CEP Discussion Paper No 1678
Revised May 2020
(Replaced February 2020 version)

The Contribution of Immigration to Local Labor Market Adjustment

Michael Amior
Abstract
Using decadal census data since 1960, I cannot reject the hypothesis that new immigrants crowd out existing residents from US commuting zones and states one-for-one. My estimate is precise and robust to numerous specifications, as well as accounting for local dynamics; and I show how it can be reconciled with apparently conflicting results in the literature. Exploiting my model’s structure, I attribute 30% of the observed effect to mismeasurement, specifically undercoverage of immigrants. Though labor demand does respond, population mobility accounts for 90% of local adjustment. These results have important implications for both structural and reduced form estimation of immigration effects.

Key words: Immigration, geographical mobility, local labor markets, employment
JEL Codes: J61; J64; R23

This paper was produced as part of the Centre’s Labour Markets Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Much of the content of this paper is taken from CEP Discussion Paper 1582 (2018): “The Contribution of Foreign Migration to Local Labor Market Adjustment”. I have split this discussion paper into two (Amior, 2020, and this paper) and included substantial new material in each. I am grateful to Christoph Albert, George Borjas, David Card, Christian Dustmann, Eric Gould, Ori Heffetz, Alan Manning, Guy Michaels, Giovanni Peri, Jonathan Portes and Jan Stuhler for helpful comments, as well as participants of the CEP, RES, OECD-CEPII “Immigration in OECD Countries”, GSE Summer Forum, IZA Annual Migration Meeting and CEPR-EBRD conferences, and seminars at IDC Herzliya, Bar Ilan, Hebrew University (Mount Scopus and Rehovot), Bank of Israel, Southampton and Tel Aviv. I thank Joe Gröver, Kin Koerber and Jordan Rappaport for sharing data.

Michael Amior, Hebrew University of Jerusalem and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© M. Amior, revised 2020.
1 Introduction

Recent years have seen growing interest in the geographical mobility of labor, as trade shocks and the collapse of manufacturing employment exacerbate regional inequalities (Autor, Dorn and Hanson, 2013; Charles, Hurst and Notowidigdo, 2016). An important tension in this discussion is the apparent inconsistency between migratory responses to different shocks. On the one hand, economists mostly agree that local population responds strongly (though not instantaneously) to labor demand (e.g. Blanchard and Katz, 1992; Monras, 2015). But, it is commonly believed that local supply shocks driven by immigration elicit little such response: see, most famously, Card (2001). As it happens, this latter result underpins much of the empirical immigration literature: many studies exploit local variation in the incidence of immigration to identify its effects (see Lewis and Peri, 2014, and Jaeger, Ruist and Stuhler, 2018, for excellent surveys); but this strategy is typically only feasible if mobility responds sluggishly (Borjas, Freeman and Katz, 1997).

This paper attempts to reconcile these results, using decadal US census data on 722 commuting zones (CZs) over 50 years. Like Amior and Manning (2018) and Amior (2020), I take a “semi-structural” approach to estimation which accounts for dynamic local adjustment. This allows me to study very general settings, without imposing heavy structural assumptions or relying on natural experiments (which restrict analysis to specific historic episodes). Identifying my model with shift-share instruments, I find that for each new foreign arrival to a CZ, 1.1 existing residents leave on net. This effect is remarkably precise and robust to numerous specification choices, and even the OLS estimate is very large (0.76). It is also educationally “balanced”: the college share of both foreign inflows and net outflows resemble the local population. And similarly to Dustmann, Schoenberg and Stuhler (2017), it is driven by a reduction in internal inflows rather than larger outflows (and hence my preference for the “crowd-out” terminology over the more typical “displacement”). I base my main estimates on CZs, but I also cannot reject perfect crowd-out across states.

Though one may expect a significant response, its sheer size is puzzling - for two reasons. First, it conflicts with much of the literature (as I describe below), and a key contribution of this paper will be to reconcile these results. Second, my estimates even lack internal consistency: I identify small but significant negative effects of foreign inflows on local employment rates, especially among the low educated. This suggests local adjustment is incomplete, which is difficult to reconcile with perfect crowd-out.

I attribute the latter inconsistency to mismeasurement - specifically, undercoverage of immigrants in the census (and the failure of census weights to account for it). The structure of my model allows me to identify the extent of undercoverage. Purging the associated bias reduces my crowd-out estimate from 1.1 to 0.8. I attribute the remaining 0.8 effect largely to the labor market and partly to native distaste for immigration. Though
labor demand does cushion the impact of immigration, I find that population mobility accounts for 90% of local labor market adjustment. This final point is consistent with the literature on local demand shocks (cited above), which typically finds local adjustment is effected almost entirely by internal mobility rather than employment.

Turning to the existing literature, I am not the first to identify large crowd-out (see Filer, 1992; Frey, 1995; 1996; Borjas, Freeman and Katz, 1997; Hatton and Tani, 2005; Borjas, 2006, 2014), though Wright, Ellis and Reibel (1997) dispute Frey’s methods, and Peri and Sparber (2011) and Card and Peri (2016) dispute Borjas’. Still, even Borjas’ (2006) estimates are smaller than mine: he finds that each immigrant crowds out 0.6 natives from US metro areas, and 0.3 from states. Monras (forthcoming) estimates large crowd-out following an unanticipated surge of Mexican migration, but finds much less over the decadal census intervals which I study. Applying a structural model to local wage data, Colas (2018) finds that crowd-out reaches 0.5 ten years after a one-off migration shock. Burstein et al. (2018) show that migrants crowd out natives from non-tradable jobs, though this is a within-CZ effect. Finally, Dustmann, Schoenberg and Stuhler (2017) find that Czech cross-border commuters crowded out German nationals one-for-one in local employment in the early 1990s, with a third of the effect (over three years) coming through internal migration.

Still, the US literature more typically reports small negative or even positive effects on native population: see Butcher and Card (1991), Wright, Ellis and Reibel (1997), Card and DiNardo (2000), Card (2001, 2005, 2009a), Card and Lewis (2007), Cortes (2008), Boustan, Fishback and Kantor (2010), Peri and Sparber (2011), Wozniak and Murray (2012), Hong and McLaren (2015), Edo and Rapoport (2017) and Piyapromdee (2017); and see Pischke and Velling (1997) on Germany. These results are commonly rationalized by an elastic local demand for labor (which absorbs the new immigrants), driven either by technological change or local consumption (Lewis, 2011; Dustmann and Glitz, 2015; Hong and McLaren, 2015). But, I find this accounts for only a small part of local adjustment.

So, how can these results be reconciled? While I pool 50 years of data, a number of papers rely on much shorter time horizons. But as Borjas, Freeman and Katz (1997) argue, such estimates may be sensitive to the particular pattern of omitted shocks in any given decade. Indeed, with no controls (except year effects), my estimates of crowd-out vary greatly by decade (consistent with Card, 2009a). However, once I control for initial conditions and observable supply/demand shocks, the estimates are much more consistent over time - and I cannot reject perfect crowd-out in any decade.

Still, while my main estimates rely on cross-CZ variation, much of the existing literature exploits variation across skill groups within locations (e.g. Card and DiNardo, 2000; Card, 2001, 2005; Borjas, 2006; Cortes, 2008; Monras, forthcoming): i.e. they study the effect of skill-specific immigration on local skill composition. But small composition
effects are not inconsistent with large crowd-out - for two reasons. First, composition effects reflect not only differential internal mobility, but also changes in the character of local birth cohorts. Indeed, I find cohort effects have historically offset the contribution of mobility (in the determination of local skill composition). And second, within-area estimates do not account for the labor market impact that new migrants exert outside their own skill group (Card, 2001; Dustmann, Schoenberg and Stuhler, 2016). This can be seen in the remarkable sensitivity of my within-area estimates to the delineation of these groups. This is not to say that within-area estimates are uninteresting: the impact on local skill composition is certainly an important question. My point is merely that they do not identify aggregate-level crowd-out.

Finally, a key contribution of this study is to highlight the importance of census undercoverage of immigrants - both for the question of crowd-out, as well as for the estimation of immigration effects more broadly. Not only will reduced form estimates be upward biased, we may also be severely overstating the skill composition of migrants - which is the key determinant of their impact in structural models (e.g. Borjas, 2003; Ottaviano and Peri, 2012). Earlier studies have relied on external evidence to assess undercoverage\(^1\), but these are not necessarily informative about undercoverage in places with larger foreign inflows (which is what matters for my application). This paper offers a strategy to estimate the bias specifically along this margin, using census data alone. Based on my model’s assumptions, I show that controlling for employment growth (in the crowd-out equation) will eliminate most of the bias: the bias in employment effectively partials out the bias in foreign inflows. An estimate of the employment elasticity is then sufficient to identify the undercount - which I place at about 30%. Variation across census years is consistent with these claims: I estimate double the crowd-out before 1980, when coverage was known to poorer. But as the model predicts, once I conditional on employment, crowd-out does not vary systematically over time.

In Section 2, I set out my model and derive my empirical specifications. Section 3 describes the data; and Section 4 presents my main estimates of crowd-out, together with numerous robustness checks. In Section 5, I estimate the impact of immigration on employment rates, wages and housing costs, and study variation by education. Section 6 assesses the implications of undercoverage for my results. Finally, Section 7 explores alternative estimates of crowd-out which exploit skill variation within areas, based on a modified version of the model.

\(^1\)See e.g. Warren and Passel (1987), Borjas, Freeman and Lang (1991), Van Hook and Bean (1998), and Marcelli and Ong (2002), who focus on undocumented migrants.
2 Model of local crowding out

2.1 Overview

Like Amior and Manning (2018), my theoretical framework has two components: (i) a Roback (1982) style model for local labor market equilibrium, conditional on population, and (ii) dynamic equations describing local population adjustment. And like Amior (2020), I distinguish between the contributions of foreign and internal mobility to population change.

The point of departure here is my choice of outcome. In Amior (2020), I study the elasticity of new immigrants’ location choices to local (demand-driven) employment shocks. A key insight is that the (large) contribution of new immigrants to local population adjustment crowds out that of internal mobility. As I show there, this effect can be summarized by the impact of local foreign inflows on net internal outflows, conditional on local employment growth. I call this “conditional” crowd-out.

In contrast, the focus of this paper is local adjustment to immigration shocks. To this end, I study “unconditional” crowd-out (i.e. without the employment control), which is the more traditional concern of the migration literature. Crucially, this unconditional effect is moderated by changes in labor demand, which may be a key margin of adjustment to immigration. As I show below, the contribution of labor demand to local adjustment can in principle be identified by the difference between the conditional and unconditional crowd-out effects (though in the presence of undercoverage, identification requires additional restrictions: see Section 6).

To ease the exposition, I assume native and migrant labor are identical and perfect substitutes. This may be justified by the existing literature, which typically finds they are close substitutes. But crucially, I do not impose this assumption on my empirical specifications. Rather, I consider its validity ex post - following a strategy proposed by Beaudry, Green and Sand (2012). In practice, it turns out that immigration does exert similar effects on native and migrant employment rates, which is consistent with the model’s assumptions. In a similar spirit, I do not account for skill heterogeneity here; but see Appendix A.3 for an exposition which does, and see Section 7 for associated empirical estimates.

In what follows, I first derive an estimable equation for conditional crowd-out, following identical steps to Amior (2020). And new to this paper, I then solve for local employment growth and derive my unconditional crowd-out specification.

\(^2\)Ottaviano and Peri (2012) estimate an elasticity of substitution between natives and migrants of 20 within skill cells. Card (2009b) finds even larger numbers, and Borjas, Grogger and Hanson (2012) and Ruist (2013) cannot reject perfect substitutability. At the aggregate level (which is the relevant point here), differences in native and migrant skill composition will also matter. Natives and migrants do have similar college shares in the US (Card, 2009b), though undercoverage may affect this conclusion (see below).
2.2 Local equilibrium conditional on population

There is a single traded good, with price $P$, and a non-traded good (housing) with price $P^h_r$ in area $r$. If preferences are homothetic, I can define a unique local price index:

$$P_r = Q(P, P^h_r)$$

Unlike Roback (1982), suppose labor supply is elastic to the real consumption wage:

$$n_r = l_r + \epsilon^s (w_r - p_r) + z^s_r$$

where lower case variables denote logs: $n_r$ is local employment, $l_r$ is population, $w_r$ is the nominal wage, and $z^s_r$ is an area $r$ supply shifter. I write labor demand as:

$$n_r = -\epsilon^d (w_r - p) + z^d_r$$

where $z^d_r$ is a demand shifter. In Appendix A.1, I set out equations for housing supply and demand. Together with (1)-(3), I can then solve for employment, wages and prices in terms of population $l_r$ alone.

