using local stimulus to labor demand, or whether local stimulus is diluted across space through a chain reaction of local spillovers. To answer this question we introduce a labor demand shock in a high-unemployment neighborhood, and use model predictions to simulate its local and surrounding effects.

As an example, we consider an increase in the number of job openings in Stratford, a high-unemployment area in East London,[15] which was the main

---

[15] In February 2005, Stratford had a ratio of claimant unemployment to resident population

31

venue of the 2012 Olympic Games. The Olympics led to both a large temporary increase in construction-related projects, and to a permanent expansion in the local retail sector, with the opening of the Westfield Stratford City shopping mall next to the Olympic Park in September 2011. In what follows we simulate the impact of a doubling in the number of vacancies in Stratford and New Town Ward in a given month, from 464 to 928. In the simulation we impose constant returns to scale, an assumption which is not rejected by our estimates. The rationale for this choice is to make the total number of applications made by workers at all locations independent of the size of the economy, and thus unaffected by the shock considered (see equation (4)). Based on the estimates of column 2 in Table 2, the model predicts a total increase in the vacancy outflow, and thus in the unemployment outflow, of 173. The implied vacancy outflow rate is 37 percent, closely in line with the average vacancy outflow rate reported in Table A1.

The spatial diffusion of this shock is illustrated in Table 4, showing the predicted absolute and percentage changes in the vacancy outflow ($\lambda(1 - e^{-\tilde{A}_b})V_b$) and the unemployment outflow ($U_a \sum_b N_{ab} P(\widetilde{A}_b)$) in different rings around Stratford, as well as the predicted percentage change in applications per job (equation (18)). In Stratford itself the vacancy outflow rises by 178 (column 1), but the unemployment outflow hardly changes at all (column 2). The absolute change in the unemployment outflow comes from a wider ring around Stratford, with 60 percent of the increase being drawn from more than 10km from Stratford (though only 20 percent coming from more than 20km). This does not mean that large numbers of unemployed workers 20km from Stratford are applying for the new jobs. Rather, what happens is that workers in or very close to Stratford re-direct their search acitvity towards Stratford and workers from slightly farther afield now apply to the jobs that receive fewer applications. The redirection of search effort towards Stratford implies a decline in the absolute vacancy outflow away from Stratford, although quantitatively crowding out effects are very modest (178.25-173.18).[16]

---

of 6%, which was nearly three times higher than the average for England and Wales.

[16]Tiny crowding out effects also imply that welfare evaluations of job creation policies

Turning to percentage changes in columns 3-5, as total applications in the economy are unchanged, applications per job on average fall. In Stratford, where vacancies double, applications per job fall by about 1.6 percent. Around Stratford, applications per job also fall, as Stratford attracts job search from surrounding areas. This spillover effect decays with distance, and the percentage change in applications per job is below 1 percent beyond 10 km from Stratford, and negligible beyond 35 km. The number of vacancies filled in Stratford virtually doubles, meaning that all extra vacancies created are filled. This is not surprising, given the localized shock. Again we find no evidence of any sharp local effect on the unemployment outflow, which is rising locally by only 0.6 percent. If anything, the unemployment outflow within 10 km rises slightly more than in Stratford,[17] and beyond this cutoff distance its change becomes negligible. As the total number of vacancies and unemployed increase with the square of the distance from Stratford, percentage changes in either outflow decline more sharply with distance than absolute changes.

The geographic diffusion of this shock is shown graphically in Figure 2, in which wards around Stratford are shaded according to the average percentage change in the unemployment outflow. As the impact on vacancy and unemployment outflows is very modest, we would also expect this policy to have limited effects on vacancy creation and residential mobility. While the location of vacancies and the unemployed is exogenous in our model, these predictions imply that incentives to relocate following local stimulus are small.

The main implication of our result is that, while labor markets are quite local, in the sense that the attractiveness of job offers strongly declines with distance, their overlapping structure means that local shocks generate wide ripple effects. Even strong local stimulus has a limited effect on the local

whose scope is small relative to the affected area may not need to take into account such search externalities.

[17]This non-monotonicity comes from the fact that the $\left[\sum_{b'} V_{b'} \left(\frac{1-e^{-\bar{A}_{b'}}}{\bar{A}_{b'}}\right)^k W_{b'}^k e^{-\delta^* d_{ab'}}\right]^{\gamma-1}$ term, capturing the extent of job competition, falls more in Stratford than elsewhere. This term determines the number of applications sent from each area a to each area b, according to (5), and thus the unemployment outflow in each area a, according to (22).

exit rate from unemployment, because a series of spatial spillovers dissipates a local shock across overlapping labor markets. In the example considered, unemployed workers living relatively close to Stratford divert some of their job search effort from their local wards towards Stratford. This reduces job competition in their local wards and attracts applications from slightly further afield, and so on. The model thus explains the spatial propagation of local shocks even in the presence of relatively high costs of distance. As a corollary, one can imagine that a local employment stimulus is likely to have sizeable effects on local unemployment if there is a "firebreak" across which few workers commute, as the firebreak prevents local stimulus to dissipate via ripple effects. Thus stimuli in more isolated areas are expected to have larger local effects, in line with findings by Briant, Lafourcade and Schmutz (2015).[18]

We next compare predictions from our simulation to actual data on job postings and the unemployment outflow around Stratford in the run-up to the 2012 Olympics. Much of the increase in labor demand took place in summer 2012 with running the Olympics itself, while some has built up steadily over time (e.g. in construction and retail). Panel A in Figure 3 presents time series for new vacancies advertised in Stratford, in wards within 3km of Stratford, and in London, all normalized to their January 2009 values.[19] There is a steady increase in job openings in Stratford since the early months of 2011, with a peak in summer 2011, associated with the opening of the Westfield Stratford City mall, and another even larger peak in spring 2012 in anticipation of the Games, with vacancy inflows running at about ten times their usual level. Other areas show no such trend. Panel B plots series for the unemployment

---

[18]From the above discussion it follows that the spatial dilution of local stimulus could also be contained by subsidies to the employment of residents of targeted areas. While ripple effects are strong enough to dilute the effect of local stimulus across surrounding areas, hiring subsidies for local residents are more effective in raising the local exit rate from unemployment, as they impose a discontinuity in the ripple effect to the advantage of those living inside the targeted area. The intervention considered in this case is more "people-based" than "place-based" and suggests that one can benefit residents of targeted areas by increasing their effectiveness in the competition for jobs.

[19]While information on the stock of vacancies in the NOMIS is not comparable before and after May 2006 (see footnote 3), the procedure for registering the inflow of newly advertised vacancies remains unchanged.

outflow in the same areas as in Panel A, and shows little or no evidence of an increase in outflows in Stratford or surrounding areas as a result of the spike in vacancies. While this simple exercise cannot provide an exhaustive analysis of the local employment effects of the 2012 Olympics and associated projects, these indications are in line with our model predictions and are consistent with negligible local effects of targeted labor demand stimulus.

Our approach has important implications for the evaluation of place-based policies. For policy evaluation one should ideally define treatment and control areas such that the whole policy impact is confined within the treatment area, while ensuring that treatment and control areas are otherwise similar. The results reported above imply that it is not feasible to cut up the country into non-overlapping areas which satisfy these criteria if the location and nature of shocks are ex ante unknown.

However, the location of intervention is often known ex-ante, in which case a "bespoke" treatment area can be defined. But ripple mechanisms imply that this may be hard to achieve. If the treatment area is defined as a relatively small area around the source of the shock, most of the impact will likely happen outside it, and would possibly contaminate the control areas if those are adjacent. The resulting estimate for the treatment effect would be biased towards zero. This problem can be avoided by defining a larger treatment area, within which most of the policy impact is indeed confined. But this area may be very large relative to the scale of the policy, in which case a test based on a comparison between treatment and control outcomes would lack statistical power. For example, in the above exercise we compute that 90 percent of the treatment is produced within 29 km from the source of the shock, suggesting that a radius of about 30 km could define an appropriate treatment area for the evaluation of this particular shock. But while the shock considered is very large relative to the size of the target ward, it only represents a 0.7 percent increase in vacancies in the suggested treatment area, making it hard to identify significant effects of intervention. These magnitudes are by no means specific of the Stratford example above. To show this, we consider a doubling of vacancies in each of the 8,850 wards in our sample, and obtain

the spatial distribution of the change in the unemployment outflow. Figure 4 reports the geographic distribution of treatment, averaged across all wards, and the 90th percentile in this distribution is roughly 31 km, thus very close to the 90th percentile of treatment in the Stratford example.

The above definition of treatment areas requires rich information on the distribution of unemployment and vacancies and the nature of local spillovers, as summarized in our structural model, and a practical issue is whether more easily available data like the distribution of commutes could help define treatment areas. Indeed ripple effects imply that treatment effects generally extend to distances beyond typical commutes. While the median commute is 6 km, half of the treatment effect of policy is only produced within 14 km; thus the median treatment distance is about 2.5 times the median commute, implying that treatment areas need to be some multiple of typical commuting distances.

B   Reduction in Transportation Costs

We next assess the importance of transportation costs by simulating the effect of a sizeable reduction in the cost of distance between a high-unemployment area and an area with relatively high supply of jobs. The idea is to evaluate whether an improved transport link can effectively reduce the degree of spatial mismatch between workers and jobs. We pick Stratford and Heathrow as the high- and low-unemployment areas, respectively.[20]   The Heathrow Villages ward is locates in West London, and surrounds Heathrow Airport. Heathrow and Stratford are about 30.7 km apart, or 1 hour and 20 minutes apart using public transport, and we simulate the impact of halving such distance.

We recalculate the utility of commuting by including in the public transport mode of the inclusive value (15) a new distance matrix in which the distance between the two wards is set at 15.35 km, and the distance between any two other wards a and b is recalculated if either the distance $a-\text{Heathrow}-\text{Stratford}-b$ or $a-\text{Stratford}-\text{Heathrow}-b$ is shorter than the original distance ab. This is equivalent to introducing a fast, non-stop service between Stratford and

---

[20]In February 2005, the ratio between registered unemployment and Jobcentre vacancies was 1.35 in Stratford and 0.17 in Heathrow.

Heathrow, as will happen when the Crossrail project is completed in 2018, allowing individuals near either node to re-optimize their travel schedule accordingly. As a consequence, we would expect some jobseekers in Stratford to choose to search for jobs in Heathrow, thus raising the unemployment outflow in Stratford, and at the same time increasing job competition and reducing the unemployment outflow around Heathrow.

