The Persistence of Local Joblessness: Online Appendix

Michael Amior and Alan Manning

A Alternative local labor market models

The model presented in the main body of the paper is deliberately kept very simple to make clear the main ideas. The aim is to show that the employment rate can serve as a sufficient statistic for local economic opportunity - the key claim which underlies our ECM empirical model. But one might be concerned that this result is not robust to alternative assumptions. In this appendix, we sketch more elaborate models to address some of these concerns. We consider a more general production structure (including intermediate goods and a non-traded goods sector that employs labor), agglomeration effects, endogenous amenities, frictional labor markets and heterogeneous skills. We also show how our industry shift-share instrument can be derived within this framework.

In general, there is a simple explanation why the sufficient statistic result is robust to these considerations. The basic argument is that utility can be expressed as a function of the employment rate and the amenity after substituting the wage curve (6) into the utility equation (7). This is because, for a given labor supply curve or "wage curve", the welfare of workers can be summarized by their position on that curve, and this position can be expressed by either the real wage or employment rate (as long as the curve is somewhat elastic). The validity of this argument is independent of how labor demand is modeled. And if amenities are endogenous, this will also be captured to the extent that they depend on variables that can be reduced to the employment rate. With this in mind, we next sketch more generalized versions of the simple model in the main text.

A.1 A more general multi-sector model

Assume that each area has S sectors, some of which produce traded goods (demanded by consumers in other areas) and some of which are non-traded (which are only demanded locally). Assume that each sector-area combination produces a distinct good. The stacked price vector $\mathbf{P}$ of all produced goods has dimension SR, where R is the number of areas. Assume for simplicity that each good is produced with constant returns to scale, possibly using intermediate goods but also labor and housing. Denote by $\mathbf{W}$ the vector of wages across areas (we assume workers are perfectly mobile across sectors within areas, so there is no wage variation within areas); and denote by $\mathbf{P^h}$ the vector of housing prices. Given the assumption of constant returns to scale, the vector of cost functions can be written as $\widetilde{c}\left(\mathbf{P}, \mathbf{W}, \mathbf{P^h}\right)$. This is written in completely general form, but the assumption that some goods are non-traded imposes restrictions on the cost function; e.g. non-traded goods' prices, wages and housing prices only affect the costs of locally produced goods. Assume that each good is produced in a perfectly competitive market so that prices are equal to marginal costs:

$$\mathbf{P} = \widetilde{c}\left(\mathbf{P}, \mathbf{W}, \mathbf{P^h}\right) \tag{A1}$$

It would be relatively simple to introduce some imperfect competition in goods markets: (A1) would simply be modified to include a mark-up representing the elasticity of the demand curve facing the firm. (A1) is a system of equations that can be solved for goods prices as a function of wages and housing costs, giving us a relationship like:

$$\mathbf{P} = c\left(\mathbf{W}, \mathbf{P^h}\right) \tag{A2}$$

where $c\left(\mathbf{W}, \mathbf{P^h}\right)$ satisfies:

$$c\left(\mathbf{W}, \mathbf{P^h}\right) = \widetilde{c}\left(c\left(\mathbf{W}, \mathbf{P^h}\right), \mathbf{W}, \mathbf{P^h}\right) \tag{A3}$$

The price for goods, both traded and non-traded will in general depend on wages and housing prices in all areas because of the presence of traded intermediate goods. If there were no traded intermediate goods, only own-area wages and housing prices would affect the prices of goods produced in an area. It is useful to differentiate (A3) to give:

$$\mathbf{c_\Lambda} = \widetilde{\mathbf{c}}_\mathbf{P}\mathbf{c_\Lambda} + \widetilde{\mathbf{c}}_\mathbf{\Lambda},\ \mathbf{\Lambda} = \mathbf{W}, \mathbf{P^h} \tag{A4}$$

We will first look for an equilibrium conditional on $\left(\mathbf{W}, \mathbf{P^h}\right)$ and assuming the demand-side of the market determines quantities. And we will then close the model by introducing supply curves for labor and housing. Denote by $\mathbf{X^O}$ the vector of gross outputs of all the goods produced. Given $\mathbf{X^O}$, the demand for intermediate inputs can be written as:

$$\mathbf{X^I} = \widetilde{\mathbf{c}}_\mathbf{P}\mathbf{X^O} \tag{A5}$$

Net output, $\mathbf{X} = \mathbf{X^O} - \mathbf{X^I}$, can then be written in terms of gross outputs:

$$\mathbf{X} = \left[\mathbf{I} - \widetilde{\mathbf{c}}_\mathbf{P}\right]\mathbf{X^O} \tag{A6}$$

Similarly to (A5), the demand for labor and housing can be written as:

$$\mathbf{N} = \widetilde{\mathbf{c}}_\mathbf{W}\mathbf{X^O} = \widetilde{\mathbf{c}}_\mathbf{W}\left[\mathbf{I} - \widetilde{\mathbf{c}}_\mathbf{P}\right]^{-1}\mathbf{X} = \mathbf{c_W}\mathbf{X} \tag{A7}$$

$$\mathbf{H} = \widetilde{\mathbf{c}}_\mathbf{P^h}\mathbf{X^O} = \widetilde{\mathbf{c}}_\mathbf{P^h}\left[\mathbf{I} - \widetilde{\mathbf{c}}_\mathbf{P}\right]^{-1}\mathbf{X} = \mathbf{c_{P^h}}\mathbf{X} \tag{A8}$$

where the second equality comes from (A6) and the final equality from inversion of (A4).

Now consider the household side of the economy. Assume that the vector of utilities across areas, $\mathbf{U}$, is given by the following indirect utility function:

$$\mathbf{U} = \mathbf{A} \circledcirc \mathbf{Y} \circledcirc \mathbf{Q^{-1}} \tag{A9}$$

where $Q\left(\mathbf{P}, \mathbf{P^h}\right)$ is the vector of consumer price indices across areas, $\mathbf{Y}$ is the vector of incomes, $\mathbf{A}$ is a vector of amenities, and $\circledcirc$ denotes a Hadamard (element-by-element) product. We assume that preferences are homothetic for simplicity, so we do not need to track the income distribution within areas. Suppose the income vector is given by:

$$\mathbf{Y} = (1-\tau)\,\mathbf{W} \circledcirc \mathbf{N} + \mathbf{B} \circledcirc (\mathbf{L} - \mathbf{N}) \tag{A10}$$

so local income derives from labor income (which is taxed) and some benefits, $\mathbf{B}$, paid to residents who are not in work. Our assumption of constant returns and perfect competition rules out the existence of profits, and we assume for simplicity that rents from housing are all taxed.

The use of Roy's identity together with (A9) and (A10) leads to the following demands for produced goods:

$$\mathbf{X} = \mathbf{Q_P}\mathbf{Y} = \mathbf{Q_P}\left[(1-\tau)\,\mathbf{W} \circledcirc \mathbf{N} + \mathbf{B} \circledcirc (\mathbf{L} - \mathbf{N})\right] \tag{A11}$$

and for housing:

$$\mathbf{X^h} = \mathbf{Q_{P^h}} \mathbf{Y} = \mathbf{Q_{P^h}} \left[(1-\tau) \mathbf{W} \circledcirc \mathbf{N} + \mathbf{B} \circledcirc (\mathbf{L} - \mathbf{N}) \right] \tag{A12}$$

We are now in a position to solve for the equilibrium conditional on $\left(\mathbf{W}, \mathbf{P^h}\right)$ and local population, $\mathbf{L}$. (A2) gives us prices. Substituting (A7) into (A11) gives the following equation to determine the level of employment:

$$\mathbf{N} = \mathbf{c_W} \mathbf{X} = \mathbf{c_W} \mathbf{Q_P} \left[(1-\tau) \mathbf{W} \circledcirc \mathbf{N} + \mathbf{B} \circledcirc (\mathbf{L} - \mathbf{N}) \right] \tag{A13}$$

We can now close the model by endogenizing $\left(\mathbf{W}, \mathbf{P^h}\right)$ using equilibrium in the labor and housing markets. For the labor market, we need a labor supply curve or "wage curve". Suppose the employment rate is a function of the local real consumer wage:

$$\boldsymbol{N} = \left(\boldsymbol{W} \oslash \boldsymbol{Q}\right)^{\theta} \circledcirc \boldsymbol{L} \tag{A14}$$

where $\oslash$ denotes a Hadamard (element-by-element) division. For given local population, this allows us to solve for the locally-determined variables as a function of the exogenous variables. For the reasons given in the main text, utility can be expressed as a function of the employment rate - which can thus be used as a sufficient statistic for utility. This result is independent of the assumptions made about the structure of production or demands.

For the housing market, one can simply augment the model to include a supply of land for development that may depend on the price of land, and the housing price is then determined to equilibrate the demands from firms and households. We do not spell this equation out here because there is no extra insight from doing so. Note that our main results do not require us to assume the housing supply elasticity is the same in all areas.

A.2 The Bartik shocks

The general specification of the model does not make clear the assumptions under which the Bartik shocks can be used as instruments. The idea behind the Bartik shocks is that areas have a comparative advantage in some sectors but that sectors are also subject to national shocks. One way of introducing these ideas would be to model the cost function for sector s in area r at time t, $\widetilde{c}_{srt}$, as:

$$\widetilde{c}_{srt} = \frac{\widetilde{c}_s \left(\mathbf{P_t^T}, P_{rt}^{NT}, W_{rt}, P_{rt}^h \right)}{Z_{sr} Z_{st}} \tag{A15}$$

where $\mathbf{P_t^T}$ are the traded goods prices at time t; P_{rt}^{NT}, W_{rt} and P_{rt}^{h} are the non-traded goods prices, wages and housing prices respectively in area r at time t, Z_{sr} is an area-sector specific cost shifter that forms the basis of the comparative (and absolute) advantage of a region, and Z_{st} is a sector-time specific cost shifter that might arise because of technical progress. This sort of structure for the cost functions will lead to regions having different mixes of sectors that persist over time. And, changes in Z_{st} will then have different consequences for different regions, even though the shocks are national in nature.

A.3 Agglomeration effects in production

The basic model can easily be modified to allow for the existence of agglomeration effects. There are many ways to model agglomeration effects, but the most common is to assume that the scale of operations (perhaps measured by aggregate employment) affects productivity, so that there is a direct effect from the aggregate level of employment on costs. This assumption can be included in the specification of cost functions, so that an extra term in employment is introduced into the right-hand side of (A13). This employment term can then be taken to the left-hand side leaving our basic equation relating utility to the employment rate completely unchanged; and qualitatively speaking, the determination of employment will look similar (though if the agglomeration effects are very large, there may be stability issues arising from a potential multiplicity of equilibria).

A.4 Endogenous amenities

It may be that the level of population or activity in an area affects the level of the amenities offered, e.g. by affecting the range of goods on offer, the crime rate, the level of social capital or population density itself (e.g. Glaeser, Kolko and Saiz, 2001; Glaeser and Redlick, 2009; Glaeser, Resseger and Tobio, 2009). The endogeneity of amenities may also amplify the impact of a given demand shock on welfare (see Diamond, 2016). But, if these endogenous responses can be summarized by the employment rate or real wage, this will still lead to equations similar to the ones we have used. In this case, we should interpret β_1 and β_2 in equation (11) as reduced form effects - in the sense that the β parameters account for *all* effects of employment on (utility and) local population growth, both the direct labor market effects and the indirect effects due to changes in local amenities such as crime.

A.5 Labor markets with frictions

Beaudry, Green and Sand (2014*b*) use a local labor market model with frictions to investigate issues closely related to the ones we have considered. In this framework, the labor demand curve is replaced by a vacancy creation curve, and market tightness is measured using the ratio of vacancies to unemployment. But, there is a one-to-one relationship between the vacancy-unemployment ratio and the employment rate. So, the wage bargaining curve - which gives a relationship between the real wage and the vacancy-unemployment ratio - can simply be translated to a wage curve like the one we have used.

On the labor demand side, the standard non-competitive model imposes constant returns to scale in production. Vacancies are created up to the point where the hiring cost, suitably amortized, is equal to the gap between the marginal product and the wage. The hiring cost is assumed to be a function of the vacancy-unemployment ratio which, as argued above, can be replaced by the employment rate; so this gives a negative relationship between the employment rate and the wage. This does differ from the labor demand curve we presented in the main body of the paper, as population appears directly in this relationship with an elasticity of 1 (see Beaudry, Green and Sand, 2014*a*). But the elasticity will be less than 1 if one introduces some diminishing returns to labor on the production side (perhaps due to imperfect competition); and we have shown above how, with non-traded goods, population will have a direct role in any case. So, models with frictions lead to very similar if not identical conclusions and specifications of empirical equations.

A.6 Heterogeneous labor

Much of the recent urban literature emphasizes differences by skill (e.g. Moretti, 2011; Notowidigdo, 2011; Diamond, 2016). If there are skill-specific labor supply curves, then one can derive equations for population growth in different skill groups as a function of skill-specific changes in employment and the lagged employment rate. Our aggregate equations can be thought of as weighted averages of the responses for the sub-groups. But, the labor demand curve for one type of labor will depend on the wages of all types of labor, reflecting the complementarity or substitutability between skill groups. This might alter the specification of the response of employment to population changes, though this is not our main focus in this study. In any case, as we show in the main text, our main object of interest - the persistence of local joblessness - cannot be explained by observed local variation in demographic composition, which includes education among other individual characteristics.

B Deriving the population growth equation

B.1 The migration response with forward-looking behavior

The specification in (9) assumes local population flows depend solely on *current* labor market conditions and amenities, while forward-looking rational agents should also pay attention to expected *future* conditions. In this section we show how that leads to an equation that is identical in form to (9), though the interpretation of the coefficient on utility needs to be changed. Our treatment follows Gallin (2004). Assume that agents' migration decisions depend on the present discounted value of being in area r at time t, denoted by $V_r(t)$. Like Gallin, we assume $V_r(t)$ can be written as:

$$\rho V_r(t) = u_r(t) + E_t \frac{\partial V_r(t)}{\partial t} \tag{A16}$$

where ρ is the interest rate and we allow for some uncertainty in the expected utility from residing in a region. This can be solved to yield:

$$V_r(t) = E_t \int_t^{\infty} e^{-\rho(s-t)} u_r(s)\, ds \tag{A17}$$

To derive an expression for this, one needs to make an assumption about the expected future path of local utility. Assume that some degree of mean reversion is expected, so the dynamic path followed by utility can be represented by:

$$\frac{\partial u_r(t)}{\partial t} = -\xi \left[u_r(t) - u(t) \right] \tag{A18}$$

where ξ is a measure of the persistence in unemployment, with a higher value representing lower persistence. If utility follows (A18), then one can derive:

$$E_t u_r(s) = e^{-\xi(s-t)} u_r(s) + \widetilde{\Gamma}(s,t) \tag{A19}$$

for some function $\widetilde{\Gamma}(s,t)$, the exact form of which will be irrelevant. Substituting (A19) into (A17) leads to:

$$V_r(t) = \frac{u_r(t)}{\rho + \xi} + \Gamma(t) \tag{A20}$$

This equation says that - up to a time-varying constant - the present discount value of being in a region is proportional to the current level of utility. This means that one can model decisions as being based solely on the current level of utility. However, the coefficient on the current level of utility does not simply come from preferences: it also involves the parameter measuring the local persistence in the level of utility, ξ. If utility is measured by employment rates and employment rate differentials are very persistent, then we would expect ξ to be close to zero. However, we do not seek to offer a deep structural interpretation of the parameters we estimate in our population growth equation, so this does not affect the validity of what we do. If one is interested in recovering preference parameters, Gallin (2004) shows how one can derive an Euler equation for the migration decision which includes future migration as an extra control. But, this approach requires good instruments both for current labor market conditions and future migration, something which is quite demanding. In addition, models with forward-looking agents typically struggle to estimate precisely the discount factor that measures the relative importance of current and future conditions and often impose a value (see, for example, Gallin, 2004; Kennan and Walker, 2011), when the assumption of forward-looking behavior is a claim one might wish to test.