I write indirect utility as:

$$v_r = w_r - p_r + a_r$$

where $w_r - p_r$ is the real consumption wage, and $a_r$ represents local amenities. Crucially, I can now substitute the labor supply equation (2) for the real wage in (4). This allows me to use the local employment rate as a sufficient statistic for labor market conditions, for given supply and amenity effects:

$$v_r = \frac{1}{\epsilon^s} (n_r - l_r - z^s_r) + a_r$$

Amior and Manning (2018) show this “sufficient statistic” result is robust to the inclusion of multiple traded and non-traded sectors (where migrants will generate their own local demand), agglomeration, endogenous amenities and labor market frictions; and Amior (2020) shows it is robust to heterogeneity in the consumer price indices of natives and migrants (as in Albert and Monras, 2018). Another concern is heterogeneous preferences for leisure: like Amior (2020), I attempt to address this empirically by adjusting employment rates for demographic composition.

2.3 Population dynamics

Long run equilibrium is characterized by spatially invariant utility $v_r$, which determines population $l_r$ in every area $r$. But I allow for sluggish population adjustment to this equilibrium, and I distinguish between the respective contributions of internal and foreign
migration:

\[ dl_r = \lambda^I_r + \lambda^F_r \]  \hspace{1cm} (6)

where \( \lambda^I_r \) is the instantaneous rate of net internal inflows (i.e. from elsewhere in the US) to \( r \), and \( \lambda^F_r \) is the foreign inflow rate, relative to local population. (6) does not account for the role of emigration, but I consider this later when interpreting the estimates. Amior (2020) shows how \( \lambda^I_r \) can be expressed as a linear function of utility \( v_r \), using a logit model of residential choice:

\[ \lambda^I_r = \gamma (n_r - l_r - z^s_r + \epsilon^s a_r) \]  \hspace{1cm} (7)

where \( \gamma \geq 0 \) is the elasticity of net internal flows. I have not included a national intercept, but the supply effect \( z^s_r \) may be redefined to include one. Mobility decisions in (7) depend only on current outcomes, so workers are implicitly myopic. However, Amior and Manning (2018) show that a model with forward-looking agents yields an equivalent expression, where the \( \gamma \) parameter depends on both mobility and the local persistence of shocks.

In Amior (2020), I also set out a parallel expression for foreign inflows, \( \lambda^F_r \). These depend partly on what I call the “foreign supply”, i.e. the foreign inflow in the absence of local utility differentials. Crucially, the foreign supply varies regionally, partly due to the size of migrant enclaves (which offer e.g. language or job access benefits). Since it does not enter internal mobility (7) directly, this yields an exclusion restriction which motivates the classic “enclave shift-share” instrument of Altonji and Card (1991) and Card (2001). But, I do not model \( \lambda^F_r \) formally in this paper: my interest here is the evolution of internal flows \( \lambda^I_r \) for a given \( \lambda^F_r \).

### 2.4 Conditional crowd-out

To study conditional crowd-out, I first substitute (7) for internal flows \( \lambda^I_r \) in (6):

\[ dl_r = \lambda^F_r + \gamma (n_r - l_r - z^s_r + \epsilon^s a_r) \]  \hspace{1cm} (8)

For estimation, I require a discrete-time expression. Suppose the supply effect \( z^s_r \), amenity \( a_r \) and employment \( n_r \) change at a constant rate within discrete intervals. For small employment shocks, Amior (2020) shows that:

\[ \lambda^I_{rt} \approx \left( 1 - \frac{1 - e^{-\gamma}}{\gamma} \right) (\Delta n_{rt} - \lambda^I_{rt-1} - \Delta z^s_{rt} + \epsilon^s \Delta a_{rt}) + (1 - e^{-\gamma}) (n_{rt-1} - l_{rt-1} - z^s_{rt-1} + \epsilon^s a_{rt-1}) \]  \hspace{1cm} (9)

where \( \lambda^I_{rt} \equiv \int_{t-1}^{t} \lambda^I_r (\tau) \, d\tau \) is the discrete-time internal response over the unit interval, and \( \lambda^F_{rt} \equiv \int_{t-1}^{t} \lambda^F_r (\tau) \, d\tau \) is the discrete foreign inflow. Conditional on employment growth \( \Delta n_{rt} \) and the initial conditions (the lagged employment rate and supply shocks), (9)
describes the impact of foreign inflows $\lambda^F_{rt}$ on net internal inflows $\lambda^I_{rt}$. This “conditional crowd-out” increases from 0 to -1 as the internal population elasticity $\gamma$ increases from 0 to $\infty$.

Equation (9) is implicitly an error correction model in population and employment. This can be seen by adding foreign inflows $\lambda^F_{rt}$ to both sides. The dependent variable then becomes the log population change $\Delta l_{rt}$, which is a linear function of the log employment change $\Delta n_{rt}$ and the lagged log employment rate $n_{rt-1} - l_{rt-1}$ (i.e. the initial steady-state deviation). As $\gamma$ becomes large, we approach full adjustment over decadal intervals: contemporaneous employment growth manifests one-for-one in population; and the coefficient on $n_{rt-1} - l_{rt-1}$ goes to -1, so any initial employment rate deviations are eliminated by population change in the subsequent interval.

Crucially, the coefficients depend only on the internal elasticity, $\gamma$. Intuitively, they represent unmediated mobility responses to changes in local utility $v_{rt}$, as summarized by the employment rate. This is because both $\lambda^F_{rt}$ and $\Delta n_{rt}$, conditional on the other, mechanically feed one-for-one into local employment rates. For the same reason, the coefficients on $\Delta n_{rt}$ and $\lambda^F_{rt}$ are identical (up to their sign). In practice, estimates of the $\lambda^F_{rt}$ coefficient will exceed that of $\Delta n_{rt}$ if there is native distaste for immigration (as in Card, Dustmann and Preston, 2012; Saiz and Wachter, 2011; Fernandez-Huertas Moraga, Ferrer-i Carbonell and Saiz, 2017): effectively, this would generate a negative correlation between $\lambda^F_{rt}$ and an (unobserved) amenity change $\Delta a_{rt}$.

### 2.5 Unconditional crowd-out

Equation (9) does not describe the unconditional impact of foreign inflows, as I am controlling for employment growth $\Delta n_{rt}$; and labor demand may be a key margin of adjustment. To derive the unconditional effect (my point of departure from Amior, 2020), I now reduce $\Delta n_{rt}$ to its determinants.

This requires a specification of the housing market, as local prices shift labor supply (2) but not demand (3). Assuming individuals spend a fixed share of their income on housing (i.e. Cobb-Douglas utility), Appendix A.1 shows that changes in local prices $p_r$ can be specified as:

$$
\Delta (p_{rt} - p_t) = \frac{1}{\kappa} \left[ \frac{1}{\epsilon_S} (\Delta n_{rt} - \Delta l_{rt} - \Delta z^S_{rt}) + \Delta n_{rt} \right]
$$

where $\kappa > 0$ goes to infinity with the elasticity of housing supply.\(^3\) Together with the

\(^3\)Specifically, $\kappa \equiv \frac{1 - \nu + \epsilon^S}{\nu}$, where $\nu$ is the (fixed) share of income spent on housing, and $\epsilon^S$ is the housing supply elasticity.
labor supply and demand equations, this implies:

$$\Delta n_{rt} = \eta (\Delta l_{rt} + \Delta z^s_{rt}) + (1 - \eta) \frac{K}{K + e^d} \Delta z^d_{rt}$$  (11)

where

$$\eta \equiv 1 - \left(1 + \frac{\kappa + 1}{\kappa + e^d} \cdot \frac{e^d}{e^s}\right)^{-1}$$  (12)

is the elasticity of employment with respect to population. This must lie between 0 and 1 if labor demand slopes down (i.e. $e^d > 0$) and labor supply slopes up ($e^s > 0$). As I show in Appendix A.2, replacing $\Delta n_{rt}$ in equation (9) with (11) yields the unconditional crowding out equation:

$$\lambda^I_{rt} = \frac{(1 - \eta) \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)}{1 - \eta \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)} \left(\frac{\kappa}{\kappa + e^d} \Delta z^d_{rt} - \lambda^F_{rt} - \Delta z^s_{rt} + \frac{e^s}{1 - \eta} \Delta a_{rt}\right)$$  (13)

As before, the crowding out effect of $\lambda^F_{rt}$ goes to -1 as internal flows become perfectly elastic ($\gamma \to \infty$). In this scenario, the employment elasticity $\eta$ has no traction, because population adjusts so swiftly.

But for finite $\gamma$, the impact of $\lambda^F_{rt}$ is now moderated by an expansion of labor demand (driven, in a more complete model, by capital mobility) and possibly housing supply. To see this, notice the effect of $\lambda^F_{rt}$ in (13) is identical to (9) if demand is inelastic, i.e. $e^d = 0$ (in which case, $\eta = 0$). But as $e^d$ grows relative to the supply elasticity $e^s$, the employment elasticity $\eta$ increases; and the effect of $\lambda^F_{rt}$ contracts. In the limit, as $\eta \to 1$, adjustment is fully effected by changes in local employment rather than population (i.e. no crowd-out).

### 2.6 Identification of local adjustment process

Following an immigration shock, both population mobility and labor demand may assist local adjustment. In principle, their respective contributions can be identified by comparing the conditional and unconditional specifications, (9) and (13). All else equal, the immigration shock reduces the employment rate (which summarizes local welfare) one-for-one. Holding employment fixed, the unmediated response of internal mobility is

---

4This moderating effect is contingent on labor demand sloping down (i.e. $e^d > 0$). Otherwise, $\eta$ may turn negative, in which case population growth would perversely reduce employment in comparative statics (though equilibrium would be unstable). Of course, agglomeration is crucial to the existence of cities. But if these returns never peter out, population adjustment cannot return us to equilibrium.

5The employment elasticity $\eta$ (and therefore the crowding out effect) is increasing in $\kappa$ (and thus in the elasticity of housing supply) if $e^d > 1$. This condition ensures that immigration causes the local wage bill (and therefore housing demand) to grow.
described by the coefficient on $\lambda_{rt}^F$ in (9). And the contribution of labor demand (to employment rate adjustment) is then identified by the difference between the $\lambda_{rt}^F$ coefficients in (9) and (13). To study the dynamics, one can perform the same analysis on the response to the lagged employment rate.

In practice though, I will argue this simple procedure is infeasible because of undercoverage of migrants in the data. I cannot identify both the employment response and the extent of undercoverage using (9) and (13) alone. However, as I discuss in Section 6, identification is feasible if I additionally exploit the restrictions implied by the employment equation (11).

## Data

I study decadal census observations between 1960 and 2010 across 722 CZs. Where possible, I use published county-level aggregates from NHGIS (Manson et al., 2017). And where necessary, I supplement this with census microdata and (for 2010) pooled American Community Survey (ACS) samples of 2009-11, taken from IPUMS (Ruggles et al., 2017). I describe the data more fully in Appendix B, but I summarize the main points here.

The first challenge is to disaggregate log population changes $\Delta l_{rt}$ into the contributions of foreign and internal mobility, $\lambda_{rt}^F$ and $\lambda_{rt}^I$. Since I only have discrete-time observations, I cannot precisely identify these components. But Amior (2020) shows they can be closely approximated by:

\[
\lambda_{rt}^F \approx \log \left( \frac{L_{rt-1} + L_{rt}^F}{L_{rt-1}} \right) 
\]

\[
\lambda_{rt}^I \approx \log \left( \frac{L_{rt} - L_{rt}^F}{L_{rt-1}} \right) 
\]

where $L_{rt}^F$ is the local foreign-born population aged 16-64 at time $t$ who arrived in the US in the previous ten years (i.e. between $t-1$ and $t$). Notice I have constructed $\lambda_{rt}^I$ as a residual population change: this accounts for the entire contribution of natives and “old” migrants (who arrived in the US before $t-1$), part of which is driven by “natural” growth and emigration (especially of the foreign-born). It is not possible to identify emigration in this data; but in an effort to exclude it, I also study the native contribution to $\lambda_{rt}^I$:

\[
\lambda_{rt}^{I,N} \approx \log \left( \frac{L_{rt-1} + \Delta L_{rt}^N}{L_{rt-1}} \right) 
\]

where $L_{rt}^N$ is the local native stock of 16-64s. An alternative approach is to take first order approximations, i.e. $\lambda_{rt}^F \approx \frac{L_{rt}^F}{L_{rt-1}}$ and $\lambda_{rt}^I \approx \frac{\Delta L_{rt} - L_{rt}^F}{L_{rt-1}}$. But as Amior (2020) shows,
(14) and (15) offer more precision.\footnote{While $\frac{L_F^t}{L_{t-1}}$ and $\frac{\Delta L_{t-1} - L_F^t}{L_{t-1}}$ converge to the true $\lambda^F$ and $\lambda^I$ as they individually become small, convergence of (14) and (15) merely requires their product become small.}

My employment sample also consists of 16-64s. Using the microdata, I adjust all employment variables for local demographic composition, controlling for age, education, ethnicity, gender, foreign-born status, and years in the US, together with a rich set of interactions. In terms of the model, this purges any local variation in the supply shocks $z_{rt}$ which is due to observable composition. The aim is to reduce the demands on the exclusion restrictions. See Appendix B.2 for methodological details.