Table 5 illustrates the impact of this improved transport link at various distances from Stratford and Heathrow, and for simplicity we only report percentage changes and omit absolute changes. As expected, applications per job rise both in Heathrow and its close vicinity (column 1, rows 1 and 2), as this location is now attracting more jobseekers from Stratford and surrounding areas. As a consequence, the vacancy outflow increases (column 2) and the locals are less likely to find jobs (column 3), as they face stronger job competition from new applicants attracted by the faster transport link. In row 4, we find that applications per job decline only very slightly in Stratford, with virtually no change in the vacancy outflow. The unemployment outflow in Stratford increases, as job opportunities are easier to reach. It should be noted that, quantitatively, the impact on the unemployment outflow is always very small, whether positive or negative, and becomes negligible beyond 10 km from either transport node (rows 7 and 8). Spillovers on the unemployment outflow around Heathrow and Stratford are illustrated in more detail on a map in Figure 5, where darker and lighter shades correspond to an increase and a decrease, respectively, in the unemployment outflow. Overall, the impact on the unemployment outflow is relatively modest, but it propagates quite widely around either target.

V   Conclusions

This paper has developed a model of job search across space that is empirically tractable and allows estimation of a labor market process with a large number of overlapping markets, avoiding drawbacks that typically arise when local labor markets are modeled as non-overlapping segments of the economy. Using

data on unemployment and vacancies to estimate a matching function at the ward (or neighborood) level, we find that unemployed workers' search efforts are strongly discouraged by distance to target jobs. Our estimates imply that the probability that a random job 5 km distant is preferred to a local random job is only 19 percent. Also, workers are significantly discouraged from applying to jobs in areas in which they expect relatively strong competition from other jobseekers. Constant returns in matching markets are not rejected, implying that the total number of job applications made in this economy does not respond to the absolute size of the vacancy pool. Our estimated model, encompassing high costs of distance, predicts actual commuting flows very accurately.

We use our estimates to simulate the impact of local development policies like local stimulus to labor demand or improved transportation links. Despite the fact that labor markets are relatively local, location-based policies turn out to be rather ineffective in raising the local unemployment outflow, because labor markets overlap and the associated ripple effects in applications largely dilute the effect of local shocks across space. This has important implications for the evaluation of place-based policies. Defining small treatment areas risks omitting most of the impact and contaminating control areas, while large treatment areas make it more difficult to identify effects for policies that are large relative to small areas but small relative to a large area in which the treatment effect is confined.

References

[1] Amior, Michael and Alan Manning. 2015. "The Persistence of Local Joblessness." CEP Discussion Paper No. 1357.

[2] Beaudry, Paul, David A. Green, and Benjamin Sand. 2012. "Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory." *Econometrica* 80: 1063–1104.

[3] Berman, Eli. 1997. "Help Wanted, Job Needed: Estimates of a Matching Function from Employment Service Data." *Journal of Labor Economics* 15: S251–292

[4] Berry, Steven, James Levinsohn and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63: 841–890.

[5] Blanchard, Olivier and Larry Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 1: 1–75.

[6] Bonhomme, Stephane and Gregory Jolivet. 2009. "The Pervasive Absence of Compensating Differentials." *Journal of Applied Econometrics*, 24: 763–795.

[7] Bound, John and Harry Holzer. 2000. "Demand Shifts, Population Adjustments and Labor Market Outcomes during the 1980s." *Journal of Labor Economics* 18: 20–54.

[8] Boustan, Leah P. and Robert Margo. 2009. "Race, Segregation, and Postal Employment: New Evidence on Spatial Mismatch." *Journal of Urban Economics* 65: 1–10.

[9] Briant, Anthony, Miren Lafourcade and Benoît Schmutz. 2015. "Can Tax Breaks Beat Geography? Lessons from the French Enterprise Zone Experience." *American Economic Journal: Economic Policy*, 7: 88–124.

[10] Burdett, Kenneth and Kenneth Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51: 955–969.

[11] Busso, Matias, Jesse Gregory and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103: 897–947.

[12] Chade, Hector and Lones Smith. 2006. "Simultaneous Search." *Econometrica* 74: 1293–1307.

[13] Criscuolo, Chiara, Ralph Martin, Henry Overman, and John Van Reenen. 2012. "The Causal Effect of an Industrial Policy". NBER Working Paper No. 17842.

[14] Diamond, Peter. 1982. "Aggregate Demand Management in Search Equilibrium". *Journal of Political Economy* 90: 881–894.

[15] Driscoll, John and Aart Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data". *Review of Economcis and Statistics* 4: 549–560.

[16] Duranton, Gilles, Laurent Gobillon and Henry Overman. 2011. "Assessing the Effects of Local Taxation Using Microgeographic Data." *Economic Journal* 121: 1017–1046.

[17] Einiö, Elias and Henry Overman. 2016. "The (Displacement) Effects of Spatially Targeted Enterprise Initiatives: Evidence from UK LEGI." CEPR Discussion Paper No. 11112.

[18] Glaeser, Edward L. and Joshua D. Gottlieb. 2008. "The Economics of Place-Making Policies." *Brookings Papers on Economic Activity,* Spring 2008, 155–253.

[19] Gregg, Paul, Stephen Machin and Alan Manning. 2004. "Mobility and Joblessness." in David Card, Richard Blundell and Richard Freeman (eds) *Seeking a Premier League Economy: The Economic Effects of British Economic Reforms, 1980–2000.* Chicago: University of Chicago Press.

[20] Hanson, Andrew. 2009. "Local Employment, Poverty, and Property Value Effects of Geographically-Targeted Tax Incentives: An Instrumental Variables Approach." *Regional Science and Urban Economics* 39: 721–731.

[21] Hellerstein, Judith K., Melissa McInerney and David Neumark. 2011. "Neighbors and Coworkers: The Importance of Residential Labor Market Networks." *Journal of Labor Economics* 29: 659–95.

[22] Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *Quarterly Journal of Economics* 82: 175–197.

[23] Kline, Pat, and Enrico Moretti. 2013. "Place-based Policies with Unemployment." *American Economic Review P&P* 103: 238–243.

[24] Krueger, Alan B., Daniel Kahneman; David Schkade; Norbert Schwarz and Arthur A. Stone. 2009. "National Time Accounting: The Currency of Life." In A. B. Krueger (ed.), *Measuring the Subjective Well-Being of Nations: National Accountf of Time Use and Well-Being.* NBER, 9–86.

[25] McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka (ed.), *Frontiers of Econometrics*, 105–142. New York: Academic Press.

[26] Manning, Alan. 2003. *Monopsony in Motion. Imperfect Competition in Labor Markets.* Princeton, Princeton University Press.

[27] Mayer, Thierry, Florian Mayneris and Loriane Py. 2015. "The Impact of Urban Enterprise Zones on Establishment Location Decisions and Labor Market Outcomes: Evidence from France." *Journal of Economic Geography* 15: 1–44.

[28] Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2015. "Commuting, Migration and Local Employment Elasticities." NBER Working Paper 21706.

[29] Moretti, Enrico. 2011. "Local Labor Markets." In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4B, 1237–1314. Amsterdam: Elsevier-NorthHolland.

[30] Neumark, David and Jed Kolko. 2010. "Do Enterprise Zones Create Jobs? Evidence from California's Enterprise Zone Program." *Journal of Urban Economics* 68: 1–29.

[31] Notowidigdo, Matthew J. 2013. "The Incidence of Local Labor Demand Shocks." Mimeo.

[32] OECD (2000). *The Public Employment Service in the United States.*

[33] Petrongolo, Barbara and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function,". *Journal of Economic Literature* 39: 390–431.

[34] Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory.* Cambridge (MA): The MIT Press.

[35] Rupert, Peter and Etienne Wasmer. 2012. "Housing and the Labor Market: Time to Move and Aggregate Unemployment." *Journal of Monetary Economics* 59: 24–36.

[36] Small, Kenneth A., Winston Clifford and Yan Jia. 2005. "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability." *Econometrica* 73: 1367–1382.

[37] Zenou, Yves. 2013. "Spatial versus Social Mismatch". *Journal of Urban Economics* 74: 113–132.

All Figures are best viewed in color print

Figure 1
Actual and predicted commutes by distance travelled.
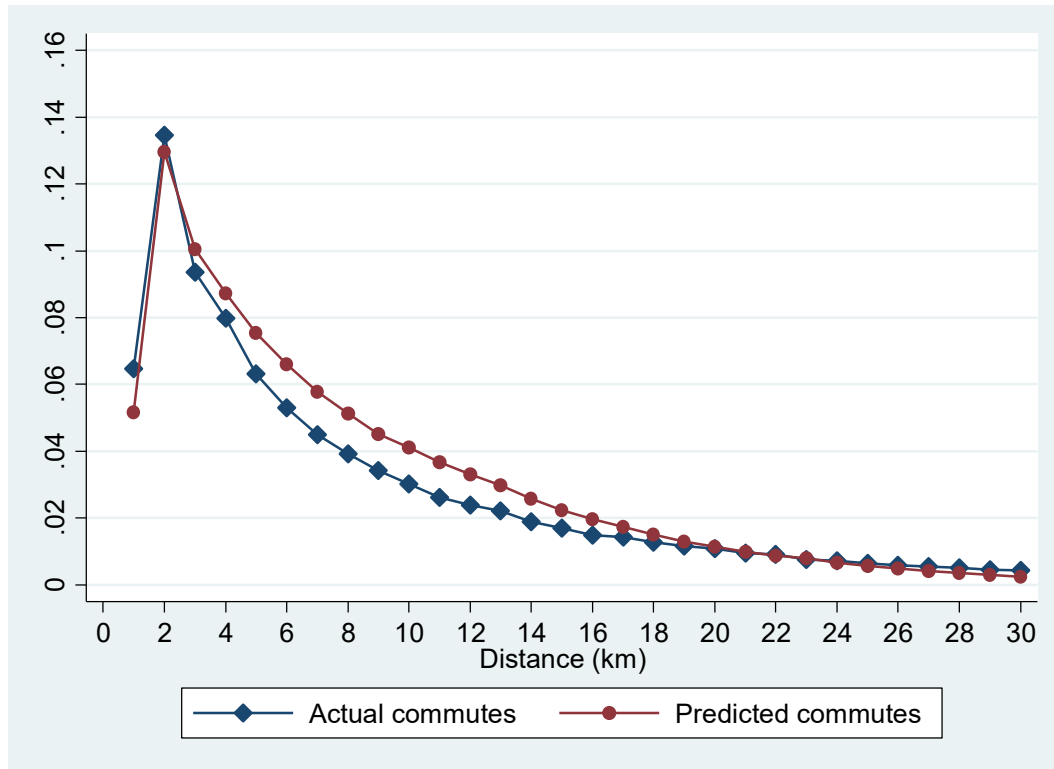
Notes. The graph plots shares of commutes by 1 km bins. Actual commutes are obtained from ASHE, using information on postcode of residence and postcode of work of employees. Predicted commutes are computed using equation (22) and the estimates of column 1 in Table 2.