B.2 Differences in utility for employed and unemployed

Equation (9) simply assumes that the average level of utility in an area can explain population changes, but neglects possible differences in incentives to migrate for the employed and unemployed - and different speeds at which they might respond to differences in utility. This section explains why this distinction may not matter. Modify (9) to take the form:

$$\frac{\partial l_r(t)}{\partial t} = a_r(t) + \frac{N_{rt}}{L_{rt}} \gamma^e \left[V_r^e(t) - V^e(t)\right] + \frac{L_{rt} - N_{rt}}{L_{rt}} \gamma^u \left[V_r^u(t) - V^u(t)\right] \qquad \text{(A21)}$$

where V_r^e is the value of being employed in area r, and V_r^u is the value of being unemployed; and V^e and V^u without subscripts are the aggregate values. (A21) allows for differences in the extent of disequilibrium for the employed and unemployed and for different degrees of responsiveness of migration to that disequilibrium. There are different ways in which one might model the values of being employed and unemployed. The simplest would be that the employment value is the wage, and the unemployment value the benefits available. Using our result that the wage can be replaced by the employment rate, this makes (A21) simply a function of the employment rate - as written in (9). This result carries over to more complicated frameworks that model transitions between employment and unemployment. Suppose the value of being employed can be written as:

$$\rho V_r^e = W_r - \delta \left(V_r^e - V_r^u \right) \tag{A22}$$

where δ is the rate of job loss. For simplicity, we assume a constant environment, though the conclusion does not depend on that assumption. Similarly, we have for V_r^u:

$$\rho V_r^u = B + \zeta \left(V_r^e - V_r^u \right) \tag{A23}$$

where ζ is the rate of finding a job. Taking the difference between (A22) and (A23) and re-arranging leads to:

$$V_r^e - V_r^u = \frac{W_r - B}{\rho + \delta + \zeta} \tag{A24}$$

Since the job finding rate ζ is typically high, the difference in the employment and unemployment values is likely to be small. But, even if the gap was large, one would still derive similar estimation equations. Substituting (A24) into (A22) and (A23) shows that both the employment and unemployment values depend on wages and the labor market transition rates that determine employment rates. Again, we can replace the wage with the employment rate using our sufficient statistic result, so everything on the right-hand side of (A21) can be written as a function of the employment rate. So a model that pays explicit attention to the different incentives for the employed and unemployed to migrate would still end up with an equation like (9).

B.3 Deriving the discrete-time equation for population growth

In this section, we show how the model (9) in continuous time can be converted to discrete time. (9) can be written as:

$$\frac{\partial e^{\gamma t} l_r \left(t \right)}{\partial t} = \gamma e^{\gamma t} \tilde{a}_r \left(t \right) + \gamma e^{\gamma t} n_r \left(t \right) \tag{A25}$$

which has as a solution:

$$e^{\gamma t} l_r \left(t \right) = l_r \left(0 \right) + \int_0^t \gamma e^{\gamma s} \left[n_r \left(s \right) + \tilde{a}_r \left(s \right) \right] ds \tag{A26}$$

which can be re-arranged to give:

$$
\begin{aligned}
l_r(t) - l_r(0) &= \int_0^t \gamma e^{\gamma(s-t)} \left[n_r(s) - n_r(0) + \tilde{a}_r(s)\right] ds \\
&\quad + \left(1 - e^{-\gamma t}\right) \left[n_r(0) - l_r(0)\right]
\end{aligned} \tag{A27}
$$

which can be written as:

$$
\begin{aligned}
l_r(t) - l_r(0) &= n_r(t) - n_r(0) + \tilde{a}_r(t) - \tilde{a}_r(0) \\
&\quad - \int_0^t e^{\gamma(s-t)} \left[\dot{n}_r(s) + \dot{\tilde{a}}_r(s)\, ds\right] ds \\
&\quad + \left(1 - e^{-\gamma t}\right) \left[n_r(0) - l_r(0) + \tilde{a}_r(0)\right]
\end{aligned} \tag{A28}
$$

If employment n_r and the supply shifter $\tilde{a}_r$ change at a constant rate over the period, this gives:

$$
\begin{aligned}
l_r(t) - l_r(0) &= \left[1 - \left(\frac{1 - e^{-\gamma t}}{\gamma t}\right)\right] \left[n_r(t) - n_r(0) + \tilde{a}_r(t) - \tilde{a}_r(0)\right] \\
&\quad + \left(1 - e^{-\gamma t}\right) \left[n_r(0) - l_r(0) + \tilde{a}_r(0)\right]
\end{aligned} \tag{A29}
$$

which is (10).

C Deriving the employment growth equation

In this appendix, we show how one can derive an error-correction model for the employment response, i.e. equation (13) in the main text. We begin by deriving an expression for the local equilibrium level of employment. Based on the general model in Online Appendix A, this is the solution to (A13). But to keep things simple, we derive the employment equilibrium using the simple model in the main text.

Housing market equilibrium can be derived by equating (3) and (4):

$$
p_r^h - p = \frac{1}{\epsilon^{hs} - \epsilon^{hd}} \left[w_r - p + l_r + \kappa \left(n_r - l_r\right)\right] \tag{A30}
$$

and labor market equilibrium using (5) and (6):

$$
n_r = \epsilon^{nd} \left(p_r - p\right) + \frac{\epsilon^{nd}}{\epsilon^{ns}} \left(n_r - l_r - z_r^s\right) + z_r^d \tag{A31}
$$

Suppose the local price index can be written as:

$$p_r = \psi p_r^h + (1-\psi)\, p \tag{A32}$$

Then, equilibrium employment can be expressed as:

$$\begin{aligned} n_r &= \frac{(1-\kappa)\,\epsilon^{nd}\epsilon^{ns}\psi - \left(\epsilon^{hs}-\epsilon^{hd}\right)\epsilon^{nd}}{\left(\epsilon^{hs}-\epsilon^{hd}\right)\left(\epsilon^{ns}-\epsilon^{nd}\right)-\psi\left(1+\kappa\epsilon^{nd}\right)\epsilon^{ns}} l_r \\ &\quad + \frac{\epsilon^{hs}-\epsilon^{hd}-\psi}{\left(\epsilon^{hs}-\epsilon^{hd}\right)\left(\epsilon^{ns}-\epsilon^{nd}\right)-\psi\left(1+\kappa\epsilon^{nd}\right)\epsilon^{ns}} \epsilon^{ns} z_r^d \\ &\quad - \frac{\epsilon^{hs}-\epsilon^{hd}}{\left(\epsilon^{hs}-\epsilon^{hd}\right)\left(\epsilon^{ns}-\epsilon^{nd}\right)-\psi\left(1+\kappa\epsilon^{nd}\right)\epsilon^{ns}} \epsilon^{nd} z_r^s \\ &\equiv n_r^* \end{aligned} \tag{A33}$$

This relates local employment n_r to the fundamentals of that region, specifically population l_r, the labor demand shifter z_r^d and the supply shifter z_r^s. Employment is increasing in population, since the latter puts downward pressure on local wages. As we have shown in Online Appendix A, the effect of population will be amplified in a world with non-traded goods - as demand for them will rise with population. This equation for employment is a static relationship because there are no dynamics of adjustment built into the production side of the model. Taken literally, this would lead to an employment equation in which the change in employment is related to the change in population, the labor demand shock (which might be proxied by the Bartik shift-share) and the supply shock. That is, there is no error correction term including the lagged employment rate. But it is reasonable to assume there is gradual adjustment of employment towards its equilibrium level n_r^*, i.e. something analogous to (9) which would take the form of:

$$\frac{\partial n_r(t)}{\partial t} = -\gamma_n \left[n_r(t) - n_r^*(t)\right] \tag{A34}$$

Using the same technique as in Online Appendix B.3 above, this yields an ECM equation for decadal changes in employment of the following form:

$$\Delta n_{rt} = \eta_1 \Delta l_{rt} + \eta_2 \left(n_{rt-1} - \nu l_{rt-1}\right) + \eta_3 \Delta z_{rt}^d + \eta_4 z_{rt-1}^d + \eta_5 \Delta z_{rt}^s + \eta_6 z_{rt-1}^s \tag{A35}$$

where $\nu = \frac{(1-\kappa)\epsilon^{nd}\epsilon^{ns}\psi - \left(\epsilon^{hs}-\epsilon^{hd}\right)\epsilon^{nd}}{\left(\epsilon^{hs}-\epsilon^{hd}\right)\left(\epsilon^{ns}-\epsilon^{nd}\right)-\psi\left(1+\kappa\epsilon^{nd}\right)\epsilon^{ns}}$. Holding the labor demand and supply shifters fixed, the long run relationship between employment and population is $n_r = \nu l_r$. In our employment ECM equation (13) in the main text, we are effectively imposing $\nu = 1$. If we do not impose this assumption, this yields the following estimating equation:

$$\Delta n_{rt} = \alpha_0 + \alpha_1 \Delta l_{rt} + \alpha_2 \left(n_{rt-1} - l_{rt-1}\right) + \alpha_2 \left(1 - \nu\right) l_{rt-1} + \alpha_3 b_{rt} + d_t + \omega_{rt} \qquad \text{(A36)}$$

where the Bartik shift-share b_{rt} proxies for the demand shock Δz_{rt}^d, the d_t are time effects which control for supply shifts common to all areas, and there is now a separate lagged population term on the right-hand side. To test whether our assumption that $\nu = 1$ is reasonable, we re-estimate equation (13) using the same instruments as before, but this time including l_{rt-1} as an exogenous regressor on the right-hand side. Our estimate of α_2 (the coefficient on the lagged employment rate) is now -0.319, while our estimate of $\alpha_2 \left(1 - \nu\right)$ (the coefficient on the lagged population l_{rt-1}) is only 0.011. This suggests that $\nu = 1$ may be a reasonable approximation.

D Data manipulation

D.1 Population and employment status

Where possible, we take our population counts and employment rates from the published county-level aggregates from the census, extracted from the National Historical Geographic Information System (NHGIS: Manson et al., 2017). Published population counts (by age and gender) are based on 100 percent samples, while employment and participation rates are usually based on samples of 15-20 percent (depending on the variable and year). The US Census Bureau did not implement a long form questionnaire in 2010, so we supplement data in that year with pooled 2009-11 American Community Survey (ACS) samples; the ACS covers a 1 percent sample each year.

NHGIS does not report data for all demographic cells of interest, in particular some employment rates before 1980 and also population disaggregations by education: see Table A1. However, micro-data samples from the US census and ACS are available for each cross-section in the Integrated Public Use Microdata Series (IPUMS: Ruggles et al., 2017), accompanied by sub-state geographical identifiers. Our strategy is to use CZ-specific shares, estimated using the IPUMS micro-data, to disaggregate the published (NHGIS) local population counts where necessary. For example, to impute the local population of college graduates aged 16-64 in a given year, we multiply the published local population count (in that age group) with the CZ-specific college graduate population share (estimated using IPUMS micro-data). And to estimate the local employment rate among college graduates, we also use the IPUMS micro-data. Table A1 describes our sources for all population and employment status data by census year, and the accompanying notes offer greater detail on the imputation process.

Unfortunately, it is not possible to perfectly identify CZs in the IPUMS micro-data samples. Each cross-section does include sub-state geographical identifiers, but these identifiers vary across years[31], and their boundaries do not coincide with those of CZs. Following Autor and Dorn (2013) and Autor, Dorn and Hanson (2013), we estimate population counts at the intersections of each CZ and the corresponding geographical identifiers. And we impute CZ outcomes by appropriately weighting outcomes for the available geographical identifiers with these population counts.[32] We make just one modification to their CZ scheme (based on Tolbert and Sizer, 1996) to enable us to construct consistent geographies over time. Specifically, we incorporate La Paz County (AZ) into the same CZ as Yuma County (AZ). Tolbert and Sizer allocated La Paz and Yuma to different CZs, but the two counties only separated in 1983.

D.2 Industry shift-shares

To construct the Bartik industry shift-shares, we require detailed local data on industrial employment composition. We take this data from the IPUMS census extracts and ACS samples, restricting our sample to workers aged 16-64. To identify industries, we use IPUMS' consistent classification based on the 1950 census scheme.[33] Given the small sample sizes at the local level, we choose to aggregate our industry data to the 2-digit level[34]: this leaves us with 57 codes. We use the same method outlined in the subsection above to impute

[31] There are 467 State Economic Areas in the continental US in 1940 and 1950, 2,287 "Mini" Public Use Microdata Areas (PUMAs) in 1960, 405 county groups in 1970, 1,148 county groups in 1980, 1,713 PUMAs in 1990, 2,057 PUMAs in the 2000 census and the ACS until 2011, and 2,336 PUMAs in the ACS of 2012.

[32] We estimate intersection population counts in 1940 and 1950 using information from the county-SEA lookup tables from IPUMS (*https://usa.ipums.org/usa/resources/volii/ sea_county_components.xls*), together with the county population counts from NHGIS. We do the same for 1970 and 1980 using lookup tables for county groups from IPUMS: see *https://usa.ipums.org/usa/resources/volii/1970cgcc.xls* and *https://usa.ipums.org/usa/resources/volii/cg98stat.xls* respectively. At the time of writing, IPUMS had not yet released population counts for intersections between 1960 Mini PUMAs and counties; so for 1960, we have relied on a preliminary version kindly shared by Joe Grover at IPUMS. And for the remaining years, we generate the population counts using the MABLE/Geocorr applications on the Missouri Census Data Center website: *http://mcdc.missouri.edu/websas/geocorr_index.shtml*.

[33] See *https://usa.ipums.org/usa/volii/occ_ind.shtml*.

[34] To address some inconsistencies over time, we further aggregate all wholesale sectors into a single category, and similarly for public administration and also finance/insurance/real estate. We also exclude the "Not specified manufacturing industries" code because of inconsistencies.

employment counts for each CZ (by industry), exploiting the available sub-state geographical identifiers.

In the group-specific models (by education, sex and age), we construct the industry shift-share instruments in exactly the same way, but with both the local industry shares and aggregate-level changes based on group-specific employment counts.

Figure A1 traces out the employment shares by major industry sector since 1940. There was a large decline in agricultural employment between 1940 and 1970, and manufacturing employment has been in secular decline since 1970. On the flip-slide, there has been a sustained expansion of professional and financial services. Given the stickiness of local industrial composition (see Section 5.4 in the main text), these persistent trends in industrial structure can account for the substantial serial correlation in the Bartik shift-shares.

D.3 Supply controls

We provide some information here on data sources and construction of our supply controls. We begin with permanent amenities. The data on coastline are borrowed from Rappaport and Sachs (2003).[35] We take county-level data on temperature from the Center for Disease Control and Prevention, based on the period 1979-2011.[36] And our relative humidity data is taken from the Natural Amenities Scale study by McGranahan (1999)[37], for the period 1941-70. All county-level climate data is aggregated to CZ-level using land area weights. Population density in 1900 is estimated using county-level population and area data from NHGIS. There have been some changes in county boundaries in the intervening period, and we impute CZ-level data using land area allocations based on shapefiles made available by NHGIS. Finally, the log distance to the closest CZ is measured between population-weighted centroids. The Missouri Census Data Center offers population-weighted centroids for counties in 1990[38], and we estimate CZ centroids by computing the population-weighted averages of the latitudes and longitudes of these county centroids.

Next, we turn to the shift-share predictor for the contribution of foreign migration to local population growth, as popularized by Altonji and Card (1991) and Card (2001). We

[35] *https://www.kansascityfed.org/~/media/files/publicat/research/journalarticles/coast_variables.zip*

[36] *http://wonder.cdc.gov*

[37] *https://www.ers.usda.gov/data-products/natural-amenities-scale*

[38] *http://mcdc.missouri.edu/websas/geocorr90.shtml*

construct this shift-share in the following way:

$$m_{rt} = \frac{\sum_o \phi^o_{rt-1} L^F_{o(-r)t}}{L_{rt-1}} \tag{A37}$$

where ϕ^o_{rt-1} is the share of population in area r at time $t-1$ which is native to origin o; and $L^F_{o(-r)t}$ is the stock of new origin-specific foreign migrants (excluding those living in area r) who arrived in the US between $t-1$ and t. The numerator of equation (A37) 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, the exclusion of area r from $L^F_{o(-r)t}$ helps allay concerns over the endogeneity of m_{rt} to the dependent variable, local population growth Δl_{rt}. For the aggregate-level models, the sample is composed of individuals aged 16-64. For the group-specific models (by education and age), we estimate the migrant inflows $L^F_{o(-r)t}$ and initial population L_{rt-1} using group-specific counts; though we always estimate the local origin shares ϕ^o_{rt-1} using the full 16-64 sample.

We use the IPUMS micro-data to construct the migrant shift-share controls. This exercise requires a panel of CZ-origin-year cells. We include 77 origin countries in this panel. For all census years t since 1970 (inclusive), we estimate the inflow of new migrants $L^F_{o(-r)t}$ using the stock of foreign-born individuals (aged 16-64) in year t who report arriving in the US in the previous ten years (i.e. since $t-1$).

