

# Commuting, Migration and Local Joblessness

Michael Amior and Alan Manning\*

October 2017

VERY PRELIMINARY AND INCOMPLETE

## Abstract

Similar to the US, the UK suffers from substantial persistence in local jobless rates. This reflects long run declines in labor demand in manufacturing heartlands, driven by secular changes in the industrial composition of employment. There is a large response from local population (similar in magnitude to the US), but it lags behind the shift in local employment. However, beyond migration, there is another local adjustment mechanism which has received little attention in the literature: changes in commuting behavior. This is likely to be especially important in a small and densely populated country like the UK. In this paper, we develop an integrated framework for analysing and estimating the migration and commuting responses to local demand shocks, and which is applicable to any level of spatial aggregation.

## 1 Introduction

The UK has very persistent regional inequalities in joblessness. This is illustrated in the first panel of Figure 1, which compares employment-population ratios (from here on, “employment rates”) in 1980 and 2010 among men aged 16-64, for the 80 largest British Travel to Work Areas (TTWAs). The correlation is 0.79. In popular discussion, these differences are often described as the “North-South divide”; and indeed, Figure 1 shows employment rates in Northern TTWAs are almost always lower than in Southern TTWAs (see Blackaby and Manning, 1990, for earnings, Henley, 2005, for output, Dorling, 2010, for a wider range of

---

\*Amior: Hebrew University of Jerusalem and Centre for Economic Performance, London School of Economics; michael.amior@mail.huji.ac.il. Manning: Centre for Economic Performance, London School of Economics; a.manning@lse.ac.uk. We are grateful to seminar participants at the ICEPR/IZA Labour Symposium and Berkeley.

variables showing a North-South difference). The conventional explanation of this North-South divide is that rates of internal migration are minimal, so migrations fails to erode spatial differences in economic opportunity. But, the second panel of Figure 1 shows a large population response to local unemployment, similar to that documented in the US by Amior and Manning (2015). One might then reasonably ask how jobless rates can persist in the face of this migratory response. As in the US, this can be explained by large persistence in the demand shocks themselves: those Northern cities which shed manufacturing jobs in the 1960s and 1970s continue to shed jobs today. This is illustrated in Figure 2, which compares local employment growth over 1971-1991 with 1991-2001.

The purpose of this paper is study how local labor markets adjust in response to these demand shocks. The main adjustment mechanisms are expected to be migration and commuting. On the one hand, we would expect higher out-migration from and lower in-migration to adversely affected areas. And on the other, those workers who do not move should increasingly switch to jobs located outside the affected area. This paper is about the effectiveness of these two adjustment mechanisms.

Most existing studies of adjustment to local demand shocks have focused on the migration channel (see e.g. Blanchard and Katz, 1992; Eichengreen, 1993; Decressin and Fatás, 1995; Obstfeld and Peri, 1998; Beyer and Smets, 2015; Dao, Furceri and Loungani, 2014; Amior and Manning, 2015), dividing countries into non-overlapping labour markets within which it is assumed that labor market opportunities are equalized. But, the UK is a relatively small, densely populated country, and it is difficult to sub-divide the country in this way. In this context, commuting behavior may be an important channel through which local economic opportunity is equalized, and infrastructure investments have been proposed to facilitate that. But there are few papers which study how commuting behavior responds to economic shocks; two exceptions are Monte, Redding and Rossi-Hansberg (2015) and Manning and Petrongolo (2017).

The aim of this paper is to develop an integrated framework for analysing and estimating both the commuting and migration responses to local demand shocks, and which is applicable to any level of spatial aggregation. The plan of the paper is as follows. The second section describes the data we use for the empirical part of the paper. We use two levels of spatial aggregation - Travel-To-Work-Areas (TTWAs) which are constructed to be, as far as possible, self-contained labour markets (hence the closest equivalent of the Commuting Zones (CZs) often used in the US) and wards which are neighborhoods. In both cases, we study decadal changes using census data from the period 1971-2011. The third section presents a model

of commuting which conditions on the distribution of population across wards and their overall employment rate. We show how one can decompose panel data on commuting into a time-invariant cost of commuting between two wards, and a time-varying ward-specific fixed effect that can be thought of as a measure of the attractiveness of working in that area, e.g. its wage. We then develop and estimate a model of the attractiveness of working in different areas.

In the fourth section, we construct a theory of the employment rate. We show how the employment rate of residents of an area would be expected to be a function of the level of population in the area and the inclusive value from commuting, up to an origin fixed effect and a time fixed effect. We estimate this model showing that it works well. We then extend the sufficient statistic result of Amior and Manning (2015), providing conditions under which the welfare of the residents of an area can be written as a function of the utility of being unemployed and the employment rate.<sup>1</sup> This model of local equilibrium for a fixed population is combined with a simple model for migration, in which people move away from areas with low utility and towards those with high utility, taken from Amior and Manning (2015). We show how this leads to an error-correction mechanism (ECM) for local population growth which responds to changes in employment growth and a lagged disequilibrium term (the log employment rate). Our results of the migration model are reported in section 5. We first estimate the ECM for population change at both TTWA and ward level. Results are very similar using both levels of spatial aggregation, as our sufficient statistic result would suggest. The model fits the British data well. In our preferred ward-level estimates, the elasticity of population to contemporaneous (decadal) employment growth is 0.61, and the elasticity to the initial local employment rate is 0.42. This implies a large but incomplete population adjustment over ten years: it corrects for about half the initial deviation in the local employment rate. These estimates are indicative of more persistence than earlier studies, such as Blanchard and Katz (1992) and Decressin and Fatás (1995), suggest. However, they are not significantly different to our earlier US results based on our ECM model. Also, like in the US, we show any sluggishness in the response is manifested on the participation margin, rather than unemployment: adjustment of the local labour force is complete over one decade.

In summary, this paper offers:

1. A model of the commuting decision, i.e. a model of the decision of residents of one

---

<sup>1</sup>This has practical advantages, as employment rates are easier to measure than real consumption wages for detailed local geographies. Also, since the employment rate is a stock measure like population, our estimates are directly informative of the speed of population adjustment.

area about the area where they work. This utility from living in one area and working in another is written as a function of the wage offered and cost of the commute.

2. A model for the determination of wages offered by employers in an area.
3. A generalization of Amior and Manning’s (2015) result that the employment rate in an area (perhaps composition-adjusted) can serve as an (easily computed) sufficient statistic for economic opportunity - to the case where workers are permitted to work outside their area of residence.
4. A simple model of migration between areas.

## 2 Data

### 2.1 Geography

We use two levels of spatial aggregation, Travel-To-Work Areas (TTWAs) and wards. Wards are the basic building blocks of the data sets used and are relatively small areas with an average population of 5,700 in 2001. Ward boundaries have changed over time - we convert all years to the 9,975 Standard Table wards of the 2001 census (excluding Northern Ireland). When boundaries in different years do not match precisely, we always allocate population or employment counts proportionally according to address count or geographical area (if address counts are unavailable) - details of how we do this are in the Appendix A.

TTWAs are areas used in official publications by the Office for National Statistics intended to be self-contained labour market areas within which people live and work. The official TTWA scheme has been updated each decade using an iterative algorithm. The number of TTWAs has declined from 334 in 1981 to and 243 in 2001.<sup>2</sup> We use the 2001 scheme for our analysis and restrict attention to the 232 TTWAs on the mainland (excluding Northern Ireland). TTWAs are the most comparable geographical units in the UK to the Commuting Zones (CZs), originally developed by Tolbert and Sizer (1996) and used in many US studies including Amior and Manning (2015). Although most of our analysis is at ward level, we include some analysis at TTWA level because the comparison with the US is instructive. We offer some comparisons between TTWAs and CZs in Appendix B. They are

---

<sup>2</sup>See <http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/other/travel-to-work-areas/index.html>.

similar in terms of population, but the British TTWAs are significantly smaller in land area (and so, more densely populated), and there is relatively more commuting them.

## 2.2 Population, employment and commuting

We take our local (TTWA and ward-level) population and employment data from the published small area decadal census aggregates of 1971-2011 inclusive.<sup>3</sup> Our estimates are based on population and employment counts for all individuals aged 16-64. We provide further details on this data in Appendix A below. The commuting data comes from the special workplace statistics and record commuting flows between every pair of wards - the data are available for the 1981-2011 censuses inclusive.

## 2.3 Amenity controls

In the population response regressions in Section 5, we control for a number of variables that might affect the attractiveness of living in an area - beyond the labor market. We use controls which are similar to those used in our earlier work on the US (Amior and Manning, 2015) to aid comparability, though one should recognize that factors like climate vary much less in the UK than the US.

First, we control for the log distance from the TTWA's population-weighted centroid to the nearest coastline, as<sup>4</sup> coastline may provide consumption or productive amenities (Rappaport and Sachs, 2003) or physical constraints on population expansion (Saiz, 2010).

Second, we control for some climate indicators. Rappaport (2007) shows that Americans have increasingly located in cities with *pleasant* weather, specifically cool summers with low humidity and warm winters. And he argues that a growing valuation of climate amenities can help explain observed trends in Southern population, driven perhaps by rising incomes.<sup>5</sup> Cheshire and Magrini (2006) find similar trends among European regions. We control for the number of heating degree days (the average number of days temperature is below 15.5°C and heating is required, per year), cooling degree days (the average number of days temperature is above 22°C and cooling is required) and rainfall intensity (average precipitation on days when

---

<sup>3</sup>Unfortunately, the results from the 1961 census have not yet been digitized - see <http://britishlibrary.typepad.co.uk/socialscience/2013/01/census-statistics-and-resources.html>.

<sup>4</sup>Population-weighted centroids for counties in 1990 are estimated by the Missouri Census Data Center: <http://mcdc.missouri.edu/websas/geocorr90.shtml>. We estimate CZ centroids by computing the population-weighted averages across the latitudes and longitudes of county centroids.

<sup>5</sup>In particular, Rappaport finds that hot humid summers have deterred population growth, *controlling for winter temperature*. This is inconsistent with an important role for air conditioning.

there is more than 1mm). This data was kindly shared by Steve Gibbons, who constructed it from Met Office statistics<sup>6</sup> (Gibbons, Overman and Resende, 2011).

Third, we control for log population density in 1921. This measure can proxy for the pull of under-developed land. Alternatively, there may be consumption or productive amenities (or disamenities) associated with population density. We use a historical measure of density to ease concerns over endogeneity.<sup>7</sup>

We also control for an index of TTWA isolation. Specifically, this is the log distance to the closest TTWA, where distance is measured between population-weighted centroids in 1990. Isolation may matter for two reasons. First, it might be considered an amenity or disamenity. And second, it limits opportunities for commuting.

## 2.4 Instrumental Variables

As explained in more detail later, credible identification of the equations we estimate requires an instrument, one on the labour demand side and one on the labour supply side. In keeping with much of the literature<sup>8</sup>, we rely on the industry shift-share variables  $b_{rt}$  originally proposed by Bartik (1991) as a demand-side instrument. The idea is to assume that, over a decade, the stock of employment in each industry  $i$  grows at the same rate in every area  $r$ , where this growth rate is estimated using national-level data. Specifically:

$$b_{rt} = \sum_i \phi_{rt-1}^i (n_{i(-r)t} - n_{i(-r)t-1}) \quad (1)$$

where  $\phi_{rt-1}^i$  is the share of workers in area  $r$  at time  $t - 1$  employed in industry  $i$ . The term  $n_{i(-r)t} - n_{i(-r)t-1}$ , expressed in logs, is the growth of employment nationally in industry  $i$ , *excluding* area  $r$ . This modification to standard practice was proposed by Autor and Duggan (2003) to address concerns about endogeneity to local employment counts. The British small-area population census data only provides an industrial disaggregation to the 1-digit level, so we construct these instruments using data from employer surveys: see Appendix A for further details. As a result, while our population and employment counts are based on local *residents*, our instruments predict employment growth among local *firms*. This is immaterial

---

<sup>6</sup>See <http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/available/annual.html>.

<sup>7</sup>These densities are estimated using estimates of local population from the 1921 census, based on local government districts in England and Wales and Scottish parishes. We impute TTWA-level data using land area allocations. All population data and shapefiles for this exercise were extracted from Great Britain Historical GIS Project, [www.visionofbritain.org.uk](http://www.visionofbritain.org.uk).