Figure 2
Effect of a doubling in the number of vacancies in Stratford
on the unemployment outflow (percentage change).
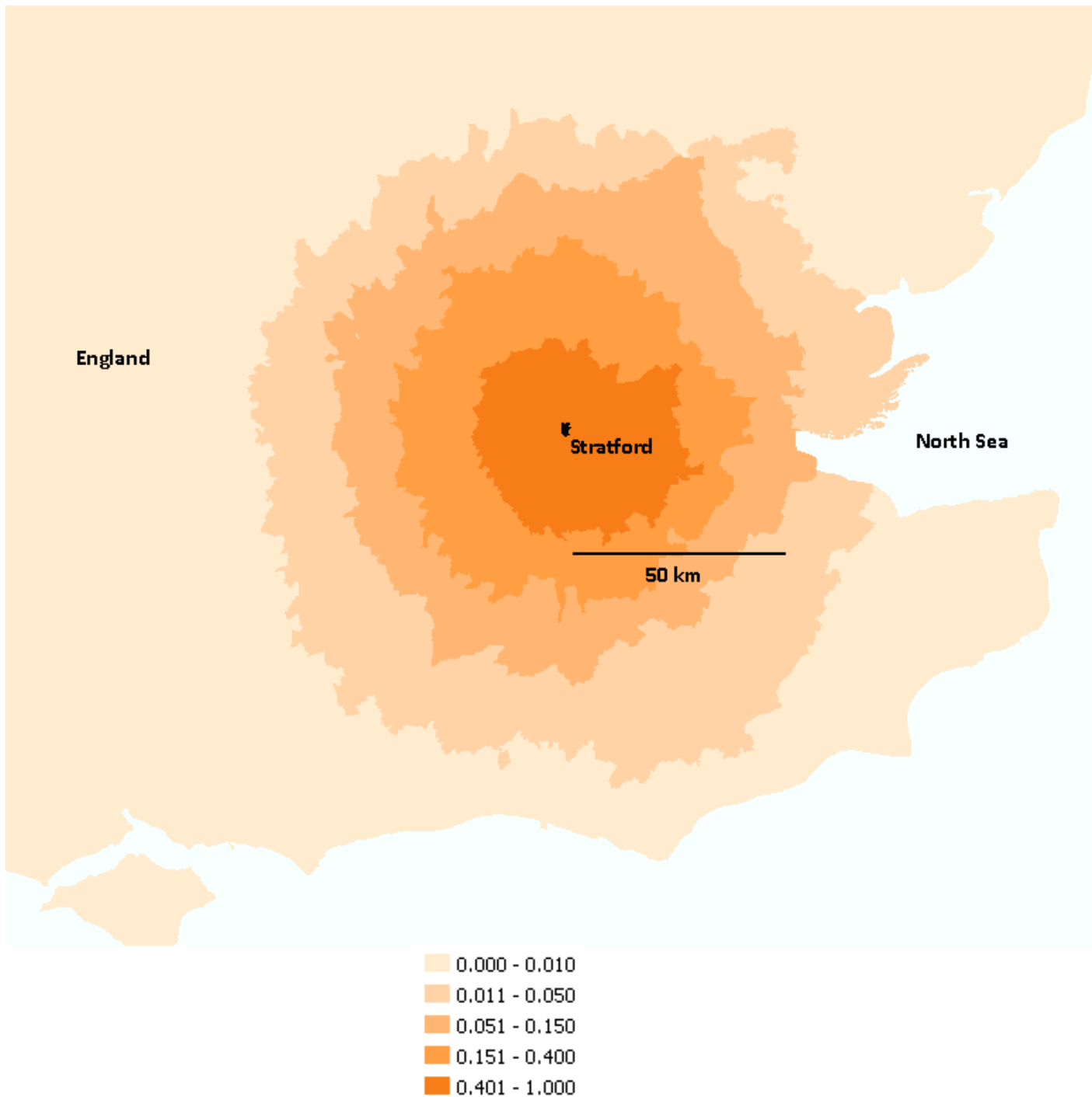
Notes. The predicted unemployment outflow by ward is given by $U_a \sum_b N_{ab} P(\tilde{A}_b)$, where N_{ab} and $\tilde{A}_b$ are evaluated using the estimates of column 2 in Table 2, for alternative values of vacancies in Stratford (464 at baseline; 928 following the shock).

Figure 3
Actual changes in the vacancy inflow (Panel A) and the unemployment outflow (Panel B) in and around Stratford
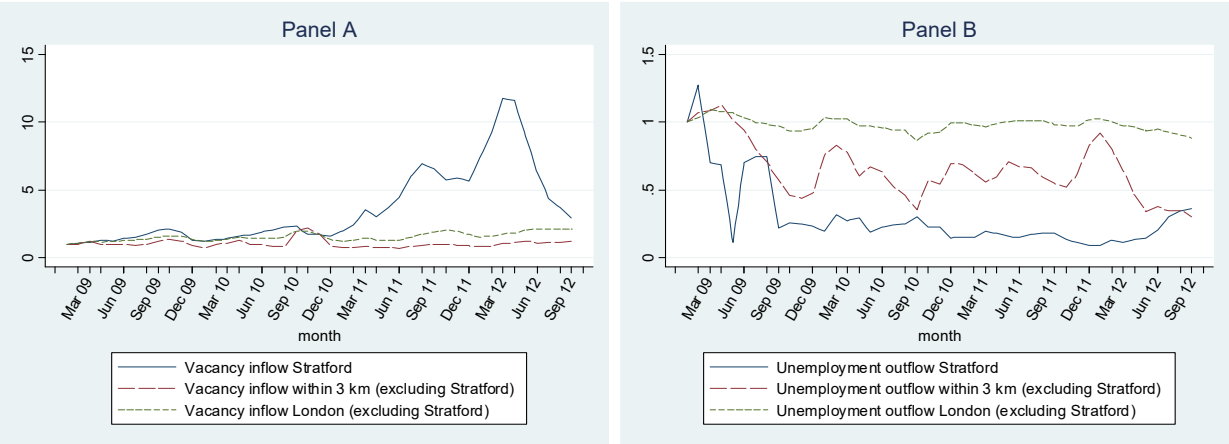
Notes: All series are smoothed using moving averages with a 3-month window and equal monthly weights, and normalized to their January 2009 values.

Figure 4
The spatial distribution of local policy impact on the unemployment outflow
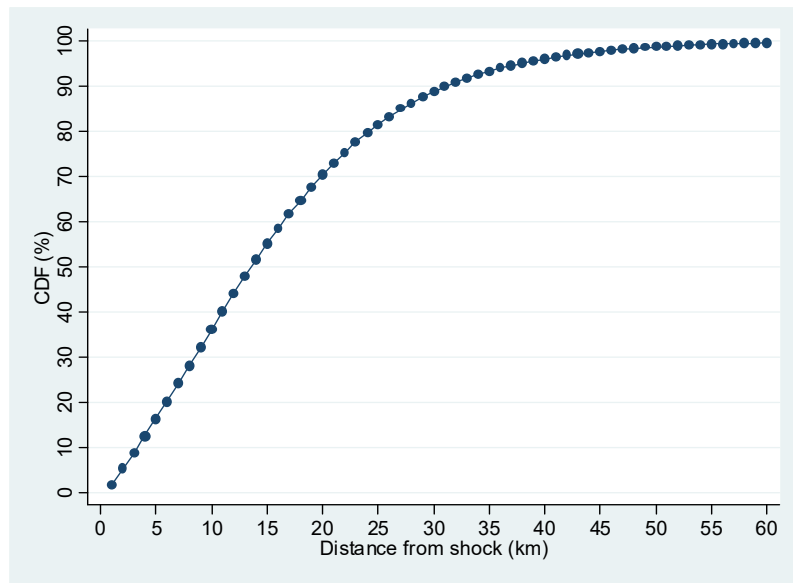
Notes. The CDF of treatment represents the percentage of the total increase in the unemployment outflow produced within a given distance from the source of a local labor demand shock. The underlying shock is a doubling in the number of local vacancies.

Figure 5
The effect of halving the cost of distance between Heathrow and Stratford
on the unemployment outflow (percentage change).

