Job search and spatial variation in local labor markets: A very disaggregated approach

Alan Manning, Barbara Petrongolo and Timothée Carayol

LSE

PSE, 2 April 2010

Motivation

- Substantial disparities in economic activity across the UK and other countries
  - from North-South divide to differences across very small areas
  - London has neighbourhoods with some of the highest unemployment rates and others with the highest vacancy rate in the country.
- Need to investigate very disaggregated aspects of the process by which the unemployed find jobs and employers fill vacancies.
- Key ingredients for this task: appropriate data and methodology
- Methodology: labor market matching function $M = m(V, U, X)$
  - popular tool that needs to be adapted to very disaggregate aspects of labor market search
- Data: recent availability of very disaggregated geographical data

Approach

- Model of unemployed job search across space
    - spatial unit: census ward
- Outcome in one area related to what's happening in 'neighboring' areas
- Hard to model these interactions in an optimizing model as every area affects every other, even if only to a small degree
    - curse of dimensionality
- But one can make progress if can find a way to reduce dimensionality

Objectives (I)

1. Provide and implement a new methodology for modelling the interdependence between different areas of the labor market
   - model for the process by which jobseekers in each area decide whether to send job applications to any other area.
2. Characterize the effective size of local labor markets
   - Whether areas *a* and *b* belong to the same labor market – how far afield the unemployed are prepared to travel when considering jobs. Policy relevance:
     - What are labor market effects of reducing transportation costs? (e.g. new train line)
     - What geographic spillovers can be expected from stimulus to job creation in a given area?

Objectives (II)

3. Provide novel evidence on the extent of returns to scale in the matching function, i.e. do thicker search markets deliver a higher matching rate?
    - recurrent question in labor/macro/ec geography
    - IRS would lead to possibility of multiple equilibria (role of policy)
    - and are potential source of agglomeration as large-scale markets then offer a more efficient matching process that can, in principle, off-set to higher land and labor costs of locating in agglomerations.

Related work (I)

- Broad literature on local interactions in various contexts (unemployment, education, crime etc.)
- Increasing emphasis on *nonmarket interactions*, via learning, social pressure, imitation etc.
    - Social interactions among neighbors may explain variation in local crime rates (Glaeser Sacerdote Scheinkman QJE 1996), welfare participation (Bertrand Luttmer Mullainathan QJE 2000), unemployment (Topa RES), commuting (Bayer Ross Topa JPE 2008), and other phenomena.
- Contribution of this paper is on *market interactions* in local variation in unemployment and commuting patterns
    - mechanism: job search across space

Related work (II)

Matching models with local spillovers

- Most existing work estimates a specification like

$$\ln M_b = \beta_0 + \beta_1 \ln U_b + \beta_2 \ln V_b + \beta_3 \ln U_b^* + \beta_4 \ln V_b^* + e_b$$

- M_b is the vacancy outflow in area b
- U_b and V_b are local vacancies and unemployment
- U_b^* and V_b^* are vacancies and unemployment within a cetrain radius from b
- Typical results: $\beta_1, \beta_2 > 0$; if spillovers: $\beta_3 > 0, \beta_4 < 0$.
  - Burda and Profit 1996 on Czech districts.
  - Burgess and Profit, 2001, on UK TTWA - much bigger areas than wards.

Related work (II)
Drawbacks

Simple and Intuitive specification, but with drawbacks:

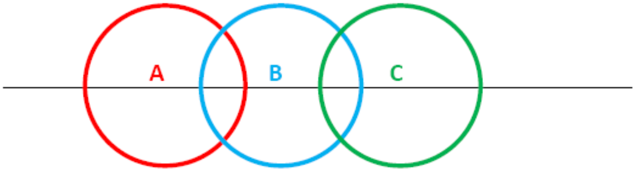
- Equal spillovers within a given radius, and arbitrary cut-off for the distance which has an effect
- Mechanisms driving spillovers uncovered
- Labor markets in areas with long-distance commuting to center may end up as being measured as misleadingly large.
- Discontinuities around boundaries.

Our contribution:

- no "structure" imposed on spillovers - rather let data characterize local labor markets.

An example

- Although labor markets for A and C are distinct, there is no way to divide the line into two without mis-measuring the labor market for B.
- So it is important to derive measures of labor markets that vary much more continuously with geographic space. Feasible when data available for very small areas.