I identify changes in local demand using Bartik (1991) industry shift-shares, and foreign inflows using the enclave shift-share of Altonji and Card (1991) and Card (2001). I will describe the exclusion restrictions when I set out the estimating equations. The Bartik predicts change in local employment, based on initial industrial composition and national-level changes by industry:

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

where $\phi_{rt}^i$ is the fraction of area $r$ individuals working in a 2-digit industry $i$ (57 categories) in $t - 1$; and $\Delta n_{i(-r)t}$ is industry $i$’s national log employment change, excluding area $r$.\footnote{Goldsmith-Pinkham, Sorkin and Swift (2018) recommend this exclusion to address possible endogeneity to local supply.}

The enclave shift-share allocates new migrants to areas proportionally to the initial size of co-patriot communities. This can be interpreted as a proxy for the “foreign supply” (in the absence of local utility differentials) which I describe in Section 2.3. Using the functional form of (14):

$$m_{rt} = \log \left( \frac{L_{t-1} + \sum_o \phi_{rt}^o L_{o(-r)t}^E}{L_{t-1}} \right)$$

where $\phi_{rt}^o$ is the fraction of origin $o$ migrants (77 countries) residing in area $r$ at time $t - 1$, and $L_{o(-r)t}^E$ is the stock of new origin $o$ migrants (excluding area $r$ residents) who arrived in the US between $t - 1$ and $t$. I construct both shift-share instruments using census microdata: see Appendix B.3.

Throughout, I control for a set of observable amenities: (i) presence of coastline\footnote{Coastline data is borrowed from Rappaport and Sachs (2003).}; (ii) climate, specifically maximum January/July temperatures and mean July relative humidity; (iii) log population density in 1900; (iv) an index of CZ isolation (log distance to closest CZ, measured between population-weighted centroids). To allow for time-varying effects, I interact each with a full set of year dummies. I do not include time-varying amenities (like crime), as these may be endogenous to the labor market (Diamond, 2001).
Thus, the estimated effects of local shocks will account for both their direct (labor market) effect and any indirect effects (via amenity changes).

Table 1 offers descriptive statistics for key variables. Since 1960, the mean foreign inflow \( \lambda_{Ft} \) has grown from 0.02 to 0.06. Unsurprisingly perhaps, the distribution of both \( \lambda_{Ft} \) and the enclave shift-share \( m_{rt} \) are very skewed. But I show my results are robust to omitting outlying observations.

4 Estimates of crowding out

4.1 Estimating equations

In line with (9), I begin by estimating the conditional crowding out equation of Amior (2020):

\[
\lambda_{rt}^c = \delta^c_0 + \delta^c_1 \lambda_{rt}^F + \delta^c_2 \Delta n_{rt} + \delta^c_3 (n_{rt-1} - l_{rt-1}) + A_r \delta^c_A r + \varepsilon_{rt}^c
\]  

(19)

where \( \delta^c_1 \) is the crowding-out effect, controlling for (composition-adjusted) employment growth, \( \Delta n_{rt} \). The lagged (composition-adjusted) employment rate, \( n_{rt-1} - l_{rt-1} \), summarizes the initial conditions: i.e. any lingering impact of historical demand or migration shocks. I account for year effects in \( \delta^c_0 \), and the \( A_r \) vector contains amenity effects (i.e. observable components of \( \Delta a_{rt} \) and \( a_{rt-1} \) in (9)), which I interact with year effects (in \( \delta^c_A r \)). Any unobserved amenity/supply effects fall into the error, \( \varepsilon_{rt}^c \). Based on (9), this equation should contain no omitted demand shocks. This is a consequence of the sufficient statistic result: the employment variables fully summarize local welfare, conditional only on supply effects.

OLS estimates of (19) are not credible: foreign inflows \( \lambda_{Ft} \), employment growth \( \Delta n_{rt} \) and the lagged employment rate are endogenous to omitted amenity/supply effects (both current and lagged). Three instruments (which exclude these) are required: I use the enclave shift-share \( m_{rt} \) for \( \lambda_{Ft} \), the current Bartik \( b_{rt} \) for \( \Delta n_{rt} \), and the lagged Bartik \( b_{rt-1} \) for \( n_{rt-1} - l_{rt-1} \). In principle, the initial employment rate will depend on a distributed lag of Bartiks; but in practice, the first lag offers sufficient power. As with all shift-share instruments, identification may be motivated by exogeneity of the initial local migrant/industry shares to the omitted shocks (Goldsmith-Pinkham, Sorkin and Swift, 2018), or by random aggregate-level shocks to migrant inflows or industries (Borusyak, Hull and Jaravel, 2018).

Next, I turn to the unconditional estimates of crowding out, which are the focus of this paper. Whereas the conditional specification identifies unmediated mobility response to changes in local welfare (given the employment control), the unconditional effect of \( \lambda_{Ft} \)
accounts for the moderating effects of labor demand. I base my empirical specification on (13):

$$\lambda_{rt} = \delta^u_0 + \delta^v_1 \lambda^F_{rt} + \delta^w_2 b_{rt} + \delta^u_3 (n_{rt-1} - l_{rt-1}) + A_r \delta^u_{At} + \varepsilon^u_{rt}$$

(20)

where I have replaced employment growth $\Delta n_{rt}$ with its Bartik instrument $b_{rt}$, to proxy for local demand. There are now just two endogenous variables ($\lambda^F_{rt}$ and $n_{rt-1} - l_{rt-1}$), so I use two instruments: the enclave shift-share $m_{rt}$ and the lagged Bartik $b_{rt-1}$. Looking at (13), the error $\varepsilon^u_{rt}$ now contains only new unobserved innovations in demand (historical demand and migration shocks are summarized by the initial employment rate), as well as unobserved supply/amenity shocks. As a result, the threats to identification are much weaker than in traditional specifications which neglect dynamics.

These dynamics pose two particular challenges. First, foreign inflows are locally very persistent (Jaeger, Ruist and Stuhler, 2018); so a naive regression of internal population flows on contemporaneous immigration may pick up a sluggish response to historical foreign inflows. Second, the location of past migrant enclaves (which underly the $m_{rt}$ instrument) depends on past local demand shocks, which are themselves highly predictive of current shocks (Amior, 2020). But given the sufficient statistic result, the initial employment rate ($n_{rt-1} - l_{rt-1}$) can partial out the entire history of immigration and labor demand shocks. This offers a theoretical basis for Pischke and Velling’s (1997) suggestion to control for the initial unemployment rate. Below, I probe the effectiveness of this approach.

### 4.2 Empirical estimates

Table 2 presents first stage estimates for the endogenous variables in (19) and (20). I weight observations by lagged local population share and cluster errors by state. Each instrument has large positive effects on its corresponding endogenous variable, with large Sanderson-Windmeijer (2016) F-stats (accounting for multiple endogenous variables) reported in Table 3.

Table 3 sets out OLS and IV estimates of (19) and (20). I focus here on the aggregate-level effects, but I consider heterogeneity by education in Section 5.2. The first two columns estimate the conditional crowd-out equation (19): these replicate results from Table 8 of Amior (2020). The OLS estimate of $\delta^c_1$ (i.e. the response to $\lambda^F_{rt}$) is just under -0.9, and the IV estimate just over (in magnitude). Their similarity can be attributed to the employment control: in principle, there are no omitted demand shocks; so any OLS-IV disparity must be due to supply shocks alone. I also estimate large responses to employment growth and the lagged employment rate.
As noted above, the model in (9) predicts equal coefficients on the foreign inflow $\lambda^F_{rt}$ and employment growth $\Delta n_{rt}$ (i.e. $\delta_1 = -\delta_2$). Intuitively, a given change in $\lambda^F_{rt}$ or $\Delta n_{rt}$ has identical implications for the local employment rate (and hence local welfare) and should trigger identical mobility responses. However, my IV estimate of $\delta_1$ significantly exceeds $\delta_2$ in magnitude. Amior (2020) argues this may reflect native distaste for immigration.

Moving beyond Amior (2020), I now turn to unconditional crowd-out. While the OLS estimate of $\delta^u_1$ in (20) is already very large (-0.76), the IV estimate in column 4 is even larger (-1.1). The IV estimate is also larger than its conditional counterpart in column 2. This is puzzling: the employment response to immigration should moderate the impact on welfare and internal mobility (assuming labor demand slopes down); and this should be reflected in a smaller unconditional effect.

One may be concerned the estimates are distorted by local dynamics. But, the evidence suggests the lagged employment rate (which is intended to summarize initial conditions) is performing its function well. First, following the recommendation of Jaeger, Ruist and Stuhler (2018), I control in column 5 for the lagged enclave shift-share, $m_{rt-1}$. Notice that $m_{rt-1}$ negatively affects the initial employment rate in the first stage (column 7 of Table 2); but reassuringly, it has no effect in the second stage (column 5 of Table 3). This is consistent with the employment rate summarizing the full history of shocks. However, once I drop the employment rate in column 6 (and replace it with its lagged Bartik instrument), $m_{rt-1}$ now picks up much of the negative effect. That is, though internal flows do respond sluggishly, the initial employment rate accounts successfully for these dynamics. Notice also the $\lambda^F_{rt}$ coefficient in column 6 is now smaller: this likely reflects a positive correlation between the instrument $m_{rt}$ and omitted historical demand shocks (which the employment rate partials out in column 5).

Peri (2016) and Goldsmith-Pinkham, Sorkin and Swift (2018) recommend testing explicitly for pre-trends. In column 7, I replace the dependent variable with its lag, $\lambda^I_{rt-1}$. Reassuringly, the current foreign inflow $\lambda^F_{rt}$ and Bartik $b_{rt}$ have no significant effect on $\lambda^I_{rt-1}$; instead, the impact is fully absorbed by the lagged enclave shift-share and Bartik. Thus, my data can tease apart the effects of current and historical shocks.

Famously, Borjas (2006) finds less crowd-out across states than metro areas. But in column 8, I cannot reject a $\delta^u_1$ of -1 using state-level data. The state data does not offer sufficient variation to identify the lagged employment rate (using the lagged Bartik); so instead, I use the specification from column 6. Crowd-out is in fact larger than in column 6, but the standard errors are also larger.

Finally, Appendix C shows the effect is entirely driven by reductions of migratory inflows to affected CZs, rather than increases in outflows - based on census respondents’ reported place of residence five years previously. This is consistent with Coen-Pirani (2010), Monras (2015), Dustmann, Schoenberg and Stuhler (2017) and Amior and Man-
ning (2018), who report a disproportionate role for inflows in driving local population adjustment.

### 4.3 Robustness of unconditional crowd-out estimates

Table 4 assesses the sensitivity of my unconditional IV estimates (column 4 of Table 3) to the choice of controls (as Goldsmith-Pinkham, Sorkin and Swift, 2018, recommend) and decadal sample. With no controls (except year effects), my $\delta_u$ estimate varies greatly by decade: I find little crowd-out before 1990, but much more after. This reflects Card’s (2009a) findings, as well as the concerns of Borjas, Freeman and Katz (1997) about the instability of spatial correlations. The average effect (column 6) increases from -0.53 to -0.75 when I control for the current Bartik and initial employment rate. And after including amenities (especially climate, a key determinant of regional migration: see Rappaport, 2007), I cannot statistically reject (at least) one-for-one crowd-out in any decade: see the penultimate row. Intuitively, both natives and migrants are attracted to places with strong labor market conditions and pleasant climate, so omitting these should bias my $\delta_u^0$ estimate in a positive direction (i.e. towards zero). And as Borjas, Freeman and Katz (1997) emphasize, the extent of this bias will depend on the peculiarities of each individual decade.

In the final row, I replace the dependent variable $\lambda_{rt}$ with the native contribution $\lambda^{L,N}_{rt}$ to local population (i.e. excluding old migrants’ contribution). Compared to the previous row, column 6 shows two thirds of the average $\delta_u$ effect is driven by natives rather than old migrants. But this overlooks some important heterogeneity: exceptionally, in the 2000s, old migrants account for the entire effect. One possible explanation is large return migration to Mexico in the 2000s (see Hanson, Liu and McIntosh, 2017), driven in part by the construction bust and recession.

In Appendix Figure A1, I plot my $\delta_u$ estimates graphically, conditional on the covariates. The pictures show the effects are not driven by outliers.