However, migrants in the 1960 census do not report year of arrival. For that year, we impute $L^F_{o(-r)t}$ using cohort changes. For example, for the aggregate-level models, we take the difference between (i) the origin-specific stock of migrants in 1960 aged 16-64 (excluding area r) and (ii) the origin-specific stock of migrants in 1950 aged 6-54 (again, excluding r). For age subgroup models, the cohort changes are constructed using appropriate age categories in the equivalent way. For education subgroup models, we derive the 1960 values of $L^F_{o(-r)t}$ using the product of (i) the aggregate-level cohort change (i.e. the difference between 16-64s and 6-54s) and (ii) the national-level education share (i.e. the college graduate or non-graduate share) of the origin o population in 1960.

D.4 Residualized house prices and housing supply elasticity

Finally, we describe how we construct the data underlying the analysis in Section 4.3. To construct residualized house prices, we use data from the IPUMS census extracts of 1960, 1970, 1980, 1990 and 2000, as well as pooled ACS cross-sections of 2009, 2010 and 2011 (for the 2010 observations). The census does not include house price data in 1950.

We restrict our sample to owner-occupied houses and apartments; and we exclude farms, units with over 10 acres of land, and units with commercial use. We further restrict attention to price observations between the 1st and 99th percentiles, within each (sub-state) geographical unit. Within each cross-section, we extract log price residuals from a hedonic regression on a range of housing characteristics, and we then estimate average log residuals within each CZ.[39] Our regression controls consist of number of rooms (9 indicators); number of bedrooms (6 indicators); an interaction between number of rooms and bedrooms; building age (up to 9 indicators, depending on cross-section); and indicators for kitchen, complete plumbing and condominium status. We also control for a house/apartment dummy, as well as interactions between this and all previously mentioned variables. These controls are similar to those used by Albouy (2008) to derive fixed-quality local housing cost measures.

We take local estimates of the elasticity of housing supply from Saiz (2010). His estimates are based on Metropolitan Statistical Areas (MSAs). We impute CZ values by weighting the MSA estimates using appropriate population allocations.[40] Since MSAs only cover a fraction of the continental US (unlike CZs which offer complete coverage), this method only yields elasticity estimates for 248 of the 722 CZs in our data.

E Further graphical illustrations of persistent joblessness

In this section, we offer further graphical illustrations of the persistence of joblessness across CZs, beyond that of Figure 1 in the main text. Figure A2 depicts the persistence over 1980-2010 of employment-population ratios, participation rates (ratio of labor force to population) and unemployment rates (unemployment to labor force), separately for men and women. And in Figure A3, we repeat this exercise for 1950-1980.

There is large persistence in joblessness in both periods, driven by both men and women and by both the unemployed and economically inactive. As Table 1 in the main text shows, the magnitude of persistence in aggregate employment rates is similar in 1980-2010 and 1950-1980. Having said that, Figure A2 shows that persistence is greater for men than women

[39] As explained in Online Appendix D.1, sub-state geographical identifiers vary by census cross-section. In general, these cannot precisely identify the CZ boundaries. As above, our strategy is to weight these log residuals using appropriate population allocations between these identifiers and the CZs.

[40] For the purposes of this exercise, we generate population counts in MSA-county interacted cells using the MABLE/Geocorr application (for 2000 geography) on the Missouri Census Data Center website: *http://mcdc.missouri.edu*. And we then aggregate up from county to CZ level.

between 1980 and 2010, and Figure A3 shows the reverse is typically true in the earlier period.

F The ACF for the local log employment rate

F.1 Construction of Table 1

Here, we provide additional information on the construction of rows 9-16 of Table 1. For row 9, we purge local employment rates of observable demographic characteristics in the following way. For each cross-section of the IPUMS micro-data (using the census for 1950-2000 and the ACS of 2009-11 for 2010), we run a logit regression of employment on a range of characteristics (age and age squared; four education indicators, each interacted with age and age squared; a gender dummy, interacted with all the earlier-mentioned variables; and black, Hispanic and foreign-born indicators) and a set of location fixed effects (where "locations" are the finest geographical indicator available in each census cross-section). Based on these estimates, we then predict the average employment rate in each location - assuming the local demographic composition in each location is identical to the national composition. We then estimate CZ-level data by weighting the location data by appropriate population allocations (see Online Appendix D.1). Just as in the other rows of Table 1, these composition-adjusted employment rates are time-demeaned before reporting the ACFs.

For row 10, we regress the log employment rate on year effects and the supply controls described in Section D.3 in the main text: climate, coastline, population density and a CZ isolation index, each of which are interacted with the full set of year effects, as well as the migrant shift-share. The purged employment rate observations are the residuals from this regression.

For row 11, we regress the log employment rate on a full set of state fixed effects, together with the usual year effects. The reported ACF corresponds to the residuals of this regression. Some CZs straddle state boundaries, so we allocate these to the state accounting for the largest population share of the CZ. In row 12, we estimate ACFs for the (time-demeaned) log employment rate of the 48 states of the continental US (rather than the 722 CZs). The Washington DC CZ is allocated to Maryland.

Row 13 purges the log employment rate of CZ fixed effects (in the same way that row 11 removes state fixed effects), together with the year effects. But given the short panel (there are only 7 observations for each CZ), these estimates are biased. We correct for this bias by following the procedure described in the section that follows. As explained below, this procedure requires one identifying assumption: we fix the ratio π of the sixth to fifth autocorrelation, and report estimates for different π in rows 14-16.

F.2 Unbiased estimator for the fixed effect ACF

In this appendix, we show how we derive an unbiased autocorrelation function for the time-demeaned log employment rate, x_{rt}, controlling for CZ fixed effects. Our data is limited to 7 time observations (over the period 1950-2010) and 722 areas, which we generalize in this exposition to T periods and R areas respectively. Suppose x_{rt} in area r is stationary with mean μ_r, which we allow to vary across areas. We are interested in modeling the average ACF across areas; so for simplicity, we assume that C^n, the nth order covariance, does not vary with area r, i.e. we have:

$$C^n = E\left[\left(x_{rt} - \mu_r\right)\left(x_{rt-n} - \mu_r\right)\right] \tag{A38}$$

for all r. But, C^n cannot be estimated directly because μ_r is unknown.

Suppose we estimate μ_r using the sample mean:

$$\hat{\mu}_r = \frac{1}{T}\sum_{t=1}^{T} x_{rt} \tag{A39}$$

We can then use $\hat{\mu}_r$ to form a sample estimate of the covariance for area r :

$$\hat{C}_r^n = \frac{1}{T-n}\sum_{t=n+1}^{T}\left(x_{rt} - \hat{\mu}_r\right)\left(x_{rt-n} - \hat{\mu}_r\right) \tag{A40}$$

for $n \leq T-1$. Since T is small, $\hat{C}_r^n$ is a biased estimator for C^n. But, we can derive the form of the bias. Specifically, taking expectations of (A40):

$$
\begin{aligned}
E\left(\hat{C}_r^n\right) &= E\left\{\frac{1}{T-n}\sum_{t=n+1}^{T}\left[(x_{rt}-\mu_r)-(\hat{\mu}_r-\mu_r)\right]\left[(x_{rt-n}-\mu_r)-(\hat{\mu}_r-\mu_r)\right]\right\} \qquad \text{(A41)}\\
&= E\left(\hat{\mu}_r-\mu_r\right)^2 + E\left\{\frac{1}{T-n}\sum_{t=n+1}^{T}\left(x_{rt}-\mu_r\right)\left(x_{rt-n}-\mu_r\right)\right\}\\
&\quad -E\left\{\frac{1}{T-n}\left(\hat{\mu}_r-\mu_r\right)\left[\sum_{t=n+1}^{T}\left(x_{rt}-\mu_r\right)+\sum_{t=n+1}^{T}\left(x_{rt-n}-\mu_r\right)\right]\right\}\\
&= C^n + E\left(\hat{\mu}_r-\mu_r\right)^2\\
&\quad -\frac{1}{T-n}E\left\{\left(\hat{\mu}_r-\mu_r\right)\left[\sum_{t=1}^{T}\left(x_{rt}-\mu_r\right)-I\left[n>0\right]\cdot\sum_{t=1}^{n}\left(x_{rt}-\mu_r\right)+\sum_{t=1}^{T-n}\left(x_{rt}-\mu_r\right)\right]\right\}\\
&= C^n - \frac{n}{T-n}E\left(\hat{\mu}_r-\mu_r\right)^2\\
&\quad +\frac{1}{T-n}E\left\{\left(\hat{\mu}_r-\mu_r\right)\left[I\left[n>0\right]\cdot\sum_{t=1}^{n}\left(x_{rt}-\mu_r\right)-\sum_{t=1}^{T-n}\left(x_{rt}-\mu_r\right)\right]\right\}
\end{aligned}
$$

where $I[n>0]$ takes 1 for $n>0$ and 0 otherwise. It is useful to express (A41) as:

$$
E\left(\hat{C}_r^n\right) = C^n - \frac{n}{T-n}E\left(\hat{\mu}_r-\mu_r\right)^2 + \frac{1}{T-n}\left\{\Omega^n - \Omega^{T-n}\right\} \qquad \text{(A42)}
$$

where

$$
\Omega^n = E\left\{\left(\hat{\mu}_r-\mu_r\right)\sum_{t=1}^{n}\left(x_{rt}-\mu_r\right)\right\} = \frac{1}{T}E\left\{\sum_{s=1}^{T}\left(x_{rs}-\mu_r\right)\sum_{t=1}^{n}\left(x_{rt}-\mu_r\right)\right\} \qquad \text{(A43)}
$$

for $n \geq 1$, and we define $\Omega^0 = 0$. As it happens, $E\left(\hat{C}_r^n\right)$ can be expressed as a function of the true covariances, C^n. This can be seen by analysing the $E\left(\hat{\mu}_r-\mu_r\right)^2$ and Ω^n terms in (A42) in turn. First, consider $E\left(\hat{\mu}_r-\mu_r\right)^2$. Notice that from (A39), we have:

$$
\hat{\mu}_r - \mu_r = \frac{1}{T}\sum_{t=1}^{T}\left(x_{rt}-\mu_r\right) \qquad \text{(A44)}
$$

so that

$$E\left(\hat{\mu}_r-\mu_r\right)^2 = \frac{1}{T^2}E\left[\sum_{t=1}^{T}\left(x_{rt}-\mu_r\right)\right]^2 \tag{A45}$$
$$= \frac{1}{T^2}E\left[\sum_{t=1}^{T}\left(x_{rt}-\mu_r\right)^2+2\sum_{s=1}^{T-1}\sum_{t=s+1}^{T}\left(x_{rt}-\mu_r\right)\left(x_{rt-s}-\mu_r\right)\right]$$
$$= \frac{1}{T}C^0+\frac{2}{T^2}\sum_{s=1}^{T-1}\left(T-s\right)C^s$$

is a linear function of the true covariances. Next, notice the Ω^n term in (A42) follows the recursion:

$$\Omega^{n+1}=\Omega^n+\frac{1}{T}E\left\{\sum_{s=1}^{T}\left(x_{rs}-\mu_r\right)\left(x_{rn}-\mu_r\right)\right\}=\Omega^n+\frac{1}{T}E\left\{\sum_{s=1}^{T}C^{|s-n|}\right\} \tag{A46}$$

for $n>0$. And given $Z^0=0$, it follows that Z^n is also a linear function of the the true covariances for all $n\geq 0$. So the same must be true of $E\left(\hat{C}_r^n\right)$ in (A42).[41] That is, there exists a $T\times T$ square matrix $\mathbf{G}$ such that:

$$E\left(\hat{\mathbf{C}}_{\mathbf{r}}\right)=\mathbf{GC} \tag{A47}$$

where $\hat{\mathbf{C}}_{\mathbf{r}}$ is a T-length vector of the sample covariances for area r, $\hat{C}_r^n$; and similarly, $\mathbf{C}$ is a vector of the true covariances C^n. In the context of (A47), a natural way to derive an unbiased estimator for $\mathbf{C}$ would be to invert the matrix $\mathbf{G}$. The problem is that $\mathbf{G}$ does not have full rank. It is easiest to see the intuition for this if $T=2$. In this case, one cannot separately identify the variance and the first-order covariance, since the only useful information is contained in $x_{r2}-x_{r1}$. Similarly, T observations are insufficient to identify $T-1$ variance/covariance parameters. One further restriction on the covariances is required for identification. We impose that $C^{T-1}=\pi C^{T-2}$, with $\pi<1$. This implies:

[41] Notice that $E\left(\hat{C}_r^0\right)=\frac{T-1}{T}C^0$ if observations are independent, from which the standard formula for deriving an unbiased estimate of the variance follows.

$$\begin{pmatrix} E\left[\hat{C}_r^0\right] \\ \vdots \\ E\left[\hat{C}_r^{T-2}\right] \end{pmatrix}_{(T-1)\times 1} = \begin{pmatrix} \mathbf{G_1} & \mathbf{g_1} \end{pmatrix}_{(T-1)\times T} \begin{pmatrix} C^0 \\ \vdots \\ C^{T-2} \\ \pi C^{T-2} \end{pmatrix}_{T\times 1} \tag{A48}$$

where

$$\mathbf{G_1} = \begin{pmatrix} \mathbf{G}[0,0] & \cdots & \mathbf{G}[0,T-2] \\ \vdots & \ddots & \vdots \\ \mathbf{G}[T-2,0] & \cdots & \mathbf{G}[T-2,T-2] \end{pmatrix}_{(T-1)\times(T-1)} \tag{A49}$$

is the top left submatrix of $\mathbf{G}$, excluding the final column and final row. And

$$\mathbf{g_1} = \begin{pmatrix} \mathbf{G}[0,T-1] \\ \vdots \\ \mathbf{G}[T-2,T-1] \end{pmatrix}_{(T-1)\times 1} \tag{A50}$$

is the final column of $\mathbf{G}$, excluding the final row. And so:

$$\begin{pmatrix} E\left[\hat{C}_r^0\right] \\ \vdots \\ E\left[\hat{C}_r^{T-2}\right] \end{pmatrix}_{(T-1)\times 1} = \begin{pmatrix} \mathbf{G_1} & \pi\mathbf{g_1} \end{pmatrix}_{(T-1)\times T} \begin{pmatrix} C^0 \\ \vdots \\ C^{T-2} \\ C^{T-2} \end{pmatrix}_{T\times 1} \tag{A51}$$

which implies:

$$\begin{pmatrix} E\left[\hat{C}_r^0\right] \\ \vdots \\ E\left[\hat{C}_r^{T-2}\right] \end{pmatrix}_{(T-1)\times 1} = \left[\mathbf{G_1} + \begin{pmatrix} 0 & \cdots & 0 & \\ \vdots & & \vdots & \pi\mathbf{g_1} \\ 0 & \cdots & 0 & \end{pmatrix}\right]_{(T-1)\times(T-1)} \begin{pmatrix} C^0 \\ \vdots \\ C^{T-2} \end{pmatrix}_{(T-1)\times 1} \tag{A52}$$

where the square matrix in (A52) is the sum of (i) $\mathbf{G_1}$ and (ii) a $(T-1)\times(T-1)$ square matrix with $\pi\mathbf{g_1}$ in the final column and 0s in the remaining columns. Inverting this expression then suggests a set of unbiased estimators $\tilde{C}^n$ for the true covariances:

$$\begin{pmatrix} \tilde{C}^0 \\ \vdots \\ \tilde{C}^{T-2} \end{pmatrix} = \left[\mathbf{G_1} + \begin{pmatrix} 0 & \cdots & 0 & \\ \vdots & & \vdots & \pi \mathbf{g_1} \\ 0 & \cdots & 0 & \end{pmatrix} \right]^{-1}_{(T-1)\times(T-1)} \frac{1}{R} \sum_{r=1}^{R} \begin{pmatrix} \hat{C}^0_r \\ \vdots \\ \hat{C}^{T-2}_r \end{pmatrix} \tag{A53}$$

which is a linear function of the biased covariances, averaged across areas r. The highest order covariance estimator $\tilde{C}^{T-1}$ is set to $\pi\tilde{C}^{T-2}$. The nth order ACF can then be estimated as:

$$ACF^n = \frac{\tilde{C}^n}{\tilde{C}^0} \tag{A54}$$

For large R and small T, this is a consistent estimate of the true ACF.