Data

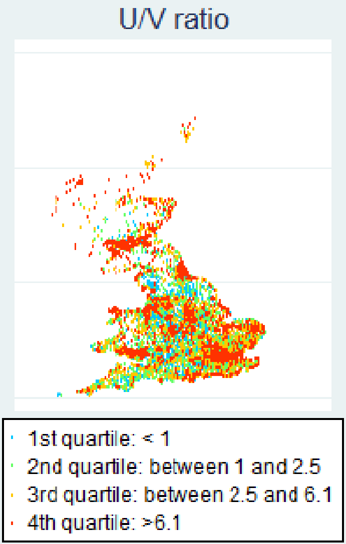
- Data on registered unemployment and job vacancies collected at UK Jobcentres (government funded employment agencies)
- Jobcentres: are close to physical "marketplaces" where employers post their job vacancies and the unemployed sample them
- Claimant unemployed need to visit Jobcentres to collect welfare benefits
- Vacancies advertised at Jobcentres represent only a fraction of available job opportuinities - about 45% during 1985-2001 (Coles and Petrongolo, 2008)
- Low-skill jobs are over-represented - but also the unskilled are over-represented among the unemployed
- All data disaggregated at the Census ward level.
- 10,072 wards in the UK. Average population about 6,000.

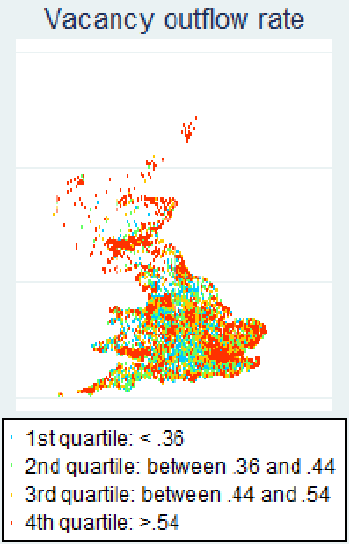
Unemployment and vacancies across UK wards

In May 2004, wards have on average 87 unemployed and 59 vacancies, with huge spatial variation

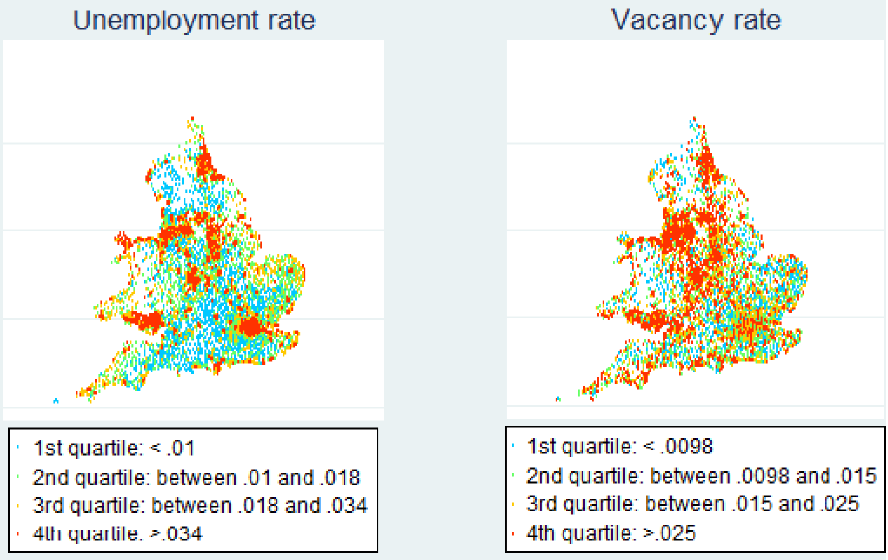
	Mean	St. Dev	Min	Max
Unemployment stock	86.7	119.1	0	1910
Unemployment inflow	20.6	24.2	0	330
Unemployment outflow	23.8	26.3	0	307
Vacancy stock	86.4	234.5	0	8700
Vacancy inflow	31.6	77.9	0	2530
Vacancy outflow	25.9	62.4	0	2201
Observations		10,072		

Descriptive evidence (I)