In Appendix D.2 (and Appendix Table A2), I show the crowding out effect is robust to numerous specification changes. I begin by considering the effect of weighting. The population weights applied in Table 3 ensure my estimates are driven mainly by variation across larger CZs. But dropping these weights makes little difference, as does dropping CZs with fewer than 100,000 individuals in 1960. Given the skew in the spatial distribution of foreign inflows (Table 1), one may be concerned the estimates are driven by CZs with unusually large inflows. But, excluding observations with enclave shift-share $m_{rt}$ above 0.1 (the maximum is 0.29) makes little difference: this is consistent with the patterns in Appendix Figure A1.
My specification of $\lambda_{rt}^I$ and $\lambda_{rt}^F$ is almost identical to Card and DiNardo (2000) and Card (2001), as recommended by Peri and Sparber (2011) and Card and Peri (2016). While they regress $\frac{\Delta L_{rt}-L_{rt-1}^F}{L_{rt-1}}$ on $\frac{L_{rt}^F}{L_{rt-1}}$, I am regressing $\log\left(\frac{L_{rt}^F-L_{rt-1}^F}{L_{rt-1}^F}\right)$ on $\log\left(\frac{L_{rt}^F+L_{rt-1}^F}{L_{rt-1}^F}\right)$, in line with my model.\textsuperscript{9} Appendix Table A2 shows this makes little difference to my $\delta_1^u$ estimate. Also, basing the enclave shift-share in (18) on 1960 origin shares in all decades (following the example of Hunt, 2017), rather than using lagged-once shares, makes little difference.

My results are also robust to an alternative specification recommended by Wozniak and Murray (2012), which casts the key variables in levels, i.e. regressing $\left(\Delta L_{rt} - L_{rt}^F\right)$ on $L_{rt}^F$, without normalizing by initial population. As Wright, Ellis and Reibel (1997) note, local population may be an important omitted variable in this specification; but in line with Wozniak and Murray, I address this by controlling for local fixed effects.\textsuperscript{10}

And even specifying the variables as I do in Table 3, I cannot reject one-for-one crowd-out when I control for CZ fixed effects. This approach is similar to the double differencing methodology of Borjas, Freeman and Katz (1997) and is recommended by Hong and McLaren (2015). The idea is to pick up time-invariant local trends in omitted supply or demand. The estimates do become less stable: this is a demanding specification, given large persistence in the enclave shift-share; and as Aydemir and Borjas (2011) note, measurement error may be a greater challenge in the presence of fixed effects. But precision does improve when I replace the lagged employment rate control with lagged Bartik and enclave shift-shares.

5 Local labor market outcomes and heterogeneity

5.1 Aggregate-level labor market effects

My result of perfect unconditional crowd-out appears to imply full local adjustment to immigration. But despite this, I find small adverse effects of foreign inflows on local employment rates, consistent with Smith (2012), Edo and Rapoport (2017), Gould (2019) and Monras (forthcoming).

In Table 5, I re-estimate (20) using the same instruments as before, but replacing the dependent variable with changes in (composition-adjusted) employment rates. In column 1, the elasticity of the native employment rate to foreign inflows is -0.21. The coefficient of -0.41 on the lagged employment rate suggests the effect is largely dissipated within

\textsuperscript{9}My $\lambda_{rt}^F$ specification in (14) shares with Peri and Sparber (2011) and Card and Peri (2016) the advantage of depending only on new foreign inflows - and not on changes in existing US residents (which might otherwise introduce a spurious correlation with $\lambda_{rt}^I$).

\textsuperscript{10}This approach can also address concerns that the Table 3 estimates are conflated with spurious correlation in local population, which appears in the denominator of both the dependent variable and regressor of interest. See Clemens and Hunt (2019).
two decades.

Again, my specification successfully disentangles the impact of current and historical shocks. As in Table 3, the lagged enclave shift-share $m_{rt-1}$ in column 2 makes little difference, which suggests the lagged employment rate is indeed controlling for initial conditions. But once I drop the employment rate in column 3, $m_{rt-1}$ now takes a positive effect (reflecting the recovery following the initial shock); and as before, the $\lambda^F_{rt}$ coefficient becomes more negative (which likely reflects omitted demand shocks, correlated with the enclave instrument). Column 4 replaces the dependent variable with its lag; reassuringly, as in Table 3, $m_{rt-1}$ picks up the entire (negative) effect on the lagged dependent, and the current inflow $\lambda^F_{rt}$ becomes insignificant.

Column 5 estimates my preferred specification (column 1) for the migrant employment rate. The effect is similar to natives, which suggests there may be no great loss (in this context) from treating natives and migrants as perfect substitutes at the aggregate level - as I do in my model.

In principle, lower employment rates should be reflected in lower real consumption wages - based on the labor supply relationship in (2). Unfortunately, local wage deflators are notoriously difficult to construct (and rely on strong theoretical assumptions), especially for the detailed geographies and long time horizons I study: see Koo, Phillips and Sigalla (2000), Albouy (2008) and Phillips and Daly (2010). Still, I can at least estimate the effects on nominal wages and housing costs separately.

I use mean residualized wages, housing rents and prices, purged of observable demographics and housing characteristics respectively: see Appendix B.4 for details. While I find no impact on native wages, this may be difficult to interpret in the context of declining employment rates, if it is the lowest paid natives who are leaving employment (Card, 2001; Bratsberg and Raaum, 2012). On the other hand, housing costs do increase (see also Saiz, 2007), though the effect is statistically insignificant.

### 5.2 Education-specific outcomes

I now study heterogeneity by education in the impact of foreign inflows. I replace the dependent variable of (20) with various outcomes $\Delta y_{grt}$ specific to education groups $g$ (college graduates, non-graduates)\(^{11}\), but keep the aggregate-level immigration shock $\lambda^F_{rt}$ and controls on the right hand side:

$$\Delta y_{grt} = \delta^u_{0g} + \delta^u_{1g} \lambda^F_{rt} + \delta^u_{2g} b_{rt} + \delta^u_{3g} (n_{rt-1} - l_{rt-1}) + A_r \delta^u_{Agt} + \varepsilon_{grt}$$  \hspace{1cm} (21)

\(^{11}\)To adjust education-specific employment rates, wages and housing costs for local composition, I apply the methods of Appendices B.2 and B.4 to the education samples.
which follows the “total effects” approach recommended by Dustmann, Schoenberg and Stuhler (2016). I report IV estimates of $\delta_{ug}$ by outcome (across columns) and education group $g$ (rows) in Table 6.

Column 1 reports the effects on education-specific population growth $\Delta l_{grt}$; and the next two columns disaggregate this change into foreign and residual contributions. I specify these analogously to (14) and (15):

$$\lambda^F_{grt} \approx \log \left( \frac{L_{grt-1} + I^F_{grt}}{L_{grt-1}} \right) \quad (22)$$

$$\lambda^I_{grt} \approx \log \left( \frac{L_{grt} - I^F_{grt}}{L_{grt-1}} \right) \quad (23)$$

where $L^F_{grt}$ is the education $g$ stock of new migrants (arriving since $t - 1$). Column 2-3 show the crowd-out is educationally “balanced”: the college share of both foreign inflows $\lambda^F_{grt}$ (elicited by the enclave shift-share) and the residual population response $\lambda^I_{grt}$ resemble the existing population. As a result, there is little change in the relative supply of college workers (column 1).

Despite this, the adverse effect of foreign inflows on native employment rates falls almost entirely on non-graduates (column 4). This suggests the (balanced) population changes in column 1 are understating the labor market pressure on low educated natives: I return to this point below.

Finally, columns 6-9 reveal a small positive effect on the wages of graduate natives; but they also face larger growth in housing expenditures. Whether this reflects changes in unobserved housing consumption or prices is open to interpretation. Coupled with the difficulty of constructing credible local wage deflators, this underscores the advantages of studying welfare effects using local employment rates (as a sufficient statistic).

6 Accounting for undercoverage

6.1 Existing evidence on undercoverage

My finding of perfect crowd-out is puzzling, even in the context of my own results. First, as I discuss in Section 2.5, theory suggests the unconditional effect, $\delta^u_1$, should be smaller than the conditional, $\delta^c_1$; but I find the reverse. Second, perfect crowd-out is indicative of

---

12The price interpretation may be relevant if housing units are imperfect substitutes within CZs. For example, Albouy and Zabek (2016) document growing house price dispersion within cities, driven mostly by changes in relative neighborhood values.
full labor market adjustment, but this is inconsistent with the adverse employment rate effects.

Undercoverage of migrants in the census can help resolve this puzzle. This is presumably a larger problem for the undocumented, who account for almost half the 1990s foreign inflow (Department of Homeland Security, 2003). Surprisingly perhaps, many of them do respond to the census (Warren and Passel, 1987), but a significant fraction do not. The problem was more severe in earlier years (Card and Lewis, 2007): based on mortality rates, Borjas, Freeman and Lang (1991) impute that 40% of undocumented Mexicans were missed by the 1980 census, and Van Hook and Bean (1998) find this shrank to 30% in 1990; and using an external Los Angeles survey, Marcelli and Ong (2002) find an undercount of 10-15% in 2000. Though the Census Bureau does itself produce estimates of the general undercount (disaggregated by race, though not by migrant status), these adjustments have not been applied to census outputs (or to microdata weights): see Williams (2011). Instead, the focus has been on improving coverage on the ground.

Of course, the studies above are not necessarily informative about undercoverage of foreign inflows elicited by the *enclave shift-share* (which is what matters for my application). But, the evidence above offers circumstantial support for its importance on this margin also. First, estimates of unconditional crowd-out are much larger before 1980 (Table 4), which matches the more severe undercoverage in those years. Second, if undercoverage is greater among low educated migrants, this can explain why low educated natives suffer more adverse employment rate effects - despite the apparently balanced immigration shock (Table 6).

### 6.2 Model for undercoverage bias

By exploiting my model’s structure, I can identify the undercoverage bias using my data alone. I will show that controlling for employment growth (as in the conditional specification) eliminates the bulk of the bias. And using an estimate of the employment elasticity, I can then identify the true unconditional crowd-out - and hence the size of the bias.

I begin by extending the model to account for undercoverage. For simplicity, I will abstract from the initial employment rate and other controls in the crowd-out equations. (One may interpret all variables in the analysis as residuals, conditional on these controls - in line with the Frisch-Waugh theorem.) Consider first the unconditional crowd-out equation, (20):

\[
\lambda^I_{rt} = \delta^u \lambda^F_{rt}
\]

Suppose I do not observe the true foreign inflow \(\lambda^F\), but rather \(\hat{\lambda}^F_{rt} = (1 - \pi) \lambda^F_{rt}\), where \(\pi \in [0,1]\) is the fraction of new migrants (arriving since \(t - 1\)) who are not covered by the census. And suppose I estimate the model with \(\hat{\lambda}^F_{rt}\) on the right hand side. The
relationship between $\lambda^I_{rt}$ and $\hat{\lambda}^F_{rt}$ can be described by:

$$\lambda^I_{rt} = \hat{\delta}^u_1 \hat{\lambda}^F_{rt}$$

(25)

where the (biased) coefficient $\hat{\delta}^u_1$ exceeds the true effect $\delta^u_1$ by fraction $\pi$:

$$\frac{\hat{\delta}^u - \delta^u}{\delta^u} = \pi$$

(26)

Now consider the conditional equation (19). Again, abstracting from the various controls, the true model is:

$$\lambda^I_{rt} = \delta^c_1 \hat{\lambda}^F_{rt} + \delta^c_2 \Delta n_{rt}$$

(27)

Recall that equation (9) in the baseline model imposes $\delta^c_1 = -\delta^c_2$. But as disamenity effects may violate this assumption, I permit $\delta^c_1$ and $\delta^c_2$ to differ here. As before, I only observe $\hat{\lambda}^F_{rt} = (1 - \pi) \lambda^F_{rt}$. But crucially, my data will also understate employment growth, $\Delta n_{rt}$. In line with the model (and as Table 5 suggests), suppose natives and migrants face identical changes in employment rates. Then, observed employment growth will be:

$$\Delta \hat{n}_{rt} = \Delta (n_{rt} - l_{rt}) + \lambda^I_{rt} + \hat{\lambda}^F_{rt} = \Delta n_{rt} - \pi \lambda^F_{rt}$$

(28)

Now, suppose I estimate the conditional equation using the observed (but mismeasured) $\hat{\lambda}^F_{rt}$ and $\Delta \hat{n}_{rt}$. This can be written as:

$$\lambda^I_{rt} = \hat{\delta}^c_1 \hat{\lambda}^F_{rt} + \hat{\delta}^c_2 \Delta \hat{n}_{rt}$$