	-0.507 - -0.300
	-0.299 - -0.150
	-0.149 - -0.050
	-0.049 - 0.030
	0.031 - 0.100
	0.101 - 0.196
	0.197 - 0.515
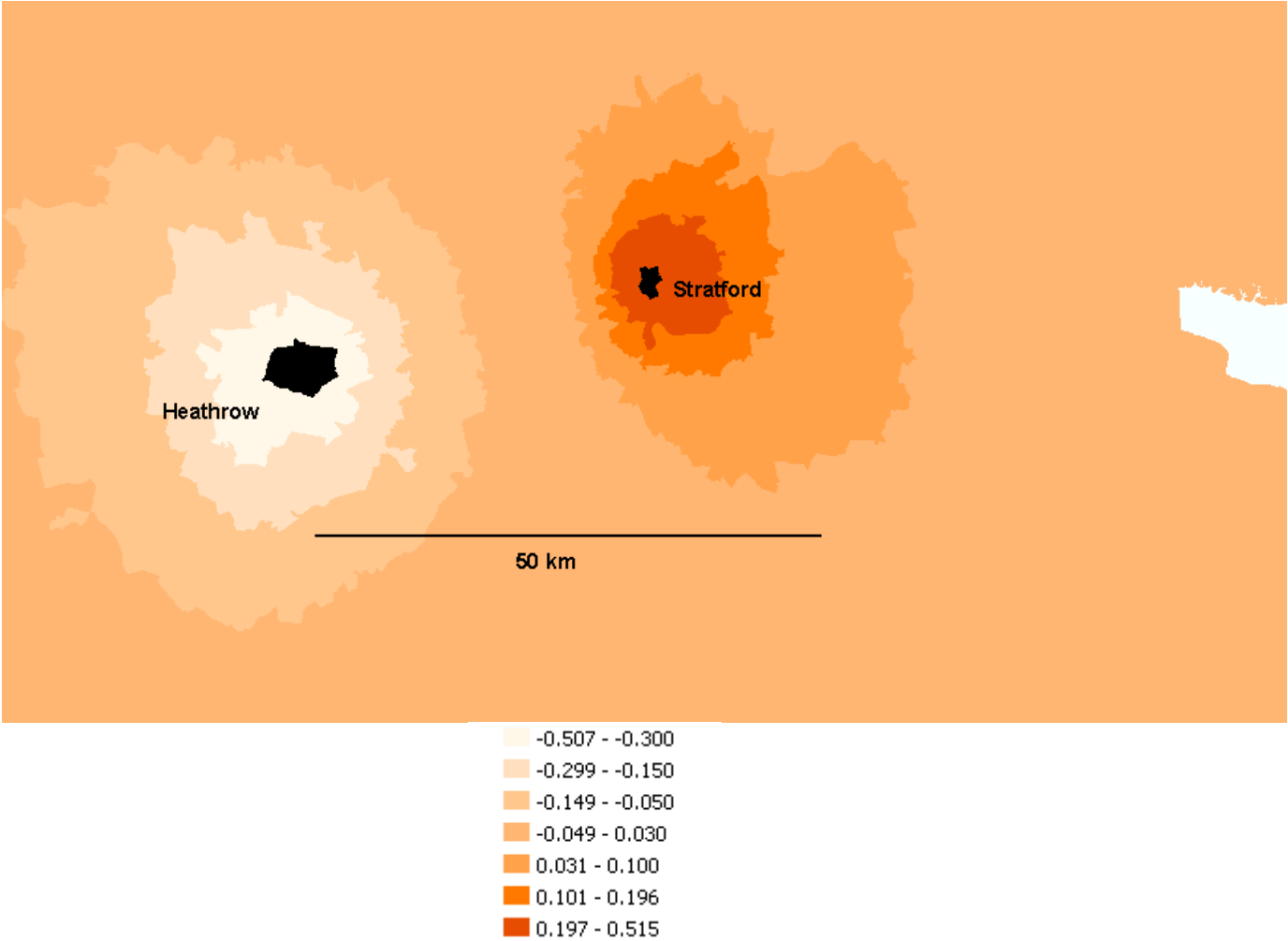
Notes. The predicted change in the unemployment outflow by ward is given by $U_a \sum_b N_{ab} P(\tilde{A}_b)$, where N_{ab} and $\tilde{A}_b$ are evaluated using the estimates of column 2 in Table 2, for alternative values of the distance between Heathrow and Stratford (30.7 km at baseline; 15.35 km following intervention).

Table 1
Estimates of the search model
May 2004-April 2006

| Parameters	Exogenous wage model			Endogenous wage model		
	(1)	(2)	(3)	(4)	(5)	(6)
	mean	s.e.$_1$	s.e.$_2$	mean	s.e.$_1$	s.e.$_2$
Cost of distance (δ^*)	0.220	0.026	0.059	0.360	0.089	0.152
Returns to scale (γ)	-0.160	0.038	0.092	0.015	0.039	0.069
Wage elasticity (ρ or ρ_1)	0.920	0.373	0.444	1.727	0.638	0.686
Matching effectiveness (λ)	0.371	0.033	0.024	0.367	0.033	0.024
Scale parameter in $\tilde{A}$ ($\tilde{c}$ or $\check{c}$)	1.185	0.207	0.382	0.902	0.357	0.397

The model specification is given by equations (18) and (21) for the exogenous wage case and equations (20) and (21) for the endogenous wage case. The reported estimates are means across the 24 months from May 2004-April 2006. The estimation method is nonlinear least squares. s.e.$_1$ is the standard variation in the monthly parameter estimates. s.e.$_2$ is the square root of the mean of the monthly parameter variances, obtained from the non-linear least squares procedure, adjusted for heteroscedasticity and possible spatial autocorrelation. Parameters ρ and $\tilde{c}$ refer to the exogenous wage model, and parameters ρ_1 and $\check{c}$ refer to the endogenous wage model.

Table 2
Estimates of a job application and matching model
Alternative specifications for February 2005

	Dependent variable: vacancy outflow rate				
	(1)	(2)	(3)	(4)	(5)
Cost of distance (δ^*)	0.200***	0.216***	0.189***	0.198***	
	(0.054)	(0.058)	(0.058)	(0.055)	
Returns to scale (γ)	-0.116		-0.136***	-0.113	
	(0.090)		(0.024)	(0.094)	
Wage elasticity (ρ)	0.821**	0.954***	0.836	0.850**	
	(0.438)	(0.395)	(0.914)	(0.433)	
Matching effectiveness (λ)	0.386***	0.384***	0.391***	0.387***	0.444***
	(0.025)	(0.023)	(0.032)	(0.024)	(0.026)
Scale parameter in $\tilde{A}$ ($\tilde{c}$)	1.002***	0.543***	1.068***	0.987***	
	(0.336)	(0.103)	(0.257)	(0.355)	
Local U_b/V_b (α_1)			0.090		
			(0.142)		
Mismatch (α_2)				0.218	
				(0.777)	
Observations	8709	8709	8709	8709	8709
Sum of squared residuals	343.3	344.0	340.3	343.2	438.7

Notes. The dependent variable is the vacancy outflow rate in ward b, M_b/V_b, where M_b denotes the vacancy outflow during February 2005, and V_b denotes the time-aggregate stock of vacancies during the same month. The model specification is given by equations (18) and (21). The estimation method is nonlinear least squares. Sandwich standard errors corrected for heteroskedasticity and spatial correlation are reported in brackets.

Table 3
Job applications and wages

	(1)	(2)	(3)
Log productivity	1.346***	1.069***	
	(0.042)	(0.034)	
Log applications per job	0.088***	-0.017***	
	(0.009)	(0.005)	
Other Controls	Year	Column (1) plus gender×age interactions and region	Column (2) plus occupation
Observations	819,771	819,771	
R^2	0.096	0.253	

Notes. The dependent variable is the log hourly wage. Productivity and applications per job are ward-level averages. Productivity is proxied by the wage prediction based on the local industry structure. Applications are obtained using equation (18) and the estimates of column 1 in Table 2. Other controls: year dummies (column 1); plus age dummies (16-65), interacted with gender, and nine region dummies (column 2); plus 3-digit occupation dummies (column 3). Standard errors are clustered at the ward level. Sample period: 2003-2008.

Table 4
The propagation of local shocks

Distance from Stratford	Absolute change		Percentage change		
	1 Vacancy outflow	2 Unemployment outflow	3 Applications per job	4 Vacancy outflow	5 Unemployment outflow
Stratford	178.25	0.42	-1.64	99.99	0.58
(0,5] km	0.00	19.25	-1.48	-0.00	0.72
(5,10] km	-0.03	49.98	-1.33	-0.00	0.73
(10,20] km	-0.44	65.58	-0.75	-0.00	0.55
(20,35] km	-2.36	28.58	-0.36	-0.02	0.28
(35,50] km	-1.88	7.07	-0.08	-0.02	0.08
50+ km	-0.34	2.30	-0.00	-0.00	0.00
Total	173.18	173.18			

Notes: The Table shows the simulated effect of a doubling in the number of vacancies in Stratford and New Town Ward. Applications per job are given by equation (18). The vacancy outflow is given by $\lambda(1 - \exp(\tilde{A}_b))V_b$. The predicted unemployment outflow is given by $U_a \sum_b N_{ab} P(\tilde{A}_b)$. All magnitudes are evaluated using the estimates of column 2 in Table 2, for alternative values of vacancies in Stratford (464 at baseline; 928 following intervention).

Table 5
The effect of reducing the cost of distance

Distance	Percentage change		
	1 Applications per job	2 Vacancy outflow	3 Unemployment outflow
Heathrow	1.10	0.04	-0.57
(0,5] km from Heathrow	0.33	0.01	-0.04
(5,10] from Heathrow	0.17	0.01	-0.25
Stratford	-0.07	-0.00	0.61
(0,5] km from Stratford	-0.07	-0.00	0.28
(5,10] from Stratford	-0.05	-0.00	0.09
(10,40] from both	-0.01	0.01	-0.02
40+ km from both	-0.00	0.00	-0.00

Notes: The Table shows the simulated effect of a halving the distance between a high-unemployment area (Stratford) and a low-unemployment area (Heathrow). Applications per job are given by equation (18). The vacancy outflow is given by $\lambda(1 - \exp(\tilde{A}_b))V_b$. The predicted unemployment outflow is given by $U_a \sum_b N_{ab} P(\tilde{A}_b)$. All magnitudes are evaluated using the estimates of column 2 in Table 2, for alternative values of the distance between Heathrow and Stratford (30.7 km at baseline; 15.35 km following intervention).

Appendices

For online publication

A Data coverage

By covering unemployment and vacancies from the UK Public Employment Service (PES), our data may not fully represent jobseekers and vacancies in the economy. On the worker side, not all jobseekers are claimant unemployed, as jobseekers may also be employed, or unemployed but not claiming benefits; and not all the claimant unemployed may be jobseekers (though they are meant to be, according to the rules for benefit entitlement). To form an idea of data coverage, we turn to the UK Labor Force Survey (LFS), which asks a direct question about job search both of those who are currently in and out of employment. In the Spring of 2005 (to give one example), according to the LFS there were about 3.1 million jobseekers in the UK, and total employment was about 28.1 million. Almost exactly half of the jobseekers were not currently employed, and at that time the official figure for the claimant count was about 875,000. In the LFS, approximately 20% of the claimant unemployed do not report looking for work in the past 4 weeks, suggesting that the claimant unemployed represent nearly a quarter of total jobseekers in the economy.

It may be argued that the claimants are among the most intensive jobseekers (see, among others, Flinn and Heckman, 1983, Jones and Riddell, 1999), and thus we weight jobseeker figures in the LFS by the number of reported search methods used. During the 2002-2007 period,[1] the unweighted share of claimants in total jobseekers was 17.6%, while the weighted share was 23.7%. The share of claimants in jobseekers also varies markedly with levels of education, being 15% among college graduates, 21.8% among high school graduates, 24.9% among those who left school at 16, and 35.2% among those with no qualifications. This means that our study is relatively more representative of

---

[1] We need to expand the sample period here in order to improve precision of the statistics reported.

low-skill labor markets, which tend to be more local.