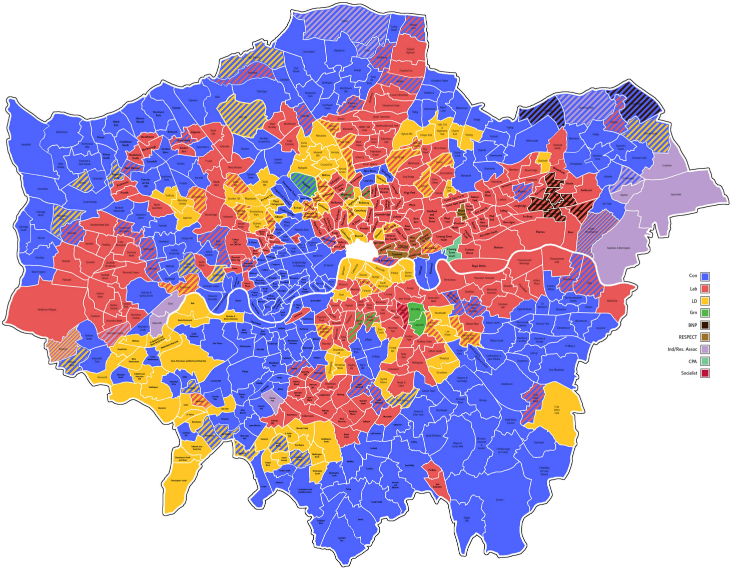
Descriptive evidence (II)



Our working sample

London (633 wards) and South East (1499 wards), cross section May 2004

London wards



A log-linear matching function for London and the SE

		Dependent variable: ln(vacancy outflow)		
Local effects	$\ln(U)$	0.107***	0.090***	0.076***
		(0.014)	(0.023)	(0.023)
	$\ln(V)$	0.821***	0.828***	0.830***
		(0.013)	(0.013)	(0.013)
Spatial spillovers	$\ln(U^*)$		0.050*	0.079**
	(up to 5km)		(0.026)	(0.031)
	$\ln(V^*)$		-0.051**	-0.053**
	(up to 5km)		(0.025)	(0.027)
	$\ln(U^*)$			0.129**
	(5-20km)			(0.053)
	$\ln(V^*)$			-0.226***
	(5-20km)			(0.062)
	$\ln(U^*)$			-0.003
	(20-35km)			(0.060)
	$\ln(V^*)$			0.009
	(20-35km)			(0.078)
	$\ln(U^*)$			0.025
	(35-50km)			(0.053)
	$\ln(V^*)$			-0.050
	(35-50km)			(0.077)
	Observations	2017	1954	1954
	R^2	0.710	0.711	0.718
	Returns to scale	0.929***	0.916***	0.816***

A model of spatial spillovers

Steps

1. Job applications - how unemployed in each area a decide to apply to jobs in any area b
   - using model obtain applications per job in each area b
2. From applications to job matches in area b

Job applications

- Unemployed workers in origin area a decide whether to apply to job vacancies in any destination area b. $a, b = 1, ..., T$.
- Goal: determine A_b (applications per job in destination area b).
- Application to vacancy i has a probability of p_i of being successful, and if so gives utility u_i. Using LLN, expected return for a worker is:

$$\sum_i D_i p_i u_i \tag{1}$$

  where $D_i = 1$ if apply to vacancy i and zero otherwise.
- Cost of making N applications:

$$\frac{c}{1+\eta} N^{1+\eta} \tag{2}$$

- Apply to vacancy i iff $p_i u_i \geq cN^{\eta}$

Job applications: more structure

- Assume:

$$\begin{aligned} p_b &= p_b(A_b), \quad p_b' < 0 \\ u_{ab} &= f_{ab}\varepsilon \end{aligned}$$

where f_{ab} is function (among other factors) of distance a to b, and idiosyncratic component ε is Pareto distributed with exponent k.

- Thus

$$\Pr(\text{unemployed in } a \text{ applies to vacancy in } b) = \left(\frac{p_b(A_b)f_{ab}}{cN_a^{\eta}}\right)^k \quad (3)$$

- Using LLN, applications sent by unemployed in a are

$$N_a = \sum_b V_b \left(\frac{p_b(A_b)f_{ab}}{cN_a^{\eta}}\right)^k = \left[\sum_b V_b \left(p_b(A_b)f_{ab}\right)^k\right]^{\gamma} \quad (4)$$

where $\gamma = 1/(1+\eta k)$

Job applications to each area b

- Applications sent by unemployed in a to vacancies in b

$$N_{ab} = c^{-k\gamma} \left(p_b(A_b) f_{ab}\right)^k \left[\sum_{b'} V_{b'} \left(p_{b'}(A_{b'}) f_{ab'}\right)^k\right]^{\gamma-1} \tag{5}$$

- Assuming $p_b(A_b) = A_b^{-\widetilde{\beta}}$:

$$\begin{aligned} A_b &= \frac{\sum_a N_{ab} U_a}{V_b} \\ &= \left\{ c^{-k\gamma} \sum_a U_a \left(f_{ab}\right)^k \left[\sum_{b'} V_{b'} \left(A_{b'}^{-\widetilde{\beta}} f_{ab}\right)^k\right]^{\gamma-1} \right\}^{1/(1+\widetilde{\beta}k)} \end{aligned}$$

Job applications to each area b (contd.)