(29)

where the (biased) coefficient estimators $\hat{\delta}^c_1$ and $\hat{\delta}^c_2$ are equal to:

$$\begin{pmatrix} \hat{\delta}^c_1 \\ \hat{\delta}^c_2 \end{pmatrix} = \left( \begin{pmatrix} \text{Var} \left( \hat{\lambda}^F_{rt} \right) & \text{Cov} \left( \Delta \hat{n}_{rt}, \hat{\lambda}^F_{rt} \right) \\ \text{Cov} \left( \Delta \hat{n}_{rt}, \hat{\lambda}^F_{rt} \right) & \text{Var} \left( \Delta \hat{n}_{rt} \right) \end{pmatrix} \right)^{-1} \begin{pmatrix} \text{Cov} \left( \hat{\lambda}^F_{rt}, \lambda^I_{rt} \right) \\ \text{Cov} \left( \Delta \hat{n}_{rt}, \lambda^I_{rt} \right) \end{pmatrix} = \left( \begin{pmatrix} \delta^c_1 + \frac{\pi}{1 - \pi} (\delta^c_1 + \delta^c_2) \\ \delta^c_2 \end{pmatrix} \right)$$

(30)

See Appendix E for more detailed steps. Notice that $\hat{\delta}^c_2$, the coefficient on $\Delta \hat{n}_{rt}$, identifies the true $\delta^c_2$. Intuitively, the bias in measured foreign inflows $\hat{\lambda}^F_{rt}$ partials out the bias in employment growth $\Delta \hat{n}_{rt}$. Under the baseline assumption that $\delta^c_1 = -\delta^c_2$ (i.e. in the absence of disamenity effects), $\hat{\delta}^c_1$ will also be unbiased - for similar reasons. More generally though, using (30), the bias in $\hat{\delta}^c_1$ can be written as:

$$\frac{\hat{\delta}^c_1 - \delta^c_1}{\delta^c_1} = \pi \left( \frac{\hat{\delta}^c_1 + \hat{\delta}^c_2}{\delta^c_1} \right)$$

(31)

In practice, column 2 of Table 3 (i.e. the IV estimates) rejects the claim that $\hat{\delta}^c_1 = -\delta^c_2$, but the relative discrepancy is small: $\frac{\hat{\delta}^c_1 + \hat{\delta}^c_2}{\delta^c_1} = \frac{0.913 - 0.743}{0.913} = 0.19$. Based on (31), this
suggests the bias in $\hat{\delta}_1^c$ is just a fifth of the bias in $\hat{\delta}_1^u$ in (26): i.e. undercoverage bias should have comparatively little effect on conditional estimates of crowd-out.

Table 7 offers suggestive evidence in favor of this claim. The first row reports unconditional crowd-out by decade (replicating the penultimate row of Table 4), and the second row conditional crowd-out: i.e. estimating (19) separately by decade. Unlike in the first row, the estimates are no longer systematically larger in the 1960s and 1970s, when undercoverage was largest: this is consistent with lower bias in the conditional estimates.

### 6.3 Identification of undercoverage bias

Using the crowd-out equations alone, it is not possible to separately identify the undercoverage bias $\pi$ from the employment response to immigration. But as I now explain, identification is feasible if I know the elasticity of employment. Abstracting from labor supply/demand shocks, employment growth in (11) collapses to:

$$\Delta n_{rt} = \eta \Delta l_{rt}$$

where the elasticity $\eta$, defined in (12), will depend on both the labor and housing market elasticities. Substituting this for $\Delta n_{rt}$ in the true conditional model, (27), I can reduce the unconditional impact $\delta_1^u$ of foreign on internal inflows to$^{13}$:

$$\lambda^l_{rt} = \frac{\delta_1^o + \eta \delta_2^c}{1 - \eta \delta_2^o} \lambda^F_{rt}$$

Note that theory predicts a negative $\delta_1^c$ and positive $\eta$ and $\delta_2^c$. Thus, a larger employment elasticity ($\eta$) and population response to employment ($\delta_2^c$) moderate the extent of crowd-out. Using (26), (30) and (31), I now replace the true $\delta_1^o$, $\delta_1^c$ and $\delta_2^o$ in (33) with the biased estimators ($\hat{\delta}_1^o$, $\hat{\delta}_1^c$, $\hat{\delta}_2^o$). Rearranging, I then have an expression for the undercoverage bias $\pi$ in terms of the three biased estimators and the employment elasticity $\eta$:

$$\pi = 1 - \frac{(1 - \eta) \hat{\delta}_2^o}{\hat{\delta}_1^o + \hat{\delta}_2^o - (1 - \eta \hat{\delta}_2^o) \hat{\delta}_1^o}$$

Since $\hat{\delta}_1^u$ is negative, (34) describes a positive relationship between $\pi$ and $\eta$, for given $\hat{\delta}_1^u$, $\hat{\delta}_1^c$ and $\hat{\delta}_2^o$. Intuitively, a larger employment elasticity $\eta$ moderates the true unconditional crowd-out $\delta_1^u$ (see (33)), so I require more bias $\pi$ to account for a given estimate of $\hat{\delta}_1^u$.

$^{13}$This expression is identical to the coefficient on $\lambda^l_{rt}$ in equation (13), under the model’s assumption that $\delta_2^o = -\delta_1^c = \left(1 - \frac{1 - \gamma}{\eta \gamma}\right)$.
For illustration, I have plotted the implied relationship between $\pi$ and $\eta$ in Figure 1, based on my IV coefficient estimates from Table 3: $\hat{\delta}_1^c = -0.913$ and $\hat{\delta}_2^c = 0.743$ (from column 2), and $\hat{\delta}_1^u = 1.096$ (column 4). Assuming the employment elasticity $\eta$ exceeds 0, the model imposes a lower bound on the undercoverage bias $\pi$ of 0.2: this is required for unconditional crowd-out $\delta_1^u$ not to exceed conditional crowd-out $\delta_1^c$. And as $\eta$ goes to 1 (implying employment takes the full burden of adjustment), $\pi$ must also go to 1 (to ensure the true unconditional response is 0).

6.4 Estimates of employment elasticity $\eta$

Using Figure 1, knowledge of the employment elasticity $\eta$ to population is sufficient to identify the undercoverage bias $\pi$. This elasticity can be estimated using the employment equation (11) directly. Amior and Manning (2018) estimate $\eta$ as 0.79, using January temperature as an instrument for population growth. Looking at Figure 1, this would imply a $\pi$ of about 0.4.

In Appendix E.2, I extend Amior and Manning’s analysis by allowing for distinct employment responses to foreign and (net) internal inflows. Using the enclave shift-share as an additional instrument, I estimate elasticities of 0.61 to the former and 0.78 to the latter. If natives and migrants supply identical labor, these should be identical. But since they are identified using divergent sources of variation (i.e. temperature and enclaves), it is perhaps unwise to over-interpret the gap between them.

In any case, since I am studying crowd-out in response to immigration, I focus on the elasticity to foreign inflows (0.61). As I show in Appendix E.3, this number is not itself immune from undercoverage bias; but the form of the bias is known. Together with the relationship in Figure 1, the model then implies a “true” employment elasticity $\eta$ of 0.44 and an undercoverage bias $\pi$ of 0.27. This lower value for $\pi$ seems more plausible in the context of the evidence described in Section 6.1. The implied “true” unconditional crowd-out $\delta^u_1$ would then be: $-1.096 \cdot (1 - 0.27) = 0.80$.

Based on these numbers, what are the respective contributions of population and employment to local adjustment? As I have explained, the conditional crowd-out effect $\delta^c_1$ represents the unmediated mobility response to immigration, conditional on employment. Using (31), the true $\delta^c_1$ can be computed as: $\left[1 - (\frac{\hat{\delta}_1^u + \hat{\delta}_1^c}{\hat{\delta}_1^c}) \pi \right] \hat{\delta}_1^c$, which I calibrate to $(1 - 0.19 \cdot 0.27) 0.913 = 0.87$. That is, with no employment response, crowd-out would have reached 0.87. Since the unconditional effect $\delta^u_1$ (0.80) is 8% smaller, this implies the employment response moderates the welfare impact (as implied by the mobility response) by just 8%.

One may be surprised the contribution of employment is so small, despite the large employment elasticity $\eta$. But as the algebra shows, the employment response may not
be so salient if population itself is very elastic. And indeed, the evidence shows the same
is true of local adjustment to labor demand shocks: here also, the burden of adjustment
falls mostly on population rather than employment (Blanchard and Katz, 1992; Hornbeck,
2012; Amior and Manning, 2018).

6.5 Implication for estimation of immigration effects

Clearly, undercoverage will bias upwards reduced form estimates of immigration effects.
But, it also has important implications for more structural methodologies, especially if
undercoverage is skill-biased.

In particular, a prominent literature (beginning with Borjas, Freeman and Katz, 1997,
and Borjas, 2003) predicts the impact of immigration by calibrating aggregate production
functions (over various skill-defined inputs) using estimated elasticities of substitution.
In a long run scenario with elastic capital, immigration in these models will raise average
wages and modify skill premia - but only to the extent that natives’ and migrants’ skill
compositions differ (Borjas, 1995; Amior and Manning, 2020). More recent iterations of
this work (such as Card, 2009b, and Ottaviano and Peri, 2012) find little effect, partly
because college shares are similar for natives and migrants in the US. But, any skill bias
in undercoverage will distort these results. In particular, if the bias is mostly due to
low educated migrants (as the evidence in Table 6 might suggest), we may be severely
overstating the migrant college share.

7 Within-area estimates of crowd-out

7.1 Empirical specification

Above, I have studied crowd-out at the aggregate CZ-level, in line with the recommen-
dations of Dustmann, Schoenberg and Stuhler (2016). But much of the literature on
crowd-out has exploited variation in immigration across skill groups within locations.
In this section, I emphasize (and show empirically) that aggregate-level and within-area
specifications identify different objects: this offers a means to reconcile the divergent
results.

Consider a within-group (superscript \(w\)) specification for unconditional crowd-out:

\[
\lambda^{F}_{g} = \delta_{0}^{w} + \delta_{1}^{w} \lambda^{F}_{g} + d_{rt} + d_{gt} + \varepsilon_{srt}
\]  

(35)

where \(\lambda^{F}_{g}\) and \(\lambda^{F}_{g}\) are the foreign and residual contributions to local population of skill
group \(g\). The \(d_{rt}\) are area-time interacted fixed effects, which absorb local shocks common
to all groups; and \(d_{gt}\) are skill-time interacted effects, which absorb group-specific national
trends.
The coefficient of interest, $\delta^w_1$, identifies the impact of skill-specific foreign inflows on (the contribution of existing US residents to) local skill composition. Existing estimates of $\delta^w_1$ are typically small and sometimes positive (Card and DiNardo, 2000; Card, 2001, 2005; Cortes, 2008), though Borjas (2006) and Monras (forthcoming) offer alternative views.\(^{14}\) Either way, a small $\delta^w_1$ is not necessarily inconsistent with large spatial crowd-out - for two reasons. First, changes in local skill composition reflect not only differential internal mobility, but also changes in the characteristics of local birth cohorts. And second, as Card (2001) and Dustmann, Schoenberg and Stuhler (2016) point out, $\delta^w_1$ does not account for the labor market impact that new immigrants exert outside their own skill group $g$.

Regarding the latter point, consider a simple example. Suppose production in area $r$, for the tradable good priced at $P_t$, has CES technology (as in e.g. Card, 2001) over skill-defined labor inputs: $Y_{rt} = \psi_{rt} \left( \sum_g \theta_{grt} N_{grt}^{\sigma} \right)^{\frac{1}{\sigma}}$, where $\psi_{rt}$ is an aggregate productivity shifter, $\frac{1}{1-\sigma}$ is the elasticity of substitution between inputs, and the exponent $\rho \leq 1$ allows for diminishing local returns. Assuming competitive markets, local wage growth for skill group $g$ is:

$$\Delta (w_{grt} - p_t) = \Delta \log \theta_{grt} - (1 - \sigma) \Delta n_{grt} + \frac{\sigma}{\rho} \Delta \log \psi_{rt} + \frac{\rho - \sigma}{\rho} \Delta y_{rt}$$

(36)

Now consider a rise in group $g$ employment $\Delta n_{grt}$, due to immigration. The area-time effects $d_{rt}$ in (35) will absorb the impact which is common to all skill groups, as encapsulated by $\Delta y_{rt}$ in (36). Conditional on the $d_{rt}$, the wage response is then the inverse elasticity of substitution, i.e. $1 - \sigma$. Intuitively, for larger $\sigma$, the wage effects are more diffused across skill groups within areas - and the same will be true of any mobility response.\(^{15}\) So even in the absence of cohort effects, $\delta^w_1$ will not in general identify an aggregate-level geographical crowd-out effect akin to $\delta^u_1$ in (20). The single exception is the case of additively separable technology (i.e. $\sigma = \rho$), which ensures no diffusion of wage effects. In Appendix A.3, I map (36) more formally onto the empirical specification (35), accounting also for local dynamics.

My purpose here is not to say that within-area estimates are uninteresting: the impact of skill-specific foreign inflows on local skill composition is certainly an important question. My point is merely that they do not generally identify aggregate-level crowd-out.