F.3 Controlling for presence of local colleges

When estimating the autocorrelations in Table 1 in the main text, we control for local differences in a range of observable demographic characteristics - but we find this has little effect. However, there may still be components of human capital quality not captured in individual education. We attempt to control for this using county-level counts of two-year and four-year colleges between 1960 and 1996, based on Currie and Moretti (2003), and kindly shared by the authors. We have taken those observations which are coincidental with the census years (i.e. 1960, 1970, 1980 and 1990), supplemented these with data for 2000 and 2010 from the Integrated Postsecondary Education Data System[42], and aggregated them up to CZ-level.

In Table A2, we estimate decadal ACFs of the log employment rate (as in Table 1), but controlling for various indicators for the presence of colleges. Given our college data only begins in 1960, we are restricted to studying five decadal lags. Row 1 reproduces the results for the basic (time-demeaned) log employment rate in Table 1, but this time for the shorter 1960-2010 sample (i.e. excluding 1950). Next, we regress the log employment rate on year effects and two dummy variables: one for the presence of a two-year college in the CZ, and the second for the presence of a four-year college. Row 2 then reports the ACF for the residuals of this regression. In row 3, we follow the same procedure, but controlling additionally for the log number of two-year and the log number of four-year colleges, each relative to the

[42] *https://nces.ed.gov/ipeds/datacenter*

population of 16-64s. In rows 4-6, we replicate this exercise for the composition-adjusted employment rate (see row 9 of Table 1), as described in Online Appendix F.1.

Several CZs have no colleges, and these observations would be lost in the log college count specifications. Our approach is to replace the log counts with zeros in these cases, while simultaneously controlling for the zero-one dummies for local presence of colleges. This ensures we retain the full CZ sample in all specifications.

The autocorrelations in rows 2-3 are slightly smaller than in row 1 (and similarly in rows 5-6 compared to 4), but the difference is small. This suggests the presence of colleges cannot account for the large persistence in local jobless rates. We show further in Online Appendix G.3 that estimates of the population response equation (11) are robust to the inclusion of these college presence controls.

G Robustness of population response

G.1 Robustness to weighting and amenity controls

In this appendix, we study the robustness of our IV estimates of the population response (Table 2 in the main text). We begin by considering the weighting of observations and our choice of amenity controls. Panel A of Table A3 reports a range of specifications, weighting the regressions by the lagged local population share (as we do in the main text). Notice that columns 7, 8 and 9 in Panel A are identical to columns 4, 5 and 6 respectively of Table 2. Panel B reports unweighted estimates of the same specifications.

The coefficient on contemporaneous employment growth varies little with the choice of controls or weighting. For the basic specification in columns 1 to 7, our estimates vary from 0.62 to 0.74 (across both panels). The coefficient on the lagged employment rate is more sensitive to the choice of controls, especially among the weighted estimates of Panel A. In particular, the estimate in column 1 with no controls is 0.23, compared to 0.39 in our preferred specification (with no fixed effects or first differencing). Most of the difference is driven by the inclusion of the climate controls of column 3. The interactions between the time-invariant amenity controls and year effects in column 7 have little effect.

In Panel B, when we do not weight by population shares, the coefficient on the lagged employment rate also varies somewhat with the choice of controls - but in this case, it generally becomes smaller as more controls are included. The estimate with no controls in column 1 is 0.55, and the full set of controls in column 7 yields an estimate of 0.34. Notice this is very similar to our weighted estimate (0.39) in Panel A. This suggests that, conditional on our controls, the response to the employment rate is not markedly different in larger cities. Interestingly also, including fixed effects or first differencing (columns 8 and 9) makes less difference in the unweighted specifications, with estimates of 0.61 and 0.60 respectively.

In Table A4, we replicate Table A3 but excluding the disequilibrium term (the lagged employment rate) and its lagged Bartik instrument. For the basic specification (columns 1 to 7), this yields a larger coefficient on the change in employment - a consequence of the serial correlation in the Bartik shift-share. For example, when we weight by population and include the complete set of controls (column 7, Panel A), the coefficient increases from 0.70 (Table A3) to 0.85 (Table A4); and the associated standard errors are small: 0.03 and 0.02 respectively. In the fixed effect and first differenced specifications however (columns 8 and 9), omitting the disequilibrium term yields a smaller coefficient on the change in employment.

G.2 Robustness to state policy controls

Next, we check whether the population response estimates are robust to controlling for state-specific welfare and labor policies. Specifically, we study the maximum AFDC/TANF benefit paid per month for a family of four[43], the log minimum wage[44] (the largest of the state or federal levels), and the proportional state supplement to the federal EITC[45]. We do not have AFDC data for 1950, so we restrict the sample to 1960-2010 in this exercise.

Table A5 shows that controlling for these policies has little effect on our results. Columns 1, 4 and 7 report estimates with no policy controls (for the basic, fixed effect and first differenced specifications): these differ somewhat from those in Table 2 because of the shorter time sample. Columns 2, 5 and 8 control for lagged levels and changes in the policies. And columns 3, 6 and 9 do the same, but expressing the ADFC/TANF benefit and minimum wage as a fraction of the local median wage (see Online Appendix J for details on our wage data). All these policy variables are differenced in the first differenced specification.

[43]Data from Robert Moffitt: *http://www.econ2.jhu.edu/people/moffitt/datasets.html*

[44]I take data for 1960-1998 from David Neumark: *https://www.socsci.uci.edu/~dneumark/datasets.html.* His dataset ends in 1998, and I use the 1998 data for my 2000 observation. I take TANF data for 2010 from the Committee of Ways and Means: *https://greenbook-waysandmeans.house.gov/sites/greenbook.waysandmeans.house.gov/files/2012/documents/Table%207-23%20RM%20TANF.pdf.*

[45]Data from Daniel Feenberg, NBER TAXSIM project: *http://users.nber.org/~taxsim/state-eitc.html* and *http://users.nber.org/~taxsim/state-eitc.2010.html*

G.3 Robustness to college presence controls

The IV population response is also robust to controls for the presence of local colleges, which might proxy for local variation in human capital not captured by individual education. The local college control variables are described in Section F.3, and the results are presented in Table A6. We restrict our sample to the period 1960-2010, as our local college data only begins in 1960.

Columns 1, 4 and 7 report estimates with no college presence controls (for the basic, fixed effect and first differenced specifications); and again, these differ somewhat from Table 2 because of the restricted sample. Columns 2, 5 and 8 control for both changes and lags of dummies for the presence of two-year and four-year colleges. Columns 3, 6 and 9 control for these variables and, in addition, lagged log counts of two and four-year colleges (each relative to the population of 16-64s) and differenced log counts of two and four-year colleges.[46] Several CZs have no colleges (see Section F.3), and our approach is to replace missing values of the differenced and lagged log counts with zeros. Note these cases are fully identified by the differenced and lagged college presence dummies. This ensures we retain the full CZ sample in all specifications.

It is clear from Table A6 that the inclusion of these controls makes little difference to the response to employment shocks.

G.4 Robustness to predicted shifts in local industry rents

One possible concern with our Bartik shift-share instrument is that it treats all industries equally. But some industries may offer workers higher rents than others (Krueger and Summers, 1988), and the population responses may then depend on which industries are affected. To address this concern, we control for what we call a "wage Bartik" - borrowing the idea from earlier work by Beaudry et al. (2012; 2014*a*; 2014*b*). The wage Bartik predicts the change in average local wages (over one decade), given the fact that some industries (at a national level) pay more than others, and assuming that each industry in all areas r grows in line with the national rate:

$$b_{rt}^{w} = \sum_{i} \left(\hat{\phi}_{rt,pred}^{i} - \phi_{rt-1}^{i} \right) w_{t-1}^{i} \tag{A55}$$

[46]We do not control for changes in the *per capita* college counts, as the change in population is itself on the left hand side of the empirical specification.

where w^i_{t-1} is industry i's wage at time $t-1$, ϕ^i_{rt-1} is industry i's share of employed individuals in area r at $t-1$, and

$$\hat{\phi}^i_{rt,pred} = \frac{\phi^i_{rt-1}\left[n_{i(-r)t} - n_{i(-r)t-1}\right]}{\sum_i \phi^i_{rt-1}\left[n_{i(-r)t} - n_{i(-r)t-1}\right]} \tag{A56}$$

is the predicted employment share of industry i in area r at time t, given the initial industrial composition of local employment and assuming industries in all areas r grow in line with the national rate (excluding area r). To attain a closer proxy for wage rents, we base w^i_{t-1} on the industry averages of residualized hourly wages, purged of observable variation in age, education, gender and ethnicity. See Online Appendix J for further details on the wage data and residualization.

We reproduce the OLS, IV and first stage estimates for the population response (Table 2 in the main text) using the current and lagged wage Bartiks, b^w_{rt} and b^w_{rt-1}, on the right hand side as additional controls. We reports our results in Table A7. These look similar to those in Table 2 in the main text (without the wage Bartiks). Surprisingly perhaps, the contemporaneous wage Bartik takes a negative coefficient in all specifications.

G.5 Robustness to Bartik instrument specification

In this section, we study the robustness of our population response estimates to the specification of our Bartik instruments. First, we consider what happens when we base our prediction of local employment growth on industrial composition fixed at 1940, rather than at the beginning of each decade (as we do in equation (12) in the main text). Table A8 reproduces Table 2 in the main text, but using this modified Bartik instrument. There is no significant change in the IV coefficients of interest, though the standard errors are somewhat larger.

Next, we consider the implications of decomposing our Bartik instruments into broad industry components, specifically:

$$b^X_{rt} = \sum_{i \in X} \phi^i_{rt-1}\left[n_{i(-r)t} - n_{i(-r)t-1}\right] \tag{A57}$$

where the $i \in X$ are two-digit industries, and X is one of three broad industry groups: (i) agriculture/mining, (ii) manufacturing and (iii) services. These of course sum up to our standard Bartik instrument described in (12) in the main text:

$$b^{AGR/MIN}_{rt} + b^{MANUF}_{rt} + b^{SERV}_{rt} = b_{rt} \tag{A58}$$

We now re-estimate the IV population response equations, but replacing the current Bartik instrument b_{rt} with its three components b_{rt}^{X}; and replacing the lagged Bartik instrument b_{rt-1} with its three lagged components b_{rt-1}^{X}. The results are reported in Table A9 - which replicates Table 2 from the main text, but using the new instruments.

The basic IV estimates (column 4, Panel A) are little affected. Interestingly, the fixed effect and first differenced estimates (columns 5 and 6) now look more similar to the basic specification (compared to Table 2).

Panel B reports the first stage estimates. It is not the case that one industry component is uniformly more important than another: this depends on whether one looks at the contemporaneous or lagged shift shares. But in any case, the estimates are difficult to interpret: there appears to be some difficulty in disentangling the respective effects of the contemporaneous and lagged shift-shares; but perhaps this is not surprising, given the number of Bartik components on the right hand side. Broadly speaking though, the coefficients on agriculture/mining tend to be larger than the others.

To summarize, these results suggest that different types of sectoral shocks may have different effects on employment growth; but allowing for these differences has little effect on our estimates of the population response.

G.6 Outliers

Finally, in Figure A4, we illustrate graphically the IV estimates from our preferred (basic) specification: column 4 of Panel A in Table 2. This exercise helps demonstrate that our estimates are not driven by outliers. These plots follow the logic of the Frisch-Waugh theorem, but applied to 2SLS.

The first panel illustrates the coefficient on employment growth, i.e. β_1. To create this plot, we first generate predictions of the two endogenous variables (the change in log employment and the lagged log employment rate), with these predictions based on the first stage regressions. On the y-axis, we then plot the residuals from a regression of population growth on the predicted lagged employment rate and all the amenity controls and year effects. And on the x-axis, we plot the residuals of a regression of the predicted change in log employment on the same set of explanatory variables.

The second panel does same the same for the coefficient on the lagged employment rate, β_2. The y-axis gives the residuals from a regression of population growth on predicted employment growth and the amenity controls and year effects. And the x-axis gives the residuals of a regression of the predicted lagged employment rate on the same explanatory variables. Notice the standard errors of the best-fit slopes do not correspond to those in Table 2; this is because this naive estimator does not account for sampling error in the first stage. In any case, it is clear that the result is not driven by outliers.

H Supplementary estimates of heterogeneity in population responses

In Table 3 of the main text, we report IV estimates of our population growth equation for various sub-groups and sub-samples. We now report the corresponding OLS and first stage estimates - in Tables A10 and A11 respectively.

Table 3 in the main text shows the population responses are somewhat larger for 25-44s (though perhaps these differences are not as large as one might have expected). But these results do not tell us whether these differences are manifested on the margin of local participation or unemployment rates. This is explored in Table A12, which shows the labor force response in each age group is similarly large and close to 1 - at least in the basic specification. This suggests that, for each age group, sluggishness in the population response is largely manifested in economic inactivity. This might be thought surprising, but there has been a well-documented decline in the participation rate among prime-age workers in recent years.

I Population inflows and outflows

This paper has focused on the overall response of population to local shocks, but we have not considered the relative contribution of migratory inflows and outflows to this response. To address this question, we exploit the fact that census respondents were asked for their place of residence 5 years previously.

We have access to gross migratory flows between all county pairs for the periods 1965-70, 1975-80, 1985-90 and 1995-2000; and we aggregate this data to CZ level. We restrict attention to individuals aged 15-64: the flow data is generally unavailable for 16-64s (our usual age sample). We are grateful to Jack DeWaard for sharing the flow data for the 1970 and 1980 census, and to Kin Koerber for sharing the 2000 data[47]. We take our 1990 data

[47] See the C2 A1 tables on the Census 2000 Migration DVD: *https://www.census.gov/population/www/cen2000/migration/mig_dvd.html*. We also incorporate information from the B4 A1 tables on counts of individuals remaining in the same county and moving from

from the Socioeconomic Data and Application Center at Columbia University[48].

Our strategy is to re-estimate the population equation (11), but replacing the dependent variable with the 5-year migratory inflow or outflow. In particular, our specification is:

$$\frac{Flow_{rt-5,t}}{L_{rt-5}} = \beta_0 + \beta_1 \left(n_{rt} - n_{rt-10}\right) + \beta_2 \left(n_{rt-10} - l_{rt-10}\right) + \beta_3 \left(\tilde{a}_{rt} - \tilde{a}_{rt-10}\right) + \beta_4 \tilde{a}_{rt-10} + \varepsilon_{rt} \tag{A59}$$

where the t subscript now designates a year, rather than a decade (as in the main text). $Flow_{rt-5,t}$ is the gross migratory flow either into or out of area r, between year $t-5$ and t. L_{rt-5} is the local population at time $t-5$; in practice, we base this on census respondents' reported previous place of residence. The right hand side is identical to our standard population equation (11): the variables of interest are the 10-year employment change and 10-year lagged employment rate.

Since we do not observe employment outcomes five years before the census cross-sections, there is necessarily a mismatch in time horizon between the left and right hand side variables. And without data on place of residence 10 years previously, we cannot address this problem. But despite this, our estimates can still shed some light on the relative contribution of migratory inflows and outflows to overall population adjustment.

We report our OLS and IV estimates in Table A13. Given data constraints, we exclude the 1950s and 2000s from our sample. Columns 1 and 5 re-estimate our standard population equation (11) using this shorter sample: the coefficients are little affected (compare Table 2 in the main text). If it were not for the mismatch in time horizon, the response of the net inflow (final column) should closely approximate these estimates. The effects on the 5-year net inflow are uniformly smaller - unsurprisingly, given the shorter horizon. In particular, in the basic IV specification, the response to the 10-year lagged employment rate is statistically insignificant. This is not unreasonable: the response to the 10-year lagged employment rate would presumably have been concentrated in the first five years of each decade (as opposed to the final five years, which constitute the time horizon for the observed flows).

In any case, the key insight of Table A13 is in the relative contributions of migratory inflows and outflows to overall adjustment. And the broad message is that the population response to local shocks is largely driven by variation in inflows (see columns 2, 3, 6 and 7). This is consistent with evidence from Coen-Pirani (2010) and Monras (2015*a*) and is an intuitive consequence of large migration costs: though it may be costly to leave a region

abroad.

[48]See the P1 STP-28 tables at *http://sedac.ciesin.columbia.edu/data/set/acrp_enhance-migration-1990*

suffering an economic downturn, it is not costly *not* to move to such a region.