For our purpose it is also important to know the fraction of jobseekers who are looking at the vacancies recorded in our data, i.e. vacancies advertised at PES Jobcentres. According to information on job-search methods used, during 2002-2007, 92% of claimants use Jobcentres, and 45.2% of them report Jobcentres as their most important job search method. These proportions fall to 44.4% and 18.3% for the non-claimant unemployed, and to 19.1% and 5.9% respectively for the employed. Thus, Jobcentres are widely used by the jobseekers in our sample. In this regard, it should be noted that the UK PES is much more widely used than the US equivalent. Manning (2003, Table 10.5) shows that only 22% of the US unemployed report using the PES compared to 75% of the UK unemployed, and OECD (2000, Table 4.2) shows that the market share of the PES in the US in vacancy coverage and total hires is substantially lower than in the UK. Hence the UK PES does play an important role in matching jobseekers and vacancies.

On the job vacancy side, to assess the representativeness of Jobcentre data we use information from the Vacancy Survey of the Office for National Statistics, which provides comprehensive estimates of the number of job vacancies in the UK, obtained from a sample of about 6,000 employers every month. Employers are asked how many job vacancies there are in their business, for which they are actively seeking recruits from outside the business. These vacancy data cover all sectors of the economy except agriculture, forestry and fishing, but are not disaggregated at the occupation or area level, so we can only make aggregate comparisons between ONS and Jobcentre vacancy series.

On average, since April 2004, the Jobcentre vacancy series in the UK is about two thirds the ONS series, but there are reasons to believe that such proportion may be overstated (Machin, 2001). In particular, in May 2002, an extra question was added to the ONS Vacancy Survey, on whether vacancies reported had also been notified at Jobcentres, and based on this information the ratio of total vacancies advertised at Jobcentres was 44%. While one should allow for sampling variation (this information is only available for May 2002, and for only 420 respondents), this 44% proportion is markedly lower that the

two thirds recorded for the post-2004 period. According to Machin (2001), the main reason for this discrepancy is that Jobcentre vacancies obtained from the computerized system may include vacancies which are "awaiting follow-up", but which have already been filled by employers, or which have been suspended by the Jobcentres, as it appears that sufficient potential recruits have already been referred. Our vacancy series obtained from Jobcentres ("live unfilled vacancies") excludes suspended vacancies, but "may still include some vacancies which have already been filled or are otherwise no longer open to recruits, due to natural lags in procedures for following up vacancies with employers",[2] thus one can still imagine that two-thirds is indeed an upper bound for the fraction of job openings that are effectively available to jobseekers at Jobcentres. As no occupation breakdown is available for the ONS vacancy series, it is not possible to determine how the skill distribution of our vacancy data compares to that of the whole economy, but it is plausible that Jobcentre vacancies over-represent less-skilled jobs.

B   Proof of contraction mappings

B.1   Exogenous Wage Model

To prove that (8) is a contraction mapping, we use Blackwell's sufficient conditions of monotonicity and discounting (Stokey and Lucas, 1989, p. 54). (8) is a function that maps one set of applications into another set, and is a valid mapping for all vectors in the positive orthant. We rewrite it in log form:

$$
\begin{aligned}
T\left(a_{b}\right) \;=\; \ln A_{b} &= \frac{1}{1+k}\left\{ k \ln \Psi\left(e^{a_{b}}\right) + \ln\left[\sum_{a} U_{a} W_{b} f_{ab}\left[\sum_{b'} V_{b'} W_{b'} f_{ab'}\left(\frac{\Psi\left(e^{a_{b'}}\right)}{e^{a_{b'}}} \right)^{k} \right]^{\gamma-1} \right] \right\} \\
&= \frac{1}{1+k}\left\{ k \ln \Psi\left(e^{a_{b}}\right) + \ln\left[\sum_{a} U_{a} W_{b} f_{ab} C_{a}\left(a\right)^{\gamma-1} \right] \right\},
\end{aligned}
$$

---

[2]https://www.nomisweb.co.uk/articles/showArticle.asp?title=warning: limitations of data&article=ref/vacs/warning-unfilled.htm

where $f_{ab} == e^{\delta_0 - \delta d_{ab}} \widehat{IV}_{ab}$ and

$$C_a(a) = \left[\sum_{b'} V_{b'} W_{b'} f_{ab'} \left(\frac{\Psi(e^{a_b})}{e^{a_b}} \right)^k \right].$$

This clearly satisfies monotonicity. As we assume $p'(A) < 0$, $\Psi'(A) > 0$, and $p(A) = \frac{\Psi(A)}{A}$ we must have $\epsilon_{\Psi A}(A) \equiv \frac{\partial \log \Psi(A)}{\partial \log A} < 1$. If $0 \leq \epsilon_{\min} \leq \epsilon_{\Psi A} \leq \epsilon_{\max} \leq 1$, we have that:

$$C_a(a + \alpha) \geq C_a(a) e^{-k(1-\epsilon_{\min})\alpha},$$

which implies:

$$\left[\sum_a U_a W_b f_{ab} C_a(a+\alpha)^{\gamma-1} \right] \leq e^{-(\gamma-1)k(1-\epsilon_{\min})\alpha} \left[\sum_a U_a W_b f_{ab} C_a(a)^{\gamma-1} \right].$$

This in turn implies

$$T(a_b + \alpha) - T(a_b) \leq \alpha \frac{k}{1+k} \left[\epsilon_{\max} + (1-\gamma)(1-\epsilon_{\min}) \right].$$

For our parameter values, this satisfies discounting.

B.2 Endogenous Wage Model

To prove that (20) is a contraction mapping, note first that it satisfies monotonicity because both its right- and left-hand sides are increasing in applications.

To prove discounting, define $Z(\widetilde{A})$ to be the log of the right-hand side of (20), i.e.:

$$Z(\widetilde{A}) = \ln\left(\widetilde{c} \widetilde{W}_b^{k\rho} \right) + \ln\left(\sum_a U_a f_{ab} \left[\sum_{b'} V_{b'} e^{-k\tilde{A}_{b'}} \widetilde{W}_{b'}^{k\rho} f_{ab'} \right]^{\gamma-1} \right),$$

which implies:

$$Z(\widetilde{A} + \alpha) = Z(\widetilde{A}) + k(1-\gamma)\alpha.$$

4

Consider the left-hand side of (20), which can be written as:

$$\ln(T) + kT = Z.$$

This can be thought of as giving a mapping $T(Z)$ where:

$$T'(Z) = \frac{T(Z)}{1 + kT(Z)}.$$

From the mean value theorem we have that:

$$
\begin{aligned}
T\left(Z\left(\widetilde{A} + \alpha\right)\right) &= T\left(Z\left(\widetilde{A}\right)\right) + T'\left(\widetilde{Z}\right)\left[Z\left(\widetilde{A} + \alpha\right) - Z\left(\widetilde{A}\right)\right] \\
&= T\left(Z\left(\widetilde{A}\right)\right) + T'\left(\widetilde{Z}\right)k(1 - \gamma)\alpha \\
&= T\left(Z\left(\widetilde{A}\right)\right) + \frac{(1 - \gamma)kT\left(\widetilde{Z}\right)}{1 + kT\left(\widetilde{Z}\right)}\alpha < \alpha,
\end{aligned}
$$

which satisfies discounting.

C Standard Errors

This section outlines in more detail the two methods we use to compute the standard errors. The first measure (s.e.$_1$) is obtained as the standard deviation of the monthly parameter estimates. Assuming that parameters are stable across months is a necessary condition for this procedure to be valid. If the data used are serially correlated, one might expect that the parameter estimates themselves may be serially correlated, and that their standard errors need to be adjusted for this fact. However Table A6 shows that the serial correlation (of first and second order) in the estimates is small and never significant so the reported standard errors are nor adjusted for serial correlation.

The second measure of standard errors (s.e.$_2$) reports the "sandwich" standard error obtained from the non-linear least squares estimator, allowing for possible heteroskedasticity and spatial correlation in the residuals, but assum-

ing that the true covariance between residuals from areas more than 100 km apart is zero. Our estimated variance-covariance matrix of the parameters is given by $\widehat{V} = \widehat{H}^{-1}(\widehat{G}'\widehat{\Omega}\widehat{G})\widehat{H}^{-1}$, where $\widehat{H}$ is the Hessian of the objective function, $\widehat{G}$ is the Jacobian, and the spatial correlation matrix $\widehat{\Omega}$ is the product matrix of the residuals after imposing the restriction that residuals from areas more than 100 km apart are uncorrelated.

D   Descriptive data analysis and link to structural parameter estimates

This Appendix complements our model estimates of Section 4 by highlighting the role of various aspects of the data in explaining specific structural parameters. We process in three steps. First, we provide descriptive evidence on local matching patterns by estimating a conventional, reduced-form, matching function, augmented for local spillovers. Second, we obtain a restricted version of our structural model, which delivers a closed-form solution for the outflow rate and thus a clearer correspondence between data and model parameters. Third, we link monthly variation in our structural parameter estimates to monthly variation in the reduced-form matching function estimates, in the spirit of Gentzkow and Shapiro (2017).

D.1   Regression Models for the Vacancy Outflow Rate

In our reduced-form matching function specification, we regress the vacancy outflow rate in a ward on the stocks of unemployment and vacancies in both the local and surrounding wards, treated as exogenous.[3]

---

[3]Existing evidence on residential migration of the unemployment is clearly in line with our assumption of exogenous jobseekers' location. Gregg, Machin and Manning (2004) show that the unemployed in the UK rarely migrate in search of better job opportunities, and evidence suggests that those who both find a job and move location in a given year typically find a job first and then seek to move home if the commute from their current location is too inconvenient (Gregg, Machin and Manning, 2004, pp. 387-395).