<sup>8</sup>See, for example, Blanchard and Katz (1992); Bound and Holzer (2000); Notowidigdo (2011); ?; Beaudry et al. (2012; 2014b; 2014a).

as long as the instruments have sufficient power, and we confirm this below.

As an instrument on the labour supply side we use the idea that immigration is an important contribution to local population growth. Of course, local inflows of foreign migrants are partly a response to local employment growth. But, as is well known, migrants are often guided in their location choice by the “amenity” of established co-patriot communities.<sup>9</sup> In the empirical migration literature, there has been a long tradition (popularized by Card, 2001) of proxying these preferences with historical local settlement patterns. Following Card, we construct a “shift-share” predictor  $m_{ct}$  for the contribution of foreign migration to local population growth:

$$m_{rt} = \frac{\sum_o \phi_{rt-1}^o M_{o(-r)t}^{new}}{L_{rt-1}} \quad (2)$$

where  $\phi_{rt-1}^o$  is the share of population in area  $r$  at time  $t - 1$  which is native to origin  $o$ .  $M_{o(-r)t}^{new}$  is the number of new migrants arriving in the US (excluding area  $r$ ) between  $t - 1$  and  $t$ . The numerator of equation (2) then gives the predicted inflow of all migrants over those ten years to area  $r$ . This is scaled by  $L_{rt-1}$ , the initial population of area  $r$ . Similarly to the Bartik industry shift-shares above, the exclusion of area  $r$  helps allay concerns over endogeneity of the shift-share measure to the dependent variable, local population growth  $\Delta l_{rt}$ . We construct this migrant shift-share variables using small area aggregates from the census data. Population is decomposed by country (or country group) of birth, though these country categories vary by census cross-section. For each pair of census years, we use the greatest possible country-level detail.<sup>10</sup>

---

<sup>9</sup>For example, because of job networks (Munshi, 2003) or cultural amenities (Gonzalez, 1998).

<sup>10</sup>For the migrant shift-share between 1971 and 1981, we use 10 birth country categories (apart from British-born): Ireland, Old Commonwealth (i.e. Canada, Australia and New Zealand), African Commonwealth, Caribbean Commonwealth, Far Eastern Commonwealth, India, Pakistan/Bangladesh, other Commonwealth, other Europe, and a residual category. For the period 1981-1991, we are able to use 12 categories: all of the above, except we are able to disaggregate Pakistan and Bangladesh into two categories, and we can split the African Commonwealth category into East African Commonwealth and other African Commonwealth. For the 1990s, we are restricted to 10 categories: these include all those for the 1980s, minus Caribbean Commonwealth and “other Commonwealth” (both of which we place into the residual category). For the 2000s, we are able to use 23 categories: Ireland, other EU members in 2001, Poland, other Europe, North Africa, Nigeria, other Central/Western Africa, Kenya, South Africa, Zimbabwe, other South/Eastern Africa, Middle East, Far East, Bangladesh, India, Pakistan, other South Asia, USA, other North America, South America, Caribbean, Oceania, and a residual category.

## 2.5 Overview of Analysis

Our analytical framework and empirical results are in three sections. First (Section 3), we present and estimate a model for the commuting decision which treats residential decisions as fixed. We show how this can be used to derive a model for the employment rate and estimate that model (Section 4). We then embed this model of the employment rate in a simple model of population change (Section 5).

# 3 Commuting

## 3.1 Theoretical Model

We assume there are  $A$  areas and individuals can live and/or work in any of them. Denote the area of residential location by  $a$ ,  $a = 1, \dots, N$  and the area of working by  $b$ ,  $b = 1, \dots, A$ . For the moment, treat the residential decisions as fixed and also condition on being in work - both these decisions are discussed later. Assume the utility available to an individual living in  $a$  but working in  $b$  at time  $t$ :

$$U_{abt} = V_{abt} + \phi_{0a} - \phi \ln \widetilde{Q}_{at} + \epsilon_{abt} \quad (3)$$

where  $V_{ab}$  is a measure of how attractive it is work in  $b$  given one lives in  $i$ ,  $\phi_{0a}$  is the amenity value of living in  $a$  (assumed to be time-invariant), and  $\ln \widetilde{Q}_{at}$  is the log of the consumer price index for the residents of  $a$  at time  $t$ . and (or not work) and  $\epsilon_{ab}$  is an idiosyncratic utility shifter. Assume that the non-idiosyncratic gain in utility (3) from living in  $a$  and working in  $b$  at time  $t$  can be written as:

$$V_{abt} = d_{ab} + \phi \ln W_{bt} \quad (4)$$

where  $d_{ab}$  is an origin-destination fixed effect (assumed to be time-invariant) that influences commuting between areas - this might be a simple function of distance though could also be influenced by transportation networks and the cost of commuting from  $a$  to  $b$  and  $\ln W_{bt}$  is the attractiveness of jobs offered by employers in  $b$  at time  $t$  (the notation reflects the fact that this might be the wage though other factors could be important).

Individuals are assumed to choose the option that gives them the highest utility. Assume that, conditional on working, the idiosyncratic error term in (3) has a simple extreme value form: this leads to a multinomial logit structure for the probability of commuting from  $a$  to



$b$  at time  $t$ ,  $c_{abt}$  that is given by:

$$c_{abt} = \frac{e^{d_{ab} + \phi \ln W_{bt}}}{\sum_i (e^{d_{ai} + \phi \ln W_{it}})} \quad (5)$$

Note that the local consumer price index and the residential amenity drop out from this expression as, while they affect the utility from living in  $a$ , they do not affect the relative attractiveness of working in different areas conditional on being in work.

### 3.2 Estimates of commuting model

We use data on commuting to estimate this model treating the 'wage' variable as an unobserved destination-time fixed effect that is a parameter to be estimated. One can only identify the origin-destination fixed effects and destination-time fixed effects up to some normalizations - for example, a doubling of  $\ln W_{bt}$  or of  $d_{ab}$  leaves the commuting probabilities unchanged. To clarify what can be identified define:

$$D_{ab} = \frac{e^{d_{ab} + \phi \ln W_{b1}}}{\sum_i (e^{d_{ai} + \phi \ln W_{i1}})} \quad (6)$$

where  $t = 1$  is the first period (other normalizations are possible) and:

$$Z_{bt} = \frac{e^{\phi(\ln W_{bt} - \ln W_{b1})}}{\sum_i e^{\phi(\ln W_{it} - \ln W_{i1})}} \quad (7)$$

By construction  $D_{ab}$  sums to one for all  $a$  and  $Z_{bt}$  to one for all  $t$ . In addition  $Z_{b1}$  is assumed identical for all  $b$ .  $D_{ab}$  and  $Z_{bt}$  represent the most that can be identified from data on commuting patterns. Using (6) and (7), (5) can be written as:

$$c_{abt} = \frac{D_{ab} Z_{bt}}{\sum_i D_{ai} Z_{it}} \quad (8)$$

We estimate this by maximum likelihood. If the actual number of commuters from  $a$  to  $b$  at time  $t$  is  $C_{abt}$  (which is the data available to us), the log-likelihood can be written (up to a constant that does not depend on parameters) as:

$$\ln L = \sum_{a,b,t} C_{abt} \ln(c_{abt}) \quad (9)$$

which can be maximized over  $(D_{ab}, Z_{bt})$  subject to the constraints that  $D_{ab}$  sums to one for all  $a$  and  $Z_{bt}$  to one for all  $t$  and  $Z_{b1} = 1/A$ . Using (8), (9) can be written as:

$$\ln L = \sum_{a,b} \left( \sum_t C_{abt} \right) \ln D_{ab} + \sum_{b,t} \left( \sum_a C_{abt} \right) \ln Z_{bt} - \sum_{a,t} \left( \sum_b C_{abt} \right) \ln \left( \sum_i D_{ai} Z_{it} \right) \quad (10)$$

If there are  $A$  areas and  $T$  time periods this likelihood function (10) contains  $A(A-1)$  parameters in  $D_{ab}$  and  $(A-1)(T-1)$  parameters in  $Z_{bt}$  (all after allowing for the normalizations), approximately 99.5m parameters, so that estimation is not straightforward in practice. But an EM-algorithm can be used as, conditional, given an initial set of parameters one can update the parameters using a simple closed-form expression and this process converges to the ML estimates. Details of this process is in Appendix C. The estimates of  $(D_{ab}, Z_{bt})$  that emerge from this model are simply a large set of fixed effects that can be thought of as one way of describing the commuting flow matrices.

### 3.3 Modeling $D_{ab}$

We first used the estimates of  $D_{ab}$  to estimate a commuting model. We would expect  $D_{ab}$  to be related to the distance between origin and destination with more distant areas having lower commuting rates. The  $D_{ab}$  matrix has a large number of zeroes in it and the normalization that  $D_{ab}$  sums to one for all  $a$  suggests that a multinomial logit model might be an appropriate functional form. But there are too many destinations for this to be feasible so we exploit the well-known equivalence between the multinomial logit model and a Poisson model when an origin fixed effect is included (see, for example, Baker, 1994). The definition of  $D_{ab}$  in (6) makes it clear that it also includes a destination fixed effect (the term  $\ln W_{b1}$ ) so a Poisson model with two-way fixed effects is required. To estimate this model we use the iterative procedure suggested by Aitkin and Francis (1992) and Guimaraes (2004) - one uses a given set of fixed effects as offsets in a standard Poisson model and estimate the coefficients on the regressors of interest. Here we use a quadratic in the log of distance between wards, a functional form that we find to fit the data well. Then, with these estimates one re-estimates the fixed effects and repeats until convergence. This process can be slow but it does eventually converge without the need to invert matrices which in our case would have a magnitude of approximately 400m elements. This process does not produce valid estimates of standard errors - we follow Guimaraes (2004) and use a likelihood ratio test. The results are reported in Table 1.

As one would expect more distant jobs are estimated to be less attractive. The estimated

coefficients in column (1) should be interpreted in the following way - *ceteris paribus*, a job a distance 5km away has only about 8% of the flows of a job 1km away. This means that, given residence, labour markets are very local. This is in line with the evidence of Manning and Petrongolo (2017). However this does not mean that localized demand shocks will necessarily have a large impact - we return to this later. The estimation of this model is very time-consuming, involving two-way fixed effects and a very large number of observations. Although the derivation of  $D_{ab}$  strongly suggests that both fixed effects are needed, one might wonder whether simpler estimation procedures produce similar results. Columns 2-4 reports estimates of a Poisson model with different combination of origin and destination fixed effects, columns 5 and 6 the results from a log-linear regression (which will drop the zeroes) with and without fixed effects and column 7a model estimated by non-linear least squares without fixed effects.

### 3.4 Modeling $Z_{bt}$

From (4), we have that  $\Delta \log Z_{bt} = \phi \Delta w_{bt}$  where lower-case letters denote logs. To model  $Z_{bt}$  requires a model of wages. Such a model is described in Appendix D. There it is shown that the changes in wages can be approximated by:

$$\Delta \log Z_{bt} = \Delta w_{bt} = \alpha_2 [I + \alpha_1 \Omega^{nw} \Omega^{nr}] \Delta m_{bt} - \alpha_3 [I + \alpha_1 \Omega^{nw} \Omega^{nr}] \Omega^{nw} \Delta l_{bt} \quad (11)$$

where  $\Delta m_{bt}$  is an exogenous change in the demand for labor in each ward caused either by changes in preferences or by changes in productivity,  $\Delta l_{bt}$  is the change in log working-age population in each ward,  $\Omega^{nw}$  is a non-negative weight matrix whose rows all sum to one and the  $j$ th column of the  $i$ th row represents the share of workers who work in area  $i$  that reside in area  $j$ . Similarly,  $\Omega^{nr}$  is a non-negative weight matrix whose rows all sum to one and the  $j$ th column of the  $i$ th row represents the share of workers who reside in area  $i$  that work in area  $j$ .