J Population response to wages

J.1 Empirical specification and data

We have argued in this study that local economic opportunity can be summarized by either the employment rate or the real consumption wage. As described in the main text, we have chosen to use the employment rate for two reasons: (i) measurement issues and (ii) because the employment rate slots naturally into our ECM framework. For completeness though, we document here the response of population to estimates of the real consumption wage. See Kline (2008) for an interesting analysis of wage and employment dynamics in response to sector-specific demand shocks, though he considers reallocation across industries (rather than geographical areas) and over shorter frequencies.

An estimating equation can be derived by substituting the real consumption wage $(w_r - p_r)$ for the employment rate in the utility equation (7), combining this with the migration response in (9), and discretizing the population response following the steps in Online Appendix B.3. We then have:

$$\Delta l_{rt} = \beta_0^w + \beta_1^w \Delta \left(w_{rt} - p_{rt}\right) + \beta_2^w \left(w_{rt-1} - p_{rt-1}\right) + \beta_3^w \Delta \tilde{a}_{rt} + \beta_4^w \tilde{a}_{rt-1} + \varepsilon_{rt} \qquad \text{(A60)}$$

where β_1^w and β_2^w measure the response of local population to changes and the lagged level of real consumption wages respectively. This is no longer an ECM specification, so these coefficients are harder to interpret in terms of the speed of adjustment.

But the largest challenge is measurement. Suppose individuals have Cobb-Douglas preferences over tradable and non-tradable goods. Then, the real consumption wage will be equal to $w_{rt} - \psi p_{rt}^h$, where w_{rt} is the local wage, p_{rt}^h is the price of non-tradable goods (which we proxy with housing rents), and the parameter ψ should equal the share of non-tradables in total expenditure. Unfortunately, the appropriate value of ψ is the subject of some debate. Davis and Ortalo-Magne (2011) suggest using the share of housing costs in total expenditure, which they place at 0.24. But Albouy (2008) proposes a much larger value of 0.65. His calculation takes into account housing's share of total expenditure (which he estimates at 21 percent), the effective federal tax rate on labor income adjusted for homeowner tax benefits (32 percent), the share of household income that depends on local wages (75 percent), and the fact that the cost-of-living differences across cities amount to a third of differences in local housing costs (net of homeowner tax benefits).

We construct residualized measures of local wages w_{rt} (purged of observable demographic characteristics) and housing rents p^h_{rt} (purged of observable housing characteristics) using the IPUMS census and ACS micro-data samples for every year since 1960: we omit 1950 because data on housing rents are unavailable. Our wage sample consists of employees aged 16-64, excluding those living in group quarters. We study hourly wages, estimated by dividing annual labor earnings by the product of weeks worked and usual hours per week. We restrict attention to wage observations between the 1st and 99th percentiles, within each (sub-state) geographical unit. For each cross-section, we extract log wage residuals from a regression on detailed demographic characteristics.[49] We then estimate the average log residual within each CZ.[50] To residualize housing rents, we use exactly the same procedure as described in Online Appendix D.4 (purging variation in observed housing characteristics), but this time for housing rents rather than prices; and our sample now consists of rented rather than owner-occupied accommodation (and again, we restrict attention to rent observations between the 1st and 99th percentiles, within each geographical unit). Housing rents are thought to offer a better approximation of the user cost of housing than prices, since the latter also account for expected appreciation; see e.g. Moretti (2013).

J.2 Results

We present our results in Table A14. The first three columns report estimates of (A60) using the nominal residualized wage in place of $(w_{rt} - p_{rt})$. In columns 4-6, we use a measure of the real consumption wage based on a ψ value of 0.24 ("real wage 1", following Davis and Ortalo-Magne, 2011); and in columns 7-9, we set ψ equal to 0.65 ("real wage 2", following Albouy, 2008). The IV estimates of (A60) are presented in Panel A and the associated first stage estimates in Panels B and C. The natural instruments to use would simply be the current and lagged Bartik shift-shares. But it turns out these instruments are weak in a number of specifications - when used alone. Inspection of the data suggests this is because of the decline of agriculture in the early part of the sample: given that agriculture typically offers low wage rents, this caused wages in rural areas to expand. To address this problem,

[49] Specifically, we control for age and age squared; four education effects, and interactions between these and the age quadratic; interactions between a gender dummy and all previously mentioned variables; and black, Hispanic and foreign-born dummies.

[50] As explained in Online Appendix D.1, sub-state geographical identifiers vary by census cross-section. In general, these cannot precisely identify the CZ boundaries. As above, our strategy is to weight these log residuals using appropriate population allocations between these identifiers and the CZs.

we also include the current and lagged wage Bartiks as two further instruments: as Section G.4 describes, these variables proxy for exactly this effect.

In almost all specifications, the current employment Bartik positively affects the current wage change (Panel B), and the lagged Bartik positively affects the lagged wage (Panel C); the one exception is the basic specification of "real wage 2". Turning to the wage Bartiks, the contemporaneous instrument has the expected positive effect on the wage change in all specifications (Panel B), though the effect of the lagged instrument on the lagged wage is usually negative (see Panel C).

Turning to the second stage (Panel A), the basic, fixed effect and first differenced specifications all yield significant positive estimates of β_1^w and β_2^w for nominal wages and "real wage 1". The basic specification estimates for "real wage 2" however are negative. This may suggest that a ψ value of 0.65 is too high. But in any case, it is clear from this exercise that it is difficult to interpret the magnitudes of the coefficients without knowledge of the true value of ψ. This is one reason why we prefer the employment rate as a sufficient statistic for local economic opportunity.

One might also be interested in estimates that include both employment and real wages on the right hand side, as follows:

$$\begin{aligned} \Delta l_{rt} &= \beta_0^{nw} + \beta_1^{nw} \Delta n_{rt} + \beta_2^{nw} \left(n_{rt-1} - l_{rt-1}\right) + \beta_3^{nw} \Delta \left(w_{rt} - p_{rt}\right) \\ &\quad + \beta_4^{nw} \left(w_{rt-1} - p_{rt-1}\right) + \beta_5^{nw} \Delta \tilde{a}_{rt} + \beta_6^{nw} \tilde{a}_{rt-1} + \varepsilon_{rt} \end{aligned} \tag{A61}$$

We already have four instruments for the four endogenous variables, so no more instruments are required. The first stages for the wage variables are reported in Panels B and C of Table A14; and the first stages for the employment variables are the same as in Panel B of Table A7, though 1950 is omitted from the sample in this case (because of the lack of data on housing costs).

We report the IV estimates of (A61) in Table A15. The coefficients on the wage variables are now mostly negative: this suggests it is difficult to empirically disentangle the effects of employment and wages. For the most part however, the employment effects are well identified, and the magnitudes of the coefficients are similar to when the wage variables are excluded (in Table 2 in the main text). Using a combination of employment and wage Bartik instruments, Beaudry, Green and Sand (2014*b*) also find that the population response to employment is more robust than the response to wages. To summarize, we think these estimates support our view that, in practice, the employment rate is a better measure of economic opportunity than wages.

K Intensive and extensive margins of the employment response

Our employment equation (13) focuses on total employment and how it responds to population and labor demand shocks. But, there is some interest in decomposing the employment response into the change in the number of establishments (extensive margin) and employment per establishment (intensive margin). We take data on local establishment counts from County Business Patterns, though our sample only begins in 1970 because earlier cross-sections have not been digitized.

We present OLS and IV estimates in Table A16 using our basic specification, without fixed effects or first differences. Columns 1 and 4 re-estimate the employment equation (13) for the smaller sample (excluding the 1950s and 1960s) for OLS and IV respectively. The results are similar to when we use the full sample in Table 5 in the main text.

In columns 2 and 5, we re-estimate this equation - but replacing the dependent variable with the change in the log number of establishments. For both OLS and IV, the establishment count responds positively to population (similarly to total employment) but also to the lagged employment rate (unlike total employment). There are a number of possible explanations, but our data does not allow us to discriminate between them. For example, there may be sizable start-up costs that make entry unattractive in low-employment areas - given the expectations of future prospects. Or alternatively, prospective entrepreneurs in more prosperous areas may be wealthier - and thus better able to afford these start-up costs.

Log employment per establishment can then be inferred as the difference between the previous two columns. The IV estimate in column 6 shows that establishment size responds negatively to both population growth and the initial employment rate, but positively to the Bartik shift-share. The latter effect is consistent with recent evidence from Dix-Carneiro and Kovak (2017), who find that both employment and establishment size declined in Brazilian regions which suffered more from trade liberalization. They argue this effect is driven by sluggish capital adjustment.

References

Albouy, David. 2008. "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas." NBER Working Paper No. 14472.

Altonji, Joseph G., and David Card. 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives." In *Immigration, Trade, and the Labor Market.* , ed. Richard B. Freeman John M. Abowd, 201–234. Chicago: University of Chicago Press.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review*, 103(5): 1553–1597.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

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.

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

Coen-Pirani, Daniele. 2010. "Understanding Gross Worker Flows across US States." *Journal of Monetary Economics*, 57(7): 769–784.

Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *Quarterly Journal of Economics*, 118(4): 1495–1532.

Davis, Morris A, and François Ortalo-Magne. 2011. "Household Expenditures, Wages, Rents." *Review of Economic Dynamics*, 14(2): 248–261.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479–524.

Dix-Carneiro, Rafael, and Brian K. Kovak. 2017. "Trade Liberalization and Regional Dynamics." *American Economic Review*, 107(10): 2908–46.

Gallin, Joshua H. 2004. "Net Migration and State Labor Market Dynamics." *Journal of Labor Economics*, 22(1): 1–21.

Glaeser, Edward L., and Charles Redlick. 2009. "Social Capital and Urban Growth."

International Regional Science Review, 32(3): 264–299.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. "Consumer City." *Journal of Economic Geography*, 1(1): 27–50.

Glaeser, Edward L., Matt Resseger, and Kristina Tobio. 2009. "Inequality in Cities." *Journal of Regional Science*, 49(4): 617–646.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, 79(1): 211–251.

Kline, Patrick. 2008. "Understanding Sectoral Labor Market Dynamics: An Equilibrium Analysis of the Oil and Gas Field Services Industry." https://eml.berkeley.edu/∼pkline.

Krueger, Alan B., and Lawrence H. Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure." *Econometrica*, 56(2): 259–293.

Manson, Steven, Jonathan Schroeder, David Van Riper, and Steven Ruggles. 2017. "IPUMS National Historical Geographic Information System: Version 12.0 [Database]." Minneapolis: University of Minnesota.

McGranahan, David A. 1999. "Natural Amenities Drive Rural Population Change." Agricultural Economic Report No. 781, U.S. Department of Agriculture.

Monras, Joan. 2015*a*. "Economic Shocks and Internal Migration." IZA Discussion Paper No. 8840.

Moretti, Enrico. 2011. "Local Labor Markets." In *Handbook of Labor Economics*. Vol. 4B, , ed. David Card and Orley Ashenfelter, 1237–1313. New York: Elsevier.

Moretti, Enrico. 2013. "Real Wage Inequality." *American Economic Journal: Applied Economics*, 5(1): 65–103.

Notowidigdo, Matthew J. 2011. "The Incidence of Local Labor Demand Shocks." NBER Working Paper No. 17167.

Rappaport, Jordan, and Jeffrey D. Sachs. 2003. "The United States as a Coastal Nation." *Journal of Economic Growth*, 8(1): 5–46.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. "Integrated Public Use Microdata Series: Version 7.0 [dataset]." Minneapolis: University of Minnesota. Minneapolis: University of Minnesota.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *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.

Tables and figures

Table A1: Data sources for population counts and employment/participation rates

Year	Population		Employment rates (i.e. emp-pop ratios)			Participation rates	
	Population counts by age/gender	College grad share (for 16-64s)	16-64s (overall and by gender)	Other age groups (16-24, 25-44, 45-64)	By education (for 16-64s)	16-64s (overall and by gender)	Other age groups (16-24, 25-44, 45-64)
1950	NHGIS (NT7, NT8)	IPUMS	IPUMS + NHGIS (NT28, NT29, NT30)	IPUMS	IPUMS	IPUMS + NHGIS (NT28)	IPUMS
1960	NHGIS (NT3, NT5)	IPUMS	IPUMS	IPUMS	IPUMS	IPUMS	IPUMS
1970	NHGIS (B58)	IPUMS	NHGIS (A67)	IPUMS	IPUMS	NHGIS (A67)	NHGIS (A67)
1980	NHGIS (B58)	IPUMS	NHGIS (A70)	NHGIS (A70)	IPUMS	NHGIS (A70)	NHGIS (A70)
1990	NHGIS (B58)	IPUMS	NHGIS (A70)	NHGIS (A70)	IPUMS	NHGIS (A70)	NHGIS (A70)
2000	NHGIS (B58)	IPUMS	NHGIS (A70)	NHGIS (A70)	IPUMS	NHGIS (A70)	NHGIS (A70)
2010	NHGIS (B58)	IPUMS	NHGIS (A70)	NHGIS (A70)	IPUMS	NHGIS (A70)	NHGIS (A70)

This table provides the data source for all county-level population counts and employment/participation rates. Employment and labor force counts are derived by multiplying the appropriate population counts and rates. NHGIS table references are reported in parentheses. The aim is to use as much information from the NHGIS as possible, as IPUMS (microdata) estimates are subject to larger sampling error and geographical mismatch (see Online Appendix D.1). Pre-1980, NHGIS data on employment and participation among 16-64s is relatively limited. Our approach in these years is the following. 1950: NHGIS reports county-level employment (and participation) rates for individuals aged 14 and over, by gender. We use this information to derive employment (and participation) rates for 16-64s, by exploiting data from IPUMS on employment (participation) rates among individuals aged 14-15 and 65+. Specifically, for each gender and overall: $EmpRate_{16-64} = [(EmpRateNHGIS_{14plus} \times PopNHGIS_{14plus}) - (EmpRateIPUMS_{14-15} \times PopNHGIS_{14-15}) - (EmpRateIPUMS_{65plus} \times PopNHGIS_{65plus})] / PopNHGIS_{16-64}$. And we use the same approach to derive participation rates among 16-64s. 1960: Unfortunately, NHGIS does not report county-level employment or participation rates in 1960, so we rely entirely on IPUMS for that year. 1970: NHGIS reports county-level labor force by gender and detailed age categories. But in terms of employment, it only offers total county counts (by gender), rather than for 16-64s specifically (or any other age category). In practice though, the unemployment rate among 16-64s is almost identical to the overall unemployment rate. So we impute the employment rate among 16-64s (separately by gender) as: $EmpRate_{16-64} = \frac{LaborForceNHGIS_{16-64}}{PopulationNHGIS_{16-64}} \times \frac{EmploymentNHGIS_{total}}{LaborForceNHGIS_{total}}$.