Geographic spillovers are captured in the following regression equation:

$$
\log\left(\frac{M_{b,t}}{V_{b,t}}\right) = \alpha_0 + \alpha_1 \log(U_{b,t} + \beta_1 U_{5b,t} + \beta_2 U_{10b,t} + \beta_3 U_{20b,t} + \beta_4 U_{35b,t}) \tag{D1}
$$
$$
+ \alpha_2 \log(V_{b,t} + \gamma_1 V_{5b,t} + \gamma_2 V_{10b,t} + \gamma_3 V_{20b,t} + \gamma_4 V_{35b,t}) + \alpha_3 \log w_b + \varepsilon_{b,t},
$$

where $M_{b,t}$ is the vacancy outflow in ward b at time t, $U_{b,t}$ is the number of unemployed in ward b, $U_{5b,t}$ is the number of unemployed in wards within 5km of b (excluding b itself), $U_{10b,t}$ is the number of unemployed in wards between 5 km and 10 km of ward b, and so on; and similarly for vacancies. w_b denotes ward-level wages relative to mean wages within 10 km, and only varies across wards. This specification implies that the probability of filling a vacancy in b depends on local unemployment and on unemployment in the surrounding areas, whereby $\beta_i < 1$ would imply that more distant unemployed workers are less effective in filling a vacancy in b than local workers. Similarly, more vacancies in b and neighboring wards are expected to reduce the vacancy outflow rate in b, whereby $\gamma_i < 1$ implies that more distant vacancies have a diminishing effect. Specifications similar to (D1) have been estimated by Burda and Profit (1996) for Czech districts, and Burgess and Profit (2001) and Patacchini and Zenou (2007) for UK TTWAs.

We next define the total number of unemployed and vacancies within 10km of b:

$$
\widetilde{U}_{10b,t} = U_{b,t} + U_{5b,t} + U_{10b,t}; \qquad \widetilde{V}_{10b,t} = V_{b,t} + V_{5b,t} + V_{10b,t},
$$

and approximate (D1) by:

$$
\log\left(\frac{M_{b,t}}{V_{b,t}}\right) \approx \alpha_0 + \alpha_1 \log \widetilde{U}_{10b,t} + \alpha_2 \log \widetilde{V}_{10b,t}
$$
$$
+ \alpha_1 \left(\frac{1-\beta_2}{\beta_2} \frac{U_{b,t}}{\widetilde{U}_{10b,t}} + \frac{\beta_1 - \beta_2}{\beta_2} \frac{U_{5b,t}}{\widetilde{U}_{10b,t}} + \frac{\beta_3 - \beta_2}{\beta_2} \frac{U_{20b,t}}{\widetilde{U}_{10b,t}} + \frac{\beta_4 - \beta_2}{\beta_2} \frac{U_{35b,t}}{\widetilde{U}_{10b,t}} \right)
$$
$$
+ \alpha_2 \left(\frac{1-\gamma_2}{\gamma_2} \frac{V_{b,t}}{\widetilde{V}_{10b,t}} + \frac{\gamma_1 - \gamma_2}{\gamma_2} \frac{V_{5b,t}}{\widetilde{V}_{10b,t}} + \frac{\gamma_3 - \gamma_2}{\gamma_2} \frac{V_{20b,t}}{\widetilde{V}_{10b,t}} + \frac{\gamma_4 - \gamma_2}{\gamma_2} \frac{V_{35b,t}}{\widetilde{V}_{10b,t}} \right)
$$
$$
+ \alpha_3 \log w_b + \varepsilon_{b,t}. \tag{D2}
$$

Specification (D2) has the advantage of being linear in parameters, so instrumental variables and ward fixed-effects can be easily introduced. Returns to scale in the matching function can be assessed by comparing coefficients on $\log \widetilde{U}_{10b,t}$ and $\log \widetilde{V}_{10b,t}$, while coefficients on share variables $U_{b,t}/\widetilde{U}_{10b,t},...,$ $U_{35b,t}/\widetilde{U}_{10b,t}$, and $V_{b,t}/\widetilde{V}_{10b,t},..., V_{35b,t}/\widetilde{V}_{10b,t}$ indicate the relative effectiveness of unemployment and vacancies at different distances. The decision to normalize by unemployment and vacancies within 10 km in (D2) is arbitrary, but it is important to choose a normalization for which β_i and γ_i are not zero, and for which the share variables are not too large. Considering this, 10 km seemed the right choice. On average, about 5% of unemployment and vacancies within 10 km are in the local ward, one-third are within 5 km. Moving beyond the 10 km ring, there are about 4.5 times the number of unemployed and vacancies between 10 and 20 km as within 10 km and 16 times as many within 35 km.

Estimates of specification (D2) are reported in Table A7. Column 1 pools all months and wards without time or ward effects. The estimates are in line with the typical matching function results in which the probability of filling any given vacancy rises with the number of unemployed and falls with the number of vacancies. The coefficients on the unemployment and vacancy variables imply a returns-to-scale parameter of 0.977 ($= 1 + 0.201 - 0.224$), suggesting (something very close to) constant returns. It is not just the level of unemployment and vacancies within 10 km that affect the outflow rate but also their geographical mix. As expected, the closer the unemployed to a ward, the higher the local vacancy matching rate. From the coefficients on $U_{b,t}/\widetilde{U}_{10b,t}$ and $U_{5b,t}/\widetilde{U}_{10b,t}$ one can derive an estimate for β_2 of 0.22 and for β_1 of 0.53, i.e. unemployed workers outside the ward but within 5 km have 53% of the matching effectiveness as those within the ward and the unemployed in the 5-10km ring have an effectiveness of 22%. Unemployed in the 20 km and 35 km rings have tiny effects on the vacancy outflow, but statistically different from zero. For vacancies, the closer they are, the lower the local outflow rate, as jobs at shorter distances are are closer substitutes to local ones. Vacancies within 5 km have 23% of the effectiveness of those within the ward, and vacancies in the 5-10 km ring have an effectiveness of 21%. Vacancies in the 20 km

8

and 35 km rings have very small effects on the vacancy outflow rate. Column 2 introduces time dummies, with a very slight attenuation of all coefficients, but virtually identical conclusions. In both columns the coefficient on relative wages is positive and highly significant.

While this is the standard approach in the empirical matching function literature, there are concerns on the identification of the parameters of interest. For example, innovations in matching efficiency in an area, as represented by $\varepsilon_{b,t}$, may affect worker location and job creation, leading to an upward bias on the coefficients on both unemployment and vacancies. Furthermore, as the dependent variable is obtained by dividing the vacancy outflow by the local stock, which also appears in the construction of some of the right-hand side variables, a division bias issue may occur if the vacancy stock is measured with error. Column 3 thus instruments all vacancy and unemployment variables using the one-month lags in the corresponding inflows. The coefficients on the unemployment variables, as expected, are now lower – specifically the coefficient on $\log \widetilde{U}_{10b,t}$ is only slightly lower, while the one on $U_{b,t}/\widetilde{U}_{10b,t}$ is markedly lower – while the coefficients on vacancy variables are higher, consistent with a division bias, rather than an endogeneity bias. And indeed the coefficient which is mostly affected is the one on $V_{b,t}/\widetilde{V}_{10b,t}$, on which the local vacancy stock has the most influence. Overall, our previous qualitative conclusions on matching elasticities α_1 and α_2, as well as on the decay of spillover effects with distance, are robust to the introduction of instrumental variables. Column 4 introduces ward fixed effects and the most noticeable change is a marked increase in standard errors on all coefficients, as within-ward variation in unemployment and vacancy variables is smaller than the cross-section variation. This is especially true for unemployment variables, as within ward variation in (log) unemployment explains less than 3% of the total variance, while for (log) vacancies the within-ward variation explains 12% of the total variance. The matching elasticities α_1 and α_2 remain firmly significant, but the spatial distribution of spillovers is no longer precisely identified. This implies that most of the useful variation in investigating spatial matching is cross-sectional. Column 5 includes region fixed effects, as opposed to ward fixed effects, and the

9

resulting magnitude and significance of local spillovers are virtually unchanged from column 3, which does not include any geographic effects.

The dependent variable in specification (D2) is not defined when the outflow rate is zero. This becomes a relevant issue when using data on very small areas, and indeed the vacancy outflow is zero in 6.2% of observations in our sample. To deal with this we estimate outflow equations like (D2) in levels instead of logs.

Column 1 in Table A8 presents estimates of a log-linear matching function, having excluded unemployment and vacancies beyond 10km, as the estimates in Table A7 suggest that their impact is negligible. Column 2 estimates the level version of this equation by non-linear least squares, excluding observations with zero vacancy outflow, thus on the same sample as in column 1. The estimates are qualitatively similar, with a considerable reduction in the size of the coefficients on all ratio variables. Column 3 estimates the levels model but includes the "zeroes", i.e. the estimation method is the same as in column 2, but with a larger sample size. The estimates obtained are very close to those reported in column 2. Columns 4 and 5 report results for the log-linear and linear models estimated for one month only (February 2005), as done for some of the estimates of Section 4.

The results of Tables A7 and A8 are consistent with a simple matching model with spatial spillovers. However, these specifications have limitations for making inference about the size of local labor markets, as they are not informative about the reasons for the spillovers. In other words, the estimated effect of the number of unemployed 10 km away on the probability of filling vacancies in b may result from both those workers directly applying to vacancies in b, and from them applying for vacancies more local to them, say 5 km away, which then become harder to obtain, and causing workers 5 km away from b to shift their search efforts towards vacancies in b. These two scenarios, while observationally equivalent in reduced-form estimates, have different implication for the size of local labor markets and the evaluation of local intervention. Our structural model provides insight into the structure of local spillovers.

10

D.2 The Model: A Special Case

One feature of our structural model that makes it hard to visualize the link between parameter estimates and specific data features is the absence of a closed-form solution. In this subsection we consider a special case which instead delivers a closed-form solution and thus a clearer correspondence between data and model parameters.