Local wages are increasing in the own-ward demand shock,  $\Delta m$ , as one would expect. This is because more labour needs to be induced to work in the own ward to produce the extra output that is demanded. But they are also increasing in the demand shocks in surrounding areas because a positive demand shock in a neighboring area causes wages to rise there, attracting labour from both this ward and other areas that supply this ward - this reduction in labour supply causes wages to rise in this ward. The Markov matrix  $\Omega^{nw} \Omega^{nr}$  measures the interaction with other wards - it is a double-convolution because workers who

reside in ward consider working in a range of wards as given by  $\Omega^{nr}$  and firms then compete for labour with firms in those wards as given by the matrix  $\Omega^{nw}$ . The impact of changes in population,  $\Delta l$ , is rather different. First, there is no special impact of an increase in the population in the own-ward, unlike for the demand shock case - this is because workers in this ward are drawn from a range of surrounding wards. The impact on wages depends on a weighted average of local population changes with the weights being the shares of different areas in the labour supply to this ward. A higher weighted labour supply depresses wages because it leads to more output being produced which reduces prices and hence the marginal revenue product of labour. There is also a double-convolution because changes in population affect the labour supply to other wards which affects wages offered there which affects wages offered in this ward. The negative effect of population does depend on the assumption of no non-traded goods. If there are local goods, more population means more consumer demand which means higher labour demand and higher employment. The model can be expanded to consider this case but the algebra becomes much more complicated for little gain in insight.

Now consider how (11) can be estimated. We estimate  $\Omega^{nw}$  and  $\Omega^{nr}$  using averages of commuting patterns over the 4 decades for which we have data so these are not time-varying. We proxy  $\Delta m_{bt}$  by the Bartik shocks described earlier.  $\Delta l_{bt}$  is measured by the change in the log of the working-age population. However, this is likely to be endogenous (our model of migration developed below suggests as much) so we instrument this using the Card instrument described earlier. When we apply a Markov matrix to the population we also apply the Markov matrix to the instrument.

The results are presented in Table 2. All estimates compute standard errors clustering on the ward level. We report specifications in which we estimate the model in first-difference form as written in (11) but also in levels form when we include area fixed effects and cumulate the Bartik shocks. We also report OLS and IV estimates. The main issue in estimating the equation by OLS is the responsiveness of population to employment opportunities so we instrument the population measures using the Card instruments described earlier and applying the same Markov matrix to the Card variable as to the population variables themselves. Panel B of Table 2 shows that the first stages are strong with each endogenous variable most strongly affected by the Card instrument that uses the same Markov matrix. Turning to the results in Panel A, there is a robust positive correlation between the estimated value of  $\log Z_{bt}$  and the current Bartik shock. The weighted average of neighbouring Bartiks does not have the expected sign in the OLS regressions but does have the expected positive sign in the IV equations. Population has, on average, a positive association with  $\log Z_{bt}$  in the OLS

regressions but this could be explained by the fact that people migrate to areas with higher wages. When we instrument population we find an overall negative impact though the two weighted averages have opposite signs and quite large magnitudes. This could be because the two weighted averages of population have a correlation coefficient of 0.93 so its is hard to distinguish between them empirically. Overall these estimates lend support to the model.

## 4 Labor Supply

So far, we have conditioned on individuals in work - this section considers the decision to work.

### 4.1 Determination of the Employment Rate

To the commuting model, we add the option of not working that we label option 0 and denote the non-idiosyncratic part of utility as  $V_{a0}$  - to keep notation to a minimum we drop the time subscript. The commuting model assumed that the idiosyncratic errors have a very particular form leading to a multinomial logit model. But, here, the theoretical result is easier to derive if we allow for a more general form. Following McFadden (1978), we assume the distribution of the idiosyncratic terms have a joint distribution function of the form::

$$F(\epsilon_{a.}) = e^{-G(e^{-\epsilon_{a0}}, e^{-\epsilon_{a1}}, \dots, e^{-\epsilon_{aA}})} \quad (12)$$

where  $G(\cdot)$  is monotonic and homogenous of degree 1 in its arguments. We will make the following assumption:

*Assumption:* The probability of choosing employment in area  $i$  relative to the probability of choosing employment in area  $j$  does not depend on the utility available if not employed,  $V_{a0}$ .

This assumption is a form of Independence of Irrelevant Alternatives assumption though not applied to all options. This implies that  $G$  must have the form:

$$G(e^{-\epsilon_{a0}}, e^{-\epsilon_{a1}}, \dots, e^{-\epsilon_{aA}}) = G(e^{-\epsilon_{a0}}, g(e^{-\epsilon_{a1}}, \dots, e^{-\epsilon_{aA}})) \quad (13)$$

where both  $G$  and  $g$  are Hod1 in their arguments - the proof of this is in Appendix E. McFadden (1978) showed that the expected level of utility of a resident of  $a$  conditional on working (what is often called the inclusive value) is given by:

$$IV_a = \ln g \left( e^{V_{a1}}, \dots, e^{V_{aA}} \right) + \gamma \quad (14)$$

McFadden (1978) also implies the probability of choosing to work (i.e. the employment rate) will be a function of the difference between  $IV_a$  and  $V_{a0}$ .

Introducing time subscripts, Using (3) and (4) and the fact that the commuting model has a multinomial logit form, we can write the inclusive value from working as:

$$IV_{at} = \phi_{0a} - \phi \ln \widetilde{Q}_{at} + \ln \sum_b e^{V_{abt}} \quad (15)$$

Now write a linearized equation for the employment rate of residents of  $a$  at time  $t$ ,  $n_{at}$ , as:

$$n_{at} = \psi_{0a} + \psi_1 (IV_{at} - V_{a0t}) \quad (16)$$

where we have allowed for an area fixed effect  $\psi_{0a}$ . Now consider how we use this to derive an estimating equation. Taking first-differences this can be written as:

$$\Delta n_{at} = \psi_1 \Delta (IV_{at} - V_{a0t}) \quad (17)$$

The specification so far has been quite general but we make more specific assumptions in order to derive an estimable model. Here we describe those assumptions and also introduce time as an extra dimension. We model the local price index  $\widetilde{Q}_{at}$  using a first-order approximation, so it is a geometric weighted average of local housing costs, denoted by  $Q_{at}^h$ , and the prices of other goods that households consume, whose price we denote by  $Q_{at}$ , i.e. we have:

$$\ln \widetilde{Q}_{at} = s^h \ln Q_{at}^h + (1 - s^h) \ln Q_{at} \quad (18)$$

where  $s^h$  is the share of total expenditure on housing. For those who live in  $a$  but are not working we assume that:

$$V_{a0t} = \phi_{0a} + \phi \left[ \ln B_{at} - \ln \widetilde{Q}_{at} \right] \quad (19)$$

where  $\ln B_{at}$  are the welfare benefits available to the unemployed in  $a$  at time  $t$  and we assume the amenity value of living in an area is the same for the employed and unemployed. To reflect the nature of the UK welfare system we assume that  $\ln B_{at}$  consists of a time-varying component that does not vary across area and a component that is partially indexed to the local cost of housing with the parameter  $\zeta$  representing the degree of indexation- this

assumption reflects the fact that housing benefit insulates many of the the unemployed from fluctuations in local housing costs that is one of the main element of differences in prices across areas. Combining this assumption about benefits with (19) and (18) leads to

$$V_{a0t} = \phi_{0a} + \phi \left( \ln B_t + (\zeta - 1) \ln Q_{at}^h \right) \quad (20)$$

$$\Delta n_{at} = \psi_1 \Delta \ln \sum_b e^{V_{abt}} - \psi_1 \zeta \phi \Delta \ln Q_{at}^h + \text{timedummies} \quad (21)$$

Note that (21) says that the local non-housing cost price index does not affect the employment rate as it affects equally the utility when in an out of work. But local housing costs do potentially affect the employment rate to the extent that housing benefit is linked to local housing costs. We do not have local housing price indices available at the ward-level so we present a simple model of the housing market.

## 4.2 The local housing market

We assume that the change in log housing supply in area  $a$ , denoted by  $\Delta \log H_{at}^s$  is given by:

$$\Delta \log H_{at}^s = \epsilon^{hs} \ln Q_{at}^h + \text{timedummies} \quad (22)$$

On the demand side we assume that the change in housing demand,  $\Delta \log H_{at}^d$ , is influenced by the change in the size of the local population,  $\Delta \log L_{at}$ , the change in local per capita income and the change in local housing costs. The change in local income will be a function of the change in the employment rate and the change in the earnings for those who are in employment. So we have something like the following for the change in housing demand:

$$\Delta \log H_{at}^{ds} = -\epsilon^{hd} \Delta \ln Q_{at}^h + \Delta \log L_{at} + \gamma_1 \Delta n_{at} + \gamma_2 \Delta \ln \sum_b e^{V_{abt}} + \text{timedummies} \quad (23)$$

Combining (22) and (23) leads to the following equation for the change in local house prices:

$$\ln Q_{at}^h = \frac{1}{\epsilon^{hd} + \epsilon^{hs}} \left[ \Delta \log L_{at} + \gamma_1 \Delta n_{at} + \gamma_2 \Delta \ln \sum_b e^{V_{abt}} \right] + \text{timedummies} \quad (24)$$

Substituting this into (21) and re-arranging leads to the form of the equation that we estimate:

$$\Delta n_{at} = \frac{\psi_1}{\epsilon^{hd} + \epsilon^{hs} + \gamma_1 \psi_1 \zeta} \left[ (\epsilon^{hd} + \epsilon^{hs} - \gamma_2 \psi_1 \zeta) \Delta \ln \sum_b e^{V_{abt}} - \zeta \Delta \log L_{at} \right] + \text{timedummies} \quad (25)$$

i.e. the effect of allowing for the endogeneity of house prices is to include the change in population as an extra regressor and to modify the interpretation of the coefficient on the term  $\Delta \ln \sum_b e^{V_{abt}}$ .

### 4.3 Employment Rate: Results

For (25) to be estimable we need an empirical form for the inclusive value for being in employment. From (4) we have that:

$$\ln \sum_b e^{V_{abt}} = \ln \sum_b \left( e^{d_{ab} + \phi \ln W_{bt}} \right) \quad (26)$$

Using (6) and (7) this can be written as:

$$\ln \sum_b e^{V_{abt}} = \ln \sum_b \left( e^{d_{ab} + \phi \ln W_{b1} + \phi (\ln W_{bt} - \ln W_{b1})} \right) = \ln \sum_b (D_{ab} Z_{bt}) + \ln \sum_b D_{ab} + \ln \sum_b (Z_{bt}) \quad (27)$$

The first term on the right-hand side of (27) can be computed from the results of the commuting model. The second term on the right-hand side is a time-invariant origin fixed effect and the final term is an origin-invariant time fixed effect. Taking first differences of (27) and putting into (25) leads to the equation:

$$\Delta n_{at} = \frac{\psi_1}{\epsilon^{hd} + \epsilon^{hs} + \gamma_1 \psi_1 \zeta} \left[ (\epsilon^{hd} + \epsilon^{hs} - \gamma_2 \psi_1 \zeta) \Delta \ln \sum_b (D_{ab} Z_{bt}) - \zeta \Delta \log L_{at} \right] + \text{timedummies} \quad (28)$$

The results from estimating this model are reported in Table 3. All estimates compute standard errors clustering on the ward level. We report specifications in which we estimate the model in first-difference form as written in (28) but also in levels form when we include area fixed effects. We report estimates both including and excluding the log of population. We sometimes estimate this in first-difference form and sometimes in level form when we also include an origin fixed effect. We also report OLS and IV estimates. There are a num-



ber of issues in estimating this equation by OLS e.g. the responsiveness of population to employment opportunities and the fact that the inclusive value is a generated variable with considerable measurement error. As instrument we use the Bartik and Card instruments described earlier. Panel B of Table 3 shows that the first stages are strong with both instruments significantly related to both endogenous variables. Turning to the results in Panel A, there is a robust positive correlation between the inclusive value and the employment rate. This is stronger in the IV estimates than the OLS estimates as would be expected given that the inclusive value is a generated variable with some measurement error. And the effect is stronger in the FE than the FD specification. Local population generally has a negative effect on the local employment rate as predicted by the model if housing benefits are partially linked to local housing costs. Overall these estimates lend support to the model.