\(^{14}\)Borjas’ (2006) methodology is disputed by Peri and Sparber (2011) and Card and Peri (2016). Monras (forthcoming) estimates a large negative $\delta^w_1$ following a sudden surge of Mexican immigration, but the effect is small over decadal census intervals.

\(^{15}\)Effectively, the $\sigma$ parameter plays an analogous role at the skill level (within areas) as the employment elasticity $\eta$, in (12), at the aggregate level.
7.2 Estimates of $\delta_1^w$

I now demonstrate empirically that cross-group spillovers and local cohort effects make an important contribution to the within-area estimate $\delta_1^w$. Of course, my estimates will be sensitive to any skill-biased undercoverage of migrants - so the actual numbers should be treated with some caution.

In practice, we do not know the “true” skill delineation. But in light of the discussion above, $\delta_1^w$ estimates are likely to be sensitive to this choice, as different delineations will artificially engender different elasticities of substitution. In Table 8, I present estimates of $\delta_1^w$ in (35) for four different education-based\textsuperscript{16} “skill” delineations: (i) college graduates / non-graduates; (ii) at least one year of college / no college (as in Monras, forthcoming); (iii) high school dropouts / all others (Card, 2005; Cortes, 2008); (iv) four groups: dropouts, high school graduates, some college and college graduates (Borjas, 2006).

To explore the role of cohort effects, I compare estimates based on (i) pooled census cross-sections and (ii) a longitudinal dimension of the census (which isolates the impact on internal mobility): respondents were asked where they lived five years previously. This question, previously exploited by Card (2001) and Borjas (2006), is available in the 1980, 1990 and 2000 census extracts, yielding information on migratory flows over 1975-1980, 1985-1990 and 1995-2000.\textsuperscript{17} To preserve comparability, I restrict the pooled cross-section sample to the same three decades: the 1970s, 1980s and 1990s.

For the longitudinal estimates, I continue to define $\lambda_{grt}^I$ and $\lambda_{grt}^F$ according to (22) and (23), but time intervals are now five years: so $L_{grt}^F$ is the stock of migrants who arrived in the US within the previous five years; and the initial population $L_{grt-1}$ is constructed according to where current respondents lived five years previously.\textsuperscript{18}

In an effort to exclude education-specific local demand shocks (the $\theta_{grt}$ in (36)), I instrument foreign inflows $\lambda_{grt}^F$ in (35) using an education-specific enclave shift-share, following Card (2001). Building on equation (18) above:

$$m_{grt} = \log \left( \frac{L_{grt-1} + \sum_o \phi_{ort-1} L_{og(-r)t}^F}{L_{grt-1}} \right)$$  \hspace{1cm} (37)
where new migrants of origin $o$ and education group $g$ are allocated proportionately to the initial co-patriot geographical distribution. As is clear from columns 1 and 4 of Table 8, $m_{grt}$ is a strong instrument.

The pooled cross-section estimates of $\delta_i^w$ are remarkably large, ranging from 1 to 1.5 for the full internal contribution in column 2 (accounting for both natives and old migrants). That is, each new immigrant in a given CZ-education cell attracts an additional 1-1.5 workers to the same cell (relative to other cells). A comparison with column 3 reveals that these positive effects are (more than) entirely driven by natives.

In contrast, the longitudinal estimates of $\delta_i^w$ in column 5 (which are not conflated with cohort effects) are universally negative. They also vary considerably in magnitude, ranging from -3.6 for the college graduate/non-graduate delineation to just -0.19 for the four-group delineation. In most cases, natives contribute substantially to these effects (column 6). The model offers a rationale for this variation: finer delineations (such as the four-group) should engender greater substitutability in production (i.e. larger $\sigma$) and consequently lower estimates of $\delta_i^w$. In particular, if high school dropouts are close substitutes with other non-college workers (Card, 2009a), the relatively low $\delta_i^w$ in the third row (-0.43) can be rationalized.

Using similar longitudinal data though, Card (2001) estimates a $\delta_i^w$ which is mildly positive. In Appendix G, I find the divergence of our estimates is mostly due to his fine delineation of skill groups (Card uses six imputed occupation groups) and choice of right hand side controls.

7.3 Direct estimates of cohort effects

The difference between the pooled cross-section and longitudinal estimates is suggestive of large cohort effects. But in Appendix F, I offer more direct evidence for this, by exploiting information on individuals’ state of birth. Specifically, using the same estimating equation (35), I show that foreign inflows to a given state exert a larger impact on the education composition of natives born in that state (i.e. the pure birth cohort effect) than on those residing in it (which accounts for both birth cohorts and mobility).

As an example, consider a local inflow of low educated migrants. Despite a large net outflow of low educated natives, the native college share will typically contract relative to elsewhere: the spatial crowd-out is more than offset by a decline in the education levels of local birth cohorts.

Certainly, one may expect low-skilled immigration to raise the return to education and stimulate greater skills acquisition (see Hunt, 2017). But the effect can go the other

---

19Card controls for various demographic means within skill-area cells (age, education, migrants’ years in US), which absorb much of the migration shock’s variation. Of course, these controls may be picking up important skill-specific shocks which I have neglected. The purpose of Appendix G is merely to show how our results can be reconciled.
way: Llull (2017) argues a fall in wages may discourage labor market attachment and the accumulation of human capital.

8 Conclusion

Using census data since 1960, I estimate that new immigrants crowd out existing residents one-for-one (or 1.1 for one, in my preferred estimates) over decadal intervals. This effect is robust to numerous specification choices; and even in OLS, I find substantial crowd-out of 0.76. I base my main estimates on CZs, but I find similar effects across states. The entire effect is driven by reduced internal inflows to the affected areas, rather than larger outflows. The crowding out result appears to conflict with much of the existing literature, but I show how these estimates can be reconciled.

The magnitude of the effect is puzzling, even in the context of my own results. First, perfect crowd-out is indicative of full labor market adjustment, but this is inconsistent with the adverse effects on employment rates (especially among the low educated). Second, crowd-out declines significantly once I condition on local employment growth; but to the extent that employment responds to population, theory predicts the opposite.

I argue that census undercoverage of migrants can resolve these puzzles. This view is consistent with the much larger estimates of unconditional (but importantly, not conditional) crowd-out before 1980, when coverage was poorer. Exploiting my model’s structure, I attribute 30% of the observed crowd-out to mismeasurement. The remaining effect is largely a consequence of the labor market impact, though I argue disamenity effects also play a role. Though labor demand does cushion the impact of immigration, population mobility accounts for 90% of local labor market adjustment. These results are consistent with what we know about local adjustment following labor demand shocks.

These findings have important methodological implications for the estimation of immigration effects. First, local variation is often exploited to identify the national impact of immigration; but the interpretation of such estimates must be cautious, in light of the near-perfect response from internal mobility. Second, undercoverage will bias upwards reduced form estimates of immigration effects; and it may also cause us to severely overstate the skill composition of migrants - which is the key determinant of the aggregate impact of immigration in structural models.

References


Hong, G., and J. McLaren. 2015. “Are Immigrants a Shot in the Arm for the Local
Economy?” NBER No. 21123.


Appendices

A Derivations of equations in theoretical model 32
A.1 Housing market specification: Derivation of (10) 32
A.2 Unconditional crowd-out: Derivation of (13) 33
A.3 Derivation of within-area empirical specification (35) 34

B Data construction 35
B.1 Population 35
B.2 Employment 36
B.3 Shift-share instruments 37
B.4 Wages and housing costs 38

C Contributions of inflows and outflows to crowd-out 39

D Robustness of crowd-out estimates 40
D.1 Graphical illustration of crowd-out estimates 40
D.2 Robustness to sample and specification 41

E Undercoverage bias: Supplementary theory and estimates 43
E.1 Bias in conditional crowd-out estimates 43
E.2 Estimates of employment elasticity 44
E.3 Derivation of true employment elasticity $\eta$ and undercoverage bias $\pi$ 45

F Cohort effects in within-area estimates 47

G Reconciliation with Card (2001) 47

A Derivations of equations in theoretical model

A.1 Housing market specification: Derivation of (10)

Suppose workers have Cobb-Douglas preferences over the traded good and housing, so they spend a fixed fraction $\nu$ of their income on housing. This implies a simple linear expression for the local price index:

$$p_{rt} = \nu p_{h_{rt}} + (1 - \nu) p_t$$  \hspace{1cm} (A1)

For simplicity, suppose there are no income transfers across areas. Then, housing demand in area $r$ can be written as:

$$H_{rt}^d = \nu \frac{W_{rt} N_{rt}}{P^h_{rt}}$$  \hspace{1cm} (A2)
and in logarithms:

\[ h^d_{rt} = \log \nu + w_{rt} + n_{rt} - p^h_{rt} \]  

(A3)

In turn, suppose housing supply can be written as:

\[ h^s_{rt} = \epsilon_{hs} (p^h_{rt} - p_t) \]  

(A4)

To ease the exposition, I have assumed in (A4) that housing production does not depend on local labor, but see the Online Appendices of Amior and Manning (2018) for such an extension. Equating supply and demand, and substituting (A1) for \( p^h_{rt} \), gives:

\[ p_{rt} - p_t = \frac{\nu}{1 - \nu + \epsilon_{hs}} \left[ \log \nu + \frac{1}{\epsilon_s} (n_{rt} - l_{rt} - z^s_{rt}) + n_{rt} \right] \]  

(A5)

And taking first differences then yields equation (10) in the main text:

\[ \Delta (p_{rt} - p_t) = \frac{1}{\kappa} \left[ \frac{1}{\epsilon_s} (\Delta n_{rt} - \Delta l_{rt} - \Delta z^s_{rt}) + \Delta n_{rt} \right] \]  

(A6)

where

\[ \kappa \equiv \frac{1 - \nu + \epsilon_{hs}}{\nu} \]  

(A7)

is increasing in the elasticity of housing supply, \( \epsilon_{hs} \).

### A.2 Unconditional crowd-out: Derivation of (13)

To move from the conditional crowd-out specification (9) to the unconditional specification (13), I require a solution for local employment. Using the labor supply and demand curves, (2) and (3), local employment growth can be expressed as:

\[ \Delta n_{rt} = \frac{\epsilon^s}{\epsilon^s + \epsilon^d} \Delta z^d_{rt} + \frac{\epsilon^d}{\epsilon^s + \epsilon^d} (\Delta l_{rt} + \Delta z^s_{rt}) - \frac{\epsilon^s \epsilon^d}{\epsilon^s + \epsilon^d} \Delta (p_{rt} - p_t) \]  

(A8)

Replacing the local price deviation \( \Delta (p_{rt} - p_t) \) with (A6):

\[ \Delta n_{rt} = \eta (\Delta l_{rt} + \Delta z^s_{rt}) + (1 - \eta) \frac{\kappa}{\kappa + \epsilon^d} \Delta z^d_{rt} \]  

(A9)

and disaggregating local population growth \( \Delta l_{rt} \) into foreign and internal contributions:

\[ \Delta n_{rt} = \eta \left( \lambda^F_{rt} + \lambda^J_{rt} + \Delta z^s_{rt} \right) + (1 - \eta) \frac{\kappa}{\kappa + \epsilon^d} \Delta z^d_{rt} \]  

(A10)

where

\[ \eta \equiv 1 - \left( 1 + \frac{\kappa + 1}{\kappa + \epsilon^d} \cdot \frac{\epsilon^d}{\epsilon^s} \right)^{-1} \]  

(A11)

Equation (13) can then be derived by substituting (A10) for \( \Delta n_{rt} \) in (9).
A.3 Derivation of within-area empirical specification (35)

In this appendix, I show how the multi-skill model described in Section 7.1 can be mapped onto the empirical specification (35), accounting for skill-specific population dynamics.

In line with (2) in Section 2, I first write an equation for labor supply to skill group $g$:

$$n_{gr} = l_{gr} + \epsilon^s (w_{gr} - p_r) + z_{gr}^s$$ \hspace{1cm} (A12)

And in line with (4), suppose that indirect utility for skill group $g$ depends on a skill-specific amenity $a_{gr}$ and real consumption wage $(w_{gr} - p_r)$, which itself can be replaced with the employment rate using (A12):

$$v_{gr} = w_{gr} - p_r + a_{gr} = \frac{1}{\epsilon^s} (n_{gr} - l_{gr} - z_{gr}^s) + a_{gr}$$ \hspace{1cm} (A13)

Notice that local labor market conditions for skill group $g$ can be fully summarized by the skill-specific employment rate $(n_{gr} - l_{gr})$: this is a skill-specific version of the sufficient statistic result in Section 2.