Table A2: ACFs of the log employment rate, controlling for local colleges

	Employment rate variant	Lag				
		1	2	3	4	5
(1)	Emp rate among 16-64s (time-demeaned)	0.86	0.78	0.72	0.58	0.58
	Controlling at CZ-level for:					
(2)	Dummies for presence of 2yr, 4yr college	0.85	0.76	0.70	0.56	0.53
(3)	+ Log no. 2yr, 4yr colleges / pop 16-64	0.83	0.74	0.67	0.52	0.48
(4)	Composition-adjusted ER (time-demeaned)	0.83	0.73	0.65	0.51	0.42
	Controlling at CZ-level for:					
(5)	Dummies for presence of 2yr, 4yr college	0.83	0.72	0.64	0.51	0.41
(6)	+ Log no. 2yr, 4yr colleges / pop 16-64	0.81	0.69	0.62	0.48	0.37

This table summarizes autocorrelation functions of the time-demeaned log employment rate, across five decadal lags, based on data between 1960 and 2010. These are estimated as the ratio of the lag-specific autocovariance to the product of the current and lagged standard deviations (weighted by CZ population share), across all CZs. Row 1 reports the ACF for the basic time-demeaned employment rate. Row 2 reports the ACF after adjusting for two dummy variables: one for the presence of a two-year college in the CZ, and the second for the presence of a four-year college. Row 3 adjusts additionally for the log number of two-year and the log number of four-year colleges, each relative to the population of 16-64s. For those observations with no colleges, we replace the log college counts with zeros. Rows 4-6 replicate the first three rows, but this time for the composition-adjusted employment rate, as described in Online Appendix F.1. See Section F.3 for further details on the local college data. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Robustness of IV population response to amenity controls and weighting

PANEL A: Weighted by lagged population share

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp 16-64	0.742***	0.737***	0.693***	0.700***	0.715***	0.712***	0.702***	0.889***	0.748***
	(0.034)	(0.038)	(0.036)	(0.035)	(0.030)	(0.030)	(0.031)	(0.052)	(0.035)
Lagged log emp rate 16-64	0.225***	0.138***	0.323***	0.360***	0.425***	0.411***	0.392***	1.223***	0.782***
	(0.049)	(0.042)	(0.055)	(0.058)	(0.062)	(0.058)	(0.056)	(0.256)	(0.165)
Migrant shift-share		0.216***	0.021	0.062	0.103*	0.078	0.056	0.529***	0.366***
		(0.055)	(0.058)	(0.061)	(0.057)	(0.057)	(0.058)	(0.157)	(0.095)
Max temp January			0.216***	0.225***	0.190***	0.194***	0.129***		
			(0.029)	(0.030)	(0.028)	(0.028)	(0.033)		
Max temp July			-0.087**	-0.118**	-0.154***	-0.149***	-0.097*		
			(0.040)	(0.046)	(0.045)	(0.044)	(0.058)		
Mean humidity July			-0.072***	-0.064***	-0.030**	-0.030**	0.048**		
			(0.017)	(0.018)	(0.015)	(0.014)	(0.021)		
Coastline dummy				-0.008**	-0.007*	-0.008**	0.010*		
				(0.003)	(0.003)	(0.003)	(0.006)		
Log pop density 1900					-0.006***	-0.004***	-0.012***		
					(0.001)	(0.001)	(0.002)		
Log distance to closest CZ						0.013**	0.005		
						(0.006)	(0.012)		
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amenity x year effects	No	No	No	No	No	No	Yes	Yes	Yes
CZ fixed effects	No	No	No	No	No	No	No	Yes	No
First differenced spec	No	No	No	No	No	No	No	No	Yes
Observations	4,332	4,332	4,332	4,332	4,332	4,332	4,332	4,332	3,610

PANEL B: Unweighted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp 16-64	0.651***	0.636***	0.620***	0.633***	0.633***	0.622***	0.663***	0.737***	0.675***
	(0.027)	(0.027)	(0.024)	(0.024)	(0.024)	(0.025)	(0.031)	(0.028)	(0.037)
Lagged log emp rate 16-64	0.547***	0.468***	0.407***	0.434***	0.433***	0.411***	0.343***	0.614***	0.603***
	(0.053)	(0.046)	(0.037)	(0.039)	(0.040)	(0.039)	(0.040)	(0.097)	(0.114)
Migrant shift-share		0.589***	0.193**	0.229***	0.229***	0.175**	0.078	0.091	0.048
		(0.076)	(0.079)	(0.080)	(0.080)	(0.077)	(0.069)	(0.087)	(0.098)
Max temp January			0.354***	0.367***	0.367***	0.359***	0.199***		
			(0.023)	(0.024)	(0.023)	(0.023)	(0.027)		
Max temp July			-0.476***	-0.509***	-0.509***	-0.473***	-0.275***		
			(0.042)	(0.043)	(0.042)	(0.041)	(0.052)		
Mean humidity July			-0.029**	-0.019	-0.020	-0.007	0.059**		
			(0.012)	(0.013)	(0.018)	(0.018)	(0.026)		
Coastline dummy				-0.014***	-0.014***	-0.018***	0.021***		
				(0.004)	(0.004)	(0.004)	(0.007)		
Log pop density 1900					0.000	0.003	-0.005**		
					(0.002)	(0.002)	(0.002)		
Log distance to closest CZ						0.034***	0.029***		
						(0.007)	(0.011)		
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amenity x year effects	No	No	No	No	No	No	Yes	Yes	Yes
CZ fixed effects	No	No	No	No	No	No	No	Yes	No
First differenced spec	No	No	No	No	No	No	No	No	Yes
Observations	4,332	4,332	4,332	4,332	4,332	4,332	4,332	4,332	3,610

This table studies the robustness of our IV estimates of the population response in Table 2 to choices of controls and weighting. As before, our sample covers the 722 CZs and six (decadal) time periods. First, we test robustness of our estimates to the weighting of observations: Panel A weights observations by the lagged population share, and Panel B applies no weighting. And second, we test robustness to the inclusion of progessively more controls. Columns 1-7 do not condition on CZ effects, and the final two columns report the fixed effect and first differenced specifications. Notice that columns 7-9 in Panel A are identical to columns 4-6 of Table 2 above. Errors are clustered by CZ, and robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Robustness of IV population response to amenity controls and weighting - excluding disequilibrium term

PANEL A: Weighted by lagged population share

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp 16-64	0.860***	0.815***	0.847***	0.848***	0.864***	0.864***	0.848***	0.810***	0.693***
	(0.032)	(0.036)	(0.022)	(0.024)	(0.019)	(0.019)	(0.018)	(0.027)	(0.049)
Migrant shift-share		0.193***	0.094**	0.096*	0.114***	0.116***	0.101***	0.209***	0.320***
		(0.048)	(0.044)	(0.052)	(0.034)	(0.039)	(0.030)	(0.051)	(0.116)
Max temp January			0.087***	0.087***	0.064***	0.063***	0.019		
			(0.016)	(0.016)	(0.013)	(0.013)	(0.031)		
Max temp July			-0.017	-0.018	-0.024	-0.024	-0.022		
			(0.020)	(0.023)	(0.020)	(0.020)	(0.059)		
Mean humidity July			-0.020**	-0.019*	-0.003	-0.002	0.070***		
			(0.009)	(0.011)	(0.009)	(0.009)	(0.022)		
Coastline dummy				0.000	0.001	0.001	0.007		
				(0.003)	(0.003)	(0.003)	(0.006)		
Log pop density 1900					-0.002***	-0.003***	-0.010***		
					(0.001)	(0.001)	(0.003)		
Log distance to closest CZ						-0.001	-0.019*		
						(0.004)	(0.010)		
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amenity x year effects	No	No	No	No	No	No	Yes	Yes	Yes
CZ fixed effects	No	No	No	No	No	No	No	Yes	No
First differenced spec	No	No	No	No	No	No	No	No	Yes
Observations	4,332	4,332	4,332	4,332	4,332	4,332	4,332	4,332	3,610

PANEL B: Unweighted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp 16-64	0.802***	0.778***	0.762***	0.757***	0.753***	0.742***	0.777***	0.687***	0.583***
	(0.018)	(0.018)	(0.018)	(0.019)	(0.020)	(0.021)	(0.023)	(0.032)	(0.055)
Migrant shift-share		0.375***	0.193***	0.184***	0.182***	0.146***	0.082**	-0.012	-0.056
		(0.042)	(0.038)	(0.039)	(0.039)	(0.037)	(0.036)	(0.070)	(0.116)
Max temp January			0.181***	0.181***	0.184***	0.185***	0.098***		
			(0.012)	(0.012)	(0.012)	(0.012)	(0.022)		
Max temp July			-0.243***	-0.239***	-0.241***	-0.225***	-0.164***		
			(0.021)	(0.022)	(0.022)	(0.021)	(0.043)		
Mean humidity July			-0.010	-0.012	-0.023*	-0.014	0.074***		
			(0.007)	(0.008)	(0.012)	(0.012)	(0.023)		
Coastline dummy				0.003	0.004	0.000	0.018***		
				(0.003)	(0.003)	(0.003)	(0.006)		
Log pop density 1900					0.001	0.003***	-0.007***		
					(0.001)	(0.001)	(0.002)		
Log distance to closest CZ						0.024***	0.015		
						(0.005)	(0.010)		
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Amenity x year effects	No	No	No	No	No	No	Yes	Yes	Yes
CZ fixed effects	No	No	No	No	No	No	No	Yes	No
First differenced spec	No	No	No	No	No	No	No	No	Yes
Observations	4,332	4,332	4,332	4,332	4,332	4,332	4,332	4,332	3,610

This table replicates Table A3 above, but excluding the disequilibrium term (the lagged employment rate) and its lagged Bartik instrument. See notes under Table A3 for further details. Errors are clustered by CZ, and robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Robustness of IV population response to welfare and labor policies

	Basic specification			CZ fixed effects			First differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp 16-64	0.625***	0.644***	0.642***	0.738***	0.731***	0.724***	0.700***	0.699***	0.663***
	(0.038)	(0.035)	(0.041)	(0.053)	(0.050)	(0.056)	(0.038)	(0.036)	(0.045)
Lagged log emp rate 16-64	0.380***	0.389***	0.370***	1.362***	1.357***	1.247***	0.921***	0.919***	0.839***
	(0.056)	(0.051)	(0.063)	(0.333)	(0.287)	(0.292)	(0.209)	(0.201)	(0.220)
Δ log max AFDC		-0.021***			-0.006			-0.013	
		(0.007)			(0.016)			(0.009)	
Lagged log max AFDC		-0.017***			-0.008			-0.008	
		(0.006)			(0.021)			(0.012)	
Δ log max AFDC (ratio)			-0.021***			-0.003			-0.014
			(0.007)			(0.017)			(0.009)
Lagged log max AFDC (ratio)			-0.020***			-0.009			-0.004
			(0.006)			(0.020)			(0.010)
Δ log min wage		-0.021			-0.100			0.017	
		(0.045)			(0.093)			(0.058)	
Lagged log min wage		-0.094*			-0.079			-0.016	
		(0.049)			(0.087)			(0.063)	
Δ log min wage (ratio)			0.009			0.032			-0.053**
			(0.031)			(0.032)			(0.027)
Lagged log min wage (ratio)			0.005			-0.007			-0.056*
			(0.024)			(0.035)			(0.033)
Δ EITC state supp		0.028**	0.034**		0.000	0.001		0.005	0.003
		(0.013)	(0.016)		(0.015)	(0.014)		(0.012)	(0.014)
Lagged EITC state supp		-0.008	-0.002		-0.031	-0.026		-0.017	-0.024
		(0.013)	(0.012)		(0.028)	(0.027)		(0.021)	(0.021)
Observations	3,610	3,610	3,610	3,610	3,610	3,610	2,888	2,888	2,888

This table tests the robustness of our IV estimates of the population response to welfare and labor policy controls. Our sample covers 722 CZs for five (decadal) time periods (1960-2010). We offer results for the basic, fixed effect and first differenced specifications. Columns 1, 4 and 7 report the IV estimates for our standard set of controls: see the notes under Table 2. The results differ somewhat from those of Table 2 because we have omitted data before 1970. Columns 2, 5 and 8 control additionally for changes and lags of three policy variables: (i) the maximum AFDC/TANF benefit paid per month for a family of four, (ii) the log minimum wage (the largest of the state or federal levels), and (iii) the proportional state supplement to the federal EITC. Columns 3, 6 and 9 repeat this exercise, but for the AFDC and minimum wage, we control for the log ratio of the policy level to the local median wage. Errors are clustered by CZ, and robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Robustness of IV population response to presence of local colleges

	Basic specification			CZ fixed effects			First differences		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp 16-64	0.625***	0.618***	0.617***	0.738***	0.737***	0.727***	0.700***	0.698***	0.685***
	(0.038)	(0.039)	(0.039)	(0.053)	(0.054)	(0.053)	(0.038)	(0.038)	(0.037)
Lagged log emp rate 16-64	0.380***	0.365***	0.357***	1.362***	1.372***	1.490***	0.921***	0.917***	0.942***
	(0.056)	(0.060)	(0.061)	(0.333)	(0.335)	(0.347)	(0.209)	(0.207)	(0.190)
Δ dummy for presence of 2yr college		0.000	-0.001		-0.011*	-0.011*		-0.001	-0.001
		(0.004)	(0.005)		(0.006)	(0.007)		(0.003)	(0.003)
Lagged dummy for presence of 2yr college		0.001	0.021		-0.005	0.013		0.006	0.072*
		(0.005)	(0.029)		(0.008)	(0.053)		(0.005)	(0.037)
Δ dummy for presence of 4yr college		0.017**	0.019**		-0.009	-0.017		-0.007	-0.008
		(0.008)	(0.008)		(0.011)	(0.013)		(0.007)	(0.008)
Lagged dummy for presence of 4yr college		0.011**	-0.026		0.010	0.472***		0.000	0.490***
		(0.005)	(0.031)		(0.021)	(0.172)		(0.014)	(0.103)
Δ log no. 2yr colleges			0.007**			0.011**			0.011***
			(0.003)			(0.005)			(0.004)
Lagged log no. 2yr colleges / pop 16-64			0.002			0.002			0.006*
			(0.003)			(0.005)			(0.003)
Δ log no. 4yr colleges			0.004			-0.003			0.016**
			(0.005)			(0.007)			(0.007)
Lagged log no. 4yr colleges / pop 16-64			-0.003			0.041***			0.043***
			(0.003)			(0.015)			(0.009)
Observations	3,610	3,610	3,610	3,610	3,610	3,610	2,888	2,888	2,888

This table tests the robustness of our IV estimates of the population response to controls for the local presence of colleges. Our sample covers 722 CZs, but just five (decadal) time periods (1960-2010) - as the local college data only begins in 1960. We offer results for the basic, fixed effect and first differenced specifications. Columns 1, 4 and 7 report the IV estimates for our standard set of controls: see the notes under Table 2. The results differ somewhat from those of Table 2 because we have omitted data before 1960. Columns 2, 5 and 8 control for both changes and lags of dummies for the presence of two and four-year colleges. Columns 3, 6 and 9 control for these variables and, in addition, lagged log counts of two and four-year colleges (each relative to the population of 16-64s) and differenced log counts of two and four-year colleges. For those observations with no colleges, we replace the differenced and lagged log counts with zeros; note these cases are fully identified by the differenced and lagged college presence dummies. See Section F.3 for further details on the local college data. Errors are clustered by CZ, and robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Estimates of population response, controlling for wage Bartiks

PANEL A: OLS and IV

	OLS			IV		
	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log emp 16-64	0.809***	0.799***	0.832***	0.707***	0.850***	0.745***
	(0.012)	(0.014)	(0.012)	(0.022)	(0.049)	(0.034)
Lagged log emp rate 16-64	0.173***	0.500***	0.953***	0.291***	0.990***	0.694***
	(0.014)	(0.032)	(0.028)	(0.044)	(0.229)	(0.162)
Current wage Bartik	-0.629***	-0.464***	-0.226***	-0.707***	-0.336***	-0.231***
	(0.088)	(0.071)	(0.066)	(0.075)	(0.104)	(0.072)
Lagged wage Bartik	0.365***	0.194**	0.143**	0.343***	0.134	0.182**
	(0.056)	(0.080)	(0.060)	(0.057)	(0.090)	(0.075)
Observations	4,332	4,332	3,610	4,332	4,332	3,610

PANEL B: First stage

	Δ log emp 16-64			Lagged log emp rate 16-64		
	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Current Bartik	0.975***	0.927***	0.762***	0.047	-0.107***	-0.003
	(0.075)	(0.082)	(0.076)	(0.039)	(0.037)	(0.029)
Lagged Bartik	0.130**	-0.028	-0.090	0.582***	0.137***	0.163***
	(0.060)	(0.059)	(0.078)	(0.061)	(0.031)	(0.020)
Current wage Bartik	0.124	0.090	0.477*	0.770***	-0.093	0.053
	(0.220)	(0.248)	(0.281)	(0.115)	(0.142)	(0.116)
Lagged wage Bartik	0.157	-0.165	-0.285	0.248	0.189*	0.379***
	(0.236)	(0.281)	(0.288)	(0.159)	(0.102)	(0.073)
Observations	4,332	4,332	3,610	4,332	4,332	3,610

This table replicates the exercise of Table 2, but controlling for current and lagged "wage Bartiks" on the right hand side - as described in Section G.4. See the notes under Table 2 for further details on the empirical specification and the right-hand side controls. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Estimates of population response with 1940-based Bartik instrument

PANEL A: OLS and IV

	OLS			IV		
	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log emp 16-64	0.814***	0.806***	0.831***	0.681***	0.801***	0.631***
	(0.012)	(0.014)	(0.012)	(0.038)	(0.063)	(0.075)
Lagged log emp rate 16-64	0.171***	0.513***	0.960***	0.351***	1.477***	0.541***
	(0.014)	(0.031)	(0.027)	(0.064)	(0.427)	(0.176)
Observations	4,332	4,332	3,610	4,332	4,332	3,610

PANEL B: First stage

	Δ log emp 16-64			Lagged log emp rate 16-64		
	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Current Bartik: 1940-based	0.480***	0.479***	0.261***	-0.013	-0.018	0.032
	(0.057)	(0.056)	(0.059)	(0.042)	(0.025)	(0.023)
Lagged Bartik: 1940-based	0.088*	-0.071	-0.245***	0.364***	0.089***	0.131***
	(0.046)	(0.053)	(0.080)	(0.044)	(0.028)	(0.019)
Observations	4,332	4,332	3,610	4,332	4,332	3,610