#### 4.4 The value of commuting

As we have previously emphasized, the ability to change work location in response to demand shocks acts to raise worker utility and provides a method of insurance against such shocks. One way of computing the value of commuting is to compare the inclusive value from the optimal choice of commuting with the expected utility from a forced pattern of commuting. In forcing a non-optimal pattern of commuting on workers one can do this either at random or in an efficient way i.e. choosing the assignment to maximize expected utility subject to constraints on the overall commuting pattern. Define  $EU(V_{ab}, p_{ab})$  to be the maximized expected utility of workers living in  $a$  if the systematic pay-offs from working in different areas are  $V_{ab}$  but  $p_{ab}$  is the fraction of individuals living in  $a$  who are forced to work in  $b$ . With the multinomial logit structure the expected utility from this can be written as:

$$EU(V_{ab}, p_{ab}) = \sum_b p_{ab} V_{ab} - \sum_b p_{ab} \ln p_{ab} \quad (29)$$

The first term in this expression is the expected utility if workers are assigned at random to different work locations. The second term can then be interpreted as the increase in expected utility from the optimal as opposed to random assignment - it takes the form of the standard entropy measure. If one maximizes (29) with respect to  $p_{ab}$  then one obtains the logit choice probabilities and the maximized function is the inclusive value as defined earlier.

To assess the value of commuting we ask how much expected utility would change if we constrain commuting patterns to be sub-optimal and then use the estimates of the previous prediction to measure how much this would affect the employment rate. How big the effect

will be depends on the nature of the constraints put on commuting. If, for example, one constrained individuals to live and work in the same area, that would typically lead to a huge cost in expected utility not least because the entropy measure is typically very large. But it might be dangerous to extrapolate to commuting patterns so far what is observed in the sample. More realistic perhaps is to consider what would happen if  $V_{ab}$  changes from one decade to the next, but that commuting patterns remain unchanged.

Table 4 contains the results of this exercise. Each number in the table represents the loss in expected utility from imposing the commuting pattern of one census year on the returns to working in different areas of another. The diagonals are zero by construction. Typical implied values are around -0.10 which, using the estimated response of the employment rate with respect to the inclusive value of 0.2 from Table 3, implies that the ability for commuting to respond causes the employment rate to be about 2 percentage points higher than it would otherwise be. This is reasonably large relative to the observed variation in employment rates over time, but much smaller than the cross-sectional variation across wards.

Table 4 reports the average benefit from being able to change commuting patterns. Of course, it may also vary systematically across areas. But the return to commuting does not have a strong relationship with the lagged employment rate, our main measure of economic opportunity.

These estimates are probably best interpreted as what would happen if an individual, chosen at random, was constrained from changing their commuting patterns as they assume that the returns to working in different areas are held fixed. If everyone was prevented from altering commuting patterns, it is quite likely that the return from working in different areas would change, i.e. there would be general equilibrium effects.

This estimated value of being able to alter commutes depends on the nature of the shocks hitting different areas. For example, if all areas to which one might commute have the same change in  $Z_b$  (including no change) then the value from being able to change commutes will be zero. And there are strong economic reasons why the shocks to neighbouring areas will be correlated. First, agglomeration means they are likely to specialize in similar sectors so sectoral shocks will affect them in similar ways. And because they compete in the same labour market, the offered wages are likely to co-vary positively and this may be the main attractor of working in different areas for workers. Appendix [TO DO] works through a simple model of local shocks to illustrate this point.

One can also use the model to predict the change that comes from a change in the economic opportunities of working in an area or a change in the cost of commuting, e.g.

through changes in infrastructure.

## 4.5 From employment rate to welfare: the 'sufficient statistic' result

**Proposition 1.** *If the IIA-type assumption at the beginning of Section 4.1 is satisfied, then the expected utility of living in area  $a$  can be written as:*

$$\bar{U}_{at} = V_{a0t} + \Psi(n_{at}) \quad (30)$$

for some function  $\Psi()$  of the employment rate  $n_{at}$ .

*Proof.* See Appendix E. □

This result can be thought of as an application of the 'conditional choice probability' result of Hotz and Miller (1993). The probability of choosing the employment option is a function of the difference in expected utility between the two options. The gain from choosing the employment option is also a function of the difference in expected utility between the two options. This implies one can write expected utility from living in an area as a function the utility obtainable when not employed and the difference in expected utility when employed and not employed. And the latter is a function of choosing the employment option which is just the employment rate. The conditions under which the result are true may seem very abstract but are satisfied by most of the commonly used functional forms in the discrete choice literature. For example, the assumption is satisfied in a simple multinomial logit or a more complicated model based on the Frechet distribution. or a nested logit specification in which one of the nests is non-employment. So most existing papers do make functional form assumptions that satisfy the assumptions of the proposition. But it is worth considering the situations in which the condition would not be satisfied. Suppose there is a sector, the probability of working in which is fixed so any variation in employment rates comes from changes in the shares of other sectors. A change in the wage of the sector with the fixed share will change the inclusive value because a job in that sector is now more valuable but will not change the overall employment rate. In this case the employment rate is not a sufficient statistic for the inclusive value.

The usefulness of this result is that it allows us to reduce measure the expected utility from residing in an area by simply two dimensions - the utility from being unemployed there and the employment rate in the area. The overall employment rate summarizes all the available

employment options in what may be a very large number of potential commuting areas. Intuitively if working in an area becomes more attractive then the probability of working overall rises and the inclusive value. This result can also be thought of as an extension of the 'sufficient statistic' result of Amior and Manning (2015) that the employment rate can be used as a sufficient statistic for economic opportunity in an area if the labour supply curve (or 'wage curve' if one prefers a non-competitive model of the labour market) is not completely inelastic, an assumption we argued to be plausible (e.g. see the wage curve literature of Blanchflower and Oswald, 1994). The intuition for the result is simple - as one moves up the labour supply curve worker welfare is increasing and one can measure how high welfare is in wither the real consumer wage or employment dimension with the latter often having the practical advantage of being easily computed. That result was derived in a model where it was assumed that everyone lives and works in the same area. The application in Amior and Manning (2015) was to US commuting zones which are intended to be closed labour markets. However even in that context there is some commuting across commuting zone boundaries and it is an even less reasonable assumption if one considers smaller geographic areas, as we would like to do in this paper.

## 5 Population Change

There is considerable existing literature about how internal migration responds to economic factors in the UK, though most of it is now quite old, perhaps surprising considering the improved data now available and the continued prominence of regional inequalities. This literature considers how migration between regions responds to economic factors such as regional differences in unemployment, vacancies, wages and the cost of living (mostly measured through house prices). Different studies have come to different conclusions about the significance of these different factors though a general theme is that migration does respond to differences in economic opportunities but at a pace that means adjustment to equilibrium will be very slow (see McCormick, 1997, for a brief survey of the findings to that date). The literature also considers how the pace of adjustment varies by working skill or housing tenure, the latter being thought important in the UK because those in social housing are often thought to be very immobile. In terms of data, some studies use annual regional panels using data on net migration drawn from the NHS registration data (using changes in doctor's address) - see, for example, Pissarides and McMaster (1990), Gordon and Molho (1998) (this uses 5 yearly intervals including Census data), Jackman and Savouri (1992),

Hatton and Tani (2005) (whose main focus is on the impact of immigration on internal migration). One limitation of this study is that the data may not be very accurate and is only available for broad statistical regions that do not correspond to any definition of a labour market. Other studies study residential mobility at the individual level, the earlier studies (Hughes and McCormick, 1981, 1987*a*, 1994; Pissarides and Wadsworth, 1989) using isolated cross-sections but the most recent studies (Henley, 1998; Böheim and Taylor, 2002; Gregg, Machin and Manning, 2004; Andrews, Clark and Whittaker, 2011; Rabe and Taylor, 2012) using longer panels like the BHPS.

## 5.1 Estimates of population response

Next, we combine our model for commuting and the employment rate (which has assumed a fixed population) with a simple specification for the migration response, where the local population responds to the gap between local utility  $u_a$  and aggregate utility  $u$  as in Amior and Manning (2015). This suggests the following equation for the change in the log population in area  $a$ ,  $l_a(t)$ :

$$\begin{aligned} \frac{\partial l_a(t)}{\partial t} &= g [\bar{U}_a(t) - \bar{U}(t)] \\ &= z_a(t) + \gamma_0 [n_a(t) - l_a(t)] - \gamma_1 l_a(t) \end{aligned} \quad (31)$$

where  $t$  denotes time and  $z_a(t)$  an amenity. The second line is a linearization of the first and uses the result that the expected utility from living in an area can be written as a function of the employment rate and the utility from non-employment which is a function of house prices which vary with the employment rate and the level of population. In a steady-state, the “spatial arbitrage” condition of the Rosen-Roback model guarantees that utility is equal in all areas. Amior and Manning (2015) contains an extensive discussion of how this equation should be interpreted and we do not reproduce it here. Amior and Manning (2015) also show how this equation in continuous time can be used to derive the following estimation equation in discrete time:

$$\Delta l_{at} = \beta_0 + \beta_1 \Delta n_{at} + \beta_2 (n_{at-1} - l_{at-1}) + \beta_3 l_{at-1} + \beta_4 \Delta \tilde{z}_{at} + \beta_5 \tilde{z}_{at-1} + d_t + \varepsilon_{rt} \quad (32)$$

(32) has the form of an ECM, with the change in log population  $\Delta l_{at}$  responding to the change in log employment  $\Delta n_{at}$  (i.e. local shocks), the lagged log employment rate

$n_{at-1} - l_{at-1}$  (which measures initial disequilibrium) and the lagged level of population. If population adjusts instantaneously to employment shocks,  $\beta_1$  would take a value of 1. And controlling for employment changes,  $\beta_2$  would equal 1 if local population adjustment over one decade is sufficient to compensate for initial deviations in the local employment rate from equilibrium. Practically though, if  $\beta_1 = 1$ , it would not be possible to identify  $\beta_2$  since there would be no observable deviations from equilibrium. We now turn to the data that we use to estimate the model presented above.

The ECM model offers an intuitive way to assess the speed of population adjustment, as a “race” against employment growth. This builds on the dynamic analysis of Blanchard and Katz (1992), by integrating contemporaneous shocks which are essential for the longer data frequencies which interest us. Suitably instrumented contemporaneous shocks are a hallmark of the modern urban literature (see, for example Notowidigdo, 2011; Autor, Dorn and Hanson, 2013; Beaudry, Green and Sand, 2014*b*), but these studies do not account for the dynamics (assuming instead that each census observations represents local equilibrium).

In estimating (32), one needs to recognize that employment growth and the lagged employment rate is endogenous but our model of the labour market suggests instrumenting it using demand shocks, specifically industry shift-share instruments following the approach of Bartik (1991).

Estimates for the US and UK Issues – frequency of data - Timing on employment change within the decade Should we look at different sorts of people: Natives/migrants, high low skill/young/old. Whether symmetry on falls/rises Commuting Is an Important Response Modifying the model to allow for commuting Once one recognizes that CZ and TTWAs are not closed labour markets it makes more sense to do things at a very local level.

We start by estimating the model (32) for population change at ward level - these results are reported in Table 5. We report a number of specifications. Panel A reports the estimates of the population growth equations while Panel B reports the first-stages. We report OLS and IV estimates for three different specifications - one in which we have model the area amenities by the amenity variables, interacted with time, a second in which we include ward fixed effects and a third in which we estimate in first-differences. The estimates imply that population responds strongly to local shocks, but not sufficiently to undo the effects of a shock within a decade.

The parameter estimates are quite similar to those we found in the US in Amior and Manning (2015). But the two sets of estimates are at very different spatial scales - wards for the UK and commuting zones for the US. Give this, it is interesting to estimate the UK

model using geographical areas that are as close as possible to those used in the US - these are TTWAs. This is also useful as providing some specification test of our underlying model. Our model can be applied at any spatial scale so the parameter estimates should be similar whether we estimate at ward or TTWA level. Appendix B a discussion of the similarities of US CZs and UK TTWAs. Table 6 reports estimates of the population growth equation at TTWA level.