Let's consider the model for job applications in (18), and impose that unemployment, vacancies and wages are equal across areas. Define $D_b = \sum_a e^{-\delta^* d_{ab}} I V_{ab}^k$ and assume that this is constant for all b i.e. that $D_b = D$. This amounts to the assumption that every areas is as well-connected as any other which would be the case if areas were regularly spaced on a sphere. Under these assumptions, vacancies in all areas receive the same number of applications. Using (18), this is given by:

$$\tilde{A} = \widetilde{c}\widehat{W}^{k\rho_1}e^{-k\tilde{A}}UD\left[Ve^{-k\tilde{A}}\widehat{W}^{k\rho_1}D\right]^{\gamma-1},$$

which can be re-arranged to give:

$$\tilde{A}e^{\gamma k\tilde{A}} = \widetilde{c}D^{\gamma}\widehat{W}^{\gamma k\rho_1}\left(\frac{U}{V}\right)V^{\gamma}. \tag{D3}$$

Equation (D3) states that applications rise with the U-V ratio. Conditional on the U-V ratio, the effect of the number of vacancies depends on the returns to scale in matching (γ), and is positive, zero or negative in case of increasing, constant or decreasing returns, respectively; and similarly for the impact of wages on applications. One can derive a similar expression if one assume that the cost of travelling between wards is infinitely high and all individuals live and work in the same ward and the cost of within-ward travel is zero - in this case one would have $D = 1$.[4]

The number of applications is unobserved, but is linked to the vacancy

---

[4]If, however, wages vary across areas, areas with higher wages attract more applications even under constant returns.

outflow rate through (17), which can be rewritten in log-linear form:

$$\ln \frac{M}{V} = \ln \lambda + \ln \left(1 - e^{-\tilde{A}} \right). \tag{D4}$$

Using (D3) and (D4), we obtain the partial effects of the U-V ratio and the number of vacancies on the vacancy outflow rate:

$$\frac{\partial \ln(M/V)}{\partial \ln(U/V)} = \frac{\tilde{A}e^{-\tilde{A}}}{1 - e^{-\tilde{A}}} \frac{1}{1 + \gamma k \tilde{A}}, \tag{D5}$$

$$\frac{\partial \ln(M/V)}{\partial \ln V} = \gamma \frac{\partial \ln(M/V)}{\partial \ln(U/V)}. \tag{D6}$$

We next show evidence on these model predictions.

D.3 Linking Structural and Non-Structural Parameter Estimates

This section aims to provide a mapping from the data to the parameters of interest following the approach of Gentzkow and Shapiro (2017). We do so by linking our structural parameter estimates to regression coefficients from reduced-form regressions, which have a clear partial-effect interpretation. Specifically, we estimate for each month in the sample a slightly modified version of the linear regression model (D2):

$$
\begin{aligned}
\log \left(\frac{M_{b,t}}{V_{b,t}} \right) = {} & \alpha_{0t} + \alpha_{1t} \log \frac{\widetilde{U}_{10b,t}}{\widetilde{V}_{10b,t}} + \alpha_{2t} \log \widetilde{V}_{10b,t} + \alpha_{3t} \left(\frac{U_{b,t}}{\widetilde{U}_{10b,t}} - \frac{V_{b,t}}{\widetilde{V}_{10b,t}} \right) \\
& + \alpha_{4t} \left(\frac{U_{5b,t}}{\widetilde{U}_{10b,t}} - \frac{V_{5b,t}}{\widetilde{V}_{10b,t}} \right) + \alpha_{5t} \log w_b + \varepsilon_{b,t},
\end{aligned}
\tag{D7}
$$

in which we have dropped unemployment and vacancies beyond 10 km and imposed equal coefficients on $U_{b,t}/\widetilde{U}_{10b,t}$ (respectively $U_{5b,t}/\widetilde{U}_{10b,t}$) and $V_{b,t}/\widetilde{V}_{10b,t}$ (respectively $V_{5b,t}/\widetilde{V}_{10b,t}$). We also use structural parameter estimates for each month (whose averages across months are reported in Table 1). Thus we

are left with 24 monthly estimates of reduced-form parameters $\alpha_{0t},...,\alpha_{5t}$, and 24 monthly estimates of structural parameters δ_t^*, γ_t, ρ_t, λ_t, $\widetilde{c}_t$, and regress each of the structural parameters (or some combination of them) on $\alpha_{0t},...,\alpha_{5t}$. The results of this exercise are reported in Table A9, and summarized in the following points:

- *Elasticity of the outflow rate to the U-V ratio*: $\frac{\tilde{A}e^{-\tilde{A}}}{1-e^{-\tilde{A}}}\frac{1}{1+\gamma k\tilde{A}}$. According to (D5), the log linear regression coefficient on the U-V ratio (α_{1t}) should be related to the number of applicants per vacancy. Column 1 in Table A8 explores this prediction by regressing the right-hand side of (D5) – as predicted by structural estimates[5] – on $\alpha_{0t},...,\alpha_{5t}$. Variation in dependent variable loads most heavily on α_{1t}, implying that job applications are closely linked to the elasticity of the outflow rate with respect to the U-V ratio, as predicted by (D5).

- *Returns to scale parameter*: γ_t. According to (D6), γ_t should be negatively related to the coefficient on the U-V ratio (α_{1t}) and positively related to the coefficient on vacancies (α_{2t}). These correlation patterns are validated by results of column 2, although α_{3t} also has a significant impact on γ_t. A plausible reason is that the unrestricted application model is more complex than the restricted model of the previous subsection.

- *Wage parameter*: ρ_t. Column 3 shows that this is mostly related, as it is to be expected, to the regression coefficient on wages, α_{5t}.

- *Scale parameter in applications*: (log) λ_t. According to (D4), λ plays the role of an intercept in an outflow rate outflow, and column 4 shows that the monthly estimate α_{0t} is indeed the parameter that has the strongest effect on λ_t.

- *Cost of distance parameter*: δ_t. The role of δ_t cannot be visualized in the restricted model of the previous subsection, which essentially assumes

---

[5]Specifically we obtain $\tilde{A}_t$ using (18) and take averages across wards for each month.

13

distance away, but – intuitively – it should be influenced by the relative importance of near and distant unemployment and vacancies in predicting vacancy outflows. Indeed column 5 shows that δ_t is more strongly influenced by α_{3t} than α_{4t}.

In summary, the correlations reported in Table A9 provide clear evidence on identification of structural parameters of our job search model by highlighting links with relevant features of the data.

References

[1] Burda, Michael and Stefan Profit. 1996. "Matching across Space: Evidence on Mobility in the Czech Republic." *Labour Economics* 3: 255-378.

[2] Burgess, Simon and Stefan Profit. 2001. "Externalities in the Matching of Workers and Firms in Britain." *Labour Economics* 8: 313-333.Flinn, Christopher and James Heckman. 1983. "Are Unemployment and Out of the Labor Force Behaviorally Distinct Labor Force States?" *Journal of Labor Economics* 1: 28-42.

[3] Gentzkow, Matthew and Jesse M. Shapiro. 2017. "Measuring the Sensitivity of Parameter Estimates to Sample Statistics." *Quarterly Journal of Economics*, forthcoming.

[4] Jones, Stephen R. G., and Craig W. Riddell. 1999. "The Measurement of Unemployment: An Empirical Approach." *Econometrica* 67: 147-161.

[5] Machin, Andrew. 2001. "The New Vacancy Survey: A New Series of National Statistics." National Statistics Features, ONS. Available at http://www.statistics.gov.uk/articles/labor_market_trends/VacancySurvey_LMT0703.pdf

[6] Patacchini, Eleonora and Yves Zenou. 2007. "Spatial Dependence in Local Unemployment Rates." *Journal of Economic Geography* 7: 169-191.

[7] Stokey, Nancy and Robert Lucas Jr. 1989. *Recursive Methods in Economic Dynamics*. Boston (MA): Harvard University Press.

D. Appendix Figures and Tables

Figure A1
Unemployment to vacancy ratios in England and Wales
Shades correspond to quartiles.

	0.000 - 0.670
	0.671 - 1.500
	1.501 - 3.570
	3.571 - 270.000

Figure A2
Scatterplot (by 100-metre bins) of postcode-based distance against ward-level distance
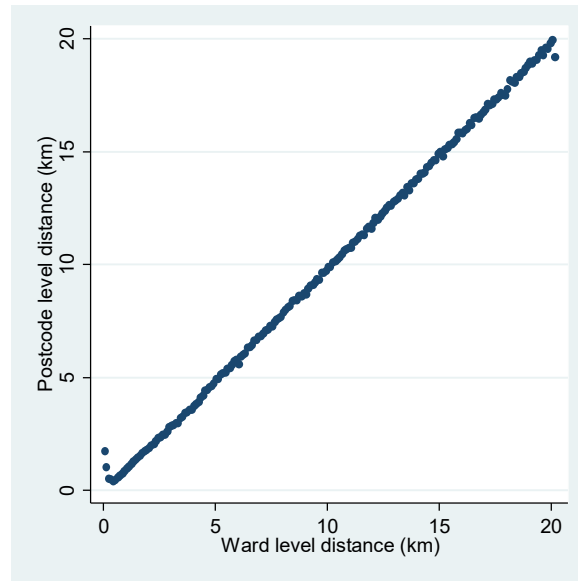
Notes. Each observation refers to a 100-meter bin. Source: ASHE.

Table A1
Descriptive statistics on local labor markets: Means and standard deviations

Variable	Mean	Standard deviation			No. Obs.
		Overall	Within ward	Within month	
Unemployment stock	106.5	147.8	14.9	147.7	208,762
Vacancy stock	91.9	228.3	61.3	228.0	208,762
Vacancy outflow	29.1	73.5	33.0	73.4	208,762
Vacancy Outflow Rate	0.331	0.201	0.184	0.198	208,762
U-V Ratio	3.69	8.03	5.23	8.00	208,762
Area	17.0	28.3	-	-	8,850
Wages	9.08	0.90	-	-	8,850

Notes. The sample includes CAS 2003 wards in England and Wales. Unemployment and vacancy variables are at monthly frequency (May 2004-April 2006) and are obtained from NOMIS. Area measures ward size in square km and is obtained from the 2001 Census. Wages are predicted based on the local industry composition of employment, combining information on the ward-level industry composition from BRES with median hourly wages by industry from ASHE. Due to small sample issues, predicted wages have been averaged at the ward-level. The overall standard deviation is across all ward-month observations. The standard deviation within ward is obtained after removing ward-level means. The standard deviation within month is obtained after removing month-level means.

Table A2
Descriptive statistics on local labor markets: Correlation Matrices

(A) Raw correlation matrix

	Unemploym.	Vacancies	Vacancy outflow	Vacancy outflow Rate	U-V Ratio
Unemployment	1				
Vacancies	0.366	1			
Vacancy outflow	0.374	0.913	1		
Vacancy Outflow Rate	0.083	-0.028	0.114	1	
U-V Ratio	0.189	-0.127	-0.118	0.115	1

(B) Correlation matrix after removing ward-level means

	Unemploym.	Vacancies	Vacancy outflow	Vacancy outflow Rate	U-V Ratio
Unemployment	1				
Vacancies	-0.144	1			
Vacancy outflow	-0.036	0.626	1		
Vacancy Outflow Rate	0.055	-0.004	0.260	1	
U-V Ratio	0.106	-0.069	-0.047	0.007	1

Notes. See notes to Table A1.