- Finally assume $f_{ab} = \exp(-\widetilde{\delta} d_{ab})$ and redefine $\beta = \widetilde{\beta} k$; $\delta = \widetilde{\delta} k$

$$A_b = \left\{ c^{-k\gamma} \sum_a U_a \exp(-zd_{ab}) \left[\sum_{b'} V_{b'} A_{b'}^{-\beta} \exp(-zd_{ab'}) \right]^{\gamma-1} \right\}^{1/(1+\beta)} \tag{6}$$

- With thousands wards, this means thousands equations in thousands unknowns.
- But under reasonable conditions, (6) is contraction mapping, thus can be solved iteratively to obtain a vector of A_b.

Job applications: Returns to scale

- A_b is homogeneous in U_a and V_a, $a = 1, ..., T$, of degree

$$\frac{\gamma}{1+\beta\gamma}$$

- $\gamma = 0$ implies that applications received by each job in b are independent of the size of all local labor markets, i.e. there are constant returns to scale
- Intuition: recall total applications sent by unemployed living in a

$$N_a = \left[\sum_b V_b \left(p_b(A_b) f_{ab} \right)^k \right]^\gamma$$

- $\gamma = 0$ implies total applications from a unchanged when size of all possible destination areas double.
- $\gamma/(1+\beta\gamma) > 0$ implies increasing returns to scale, and viceversa

Returns to scale: An example

- Special case in which areas are isolated, so $f_{ab} = f > 0$ if $a = b$, and $f_{ab} = 0$ if $a \neq b$.
- Equation (6) becomes

$$\log A_b = \frac{1}{1+\beta\gamma} \log\left(\frac{U_b}{V_b}\right) + \frac{\gamma}{1+\beta\gamma}\left[\log(V_b) + \log(f)\right]$$

- $\gamma = 0$ means that applications to an area do not depend on the size of the area (V_b), thus CRS.

From applications to job matches

- Job matches (vacancy outflow) in area b :

$$\frac{M_b}{V_b} = \Psi(A_b) + e_b$$

- Use urn-ball model. Denote by π the probability that upon contact, a match is made. The probability that a given vacancy is not filled by a worker is $(1-\pi)^{A_b}$.
- Thus the total vacancy outflow in area b is $M_b = V_b\left[1-(1-\pi)^{A_b}\right]$

Empirical specifications

- For small enough π: $(1-\pi)^{A_b} \simeq \exp(-\pi A_b)$, thus estimate

$$\ln M_b = \ln V_b + \ln \left\{1 - \exp\left[-\exp(b_0 + b_1 \ln A_b)\right]\right\} + e_b$$

- Alternatively: log linear specification

$$\ln M_b = \ln V_b + b_0 + b_1 \ln A_b + e_b$$

- Recall A_b is given by

$$A_b = \left\{ c^{-k\gamma} \sum_a U_a \exp(-zd_{ab}) \left[\sum_{b'} V_{b'} A_{b'}^{-\beta} \exp(-zd_{ab'}) \right]^{\gamma-1} \right\}^{1/(1+\beta)}$$

so β, γ, z are also parameters to be estimated.

Main estimates

Model:	urn-ball		log-linear	
z	0.273^{***} (0.035)		0.265^{***} (0.040)	
γ	-0.501^{***} (0.042)		-0.502^{***} (0.043)	
β	0.442 (0.329)		0.420 (0.332)	
const	0.108^{***} (0.217)	-0.408^{***} (0.155)	-0.239 (0.182)	-0.686^{***} (0.132)
$\ln(A_b)$	0.306^{***} (0.047)	0.209^{***} (0.032)	0.265^{***} (0.040)	0.180^{***} (0.028)
$\ln(U_b/V_b)$		0.139^{***} (0.014)		0.123^{***} (0.012)
Obs.:	2022	2022	2022	2022
R^2	0.940	0.943	0.940	0.943