This similarity in the adjustment process in the UK and the US might be thought surprising as there seems to be a broad consensus that migratory responses are larger in the US than the UK or Continental Europe, although there is some disagreement over the magnitudes. The bulk of this literature has deployed the empirical model of Blanchard and Katz (1992). This is a vector-autoregression (VAR) using annual data on local employment growth, employment rate and participation rate, usually with two lags and controlling for local-specific trends. They concluded that the migration mechanism was sufficiently strong to dissipate completely the impact of local demand shocks on the employment rate within 5 years. Decressin and Fatás (1995), Jimeno and Bentolila (1998), Dao, Furceri and Loungani (2014) and Beyer and Smets (2015) find a slower adjustment in Europe (and the UK in particular) of the order of 10 years. Obstfeld and Peri (1998) estimate a similar model, but without controlling for region-specific trends: they find similar results for the US, though demand shocks do persist beyond ten years in Europe - especially in Germany and Italy, but also in the UK to some extent. A number of explanations have been suggested for these mobility differences. Bertola and Ichino (1995) argues that European labour market institutions play an important role, by compressing geographical wage differentials and reducing turnover. Decressin and Fatás (1995) argues that much of the impact of local job loss in Europe is manifested in early retirement and disability benefit claims. Rupert and Wasmer (2012) emphasize the role of housing market frictions, and Hughes and McCormick (1987*b*) point to the decline of the British private rental sector in particular.<sup>11</sup> Certainly, the US also has a high rate of homeownership; but Obstfeld and Peri (1998) argue that American mortgage markets are more efficient and transaction costs are lower.

However, there is accumulating evidence that there are very persistent spatial differences in joblessness even in the US (see Overman and Puga, 2002; OECD, 2005; Rappaport, 2012; Kline and Moretti, 2013) and that local demand shocks have very long-lasting impacts on economic opportunity - for example, Autor, Dorn and Hanson (2013); Yagan (2014). Amior

---

<sup>11</sup>Related to this, Oswald (1996) argues that high rates of homeownership reduce geographical mobility and thereby sustain high levels of unemployment.

and Manning (2015) document that joblessness is very persistent over decades across US commuting zones and offer the interpretation that the demand shocks themselves are very persistent so that economic opportunity in an area is the result of a race between the demand shocks and the migration response. However, given the short lag structure, Obstfeld and Peri (1998) note the Blanchard-Katz model may not be suitable for making long run predictions<sup>12</sup>; and it is the long run which concerns us most in this study. Rather than using annual data, we study decadal changes between census observations.

## 6 Conclusion

This paper is primarily about the role that commuting and migration play in the UK in equalizing economic opportunity across areas. We find that both commuting and migration respond to economic shocks, but that these responses are insufficient to equalize opportunity in the face of demand shocks that are very persistent. The paper also makes a number of methodological contributions

1. A model of the commuting decision i.e. a model of the decision of residents of one area about the area where they work. This utility from living in one area and working in another is written as a function of the wage offered and cost of the commute.
2. A model for the determination of wages offered by employers in an area.
3. A generalization of Amior and Manning's (2015) result that the employment rate (perhaps composition-adjusted) in an area can be used as an easily computed sufficient statistic for economic opportunity to the case where workers are not required to work in the area where they live.
4. A simple model of migration between areas

and this framework may be of use in other applications.

---

<sup>12</sup>An impulse response function may over-state the long-run pace of adjustment if, for example, it is the most mobile workers who move first.



# Appendices

## A Further details on data

### A.1 Population census data

We extract the small area aggregates of the English and Welsh 2011 census data from Nomis (<https://www.nomisweb.co.uk/census/2011>) and Scottish 2011 data from National Records for Scotland (<http://www.scotlandscensus.gov.uk>). Scottish 2001 data was provided on DVD by the General Register Office for Scotland (<http://www.scrol.gov.uk>). We downloaded all other UK census data from UK Data Service Census Support (<http://casweb.mimas.ac.uk>).

All the small area aggregates are based on 100% samples, though some noise has been artificially injected into some small cells to preserve anonymity. As mentioned above, TTWA-level outcomes were imputed using geographical look-up tables based on address counts from the National Postcode Directory and area shapefiles. To ensure consistency, we have used the geographically finest census data available to construct our TTWA aggregates. This is at the level of enumeration districts or output areas, of which there are between 120,000 and 230,000 in the country (depending on census year). Given this level of geographical detail, we are confident in the comparability of our TTWA series over time.

### A.2 Industry data

NOMIS provides annual local employment statistics (by workplace geography) going back to 1971.<sup>13</sup> These were based on the Census of Employment between 1971 and 1991 (3-digit SIC 1968), the Annual Employment Survey (AES) between 1991 and 1998 (4-digit SIC 1992), the Annual Business Inquiry (ABI) between 1998 and 2008 (4-digit SIC 1992 and 2003), and the Business Register Employment Survey (BRES) since then (5-digit SIC 2007).<sup>14</sup> All of these provide counts of paid employees from administrative data; though the AES, ABI and BRES are based on surveys. Data on farm employment has not been available at TTWA or ward level since 1981, so we impute these missing cells using supplementary data from the population census. Unfortunately, administrative employment at the local area prior to 1971 have not been digitized. So we impute local industrial composition in 1961 by applying

---

<sup>13</sup><https://www.nomisweb.co.uk/>

<sup>14</sup><http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/business-register-and-employment-survey-bres-/history-and-background/index.html>

national employment growth rates by industry (compiled by Department of Employment, 1975) to the 1971 local shares from the Census of Employment.

We construct industry look-up tables with proportional allocations to convert all the data above to a 3-digit SIC 1992 classification with 212 industries. We estimate these allocations using longitudinal micro-data from the Annual Survey of Hours and Earnings (formerly the New Earnings Survey); this is administrative data based on a 1% sample of employees. Specifically, in those years where there was a change of classification, we estimate transitions between industry codes - for those workers who remained in the same job.

Geographical changes are an important concern for earlier cross-sections of the local industry data. The 2011 data are available in very fine geographical detail (by output area, of which there are 230,000 in 2011), so a precise approximation of the boundaries of the much larger TTWAs is feasible. In 1991 and 2001, the finest geographical classification is the 10,764 wards of the 1991 census; this still allows for a reasonable approximation of the 232 TTWAs in our data. The match in 1971 and 1981 is problematic: in this case, we use employment estimates for the TTWAs of the 1981 census; there are only 309 of these in our data. We believe a simple match between the TTWAs of 1981 and 2001 would be problematic. Instead, we also exploit the ward-level data from the 1991 cross-section. This procedure consists of three steps: (1) we estimate the growth of each industry within each 1981-definition TTWAs between 1971 and 1991; (2) we impute local industry composition for 1991-definition wards by applying the local growth rates from step 1 to the 1991 data; and (3) we convert the 1971 and 1981 data (now in terms of 1991-definition wards) to 2001-definition TTWAs using our mapping based on address counts from the National Postcode Directory.

## B Comparison of US CZs and UK TTWAs

In Table 7, for the sake of comparison, we report percentiles of some key statistics on the distribution of the 232 British TTWAs (based on the census of 2001) and the 722 American CZs (based on the census of 2000). The populations of TTWAs and CZs are similar, with a median of 123,000 in the UK and 107,000 in the US. But, American CZs are much larger in terms of land area, with a median of 8km<sup>2</sup> compared to just 0.7km<sup>2</sup> in the UK. This reflects the fact that the UK is much more densely populated than the US. Perhaps a more useful measure of population density is the “weighted” density, which is intended to measure the average density *experienced by residents*. For a given TTWA or CZ  $r$ , the weighted density

$WD_r$  is the population-weighted average of the densities of the composite neighborhoods  $n \in r$ :

$$WD_r = \sum_{n \in r} \left( \frac{P_n}{\sum_{n \in r} P_n} \right) \frac{P_n}{A_n} \quad (33)$$

where  $P_n$  is the population of neighbourhood  $n$  and  $A_n$  is its area. Identifying “neighborhoods” in equation (33) with wards for the UK (average population of 5700 in 2001) and census tracts for the US <sup>15</sup> (average population is 4300 in 2000). While the TTWAs have much larger weighted densities, notice that the proportional gap narrows considerably as one moves up the distribution (i.e. when larger cities are compared). Indeed, the weighted density across the entire US is 2,170 residents/km<sup>2</sup>, which is not much smaller than the British value of 2,806 residents/km<sup>2</sup>.<sup>16</sup> Notice that these numbers are simply population-weighted averages of the TTWA or CZ-specific weighted densities reported in Table 7. So, applying local population weights to the analysis below may help create a more comparable sample of commuting areas.

The final two rows of Table 7 relate to commuting patterns, based on census flow data. The first reports the share of employed individuals residing in a given TTWA or CZ who also work in that area. This tends to be somewhat smaller for TTWAs than CZs, with a median of 74% compared to 90%. The same is true for the share of individuals working in a given locality who also live in that area (final column): the median is 80% for TTWAs, compared to 91% for CZs. The of course reflects to some extent differences in the algorithms used to define TTWAs and CZs. But also, the relative compactness of the UK is likely to play a part: the distance between the largest cities is much smaller than in the US, and this must encourage more commuting. This makes the use of TTWAs as self-contained labour markets more problematic in the UK than CZs are in the US and is one of the motivations for developing a framework that does not rely on self-contained labor markets.

---

<sup>15</sup>The Census Bureau has recently been compiling weighted density for Metropolitan Statistical Areas, and they also choose tracts as their “neighborhood” identifier.

<sup>16</sup>These densities for the entire US and entire UK again identify neighborhoods with census tracts and wards respectively.

## C Details of estimation procedure for commuting model

Define a multiplier  $\mu_a^d$  for the constraint  $\sum_b D_{ab} = 1$ . Then the first-order condition for the maximization of (10) with respect to  $D_{ab}$  can be written as:

$$\frac{1}{D_{ab}} \sum_t C_{abt} - \sum_{i,t} \frac{Z_{bt}}{\sum_j D_{aj} Z_{ij}} C_{ait} - \mu_a^d = 0 \quad (34)$$

Multiplying every term by  $D_{ab}$ , re-arranging and summing over  $b$  leads to:

$$\mu_a^d \sum_b D_{ab} = \sum_{b,t} C_{abt} - \sum_{i,b,t} \frac{D_{ab} Z_{bt}}{\sum_j D_{aj} Z_{ij}} C_{ait} = \sum_{b,t} C_{abt} - \sum_{i,t} \frac{\sum_b D_{ab} Z_{bt}}{\sum_j D_{aj} Z_{ij}} C_{ait} = \sum_{b,t} C_{abt} - \sum_{i,t} C_{ait} = 0 \quad (35)$$

which implies that  $\mu_a^d = 0$ . Using this in (34) and re-arranging leads to the following expression for the ML estimate of  $D_{ab}$ :

$$D_{ab} = \frac{\sum_t C_{abt}}{\sum_{i,t} \frac{Z_{bt}}{\sum_j D_{aj} Z_{ij}} C_{ait}} \quad (36)$$

Now define a multiplier  $\mu_t^z$  for the constraint  $\sum_b Z_{bt} = 1$ . Then the first-order condition for the maximization of (10) with respect to  $Z_{bt}$  can be written as:

$$\frac{1}{Z_{bt}} \sum_a C_{abt} - \sum_{a,i} \frac{D_{ab}}{\sum_j D_{aj} Z_{ij}} C_{ait} - \mu_t^z = 0 \quad (37)$$

Multiplying every term by  $Z_{bt}$ , re-arranging and summing over  $b$  leads to:

$$\mu_t^z \sum_b Z_{bt} = \sum_{a,b} C_{abt} - \sum_{a,i,b} \frac{D_{ab} Z_{bt}}{\sum_j D_{aj} Z_{ij}} C_{ait} = \sum_{a,b} C_{abt} - \sum_{a,i} \frac{\sum_b D_{ab} Z_{bt}}{\sum_j D_{aj} Z_{ij}} C_{ait} = \sum_{a,b} C_{abt} - \sum_{a,i} C_{ait} = 0 \quad (38)$$

which implies that  $\mu_t^z = 0$ . Using this in (37) and re-arranging leads to the following expression for the ML estimate of  $D_{ab}$ :

$$Z_{bt} = \frac{\sum_a C_{abt}}{\sum_{a,i} \frac{D_{ab}}{\sum_j D_{aj} Z_{ij}} C_{ait}} \quad (39)$$

The equations (36) and (39) can be thought of as updates on the parameter estimates given an initial set. If this process converges (and it does) the limit must be the ML estimates.