Group $g$ subscripts can also be applied to the internal migratory response, equation (7). For simplicity, suppose the elasticity $\gamma$ is common to all skill groups. So, the resident group $g$ population adjusts (sluggishly) with elasticity $\gamma$ to skill-specific differentials in local utility $v_{gr}$.

$$\lambda^I_{gr} = \gamma \left( n_{gr} - l_{gr} - z_{gr}^s + \epsilon^s a_{gr} \right)$$ \hspace{1cm} (A14)

By symmetry with the model in Section 2, these equations can be discretized to yield a skill-specific version of (9):

$$\lambda^I_{grt} = \left( 1 - \frac{e^{-\gamma}}{\gamma^I} \right) \left( \Delta n_{grt} - \lambda^F_{grt} - \Delta z_{grt}^s + \epsilon^s \Delta a_{grt} \right)$$

$$+ \left( 1 - e^{-\gamma} \right) \left( n_{grt-1} - l_{grt-1} - z_{grt-1}^s + \epsilon^s a_{grt-1} \right)$$ \hspace{1cm} (A15)

where $\lambda^F_{grt}$ is the skill-specific foreign inflow. To derive the unconditional crowding out effect, I require a solution for local skill-specific employment $\Delta n_{grt}$. Given (A12) and the skill demand relationship in (36), this can be characterized as:

$$\Delta n_{grt} = \frac{\epsilon^s}{1 + \epsilon^s (1 - \sigma)} \left( \Delta \log \theta_{grt} + \frac{\sigma}{\rho} \Delta \log \psi_{rt} + \frac{\rho - \sigma}{\rho} \Delta y_{rt} - \Delta p_{rt} + \Delta p_t \right)$$

$$+ \frac{1}{1 + \epsilon^s (1 - \sigma)} \left( \Delta l_{grt} + \Delta z_{grt}^s \right)$$ \hspace{1cm} (A16)
Substituting this for $\Delta n_{grt}$ in (A15) yields:

$$
\lambda^{I}_{grt} = \frac{(1 - \eta^w) \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)}{1 - \eta^w \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)} \left(1 - \sigma\Delta \log \theta_{grt} - \lambda^{F}_{grt} - \Delta z_{grt}^s + \frac{\varepsilon^s}{1 - \eta^w} \Delta a_{grt}\right) (A17)
$$

$$
+ \frac{(1 - \eta^w) \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)}{1 - \eta^w \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)} \cdot \frac{1}{1 - \sigma} \left(\frac{\sigma}{\rho} \Delta \log \psi_{rt} + \frac{\rho - \sigma}{\rho} \Delta y_{rt} - \Delta p_{rt} + \Delta p_t\right)
$$

$$
+ \frac{1 - e^{-\gamma}}{1 - \eta^w \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)} \left(n_{grt-1} - l_{grt-1} - z_{grt-1}^s + \varepsilon^s a_{grt-1}\right)
$$

where

$$
\eta^w \equiv [\varepsilon^s (1 - \sigma)]^{-1} \quad (A18)
$$

is the within-area elasticity of employment with respect to population, analogous to the aggregate-level $\eta$ in (A11).

Now consider how this maps onto the within-area empirical specification (35). The area-time fixed effects $d_{rt}$ will absorb the contents of the second line of (A17). The skill-time fixed effects $d_{gt}$ will absorb any skill-time varying components of: (i) the skill-specific demand shock $\Delta \log \theta_{grt}$, (ii) the skill-specific supply shock $\Delta z_{grt}^s$, (iii) the skill-specific amenity shock $\Delta a_{grt}$, and (iv) the initial conditions on the final line of (A17). All remaining variation will fall into the error term $\varepsilon_{grt}$, so the IV exclusion restriction requires that it is uncorrelated with the skill-specific enclave shift-share $m_{grt}$ in (37). Under these conditions, the coefficient of interest $\delta^w_1$ will identify the coefficient on $\lambda^{F}_{grt}$ in (A17):

$$
\delta^w_1 = \frac{(1 - \eta^w) \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)}{1 - \eta^w \left(1 - \frac{1-e^{-\gamma}}{\gamma}\right)}
$$

As I explain in Section 7.1, $\delta^w_1$ is increasing in the internal mobility response $\gamma$, but decreasing in the elasticity of substitution $\sigma$ between skill groups in production.

**B Data construction**

**B.1 Population**

I use the same data as Amior (2020), building on Amior and Manning (2018). Local population counts for 16-64s are based on published county-level statistics, taken from the National Historical Geographic Information System (NHGIS: Manson et al., 2017). The precise tables references are given in the Online Appendices of Amior and Manning (2018). I group counties together to form Commuting Zones (CZs), following the scheme...
of Tolbert and Sizer (1996).\textsuperscript{20}

I disaggregate these local population stocks into demographic components: “new” migrants (in the US for 10 years or less), “old” migrants, natives, and education-by-nativity groups. To this end, I rely on local population shares computed from the Integrated Public Use Microdata Series (IPUMS; Ruggles et al., 2017) microdata samples. For the 2010 cross-section, I use the American Community Surveys (ACS) of 2009, 2010 and 2011 (pooled together); I use the 5% census extracts for 2000, 1990, 1980 and 1960; and the (pooled) forms 1 and 2 metro 1% extracts of 1970. The main difficulty here is that the sub-state geographical identifiers in the IPUMS microdata do not precisely identify CZs, and these identifiers also vary across years. Following Autor and Dorn (2013) and Autor, Dorn and Hanson (2013), I impute CZ-level data by weighting estimates with population counts at the intersection of CZs and these identifiers. The sources for these intersection population counts can be found in the Online Appendices of Amior and Manning (2018).

\section*{B.2 Employment}

Local employment stocks (for 16-64s) are adjusted for demographic composition. I follow the same strategy as Amior (2020). The first step is to run individual-level probit regressions of a binary employment variable on detailed demographic characteristics (listed below) and local fixed effects, separately for each microdata census cross-section (for 1960-2000) and the pooled ACS sample (of 2009-11). The local fixed effects correspond to the finest available geographical units available in each cross-section; but where geographical units are subsumed within the same CZ, I aggregate them together (to reduce computing demands).

Separately for each cross-section, I then compute the mean predicted employment rate in each geographical unit, for a distribution of demographic characteristics identical to the full national sample:

\begin{equation}
\text{EmpRateAdj}_{rt} = \int_\Omega \left( X_{it} \hat{\theta}_t + \hat{\theta}_rt \right) g (X_{it}) \, di \tag{A20}
\end{equation}

where $\Omega$ is the normal c.d.f., $\hat{\theta}_rt$ are the local fixed effects, $\hat{\theta}_t$ is the vector of coefficients on the individual characteristics $X_{it}$, and $g (X_{it})$ is the density of individuals with these characteristics. I aggregate to CZ-level by taking averages across the available geographical units, weighted by population counts at the intersections (as described in Section B.1).

In the empirical application, I identify $n_{rt} - l_{rt}$ with the log of the composition-adjusted employment rate, $\text{EmpRateAdj}_{rt}$. And I identify the log employment level $n_{rt}$ with the

\textsuperscript{20}I make one minor modification to this scheme, to facilitate consistent definitions over time: I include La Paz County (AZ) in the same CZ as Yuma County (AZ).
sum of the log adjusted employment rate and log population: \[ \log EmpRateAdj + \log rt. \]
See Amior (2020) for a more formal exposition.

In the probit regressions, I control for the following characteristics: age, age squared, four education indicators\(^{21}\) (each interacted with age and age squared), a gender dummy (interacted with all previously-mentioned variables), black/Asian/Hispanic indicators (interacted with all previously-mentioned variables), and a foreign-born indicator (also interacted with all previously-mentioned variables). In each cross-section, I also control for years in the US (among the foreign-born), again interacted with all previously mentioned variables. Information on years in the US is not consistently reported in each year, so I use different variables in each year:

- **ACS 2009-11**: Years in US, years in US squared.
- **Census 2000**: Years in US, years in US squared.
- **Census 1990**: The census only reports years in US as a categorical variable. I take the mid-point of each category (and its square), and I also include a dummy for top-category cases.
- **Census 1980**: Same as 1990. Except those who were citizens at birth do not report years in US: I code all these cases with a dummy variable.
- **Census 1970**: Same as 1980. Except some respondents do not report years in US: I code all these non-response cases with a dummy variable. I also include an additional binary indicator for migrants who report living abroad five years previously (based on a different census question), which is available for the full sample. (I exclude foreign-born respondents in the form 2 sample, as they do not report years in US.)
- **Census 1960**: No information on years in US is available.

### B.3 Shift-share instruments

In this section, I offer further details on the data underlying the Bartik industry and enclave shift-shares. The Bartik shift-share \(b_{rt}\) is based on a panel of CZ-by-industry employment. My sample for this exercise consists of employed individuals aged 16-64. For industries, I use the IPUMS consistent classification (based on the 1950 census scheme), aggregated to the 2-digit level\(^{22}\) (with 57 codes). For each cross-section, I collapse the data to industry and the finest available sub-state geographical unit; and I then aggregate to CZ-level using the population weights described in Section B.1.

In turn, I construct the enclave shift-share \(m_{rt}\) using a panel of migrant population counts by CZ and 77 countries of origin. Again, I aggregate from the various sub-state

---

\(^{21}\)High school graduate (12 years of education), some college education (1 to 3 years of college), undergraduate degree (4 years of college) and postgraduate degree (more than 4 years of college). High school dropouts (less than 12 years of education) are the omitted category.

\(^{22}\)See https://usa.ipums.org/usa/volii/occ_ind.shtml. To address inconsistencies between census years, I group all wholesale sectors in a single category, and similarly for public administration and finance/insurance/real estate. I also drop individuals coded to “Not specified manufacturing industries”.

37
geographical identifiers to CZ-level, using appropriate population weights. An important component of \( m_{r,t} \) is the stock of new migrants (by origin \( o \)) who arrived in the US in the previous ten years (outside area \( r \)): i.e. \( L^C_{o(-r),t} \) in equation (18). This can be constructed directly in all cross-sections from 1970 inclusive, which covers foreign inflows from the 1960s. But, for columns 5-8 of Table 3 (which condition on the lagged enclave shift-share), I need to construct \( m_{r,t} \) for 1960 (i.e. covering the 1950s inflow). My strategy is to impute the 1950s inflows using cohort changes: this is the difference between (i) the stock of origin \( o \) migrants in 1960 aged 16-64 (outside area \( r \)) and (ii) the stock of origin \( o \) migrants in 1950 aged 6-54 (again, outside \( r \)).

B.4 Wages and housing costs

In Section 5, I study the impact of immigration on local residualized wages, housing rents and housing prices, all of which I adjust for local (demographic or housing) composition.

I compute hourly wages as the ratio of annual labor earnings to the product of weeks worked and usual hours per week, in the census and ACS microdata. I restrict my wage sample to employees aged 16-64, excluding those in group quarters; and I also exclude wage observations below the 1st and above the 99th percentiles within each geographical unit in the microdata. For each census cross-section, I then regress log hourly wages on a rich set of demographic controls\(^{23}\), and I compute the mean residual within each geographical unit (for the nativity group of interest). I then impute CZ-level wages by taking weighted averages across these units.

My housing sample consists of houses and apartments: I exclude farms, units with over 10 acres of land, and units with commercial use. To construct the rental index, I regress the monthly rents of privately rented units on a rich set of housing characteristics\(^{24}\) (restricting attention to prices between the 1st and 99th percentiles, within each geographical unit), separately for each census cross-section. And I compute the local mean of the residuals within each geographical unit. I residualize local housing prices in the same way, though the sample is now restricted to owner-occupied units. As with wages, I impute CZ-level housing costs by taking weighted averages across the geographical units in each microdata sample.

\(^{23}\)These are the same controls I use for adjusting local employment rates: age, age squared, five education indicators, black/Asian/Hispanic indicators, gender, foreign-born status, and where available, years in US and its square for migrants, together with a rich set of interactions.

\(^{24}\)Number of rooms (9 indicators) and bedrooms (6 indicators), an interaction between number of rooms and bedrooms, building age (up to 9 indicators, depending on cross-section), presence of kitchen, complete plumbing and condominium status. I also control for a house/apartment dummy, together with interactions between this and all previously-mentioned variables.
C Contributions of inflows and outflows to crowd-out

It turns out that geographical crowd-out is entirely driven by a reduction in migratory inflows to the affected CZ, rather than an increase in migratory outflows. I present the evidence in this appendix.