Further results

Model:	urn-ball			
	Scotland		Wales	
z	0.551^{***} (0.128)		0.226^{***} (0.054)	
γ	-0.673^{***} (0.252)		-0.318^{**} (0.162)	
β	-0.329 (0.426)		0.314 (0.412)	
const	-0.432^{***} (0.083)	-0.543^{***} (0.055)	-0.680^{***} (0.158)	-0.939^{***} (0.080)
$\ln(A_b)$	0.151^{***} (0.033)	0.127^{***} (0.018)	0.220^{***} (0.051)	0.129^{***} (0.043)
$\ln(U_b/V_b)$		0.064^{***} (0.017)		0.099^{***} (0.023)
Obs.:	1155	1155	798	798
R^2	0.962	0.963	0.933	0.935

Estimated returns to scale for London and South East

Consider

$$
\begin{aligned}
A_b &= \left\{ c^{-k\gamma} \sum_a U_a \exp(-z d_{ab}) \left[\sum_{b'} V_{b'} A_{b'}^{-\beta} \exp(-z d_{ab'}) \right]^{\gamma - 1} \right\}^{1/(1+\beta)} \\
\ln M_b &= \ln V_b + b_0 + b_1 \ln A_b + e_b
\end{aligned}
$$

- A_b is homogeneous of degree $\gamma/(1+\beta\gamma)$ in U_a and V_a
- Thus M_b is homogeneous of degree $1 + b_1\gamma/(1+\beta\gamma)$
- With above estimates $1 + b_1\gamma/(1+\beta\gamma) = 0.831$
- Evidence of decreasing returns

Predicted commuting patterns

- Share of applications to area b that come from area a

$$\frac{U_a N_{ab}}{A_b V_b}$$

- Number of vacancies in area b filled by someone in area a

$$\frac{U_a N_{ab}}{A_b V_b} M_b$$

- Share of workers who live in a and work in b

$$\frac{N_{ab} M_b / A_b V_b}{\sum_{b'} N_{ab'} M_{b'} / A_{b'} V_{b'}}$$

This gives the distribution of commutes predicted by the model, i.e. the share of workers who live in a that work in b.

Data on commuting

- Data on actual commutes: Special Workforce Statistics 2001 Census – count of all those who live in one ward and work in another
- Census data cover all workers, while model predicts commutes associated to newly-filled vacancies by the registered unemployed.
- Commutes for these two groups may not coincide, but the UK LFS contains data on commuting times for those in new jobs (tenure $\leq$3 months) and those in continuing jobs, and, for those in new jobs, how that job was obtained.

Commuting times in the UK Labour Force Survey

One-way daily commuting times in minutes

	Mean	St. Dev	No. Obs.
Not on new job	24.5	22.2	612787
On new job, found via:			
Reply to advert	24.5	21.6	16059
Jobcentre	**24.5**	20.2	4491
Careers Office	30.2	26.1	453
Jobclub	25.6	25.6	61
Private agency	34.6	26.4	4859
Personal contact	23.2	23.0	15523
Direct application	22.4	21.7	9646
Some other method	27.5	26.7	5618
Total	24.5	22.3	669497

Model predictions for commuting patterns

Estimated correlation between the commuting flows predicted by our model and the actual flows provided by the Census:

- 0.51^{***} for the urn-ball model.
  - rises to 0.66^{***} if one excludes the locals (people who work and live in the same ward)
  - our model matches the data fairly well in general but underpredicts locals
  - possible extension is $f_{ab} = \exp(-\widetilde{\delta} d_{ab} + f)$, where $f > 0$ for $a = b$ and $f = 0$ otherwise
- for the log-linear model correlations are 0.50^{***} and 0.65^{***}

Conclusions

- Paper proposed model for interactions among areas that are numerous and small – without imposing boundaries on spillovers, but rather letting data characterize local labor markets.
- A spatial model for job applications performs better than the simpler log-linear specifications.
- Evidence of spillovers - but they decay rapidly with distance. Attractiveness of a job falls by factor of about 3 when moving job from 5 to 10 km away.
    - important implications for policies aimed at geographic aspects of disadvantage

Conclusions (contd.)

- No evidence of IRS - if anything estimates predict DRS.
- Evidence of congestion effects stemming from job competition in a given area - but imprecisely estimated.
- Model predicts fairly good correlation with observed commuting patterns

Extensions

- Enrich this model - This model underpredicts "locals", i.e. people who work and live in the same ward.
  - allow for "unrestricted" local job applications to correct failure of the model
- Job heterogeneity
  - introduce occupational dimension