## D Wage Determination

### D.1 Labour Supply

Given the assumptions on commuting in (5), the employment rate (21), and (4) the number of residents of  $a$  working in  $b$ ,  $N_{ab}^s$  is given by:

$$N_{ab}^s = e^{\eta_{0a}} \frac{D_{ab} W_b^\phi}{\left[ \sum_i D_{ai} W_i^\phi \right]^{1-\psi_1}} \left( Q_a^h \right)^{-\phi \zeta \psi_1} L_a \quad (40)$$

where  $L_a$  is the resident population in  $a$ . Hence total labour supply of workers to area  $b$  will be given by:

$$N_b^{sw} = \sum_i N_{ib}^s \quad (41)$$

### D.2 Product Demands

Although our ultimate aim is to derive wages through the interaction between the demand for and the supply of labour, we also need to specify product demands. Assume that demands are homothetic so that one does not have to worry about the distribution of income within areas. Assume that the non-housing part of the price index for the residents of area  $a$ ,  $Q_a$ , (from (18)) is given by a CES function of a price index for domestic goods  $Q_a^d$  foreign goods,  $Q^f$  (which will be treated as exogenous) according to:

$$Q_a = \left[ Q_a^{d\gamma} + \gamma_f Q^f \gamma \right]^{\frac{1}{\gamma}} \quad (42)$$

In turn, the price index for local goods is assumed to be given by another CES index:

$$Q_a^d = \left[ \sum_i M_i \Gamma_{ai} P_i^{1-\theta} \right]^{\frac{1}{1-\theta}} \quad (43)$$

where  $P_i$  is the price of goods produced in area  $i$ ,  $\Gamma_{ai}$  represents the demand for the residents of area  $a$  for goods produced in area  $i$  (the specification allows for the possibility that there is some stronger demand for local non-traded goods) and  $M_i$  is a demand shifter assumed to affect consumers in all areas equally. Changes in  $B_i$  will be one possible source of shocks to the economy.

Using these price indices, the demand for goods produced in  $b$  by residents of  $a$ ,  $X_{ab}^d$ , can

be written as:

$$X_{ab}^d = M_b \Gamma_{ab} \left( \frac{P_b}{Q_a^d} \right)^{-\theta} \left( \frac{Q_a^d}{Q_a} \right)^\gamma \left( \frac{Q_a}{Q_a} \right)^{-\epsilon_{hd}} Y_a \quad (44)$$

where  $Y_a$  is the total income of the residents of  $a$ , which can be written as:

$$Y_a = B (Q_a^h)^\zeta L_a + \sum_i N_{ai}^s [W_i - B (Q_a^h)^\zeta] \quad (45)$$

the specification of which embodies the assumption that the real income of the non-employed, is assumed to be partial indexed to local house prices through housing benefits - see (20). Total demand for goods produced in area  $b$  is then given by:

$$X_b^d = \sum_i X_{ib}^d + X_b^f \quad (46)$$

where  $X_b^f$  is demand for goods from foreign consumers who we assume to have the same price elasticity as domestic consumers.

### D.3 Housing Prices

From the demand and supply of housing we have that local house prices are given by:

$$\ln Q_a^h = \frac{\ln Y_a}{\epsilon_{hd} + \epsilon_{hs}} \quad (47)$$

### D.4 The Production Function

Assume there is constant returns to scale in production so that output is given by  $\tilde{A}_b N_b$ . However we allow for the possibility that there is some agglomeration externality exogenous to the individual firm so that  $\tilde{A}_b = A_b N_b^\varphi$ . If we assume that prices are equal to marginal costs (a mark-up would make no difference). we have that :

$$W_b = A_b N_b^\varphi P_b \quad (48)$$

### D.5 Equilibrium Wages

Putting together these equations we can find the equilibrium. Wages in an area will be a function of the exogenous variables, the demand shocks  $M_b$ , productivity shocks,  $A_b$ , and the distribution of population. In order to consider the response of log wages  $dw$  to

log product demand shocks  $dm$  and log population shocks  $dl$ , we will consider a special case where the only good for which there are local preferences is housing i.e. we assume that each row in  $\Gamma_{ai}$  in (43) is identical. In this case local income is not relevant for demand for locally produced goods. The endogenous variables are the changes in wages  $dw$ , employment,  $dn^s$ , incomes,  $dy$ , output,  $dx$ , prices of goods,  $dp$ , and house prices,  $dq^h$  (all in logs). In the equations that follow we omit constants that are common to all areas to keep notation to a minimum.

From (48) we have that:

$$dp = -da + dw - \varphi dn^s \quad (49)$$

and, from the production function we have:

$$dx = da + (1 + \varphi) dn^s \quad (50)$$

From consumer demand we have that:

$$dx = db - \theta dp \quad (51)$$

From (40) and (41) we have that:

$$dn^s = \phi [I - (1 - \eta) \Omega^{nw} \Omega^{nr}] dw + \Omega^{nw} [dl - \phi \zeta \psi_1 dq^h] \quad (52)$$

where  $\Omega^{nw}$  is a non-negative weight matrix whose rows all sum to one and the  $j$ th column of the  $i$ th row represents the share of workers who work in area  $i$  that reside in area  $j$ . Similarly,  $\Omega^{nr}$  is a non-negative weight matrix whose rows all sum to one and the  $j$ th column of the  $i$ th row represents the share of workers who reside in area  $i$  that work in area  $j$ .

Next consider the change in house prices which is given by:

$$dq^h = \frac{1}{\epsilon^h + \epsilon^s} dy \quad (53)$$

And, finally consider the change in local incomes: From (45) we have that:

$$dy = dl + \beta \Omega^{yr} dw + (1 - \beta) \zeta dq^h \quad (54)$$

where  $\Omega^{yr}$  is a weight matrix whose rows all sum to one and the  $j$ th column of the  $i$ th row represents the share of total labour income for residents of area  $i$  that comes from area  $j$ , and  $\beta$  is the share of earned income in area income.

Using (49)-(51) we can derive the following expression for the relationship between wages and employment from the demand side:

$$\theta dw = (\theta - 1) da + db + [\theta\varphi - (1 + \varphi)] dn^s \quad (55)$$

Combining (53) and (54) we can write house prices as:

$$\left[ (\epsilon^h + \epsilon^s) - (1 - \beta)\zeta \right] dq^h = dl + \beta\Omega^{yr} dw \quad (56)$$

Substituting (56) into (52) leads to the following expression for the relationship between wages and employment from the supply side:

$$\begin{aligned} h_r^d = & \phi [I - (1 - \eta) \Omega^{nw} \Omega^{nr}] dw - \frac{\beta\phi\zeta\psi_1}{[(\epsilon^h + \epsilon^s) - (1 - \beta)\zeta]} \Omega^{nw} \Omega^{yr} dw \\ & + \left[ 1 - \frac{\phi\zeta\psi_1}{[(\epsilon^h + \epsilon^s) - (1 - \beta)\zeta]} \right] \Omega^{nw} dl \end{aligned} \quad (57)$$

Using (55) and (57) to eliminate employment we end up with the following expression for wages:

$$\begin{aligned} \theta dw - (\theta - 1) da - db = & [\theta\varphi - (1 + \varphi)] \phi [I - (1 - \eta) \Omega^{nw} \Omega^{nr}] dw \\ & - \frac{\beta\phi\zeta\psi_1 [\theta\varphi - (1 + \varphi)]}{[(\epsilon^h + \epsilon^s) - (1 - \beta)\zeta]} \Omega^{nw} \Omega^{yr} dw \\ & + [\theta\varphi - (1 + \varphi)] \left[ 1 - \frac{\phi\zeta\psi_1}{[(\epsilon^h + \epsilon^s) - (1 - \beta)\zeta]} \right] \Omega^{nw} dl \end{aligned} \quad (58)$$

Simplifying (58) by assuming that  $\Omega^{yr} = \Omega^{nr}$  leads to an expression of the form:

$$dw = \alpha_1 \Omega^{nw} \Omega^{nr} dw + \alpha_2 [db + (1 - \theta) da] - \alpha_3 \Omega^{nw} dl \quad (59)$$

where  $(\alpha_1, \alpha_2, \alpha_3)$  are functions of the underlying parameters. Can write the term for wages as a function of the own shock and the weighted average. and the own labour supply and labour supply in surrounding areas.

There are a number of points that can be made from (59). First, wages in a ward are increasing in the wages offered in surrounding wards. This is because firms compete for labour with surrounding wards and labour supply to this ward decreases if wages in



neighbouring wards increase. The matrix  $\Omega^{nw}\Omega^{nr}$  measures the interaction with other wards - it is a double-convolution because workers who reside in ward consider working in a range of wards as given by  $\Omega^{nr}$  and firms then compete for labour with firms in those wards as given by the matrix  $\Omega^{nw}$ . Secondly local wages are increasing in the own-ward demand shock,  $db$ , as one would expect. This is because more labour needs to be induced to work in the own ward to produce the extra output that is demanded. How much wages need to change depends on the wage elasticity of the labour supply curve to a ward, the price elasticity of demand, and the extent of agglomeration externalities. Local productivity shocks,  $da$ , causes wages to fall or rise according to whether the price elasticity of demand is greater than or less than one. The impact of changes in population,  $dl$ , is rather different. First, there is no special impact of an increase in the population in the own-ward, unlike for the demand shock case. The impact on wages depends on a weighted average of local population changes with the weights being the shares of different areas in the labour supply to this ward. A higher weighted labour supply depresses wages because it leads to more output being produced which reduces prices and hence the marginal revenue product of labour. However, the strength of this result does depend on the assumption of no non-traded goods. If there are local goods, more population means more consumer demand which means higher labour demand and higher employment. The model can be expanded to consider this case but the algebra becomes much more complicated for little gain in insight.

(59) can be re-arranged to give the following 'reduced-form' expression for the change in wages:

$$dw = \alpha_2 [I - \alpha_1 \Omega^{nw} \Omega^{nr}]^{-1} [db + (1 - \theta) da] - \alpha_3 [I - \alpha_1 \Omega^{nw} \Omega^{nr}]^{-1} \Omega^{nw} dl \quad (60)$$

In deriving the estimation equation we take a first-order approximation to this:

$$dw = \alpha_2 [I + \alpha_1 \Omega^{nw} \Omega^{nr}] [db + (1 - \theta) da] - \alpha_3 [I + \alpha_1 \Omega^{nw} \Omega^{nr}] \Omega^{nw} dl \quad (61)$$

## E Proof of Propositions

### E.1 Proof of Proposition 1

Denote by  $G_b$  the derivative  $G$ , with respect to its  $b$ th argument. Theorem 1 in McFadden (1978) shows that the probability of a resident of  $a$  choosing option  $b$ ,  $p_{ab}$ , can be written as:

$$p_{ab} = \frac{e^{V_{ab}} G_b(e^{V_{a0}}, e^{V_{a1}}, \dots, e^{V_{aA}})}{G(e^{V_{a0}}, e^{V_{a1}}, \dots, e^{V_{aA}})} = \frac{e^{V_{ab}-V_{a0}} G_b(1, e^{V_{a1}-V_{a0}}, \dots, e^{V_{aA}-V_{a0}})}{G(1, e^{V_{a1}-V_{a0}}, \dots, e^{V_{aA}-V_{a0}})} \quad (62)$$

where the second inequality follows from the assumption that  $G$  is Hod1. McFadden (1978) also shows that the expected level of utility of a resident of  $a$  (what is often called the inclusive value) is given by:

$$\bar{U}_a = \ln G(e^{V_{a0}}, e^{V_{a1}}, \dots, e^{V_{aA}}) + \gamma = V_{a0} + \ln G(1, e^{V_{a1}-V_{a0}}, \dots, e^{V_{aA}-V_{a0}}) + \gamma \quad (63)$$

where  $\gamma$  is Euler's constant and the second equality follows from the fact that  $G$  is Hod1. Using (62) the assumption implies that  $G_j/G_i$  does not depend on  $V_0$ . This implies that  $G$  must have the form:

$$G(e^{V_0}, e^{V_1}, \dots, e^{V_A}) = G(e^{V_0}, g(e^{V_1}, \dots, e^{V_A})) \quad (64)$$

for some function  $g$  that is Hod1 in its arguments. That this restriction satisfies the assumption can be seen from the fact that it implies that:

$$p_j = \frac{e^{V_j} G_g(e^{V_0}, g) g_j(e^{V_1}, \dots, e^{V_A})}{G(e^{V_0}, e^{V_1}, \dots, e^{V_A})} \quad (65)$$

so that the relative employment probabilities do not depend on  $V_0$ . Using (64) the expected maximum level of utility can be written as:

$$\bar{U}_i = V_0 + \ln G(1, g) + \gamma \quad (66)$$

i.e. is a function of  $V_0$  and  $g$  alone. Using (62) and the restriction in (64) the employment rate  $n_i = 1 - p_0$  can be written as:

$$1 - n_i = \frac{e^{V_0} G_0(e^{V_0}, g(e^{V_1}, \dots, e^{V_A}))}{G(e^{V_0}, g(e^{V_1}, \dots, e^{V_A}))} = \frac{G_0(1, g(e^{V_1-V_0}, \dots, e^{V_A-V_0}))}{G(1, g(e^{V_1-V_0}, \dots, e^{V_A-V_0}))} \quad (67)$$

where the second equality follows from the assumptions that both  $G$  and  $g$  are Hod1. (67) is a mapping from the value of  $g$  to the employment rate i.e. all combinations of outside alternatives that have the same value of  $g$  will have the same employment rate. If this mapping is monotonic then it can be inverted to go from the employment rate to the value of  $g$ . Under the conditions of the proposition, the right-hand side of (5) is monotonic in  $g$  as using the Hod1 property we can write (5) as:

$$1 - n_i = \frac{G(1, g) - gG_1(1, g)}{G(1, g)} = 1 - \frac{gG_1(1, g)}{G(1, g)} \quad (68)$$

and the right-hand side of (6) is the elasticity of  $G$  with respect to  $g$ . This means we can derive a function  $g(n)$  relating  $g$  to the employment rate. Now the inclusive value in (66) can be written as:

$$\bar{U}_i = V_{0i} + \ln G(1, g) + \gamma = V_{0i} + \ln G(1, g(n_i)) + \gamma \quad (69)$$

which is of the form of (30).

# Tables and figures

Table 1: Models for the Cost of Commuting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	FE Poisson	FE Poisson	FE Poisson	FE Poisson	Log-Linear	Log-Linear	NLS
Log Distance	-1.23*** (0.030)	-1.033*** (0.0223)	-0.937*** (0.0222)	-0.963*** (0.0194)	-2.872*** (0.00324)	-3.103*** (0.00245)	-2.26*** (0.0004)
Log Distance Squared	-0.22*** (0.005)	-0.228*** (0.00504)	-0.228*** (0.00503)	-0.206*** (0.00465)	0.289*** (0.000443)	0.290*** (0.000338)	
Origin Fixed Effects	yes	no	yes	no			
Destination Fixed Effects	yes	no	no	yes			
Observations	99.5m	99.5m	99.5m	99.5m	99.5m	99.5m	99.5m

Table 2: The Determinants of  $\log Z_{bt}$

PANEL A: OLS and IV				
	OLS		IV	
	FE (1)	FD (2)	FE (3)	FD (4)
Bartik	0.407*** (0.025)	0.343*** (0.025)	0.382*** (0.035)	0.346*** (0.028)
$\Omega^{nw}\Omega^{nr}$ *Bartik	-0.089*** (0.023)	-0.020 (0.020)	0.168*** (0.046)	0.172*** (0.030)
$\Omega^{nw}$ *Log Population	0.538*** (0.067)	0.722*** (0.074)	-5.090*** (0.391)	-2.601*** (0.250)
$\Omega^{nw}\Omega^{nr}\Omega^{nw}$ *Log Population	-0.041 (0.116)	-0.427*** (0.122)	3.550*** (0.517)	1.082*** (0.334)
Observations	39,900	29,925	39,900	29,925

PANEL B: First stage				
	$\Omega^{nw}$ *Log Population		$\Omega^{nw}\Omega^{nr}\Omega^{nw}$ *Log Population	
	FE (1)	FD (2)	FE (3)	FD (4)
$\Omega^{nw}$ *Card Instrument	0.653*** (0.069)	1.050*** (0.062)	-0.202*** (0.028)	0.177*** (0.023)
$\Omega^{nw}\Omega^{nr}\Omega^{nw}$ *Card Instrument	-0.171** (0.085)	-0.829*** (0.064)	0.773*** (0.037)	0.174*** (0.026)
Observations	39,900	29,925	39,900	39,900

Table 3: Employment Rate, Inclusive Value and Population

PANEL A: OLS and IV								
	OLS				IV			
	FE (1)	FE (2)	FD (3)	FD (4)	FE (5)	FE (6)	FD (7)	FD (8)
Inclusive Value	0.076*** (0.003)	0.075*** (0.003)	0.038*** (0.002)	0.038*** (0.003)	0.271*** (0.020)	0.325*** (0.018)	0.121*** (0.021)	0.131*** (0.020)
Log Population	-0.024*** (0.008)		0.001 (0.008)		-0.331*** (0.022)		-0.164*** (0.022)	
Observations	39,900	39,900	29,925		39,900	39,900	29,925	29,925

PANEL B: First stage					
	Inclusive Value		Log Population		
	FE (1)	FD (2)	FE (3)	FD (4)	
Bartik Instrument	0.411*** (0.020)	0.426*** (0.019)	0.230*** (0.019)	0.228*** (0.018)	
Card Instrument	-1.187*** (0.054)	-0.880*** (0.050)	1.025*** (0.059)	0.895*** (0.060)	
Observations	39,900	39,900	29,925		39,900 39,900 29,925

Table 4: The Value of Commuting: Change in Expected Utility from Imposing Sub-Optimal Commuting Patterns

Percentile:	Returns				
	1981	1991	2001	2011	
Commuting Pattern	1981	0	-0.11	-0.15	-0.18
	1991	-0.11	0	-0.11	-0.13
	2001	-0.15	-0.10	0	-0.03
	2011	-0.18	-0.13	-0.03	0

The numbers in this Table represent the computed change in inclusive value from imposing sub-optimal commuting patterns for a particular set of returns to working in different areas. So, for example, the row labelled 2001 and column labelled 2011 represents the loss in inclusive value from imposing the commuting pattern of 2001 on the returns from 2011.

Table 5: Population response: ward-level

PANEL A: OLS and IV						
	OLS			IV		
	Basic (1)	FE (2)	FD (3)	Basic (4)	FE (5)	FD (6)
$\Delta$ log emp 16-64	0.907*** (0.007)	0.939*** (0.005)	0.933*** (0.005)	0.607*** (0.150)	0.721*** (0.134)	0.678*** (0.158)
Lagged log emp rate 16-64	0.183*** (0.018)	0.606*** (0.021)	0.945*** (0.029)	0.419*** (0.096)	0.424*** (0.085)	0.963*** (0.141)
Observations	39,900	39,900	29,925	39,900	39,900	29,925

PANEL B: First stage						
	$\Delta$ log emp 16-64			Lagged log emp rate 16-64		
	Basic (1)	FE (2)	FD (3)	Basic (4)	FE (5)	FD (6)
$\Delta$ inc value	0.673*** (0.197)	1.456*** (0.301)	1.501*** (0.301)	-0.455*** (0.111)	-0.093 (0.156)	0.012 (0.157)
Lagged $\Delta$ inc value	0.003 (0.169)	0.160 (0.314)	0.873*** (0.304)	0.910*** (0.131)	1.109*** (0.145)	0.581*** (0.137)
Observations	39,900	39,900	29,925	39,900	39,900	29,925

Table 6: Population response: TTWA-level

PANEL A: OLS and IV						
	OLS			IV		
	Basic (1)	FE (2)	FD (3)	Basic (4)	FE (5)	FD (6)
$\Delta$ log emp 16-64	0.553*** (0.022)	0.560*** (0.028)	0.543*** (0.022)	0.490*** (0.064)	0.679*** (0.084)	0.636*** (0.051)
Lagged log emp rate 16-64	0.222*** (0.022)	0.397*** (0.045)	0.667*** (0.050)	0.473*** (0.110)	0.837*** (0.149)	1.136*** (0.176)
Observations	928	928	696	928	928	696

PANEL B: First stage						
	$\Delta$ log emp 16-64			Lagged log emp rate 16-64		
	Basic (1)	FE (2)	FD (3)	Basic (4)	FE (5)	FD (6)
$\Delta$ inc value	0.858*** (0.110)	1.307*** (0.206)	1.640*** (0.246)	-0.326*** (0.063)	-0.304** (0.132)	-0.164* (0.096)
Lagged $\Delta$ inc value	-0.098 (0.096)	0.424* (0.217)	0.641*** (0.227)	0.480*** (0.075)	0.341*** (0.126)	0.284*** (0.087)
Observations	928	928	696	928	928	696

Table 7: Comparison of US Commuting Zones and UK Travel-to-Work Areas

Percentile:	TTWAs (UK, 2001)			CZs (US, 2000)		
	10th	50th	90th	10th	50th	90th
Population (000s)	20	123	512	12	107	807
Land area (000s km <sup>2</sup> )	0.30	0.74	1.96	3.32	7.82	20.15
Population density (resident/km <sup>2</sup> )	23	188	850	1	16	87
Weighted population density (resident/km <sup>2</sup> )	307	1485	2713	16	215	962
Share of residents working locally	0.68	0.74	0.87	0.78	0.90	0.96
Share of workforce residing locally	0.72	0.80	0.87	0.85	0.91	0.96

This table summaries key statistics on the distribution of the 232 British Travel-To-Work-Areas in our sample and the 722 American Commuting Zones from our US study, reporting the 10th, 50th and 90th percentiles for a number of variables for cities in each country. Weighted population density measures the average neighbourhood-level density experienced by local residents, where we define neighbourhoods as wards in the UK and census tracts in the US. The specific formula is given in equation (33). The share of residents working locally is the proportion of workers residing in the TTWA or CZ who work in the same area. And the share of workforce residing locally is the proportion of individuals working in the area who also live in it. All population data are based on census data of 2001 for the UK and 2000 for the US.



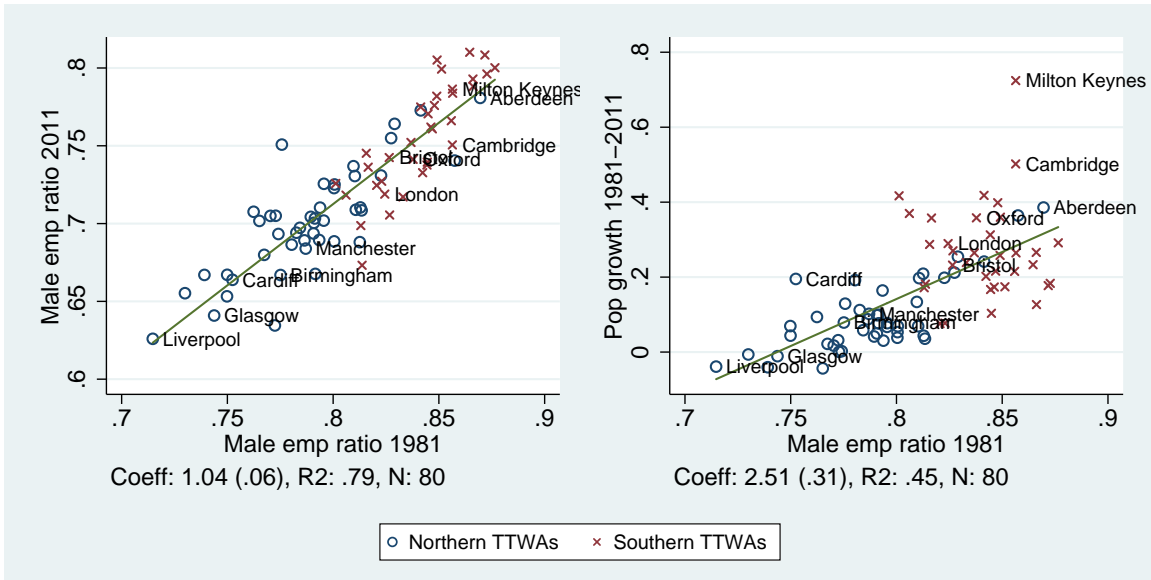


Figure 1: Persistence in male employment ratio and population response

Note: Data-points denote Travel-To-Work-Areas (TTWAs). Sample is restricted to the 80 largest commuting zones in 1981, for individuals aged 16-64. TTWAs are divided into “North” and “South”, where the latter consists of the South West, South East, East of England and East Midlands regions.