Table A3
Conditional logit estimates for the choice of transport mode

	Mode of transport		
	Walking	Cycling	Public Transport
Constant	-1.0563	-2.5607	-1.527
	(0.00102)	(0.00175)	(0.00073)
Distance	-0.1288	-0.0787	0.0043
	(0.00017)	(0.00024)	(0.00003)

Notes. The coefficients reported are obtained from a conditional logit model where the omitted (base) category is driving. The number of observations is 8.5 million. Source: 2001 Census Special Workplace Statistics.

Table A4
Estimates of a job application and matching model
Sample averages for May 2004-April 2005

	(1)	(2)	(3)	(4)	(5)
Cost of distance (δ^*)	0.220***	0.248***	0.215***	0.220***	
	(0.026)	(0.034)	(0.027)	(0.028)	
Returns to scale (γ)	-0.160***		-0.176***	-0.145***	
	(0.038)		(0.039)	(0.033)	
Wage elasticity (ρ)	0.920***	1.150***	0.942**	1.025***	
	(0.373)	(0.410)	(0.378)	(0.261)	
Matching effectiveness (λ)	0.371***	0.366***	0.371***	0.383***	0.425***
	(0.033)	(0.033)	(0.033)	(0.032)	(0.039)
Scale parameter in $\tilde{A}$ ($\tilde{c}$)	1.185***	0.557***	1.239***	1.075***	
	(0.207)	(0.083)	(0.202)	(0.131)	
Local U_b/V_b (α_1)			0.051		
			(0.043)		
Mismatch (α_2)				0.715*	
				(0.412)	
Number of months	24	24	24	15	24

Notes. Model specifications are the same as in Table 2. Coefficients reported are averages across monthly estimates, with standard deviations reported in brackets. Specification (4) is estimated on the months February 2005-May 2006, as unemployment data by occupation become available in January 2005.

Table A5
Average commuting times in the UK

	Mean	Std. Dev.	No. Obs.
Not on new job	24.4	22.2	620824
On new job, found via:			
Reply to advert	24.5	21.6	16117
Job centre	24.5	20.2	4499
Careers office	30.2	26.1	453
Jobclub	25.6	25.6	61
Private agency	34.6	26.3	4869
Personal contact	23.2	23.0	15639
Direct application	22.4	21.7	9673
Some other method	27.6	26.6	5708
Total	24.5	22.3	677843

Notes. Figures report one-way daily commuting times (in minutes). New jobs are defined by tenure up to three months. Source: Labour Force Survey, 1993-2007.

Table A6
Test for serial correlation in structural parameter estimates

	Dependent variable				
	δ_t^*	γ_t	ρ_t	λ_t	$\tilde{c}_t$
1st order lag	0.115	-0.030	0.075	-0.195	0.219
	(0.220)	(0.213)	(0.212)	(0.269)	(0.210)
Observations	23	23	23	23	23
1st order lag	0.166	-0.057	0.026	-0.103	0.129
	(0.229)	(0.203)	(0.232)	(0.287)	(0.209)
2nd order lag	-0.255	0.442	-0.026	0.277	0.421[*]
	(0.239)	(0.200)	(0.221)	(0.289)	(0.211)
Observations	22	22	22	22	22

Notes. The estimates reported are obtained from first- and second-order autoregressive models for the monthly parameter estimates summarized in Table 1.

Table A7
Log-linear matching functions with local spillovers

	Dependent variable: (log) vacancy outflow rate				
	(1)	(2)	(3)	(4)	(5)
Estimation method	OLS	OLS	IV	IV	IV
$\log \widetilde{U}_{10b}$	0.201***	0.193***	0.178***	0.120**	0.168***
	(0.00440)	(0.00477)	(0.00612)	(0.0581)	(0.00459)
$\log \widetilde{V}_{10b}$	-0.224***	-0.214***	-0.191***	-0.159***	-0.190***
	(0.00528)	(0.00573)	(0.00751)	(0.0311)	(0.00521)
$U_b / \widetilde{U}_{10b}$	0.714***	0.711***	0.317***	-0.155	0.222***
	(0.0563)	(0.0559)	(0.0798)	(0.415)	(0.0491)
$U_{5b} / \widetilde{U}_{10b}$	0.287***	0.281***	0.193***	0.0764	0.196***
	(0.0267)	(0.0267)	(0.0308)	(0.265)	(0.0177)
$U_{20b} / \widetilde{U}_{10b}$	-0.00135*	-0.00158**	-4.21e-05	-0.0103	-0.000530
	(0.000770)	(0.000755)	(0.00135)	(0.00982)	(0.000853)
$U_{35b} / \widetilde{U}_{10b}$	0.000262**	0.000233**	0.000683**	0.000350	0.000747***
	(0.000108)	(0.000108)	(0.000318)	(0.00221)	(0.000183)
$V_b / \widetilde{V}_{10b}$	-0.852***	-0.842***	-0.164***	-0.371**	-0.175***
	(0.0445)	(0.0444)	(0.0508)	(0.144)	(0.0308)
$V_{5b} / \widetilde{V}_{10b}$	-0.120***	-0.116***	-0.0588**	-0.0374	-0.0530***
	(0.0240)	(0.0240)	(0.0277)	(0.0882)	(0.0160)
$V_{20b} / \widetilde{V}_{10b}$	-0.000699	-0.000531	-0.00618***	-0.00559	-0.00651***
	(0.00106)	(0.00100)	(0.00153)	(0.00408)	(0.000971)
$V_{35b} / \widetilde{V}_{10b}$	-2.42e-05	-4.42e-06	-0.00101**	-0.000399	-0.000884***
	(0.000121)	(0.000116)	(0.000434)	(0.00181)	(0.000279)
$\log w_b$	0.169***	0.168***	0.178***		0.180***
	(0.0318)	(0.0318)	(0.0333)		(0.0158)
Observations	197,579	197,579	175,157	175,157	175,157
Time fixed effects	No	Yes	Yes	Yes	Yes
Ward fixed effects	No	No	No	Yes	No
Region fixed effects	No	No	No	No	Yes

Notes. The Table provides estimates for equation (D2). The relative wage coefficient cannot be estimated when ward fixed effects are included as it only varies across wards. Standard errors are clustered by ward and reported in brackets. Sample period: May 2004-April 2006.

Table A8
Matching functions in log and level

	(1)	(2)	(3)	(4)	(5)
Estimation method	OLS	NLLS	NLLS	OLS	NLLS
$\log \widetilde{U}_{10b}$	0.192***	0.178***	0.172***	0.191***	0.167***
	(0.004)	(0.004)	(0.004)	(0.010)	(0.011)
$\log \widetilde{V}_{10b}$	-0.212***	-0.220***	-0.171***	-0.205***	-0.162***
	(0.00510)	(0.005)	(0.005)	(-0.012)	(0.013)
$U_b / \widetilde{U}_{10b}$	0.706***	0.502***	0.332***	0.532***	0.196
	(0.0560)	(0.050)	(0.045)	(0.141)	(0.139)
$U_{5b} / \widetilde{U}_{10b}$	0.290***	0.203**	0.201*	0.231***	0.149**
	(0.0264)	(0.024)	(0.023)	(0.065)	(0.062)
$V_b / \widetilde{V}_{10b}$	-0.843***	-1.025***	-0.429***	-0.705***	-0.365***
	(0.0444)	(0.040)	(0.035)	(0.102)	(0.102)
$V_{5b} / \widetilde{V}_{10b}$	-0.117***	-0.091	-0.031	-0.046	0.045
	(0.0240)	(0.022)	(0.021)	(0.060)	(0.059)
$\log w_b$	0.168***	0.139***	0.0135***	0.176**	0.073
	(0.0319)	(0.029)	(0.028)	(0.072)	(0.069)
Observations	197579	197579	208717	8282	8708
Functional form	Log	Level	Level	Log	Level
Time effects	Yes	Yes	Yes	No	No
Sample	Non-zero Outflow	Non-zero Outflow	All	Feb 2005; Non-zero outflow	Feb 2005; All

Notes. Columns (1) and (4) provide estimates for equation (D2). Columns (2), (3) and (5) provide estimates for the exponential of equation (D2). Standard errors are clustered by ward and reported in brackets. Sample: May 2004-April 2006 in columns (1)-(3) and February 2005 in column (5).

Table A9
The relationship between structural parameters
and coefficients from log-linear regression models

	Dependent variable: Estimates from structural model				
	(1)	(2)	(3)	(4)	(5)
	Elasticity of vacancy outflow rate w.r.t U-V	Returns to scale (γ_t)	Coefficient on wages (ρ_t)	Matching effectiveness (λ_t)	Cost of Distance (δ_t^*)
Constant (α_{0t})	-0.085 (0.068)	-0.116 (0.081)	0.006 (0.787)	0.997*** (0.056)	0.045 (0.062)
Coef on $\log(\tilde{U}_{10b,t}/\tilde{V}_{10b,t})$ (α_{1t})	1.527*** (0.207)	-0.551** (0.246)	1.678 (2.389)	0.512*** (0.171)	0.200 (0.188)
Coef on $\log(\tilde{V}_{10b,t})$ (α_{2t})	0.768 (0.751)	2.123** (0.894)	10.200 (8.683)	0.116 (0.621)	-0.007 (0.683)
Coef on $\left(\frac{U_{b,t}}{\tilde{U}_{10b,t}} - \frac{V_{b,t}}{\tilde{V}_{10b,t}}\right)$ (α_{3t})	-0.042 (0.035)	0.100** (0.042)	0.377 (0.406)	0.102*** (0.029)	0.083** (0.032)
Coef on $\left(\frac{U_{5b,t}}{\tilde{U}_{10b,t}} - \frac{V_{5b,t}}{\tilde{V}_{10b,t}}\right)$ (α_{4t})	-0.119 (0.097)	-0.064 (0.116)	-0.861 (1.123)	0.031 (0.080)	0.033 (0.088)
Coef on $\log(w_{b,t})$ (α_{4t})	-0.072	0.202*	3.650***	0.017	0.011
Observations	24	24	24	24	24
R-squared	0.820	0.548	0.563	0.959	0.445

Notes. The Table reports estimates of linear regression models in which the dependent variable is the monthly estimate of a given parameter of the structural model (or, in the case of column (1), a function of parameters), and the independent variables are regression coefficients from the monthly estimates of the reduced-form model for the vacancy outflow rate reported in equation (D7).