This table replicates the exercise of Table 2, but using a modified Bartik instrument. Specifically, we base our prediction of local employment growth on industrial composition in 1940, rather than at the beginning of each decade - as we do in equation (12) in the main text. See the notes under Table 2 for further details on the empirical specification and the right-hand side controls. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A9: Estimates of population response with disaggregated Bartik instruments

PANEL A: OLS and IV

	OLS			IV		
	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log emp 16-64	0.814***	0.806***	0.831***	0.714***	0.829***	0.773***
	(0.012)	(0.014)	(0.012)	(0.025)	(0.025)	(0.040)
Lagged log emp rate 16-64	0.171***	0.513***	0.960***	0.337***	0.470***	0.587***
	(0.014)	(0.031)	(0.027)	(0.031)	(0.116)	(0.159)
Observations	4,332	4,332	3,610	4,332	4,332	3,610

PANEL B: First stage

	Δ log emp 16-64			Lagged log emp rate 16-64		
	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)
Current Bartik: agr/mining	1.174***	1.130***	1.163***	0.196***	-0.222***	-0.171***
	(0.073)	(0.087)	(0.088)	(0.060)	(0.048)	(0.034)
Current Bartik: manufac	0.655***	0.906***	0.506***	-0.272***	0.157***	0.172**
	(0.119)	(0.136)	(0.166)	(0.079)	(0.056)	(0.080)
Current Bartik: services	0.250	-0.034	-0.440*	-0.067	-0.082	0.165***
	(0.254)	(0.282)	(0.232)	(0.112)	(0.096)	(0.063)
Lagged Bartik: agr/mining	0.005	-0.098	-0.132*	0.581***	0.132***	0.167***
	(0.087)	(0.087)	(0.079)	(0.069)	(0.038)	(0.032)
Lagged Bartik: manufac	-0.224***	-0.160	-0.398**	0.152*	0.176***	0.179***
	(0.086)	(0.116)	(0.194)	(0.087)	(0.054)	(0.033)
Lagged Bartik: services	0.885***	0.423***	0.343**	0.568***	0.021	0.039
	(0.149)	(0.164)	(0.163)	(0.126)	(0.073)	(0.055)
Observations	4,332	4,332	3,610	4,332	4,332	3,610

This table replicates the exercise of Table 2, but using modified Bartik instruments. Specifically, following equation (A57), we decompose the current and lagged Bartik variables (separately) into three broad industry groups: (i) agriculture/mining, (ii) manufacturing and (iii) services. See the notes under Table 2 for further details on the empirical specification and the right-hand side controls. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: Heterogeneity in population responses - OLS

	1950-1980 (1)	1980-2010 (2)	Lab force (3)	Coll grad (4)	Non grad (5)	16-24s (6)	25-44s (7)	45-64s (8)
Basic specification								
Δ log emp	0.839***	0.771***	0.947***	0.979***	0.780***	0.658***	0.878***	0.770***
	(0.010)	(0.022)	(0.006)	(0.003)	(0.015)	(0.012)	(0.009)	(0.013)
Lagged log emp rate	0.135***	0.205***	0.434***	0.831***	0.201***	0.218***	0.215***	0.183***
	(0.015)	(0.018)	(0.023)	(0.017)	(0.015)	(0.013)	(0.016)	(0.016)
CZ fixed effects								
Δ log emp	0.861***	0.696***	0.954***	0.981***	0.774***	0.665***	0.888***	0.729***
	(0.017)	(0.024)	(0.007)	(0.004)	(0.016)	(0.013)	(0.011)	(0.017)
Lagged log emp rate	0.788***	0.714***	0.921***	0.912***	0.584***	0.448***	0.606***	0.529***
	(0.040)	(0.060)	(0.050)	(0.015)	(0.031)	(0.033)	(0.039)	(0.023)
First differences								
Δ log emp	0.882***	0.758***	0.955***	0.991***	0.803***	0.734***	0.898***	0.699***
	(0.014)	(0.021)	(0.006)	(0.003)	(0.013)	(0.010)	(0.009)	(0.013)
Lagged log emp rate	1.005***	0.845***	1.305***	1.050***	1.021***	0.841***	1.171***	0.869***
	(0.032)	(0.051)	(0.037)	(0.011)	(0.027)	(0.030)	(0.026)	(0.022)
Obs (basic, FE)	2,166	2,166	4,332	4,331	4,332	4,332	4,332	4,332

Each column reports OLS estimates of β_1 and β_2 in the population response equation (11) for a different subsample. Columns 1 and 2 report estimates for 1950-1980 and 1980-2010 respectively. In column 3, population is replaced by labor force, and the employment rate is measured as a share of labor force participants only (i.e. excluding the inactive). Columns 4 and 5 split the sample by education, and columns 6-8 by age. In columns 4-8, all variables and instruments are constructed using group-specific data. For other columns, variables are based on individuals aged 16-64. Observation counts for the basic and fixed effect specifications are given in the final row. The first differenced sample is 772 fewer in each case. Column 4 is missing one observation, because in one largely rural CZ (centered around Mecosta County MI), there were no working-age employed graduates in the micro-data extract of the 1950 census. See notes under Table 2 for further details on empirical specification and right hand side controls. Errors clustered by CZ, and robust standard errors reported in parentheses. Each observation is weighted by lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Heterogeneity in population responses - first stage

	1950-1980 (1)	1980-2010 (2)	Lab force (3)	Coll grad (4)	Non grad (5)	16-24s (6)	25-44s (7)	45-64s (8)
Δ log emp: Basic specification								
Δ log emp	0.965***	1.049***	0.972***	0.699***	0.930***	1.028***	0.942***	0.726***
	(0.083)	(0.119)	(0.074)	(0.147)	(0.071)	(0.062)	(0.064)	(0.088)
Lagged log emp rate	0.096	0.126	0.094	0.209**	0.011	0.078	0.084	0.153***
	(0.073)	(0.141)	(0.059)	(0.103)	(0.057)	(0.059)	(0.057)	(0.056)
Δ log emp: CZ fixed effects								
Δ log emp	0.874***	0.675***	0.930***	0.544***	0.956***	1.026***	0.903***	0.693***
	(0.118)	(0.222)	(0.079)	(0.164)	(0.079)	(0.075)	(0.069)	(0.094)
Lagged log emp rate	0.060	-0.637***	-0.024	0.013	-0.020	0.040	-0.038	0.081
	(0.110)	(0.221)	(0.059)	(0.114)	(0.061)	(0.075)	(0.049)	(0.074)
Δ log emp: First differences								
Δ log emp	0.763***	0.387**	0.756***	0.525***	0.778***	1.083***	0.734***	0.536***
	(0.084)	(0.181)	(0.071)	(0.151)	(0.071)	(0.081)	(0.074)	(0.078)
Lagged log emp rate	0.175**	-0.578***	-0.118*	0.101	-0.175**	0.017	-0.119*	0.050
	(0.081)	(0.192)	(0.072)	(0.111)	(0.068)	(0.073)	(0.068)	(0.096)
Lagged log emp rate: Basic specification								
Δ log emp	0.035	0.090	0.007	0.071	0.027	-0.034	0.130***	0.022
	(0.033)	(0.116)	(0.009)	(0.045)	(0.044)	(0.041)	(0.028)	(0.050)
Lagged log emp rate	0.392***	0.821***	0.057***	0.086***	0.412***	0.488***	0.260***	0.388***
	(0.054)	(0.112)	(0.011)	(0.030)	(0.046)	(0.045)	(0.032)	(0.059)
Lagged log emp rate: CZ fixed effects								
Δ log emp	-0.091**	0.244**	-0.029***	0.175***	-0.084**	-0.226***	-0.002	-0.092
	(0.040)	(0.103)	(0.010)	(0.049)	(0.035)	(0.038)	(0.022)	(0.056)
Lagged log emp rate	0.112**	0.627***	0.016	0.155***	0.114***	0.174***	0.070**	0.069*
	(0.050)	(0.082)	(0.015)	(0.038)	(0.039)	(0.046)	(0.030)	(0.037)
Lagged log emp rate: First differences								
Δ log emp	-0.047	0.256***	0.005	0.142***	0.013	-0.163***	0.066***	0.002
	(0.029)	(0.076)	(0.014)	(0.036)	(0.030)	(0.031)	(0.022)	(0.033)
Lagged log emp rate	0.067	0.550***	0.047***	0.102***	0.162***	0.192***	0.102***	0.103***
	(0.043)	(0.058)	(0.014)	(0.028)	(0.028)	(0.035)	(0.021)	(0.032)
Obs (basic, FE)	2,166	2,166	4,332	4,331	4,332	4,332	4,332	4,332

This table reports first stage results corresponding to the IV estimates in Table 3 of the population response equation (11). The first three panels report first stage estimates for the contemporaneous employment change, and the final three for the lagged log employment rate. Each column reports estimates for a different subsample. Bartik instruments are constructed using individuals aged 16-64 in columns 1-3, and using group-specific data in columns 4-8. Observation counts for the basic and fixed effect specifications are given in the final row. The first differenced sample is 772 fewer in each case. Column 4 is missing one observation, because in one largely rural CZ (centered around Mecosta County MI), there were no working-age employed graduates in the micro-data extract of the 1950 census. See notes under Table 2 for further details on empirical specification and right hand side controls. Errors clustered by CZ, and robust standard errors reported in parentheses. Each observation is weighted by lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: IV estimates of labor force response by age

	Basic specification			CZ fixed effects			First differences		
	16-24	25-44	45-64	16-24	25-44	45-64	16-24	25-44	45-64
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp	0.901***	0.893***	0.951***	1.351***	1.026***	0.970***	0.970***	0.923***	1.066***
	(0.015)	(0.017)	(0.020)	(0.222)	(0.059)	(0.118)	(0.032)	(0.019)	(0.090)
Lagged log emp rate	1.084***	1.441***	0.870	5.252**	3.511**	0.671	1.869***	2.073***	2.023
	(0.198)	(0.293)	(0.785)	(2.353)	(1.648)	(4.540)	(0.430)	(0.507)	(2.051)
Observations	4,332	4,332	4,332	4,332	4,332	4,332	3,610	3,610	3,610

This table reports IV estimates of the labor force response across 722 CZs and six (decadal) time periods, separately by age category. The estimates are based on the empirical specification (11), except population is replaced by labor force, and the employment rate is measured as a share of labor force participants only (i.e. excluding the inactive). All variables and instruments (Bartik shift-shares) are constructed using age group-specific data. See notes under Table 2 for further details on empirical specification and right hand side controls. Errors clustered by CZ, and robust standard errors reported in parentheses. Each observation is weighted by lagged local population share. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A13: Population inflows and outflows

	OLS				IV			
	Δ log pop	Inflow	Outflow	Net inflow	Δ log pop	Inflow	Outflow	Net inflow
	10yr	5yr	5yr	5yr	10yr	5yr	5yr	5yr
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Basic specification								
Δ log emp (10yr)	0.836***	0.342***	-0.041***	0.383***	0.668***	0.595***	0.128***	0.466***
	(0.011)	(0.019)	(0.014)	(0.012)	(0.042)	(0.066)	(0.045)	(0.033)
Lagged log emp rate (10yr)	0.135***	0.116***	0.044*	0.072***	0.351***	-0.080	-0.141*	0.061
	(0.013)	(0.034)	(0.025)	(0.018)	(0.066)	(0.112)	(0.076)	(0.055)
CZ fixed effects								
Δ log emp (10yr)	0.822***	0.285***	-0.058***	0.344***	0.800***	0.343***	-0.090***	0.432***
	(0.016)	(0.017)	(0.009)	(0.018)	(0.063)	(0.039)	(0.019)	(0.038)
Lagged log emp rate (10yr)	0.583***	0.109***	0.027	0.081**	1.467***	0.335	-0.020	0.355*
	(0.036)	(0.041)	(0.025)	(0.038)	(0.399)	(0.222)	(0.124)	(0.212)
First differences								
Δ log emp (10yr)	0.836***	0.281***	-0.074***	0.355***	0.741***	0.316***	-0.143***	0.458***
	(0.013)	(0.015)	(0.008)	(0.018)	(0.046)	(0.039)	(0.023)	(0.044)
Lagged log emp rate (10yr)	0.920***	0.141***	0.023	0.118***	0.765***	0.265	-0.198	0.463
	(0.032)	(0.036)	(0.022)	(0.043)	(0.230)	(0.226)	(0.159)	(0.282)
Observations (basic, FE)	2,888	2,888	2,888	2,888	2,888	2,888	2,888	2,888

This table reports OLS and IV estimates of β_1 and β_2 in the population flow response equation (A59), separately for the basic, fixed effect and first differenced specifications. There is an unfortunate mismatch in time horizons between the left and right hand side variables, necessitated by data constraints. In columns 1 and 5, the dependent variable is the decadal change in log population (aged 16-64). In columns 2 and 6, the dependent is the ratio of 5-year population inflows to the lagged population 5 years previously, for individuals aged 15-64 in the current year. Columns 3 and 7 gives the estimates for the population outflows, and columns 4 and 8 the net inflows (the difference between the two). The right hand side is identical to our standard population equation (11), based on a 10-year time horizon. See the notes under Table 2 for further details on the empirical specification and the right-hand side controls and instruments. The sample is restricted here to 1960-2000 (given the availability of the flow data). Observation counts for the basic and fixed effect specifications are given in the final row. The first differenced sample is 772 fewer in each case. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.

Table A14: IV estimates of population response to local wages

PANEL A: IV estimates

	Nominal wage			Real wage 1 (ψ=0.24)			Real wage 2 (ψ=0.65)		
	Basic	FE	FD	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log wage	1.182***	1.282***	1.224***	2.520***	1.446***	1.354***	-2.232***	2.388***	1.988***
	(0.175)	(0.178)	(0.141)	(0.470)	(0.197)	(0.168)	(0.362)	(0.594)	(0.383)
Lagged log wage	0.753***	1.193***	1.226***	1.485***	1.115***	1.111***	-1.140***	1.201***	0.829***
	(0.095)	(0.178)	(0.154)	(0.259)	(0.150)	(0.147)	(0.226)	(0.239)	(0.170)
Observations	3,610	3,610	2,888	3,610	3,610	2,888	3,610	3,610	2,888

PANEL B: First stage for Δ log wage

	Nominal wage			Real wage 1 (ψ=0.24)			Real wage 2 (ψ=0.65)		
	Basic	FE	FD	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Current Bartik	0.288***	0.455***	0.515***	0.156***	0.327***	0.411***	-0.069	0.109	0.232***
	(0.064)	(0.075)	(0.098)	(0.051)	(0.063)	(0.079)	(0.044)	(0.067)	(0.061)
Lagged Bartik	-0.110*	0.016	0.029	-0.101*	0.027	0.052	-0.085*	0.046	0.092
	(0.065)	(0.103)	(0.125)	(0.056)	(0.086)	(0.101)	(0.047)	(0.066)	(0.070)
Current wage Bartik	3.841***	4.290***	4.682***	2.868***	3.313***	3.569***	1.205***	1.643***	1.667***
	(0.228)	(0.272)	(0.319)	(0.203)	(0.255)	(0.281)	(0.214)	(0.289)	(0.277)
Lagged wage Bartik	-0.435***	-0.332	-0.180	-0.005	-0.019	-0.002	0.728***	0.515***	0.303
	(0.142)	(0.207)	(0.302)	(0.133)	(0.172)	(0.245)	(0.153)	(0.152)	(0.189)
Observations	3,610	3,610	2,888	3,610	3,610	2,888	3,610	3,610	2,888

PANEL C: First stage for lagged log wage

	Nominal wage			Real wage 1 (ψ=0.24)			Real wage 2 (ψ=0.65)		
	Basic	FE	FD	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Current Bartik	-0.095	-0.138**	-0.067	-0.198***	-0.056	-0.037	-0.374***	0.086	0.013
	(0.070)	(0.057)	(0.043)	(0.058)	(0.053)	(0.043)	(0.065)	(0.071)	(0.052)
Lagged Bartik	0.628***	0.227***	0.160**	0.364***	0.217***	0.133**	-0.088*	0.200***	0.087*
	(0.057)	(0.051)	(0.068)	(0.040)	(0.046)	(0.059)	(0.048)	(0.051)	(0.050)
Current wage Bartik	-3.792***	-3.837***	-3.856***	-3.703***	-3.396***	-3.396***	-3.552***	-2.643***	-2.610***
	(0.301)	(0.231)	(0.173)	(0.243)	(0.213)	(0.165)	(0.235)	(0.254)	(0.188)
Lagged wage Bartik	-0.776***	0.078	0.189	-0.881***	-0.343**	-0.209	-1.061***	-1.063***	-0.888***
	(0.196)	(0.147)	(0.160)	(0.163)	(0.136)	(0.148)	(0.160)	(0.170)	(0.161)
Observations	3,610	3,610	2,888	3,610	3,610	2,888	3,610	3,610	2,888

This table reports estimates of β_1^w and β_2^w in equation (A60). Panel A reports IV estimates, and Panels B and C report the first stage - for the wage change and lagged wage respectively. We use four instruments for the two endogenous variables: the standard Bartik shift-share and the "wage Bartik" (see Section G.4), both current and lagged. In the first three columns, we use residualized wages as the endogenous variables. In columns 4-6, we construct a real wage measure ("real wage 1") based on the difference between the residualized wage and 0.24 times the residualized housing rent. And in columns 7-9, we take the difference between the residualized wage and 0.65 times the residualized housing rent ("real wage 2"). The sample covers 722 CZs and five (decadal) time periods beginning in 1960. We omit 1950 because there is no housing rents data. This yields 3,610 observations for the basic and fixed effect specifications and 2,888 observations for the first differenced specification. See the notes under Table 2 for details on the right-hand side controls. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.