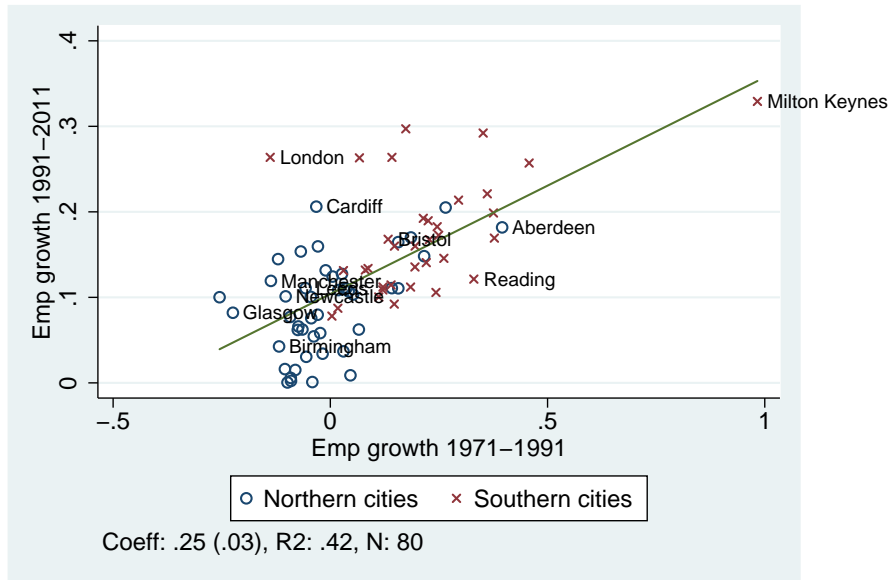


Figure 2: Persistence in local employment growth

Note: Data-points denote Travel-To-Work-Areas (TTWAs). Sample is restricted to the 80 largest commuting zones in 1981, for individuals aged 16-64

## Bibliography

- Aitkin, Murray, and Brian Francis.** 1992. “Fitting the Multinomial Logit Model with Continuous Covariates in GLIM.” *Computational Statistics & Data Analysis*, 14(1): 89–97.
- Amior, Michael, and Alan Manning.** 2015. “The Persistence of Local Joblessness.” CEP Discussion Paper 1357.
- Andrews, Martyn, Ken Clark, and William Whittaker.** 2011. “The Determinants of Regional Migration in Great Britain: A Duration Approach.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(1): 127–153.
- Autor, David H., and Mark G. Duggan.** 2003. “The Rise in the Disability Rolls and the Decline in Unemployment.” *The Quarterly Journal of Economics*, 157–205.
- Autor, David H., David Dorn, and Gordon H. Hanson.** 2013. “The China Syndrome: Local Labor Market Effects of Import Competition in the United States.” *The American Economic Review*, 103(6): 2121–2168.
- Baker, Stuart G.** 1994. “The Multinomial-Poisson Transformation.” *The Statistician*, 495–504.
- Bartik, Timothy J.** 1991. *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.
- Beaudry, Paul, David A. Green, and Benjamin M. Sand.** 2012. “Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory.” *Econometrica*, 80(3): 1063–1104.
- Beaudry, Paul, David A. Green, and Benjamin M. Sand.** 2014*a*. “In Search of Labor Demand.” NBER Working Paper No. 20568.
- Beaudry, Paul, David A. Green, and Benjamin M. Sand.** 2014*b*. “Spatial Equilibrium with Unemployment and Wage Bargaining: Theory and Estimation.” *Journal of Urban Economics*, 79: 2–19.
- Bertola, Giuseppe, and Andrea Ichino.** 1995. “Wage inequality and unemployment: United States versus Europe.” In *NBER Macroeconomics Annual 1995, Volume 10*. 13–66. MIT Press.
- Beyer, Robert C.M., and Frank Smets.** 2015. “Labour Market Adjustments and Migration in Europe and the U.S.: How Different?” ECB Working Paper No. 1767.
- Blackaby, D. H., and D. N. Manning.** 1990. “The North-South Divide: Questions of Existence and Stability.” *The Economic Journal*, 510–527.
- Blanchard, Olivier J., and Lawrence F. Katz.** 1992. “Regional Evolutions.” *Brookings Papers on Economic Activity*, 23(1): 1–76.

- Blanchflower, David G., and Andrew J. Oswald.** 1994. *The Wage Curve*. Cambridge: MIT Press.
- Böheim, René, and Mark P Taylor.** 2002. “Tied Down or Room to Move? Investigating the Relationships between Housing Tenure, Employment Status and Residential Mobility in Britain.” *Scottish Journal of Political Economy*, 49(4): 369–392.
- Bound, John, and Harry J. Holzer.** 2000. “Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s.” *Journal of Labor Economics*, 18(1): 20–54.
- Card, David.** 2001. “Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration.” *Journal of Labor Economics*, 19(1): 22–64.
- Cheshire, Paul C., and Stefano Magrini.** 2006. “Population Growth in European cities: Weather Matters – But Only Nationally.” *Regional Studies*, 40(1): 23–37.
- Dao, Mai, Davide Furceri, and Prakash Loungani.** 2014. “Regional Labor Market Adjustments in the United States.” IMF Working Paper No. 1426.
- Decressin, Jörg, and Antonio Fatás.** 1995. “Regional Labor Market Dynamics in Europe.” *European Economic Review*, 39(9): 1627–1655.
- Dorling, Danny.** 2010. “Persistent North-South Divides.” In *The Economic Geography of the UK*, ed. N.M. Coe and A. Jones, 12–28. London: Sage.
- Eichengreen, Barry.** 1993. “Policy Issues in the Operation of Currency Unions.” , ed. Mark P. Taylor Paul R. Masson, Chapter Labor markets and European monetary unification, 130–162. Cambridge University Press, Cambridge.
- Gibbons, Stephen, Henry G. Overman, and Guilherme Resende.** 2011. “Real earnings disparities in Britain.” Spatial Economics Research Centre.
- Gonzalez, Arturo.** 1998. “Mexican Enclaves and the Price of Culture.” *Journal of Urban Economics*, 43(2): 273–291.
- Gordon, Ian R., and Ian Molho.** 1998. “A Multi-stream Analysis of the Changing Pattern of Interregional Migration in Great Britain, 1960-1991.” *Regional Studies*, 32(4): 309.
- Gregg, Paul, Stephen Machin, and Alan Manning.** 2004. “Mobility and Joblessness.” In *Seeking a Premier League Economy*, ed. David Card, Richard Blundell and Richard B. Freeman. Chicago: University of Chicago Press.
- Guimaraes, Paulo.** 2004. “Understanding the Multinomial-Poisson Transformation.” *Stata Journal*, 4: 265–273.
- Hatton, T. J., and M. Tani.** 2005. “Immigration and inter-regional mobility in the UK, 1982-2000.” *Economic Journal*, 115(507): F342–F358.
- Henley, Andrew.** 1998. “Residential Mobility, Housing Equity and the Labour Market.”

- The Economic Journal*, 108(447): 414–427.
- Henley, Andrew.** 2005. “On regional Growth Convergence in Great Britain.” *Regional Studies*, 39(9): 1245–1260.
- Hotz, V Joseph, and Robert A Miller.** 1993. “Conditional Choice Probabilities and the Estimation of Dynamic Models.” *The Review of Economic Studies*, 60(3): 497–529.
- Hughes, G., and B. McCormick.** 1994. “Did migration in the 1980s narrow the North-South divide?” *Economica*, 509–527.
- Hughes, Gordon, and Barry McCormick.** 1981. “Do Council Housing Policies Reduce Migration between eRegions?” *The Economic Journal*, 91(364): 919–937.
- Hughes, Gordon, and Barry McCormick.** 1987a. “Housing Markets, Unemployment and Labour Market Flexibility in the UK.” *European Economic Review*, 31(3): 615–641.
- Hughes, Gordon, and Barry McCormick.** 1987b. “Housing markets, unemployment and labour market flexibility in the UK.” *European Economic Review*, 31(3): 615–641.
- Jackman, Richard, and Savvas Savouri.** 1992. “Regional Migration in Britain: An Analysis of Gross Flows Using NHS Central Register Data.” *The Economic Journal*, 102(415): 1433–1450.
- Jimeno, Juan F., and Samuel Bentolila.** 1998. “Regional Unemployment Persistence (Spain, 1976–1994).” *Labour Economics*, 5(1): 25–51.
- Kline, Patrick, and Enrico Moretti.** 2013. “Place Based Policies with Unemployment.” *The American Economic Review*, 103(3): 238–243.
- Manning, Alan, and Barbara Petrongolo.** 2017. “How Local are Labor Markets? Evidence from a Spatial Job Search Model.” *American Economic Review*, 107(10): 2877–2907.
- McCormick, Barry.** 1997. “Regional unemployment and labour mobility in the UK.” *European Economic Review*, 41(3-5): 581–589.
- McFadden, Daniel.** 1978. “Modeling the Choice of Residential Location.” *Transportation Research Record*, , (673).
- Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg.** 2015. “Commuting, Migration and Local Employment Elasticities.” <http://www.princeton.edu/~erossi/CMLEE.pdf>.
- Munshi, Kaivan.** 2003. “Networks in the Modern Economy: Mexican Migrants in the US Labor Market.” *The Quarterly Journal of Economics*, 118(2): 549–599.
- Notowidigdo, Matthew J.** 2011. “The Incidence of Local Labor Demand Shocks.” NBER Working Paper No. 17167.
- Obstfeld, Maurice, and Giovanni Peri.** 1998. “Regional Non-Adjustment and Fiscal

- Policy.” *Economic Policy*, 207–259.
- OECD.** 2005. “How Persistent Are Regional Disparities in Employment? The Role of Geographic Mobility.” OECD Employment Outlook.
- of Employment, Department.** 1975. “New Estimates of Employment on a Continuous Basis: Employees In Employment by Industry 1959-7.” *Department of Employment Gazette*, March 1975: 193–202.
- Oswald, Andrew J.** 1996. “A conjecture on the explanation for high unemployment in the industrialized nations: part 1.” Working or Discussion Paper.
- Overman, Henry G., and Diego Puga.** 2002. “Unemployment Clusters across Europe’s Regions and Countries.” *Economic Policy*, 17(34): 115–148.
- Pissarides, C. A., and I. McMaster.** 1990. “Regional migration, wages and unemployment: empirical evidence and implications for policy.” *Oxford Economic Papers*, 42(4): 812–831.
- Pissarides, Christopher A., and Jonathan Wadsworth.** 1989. “Unemployment and the Inter-regional Mobility of Labour.” *Economic Journal*, 99(397): 739–55.
- Rabe, Birgitta, and Mark P Taylor.** 2012. “Differences in Opportunities? Wage, Employment and House-Price Effects on Migration.” *Oxford Bulletin of Economics and Statistics*, 74(6): 831–855.
- Rappaport, Jordan.** 2007. “Moving to Nice Weather.” *Regional Science and Urban Economics*, 37(3): 375–398.
- Rappaport, Jordan.** 2012. “Why Does Unemployment Differ Persistently Across Metro Areas?” *Federal Reserve Bank of Kansas City, Economic Review*, 97(2): 5–35.
- Rappaport, Jordan, and Jeffrey D. Sachs.** 2003. “The United States as a Coastal Nation.” *Journal of Economic Growth*, 8(1): 5–46.
- Rupert, Peter, and Etienne Wasmer.** 2012. “Housing and the labor market: Time to move and aggregate unemployment.” *Journal of Monetary Economics*, 59(1): 24–36.
- Saiz, Albert.** 2010. “The Geographic Determinants of Housing Supply.” *The Quarterly Journal of Economics*, 125(3): 1253–1296.
- Tolbert, Charles M., and Molly Sizer.** 1996. “U.S. Commuting Zones and Labor Market Areas: A 1990 Update.” Economic Research Service Staff Paper No. 9614.
- Yagan, Danny.** 2014. “Moving to Opportunity? Migratory Insurance Over the Great Recession.” <http://eml.berkeley.edu/~yagan/>.