Table A15: IV estimates of population response to local employment and wages

	Nominal wage			Real wage 1 (ψ=0.24)			Real wage 2 (ψ=0.65)		
	Basic	FE	FD	Basic	FE	FD	Basic	FE	FD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ log emp	0.683***	0.846***	0.956***	0.653***	0.858***	0.862***	0.886***	1.952	0.927***
	(0.040)	(0.198)	(0.243)	(0.043)	(0.171)	(0.120)	(0.119)	(4.682)	(0.183)
Lagged log emp rate	0.389***	1.538**	1.530**	0.390***	1.611**	1.282***	0.468***	6.986	1.619**
	(0.083)	(0.724)	(0.671)	(0.094)	(0.643)	(0.340)	(0.095)	(22.493)	(0.646)
Δ log wage	-0.224***	-0.337	-0.499	-0.427**	-0.442	-0.380	0.869**	-5.978	-0.829
	(0.083)	(0.413)	(0.462)	(0.180)	(0.403)	(0.258)	(0.426)	(22.037)	(0.635)
Lagged log wage	-0.089*	-0.215	-0.424	-0.193**	-0.241	-0.237	0.493**	-2.703	-0.268
	(0.049)	(0.389)	(0.453)	(0.098)	(0.317)	(0.210)	(0.219)	(10.212)	(0.250)
Observations	3,610	3,610	2,888	3,610	3,610	2,888	3,610	3,610	2,888

This table reports IV estimates of β_1^{nw}, β_2^{nw}, β_3^{nw} and β_4^{nw} in equation (A61). The first stage for the employment variables is the same as in Panel B of Table A7 (though the time sample for this table is smaller). The first stage for the wage variables are contained in Panels B and C of Table A14. The sample covers 722 CZs and five (decadal) time periods beginning in 1960. We omit 1950 because there is no housing rents data. This yields 3,610 observations for the basic and fixed effect specifications and 2,888 observations for the first differenced specification. See the notes under Table 2 for details on the right-hand side controls. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.

Table A16: Intensive and extensive margins of employment response 1970-2010

PANEL A: OLS and IV

	OLS			IV		
	Δ log emp	Δ log estab	Δ log emp/estab	Δ log emp	Δ log estab	Δ log emp/estab
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log pop	1.012***	0.952***	0.060***	0.780***	1.026***	-0.246***
	(0.013)	(0.017)	(0.021)	(0.071)	(0.069)	(0.085)
Lagged log emp rate	-0.120***	0.093***	-0.213***	-0.197***	0.423***	-0.621***
	(0.014)	(0.026)	(0.025)	(0.055)	(0.073)	(0.077)
Current Bartik	0.410***	0.159	0.251***	0.609***	-0.013	0.622***
	(0.043)	(0.115)	(0.096)	(0.074)	(0.075)	(0.092)
Observations	2,888	2,888	2,888	2,888	2,888	2,888

PANEL B: First stage

	Δ log pop	Lagged log emp rate
	(1)	(2)
Max temp January	0.672***	-0.034
	(0.144)	(0.091)
Max temp January * time	-0.057***	-0.038**
	(0.021)	(0.017)
Lagged Bartik	0.297***	0.653***
	(0.078)	(0.067)
Current Bartik	0.463***	0.039
	(0.129)	(0.131)
Observations	2,888	2,888

This table breaks down the employment response (i.e. equation (13)) in Table 5 into "intensive" and "extensive" margins. We restrict attention here to the "basic" specification (no fixed effects). Since we only have data on local establishment counts since 1970, we now restrict our sample to 1970-2010. In columns 1 and 4 of Panel A, we re-estimate the OLS and IV employment responses (basic specification) in Table 5 for the restricted sample. In columns 2 and 5, we re-estimate this equation, but change the dependent variable to the change in the log number of establishments (extensive margin). In columns 3 and 6, we change the dependent variable to the log employment per establishment (i.e. the difference between the previous two columns: the intensive margin). Panel B reports the associated first stage estimates. See the notes under Table 5 for details on the right-hand side controls. Errors are clustered by CZ, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged local population share. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

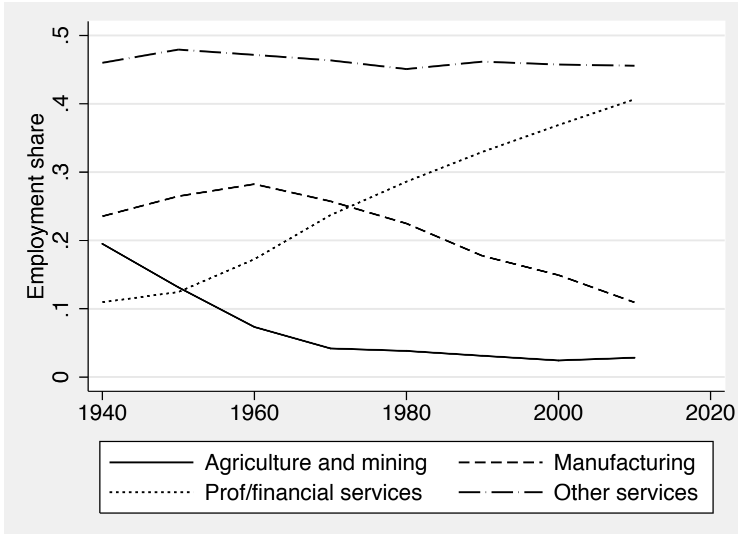

Figure A1: Employment shares by industry

We define professional and financial services as census 1950 codes 716-756, 806-808 and 868-899. See *https://usa.ipums.org/usa-action/variables/IND1950*.

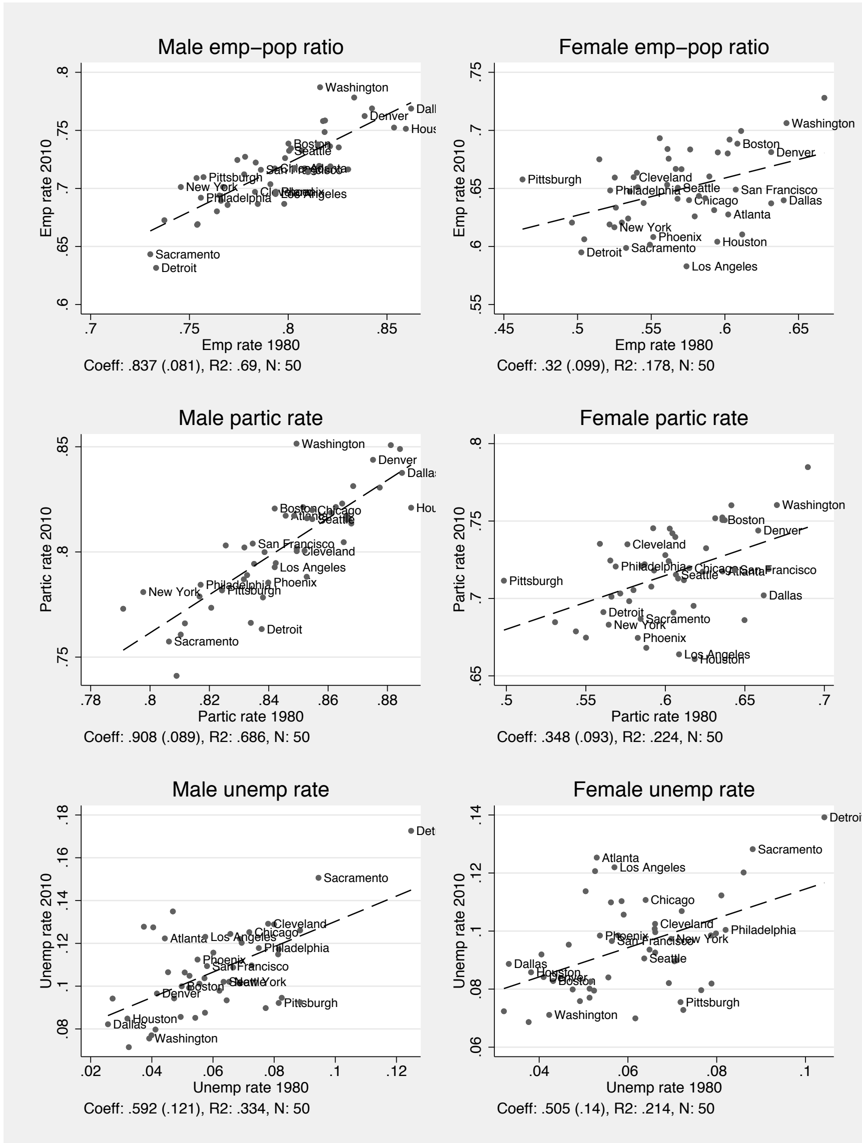

Figure A2: Persistence of joblessness across CZs: 1980-2010

Data-points denote commuting zones (CZs). Sample is restricted to 50 largest CZs in 1980, for individuals aged 16-64. "Participation rate" is ratio of labor force to population, and "unemployment rate" is ratio of unemployment to labor force.

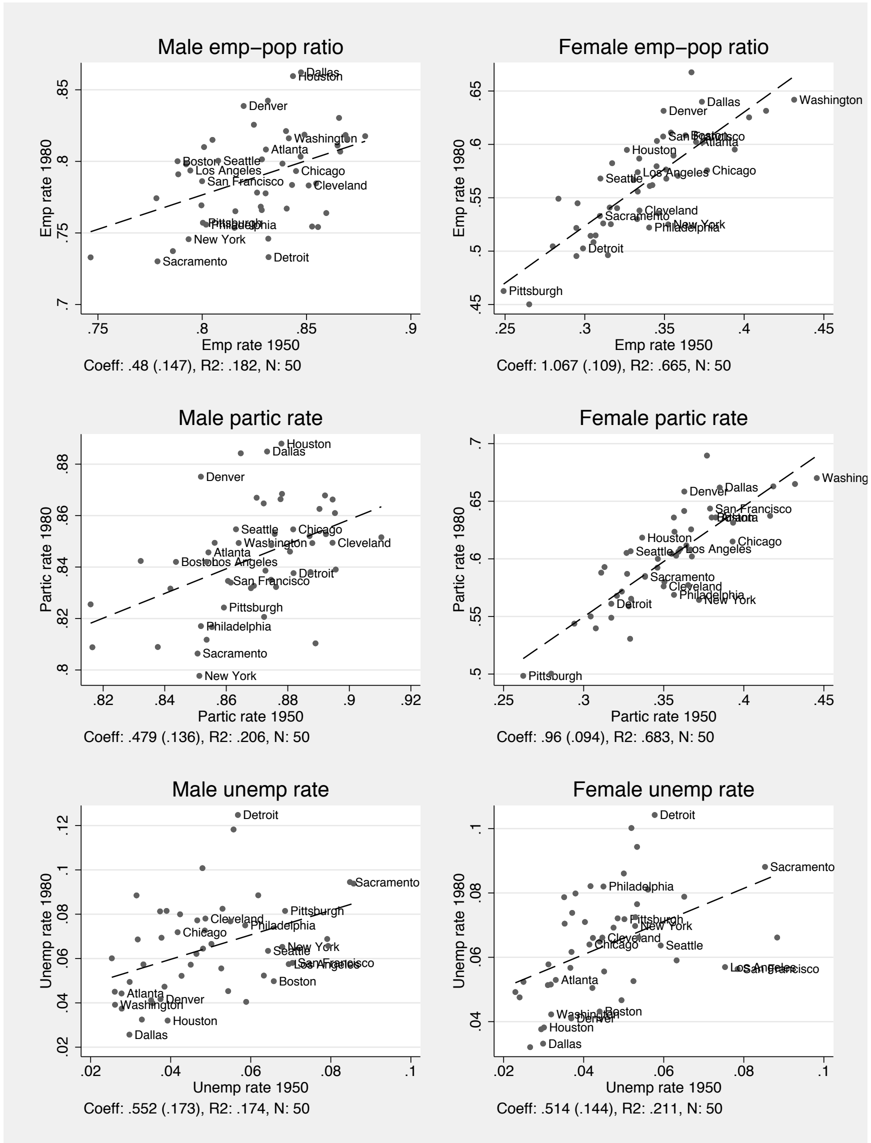

Figure A3: Persistence of joblessness across CZs: 1950-1980

Data-points denote commuting zones (CZs). Sample is restricted to 50 largest CZs in 1950, for individuals aged 16-64. "Participation rate" is ratio of labor force to population, and "unemployment rate" is ratio of unemployment to labor force.

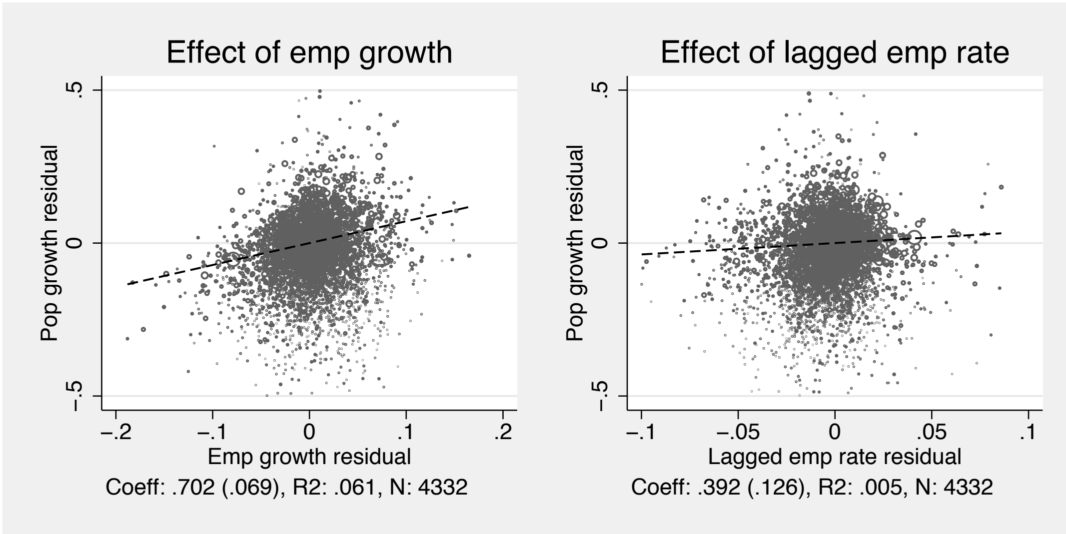

Figure A4: Graphical illustration of 2SLS coefficients

Following the Frisch-Waugh theorem, this figure graphically depicts the 2SLS coefficients of the basic specification (column 4, Table 2). The first panel illustrates the coefficient on employment growth, β_1; and the second panel illustrates the coefficient on the lagged employment rate, β_2. See Online Appendix G.6 for details on the methodology. Data-points denote CZ-year observations, with size corresponding to lagged population share. Standard errors for the best-fit slope are robust and clustered by CZ. Notice the standard errors of the best-fit slopes do not correspond to those in Table 2; this is because this naive estimator does not account for sampling error in the first stage. A small number of outlying data points (no more than 25 observations of minor CZs in each panel) have been excluded because of our choice of axis range.