Similarly to Section 7, I exploit the longitudinal dimension of the census: respondents were asked where they were living five years previously. One can construct the relevant variables using microdata, as I do in Section 7.2. But since I do not need to disaggregate by education for this exercise, I instead use published statistics on gross migratory flows between all country pairs: these are based on larger samples and require no geographical imputation. I use data for the periods 1965-70, 1975-80, 1985-90 and 1995-2000, and I aggregate all flows to CZ level.\footnote{I thank Jack DeWaard for sharing the 1965-70 and 1975-80 flow data. I take the 1985-90 flow data from the Socioeconomic Data and Application Center at Columbia University; see the P1 STP-28 tables at \url{http://sedac.ciesin.columbia.edu/data/set/acrp_enhance-migration-1990}. I construct the 1995-2000 flows using the C2 A1 and B4 A1 tables on the Census 2000 Migration DVD (kindly made available by Kin Koerber); see \url{https://www.census.gov/population/www/cen2000/migration/mig_dvd.html}. Comparable data is not available for 2005-2010. Note the published flow data is available for individuals aged 15-64, rather than my usual 16-64 sample.}

My strategy is to re-estimate the unconditional crowd-out equation (20), but replacing the decadal foreign and internal contributions with 5-year flows. In particular, my specification is:

\[
\lambda_{rt}^{F5} = \delta_{0t}^{u} + \delta_{1}^{u} \lambda_{rt}^{F5} + \delta_{2}^{u} b_{rt} + \delta_{3}^{u} (n_{rt-10} - l_{rt-10}) + A_{t} \delta_{At}^{u} + \varepsilon_{rt} \tag{A21}
\]

where the \( t \) subscript now designates years, rather than decades (as in the main text), and \( \lambda_{rt}^{F5} \) and \( \lambda_{rt}^{I5} \) are respectively the 5-year foreign and internal contributions to the change in log population. These are constructed in line with equations (14) and (15). Specifically, \( \lambda_{rt}^{F5} \) is approximated as \( \log \left( \frac{L_{rt-5} + L_{rt}^{F5}}{L_{rt-5}} \right) \), where \( L_{rt}^{F5} \) is the 5-year flow into area \( r \) from abroad, and \( L_{rt-5} \) is the local population at time \( t - 5 \) (based on census respondents’ reported place of residence five years previously). And in turn, \( \lambda_{rt}^{I5} \) is approximated as \( \log \left( \frac{L_{rt-5} + L_{rt}^{I5} - L_{rt}^{O5}}{L_{rt-5}} \right) \), where \( L_{rt}^{I5} \) and \( L_{rt}^{O5} \) are respectively the 5-year inflows and outflows to/from others parts of the US. Notice that, by construction, \( L_{rt} = L_{rt-5} + L_{rt}^{F5} + L_{rt}^{I5} - L_{rt}^{O5} \). Given that the flows are based on the reports of time \( t \) residents, individuals who emigrated from the US between \( t - 5 \) and \( t \) are excluded from this data.

I do not observe employment outcomes between census years (i.e. at 5 year intervals), so I choose to use the same right hand side variables as in equation (20): the decadal Bartik shift-share \( b_{rt} \) (which predicts employment growth between \( t - 10 \) and \( t \)), the employment rate lagged ten years, and the amenity controls. The mismatch in time periods is not ideal, and one should keep this in mind when interpreting the estimates.
I report OLS and IV estimates in Table A1. I instrument $\lambda_{rt}^{I5}$ using a 5-year enclave shift-share, constructed to predict the 5-year flow and based on migrant settlement patterns in $t-5$. I construct these settlement patterns using migrants’ reported historical residence in the census microdata of year $t$ (i.e. following a similar procedure to the longitudinal estimates of Section 7.2). I instrument the lagged employment rate using the lagged decadal Bartik shift-share.

The standard errors on the OLS estimates are too large to make definitive statements. But the IV estimates tell a much clearer story. Column 4 reports the basic $\delta^u_1$ estimate, based on equation (A21). This points to a large crowding out effect (-1.6), somewhat in excess of one-for-one (though not significantly different). In the next two columns, I disaggregate the effect into (approximate) contributions from internal inflows and outflows: column 5 replaces the dependent variable with $\lambda_{rt}^{I5}$, approximated as $\log\left(\frac{L_{rt+5}}{L_{rt}+L_{rt+5}}\right)$; and column 6 replaces it with $\lambda_{rt}^{Io5}$, approximated as $\log\left(\frac{L_{rt+5}+L_{rt}^{Io5}}{L_{rt+5}}\right)$. The crowding out effect is entirely driven by variation in inflows rather than outflows.

## D Robustness of crowd-out estimates

### D.1 Graphical illustration of crowd-out estimates

I now consider the robustness of my estimates of unconditional crowd-out, $\delta^u_1$, in equation (20). One concern is that it may be driven by outliers. To address this question, Figure A1 illustrates the correlation underlying the basic OLS and IV estimates of $\delta^u_1$ (in columns 3 and 4 of Table 3).

For the OLS plot, I extract residuals from regressions of both the residual and foreign contributions to local population (i.e. $\lambda_{rt}^{I}$ and $\lambda_{rt}^{F}$ respectively) on the remaining controls: the initial employment rate, the current Bartik shift-share, year effects and the amenity variables (interacted with year effects). And I then plot the $\lambda_{rt}^{I}$ residuals against those of $\lambda_{rt}^{F}$.

For IV, I apply the same Frisch-Waugh logic to two-stage least squares. I first generate predictions of the two endogenous variables (the foreign inflow, $\lambda_{rt}^{F}$, and the initial employment rate, $n_{rt-1} - l_{rt-1}$), based on the first stage regressions (using the enclave shift-share $m_{rt}$ and lagged Bartik $b_{rt-1}$ instruments). I then extract residuals from regressions of both $\lambda_{rt}^{I}$ and the predicted $\lambda_{rt}^{F}$ on the remaining controls: the predicted initial employment rate, the current Bartik shift-share, year effects and the amenity variables (interacted with year effects). And as before, I then plot the $\lambda_{rt}^{I}$ residuals against those of $\lambda_{rt}^{F}$.
The marker size in these figures is proportional to the initial population share weights. The fit lines’ (weighted) slopes are identical to the $\delta_t^n$ estimates in columns 3 and 4 in Table 3. Of course, the standard errors are different: I do not cluster errors in Figure A1; and for IV, the naive two-step estimation does not account for sampling error in the first stage. In any case, the main take-away is that the $\delta_t^n$ estimates are not visibly driven by outliers.

D.2 Robustness to sample and specification

In the main text, I have already considered the implications of local dynamics, decadal sample, and the right hand side controls, for my $\delta_t^n$ estimates. In Table A2, I now consider robustness to other considerations.

Panel A of Table A2 offers estimates weighted by lagged population share, in line with the main text. Column 1 reproduces the basic IV estimate of $\delta_t^n$ in column 4 of Table 3, based on equation (20). I instrument the foreign inflow $\lambda_F^{rt}$ using the enclave shift-share $m_{rt}$ and the lagged employment rate using a lagged Bartik $b_{rt-1}$. Panel B offers unweighted estimates of $\delta_t^n$: the coefficient is not much different (-0.94 compared to -1.1), though the standard error is somewhat larger. This suggests the effects are not merely driven by large CZs, consistent with the patterns in Figure A1.

Alternatively, one may prefer to see estimates which exclude smaller CZs - in line with much of the literature, which often restricts attention to metro areas. In column 2, I re-estimate the model for the sample of CZs which have at least 100,000 individuals aged 16-64 in 1960. But this makes little difference to the results.

Given the skew in the spatial distribution of foreign inflows, one may also be concerned that the estimates are driven by CZs facing unusually high inflows. In column 3, I exclude observations with values of the enclave shift-share $m_{rt}$ above 0.1, which is the 98th percentile (the maximum value is 0.29: see Table 1). But again, this makes little difference.

Recall the enclave instrument $m_{rt}$ is given by $\log \left( \frac{L_{rt-1} + \Lambda_{rt-1}^F}{L_{rt-1}^F} \right)$, where $\Lambda_{rt-1}^F \equiv \sum_o \phi_{rt-1}^o L_{ot(-r)t}^F$ is shorthand for the predicted number of incoming migrants between $t-1$ and $t$: see equation (18). Notice I am using the $t-1$ migrant settlement patterns (in $\phi_{rt-1}^o$) to predict foreign inflows in each subsequent decade. But other studies have taken a different approach: for example, Hunt (2017) predicts inflows in all decades from 1940 to 2010 using the 1940
settlement patterns. In column 4, I replace my instrument with \( \log \left( \frac{L_{rt-1} + \Lambda_{F60}}{L_{rt-1}} \right) \), where \( \Lambda_{F60} \equiv \sum_o \phi_o L_{o(-r)t} \) predicts the migrant inflow based on 1960 settlement patterns, \( \phi_o \), for every decade. The weighted and unweighted estimates are now somewhat larger (-1.4 and -1.5 respectively), though they are not significantly different from -1.

In my basic specification (20), I approximate the foreign and internal contributions (to the change in log population) as \( \log \left( \frac{L_{rt-1} + L_{Frt}}{L_{rt-1}} \right) \) and \( \log \left( \frac{L_{rt-1} - L_{Frt} - L_{rt-1}}{L_{rt-1}} \right) \) respectively. But much of the literature has taken a first order approximation, defining them as \( L_{Frt} \) and \( \Delta L_{rt-1} \): see e.g. Card (2001), Peri and Sparber (2011) and Card and Peri (2016). Column 5 re-estimates (20) using these definitions; and to maintain symmetry, I replace the instrument with \( \frac{\Lambda_{F60}}{L_{rt-1}} \). But this makes little difference to the estimate.

Another possible concern is the predictive power of the instrument. Suppose the predicted number of incoming migrants, \( \Lambda_{Frt} \), is largely noise. Then variation in \( L_{rt-1} \) may generate artificial positive correlation between the endogenous variable and the instrument: see Clemens and Hunt (2019). This problem becomes worse if the \( L_{Frt} \) component of the endogenous variable; \( \frac{L_{Frt} - L_{rt-1}}{L_{rt-1}} \), is itself also noisy. Indeed, Aydemir and Borjas (2011) argue that measurement error in the local migrant share can result in substantial attenuation bias, especially in the presence of fixed effects (which may absorb much of the meaningful variation). To address this concern, in column 6, I replace the instrument (which is expressed relative to the initial population) with the predicted inflow of new migrants in levels, \( \Lambda_{F} \). But again, this has little effect on the crowding out estimate or even its standard error.

An important reference in this context is Wozniak and Murray (2012), who estimate geographical crowd-out using a specification entirely expressed in levels. Building on equation (20), a specification in levels would be:

\[
\Delta L_{rt} - L_{rt} = \delta_{0L} + \delta_{1L} L_{rt} + \delta_{2L} b_{rt} + \delta_{3L} (n_{rt-1} - l_{rt-1}) + A_r \delta_{uL} + \varepsilon_{rt} \quad (A22)
\]

where the dependent variable is the change in local population, less the stock of new immigrants; and the key regressor \( L_{rt} \) is simply the number of new immigrants. I estimate \( \delta_{1L} \) in column 7, yielding a coefficient of just \(-0.23\) in Panel A. However, local population is an important omitted variable in this specification (Wright, Ellis and Reibel, 1997; Peri and Sparber, 2011; Wozniak and Murray, 2012): local population may be correlated with both the inflow of new migrants and subsequent population change. To address this concern, Wozniak and Murray recommend controlling for local fixed effects. Once I include CZ fixed effects (column 8), my estimate of \( \delta_{1L} \) is again remarkably close to -1, irrespective of weighting. Notice this specification is also immune to the criticism of Clemens and Hunt (2019), described in the previous paragraph, of spurious correlation in the population denominator.

In column 9, I apply CZ fixed effects directly to the basic specification in column
1. These effectively partial out CZ-specific linear trends in population. This approach is similar in spirit to the double differencing methodology (comparing changes before and after 1970) of Borjas, Freeman and Katz (1997) and is recommended by Hong and McLaren (2015). Theoretically, the purpose of the fixed effects is to account for time-invariant unobserved components of the amenity, supply or demand changes in equation (20). However, their inclusion is empirically demanding in such a short panel, especially given the strong persistence in the enclave instrument $m_{rt}$. And as Aydemir and Borjas (2011) argue, measurement error may be more of a problem here. With population weights, I estimate a $\delta_1$ of -0.63 with a very large standard error (0.61). In column 10, to ease the demands of the specification, I replace the lagged employment rate (i.e. the initial conditions) with historical shocks: a lagged Bartik $b_{rt-1}$ (originally used as an instrument) and a lagged enclave shift-share $m_{rt-1}$. I now estimate a much larger $\delta_1$ of -1.35, with a standard error of just 0.26. Without population weights, I attain perversely large estimates of $\delta_1$ (in excess of -2) in both columns 9 and 10, though the standard errors are also large. This may reflect a lack of power: the first stage F-statistics for the foreign inflow $\lambda^F_{rt}$ are small in the unweighted specifications (about 6 in each case).

E Undercoverage bias: Supplementary theory and estimates

E.1 Bias in conditional crowd-out estimates

In this section, I offer more complete steps for equation (30), where I derive expressions for the (biased) conditional crowd-out estimators, $\hat{\delta}_1^c$ and $\hat{\delta}_2^c$. From Section 6.2, I have the following expressions for the observed (biased) foreign inflow $\hat{\lambda}^F_{rt}$ and employment growth $\Delta\hat{n}_{rt}$:

$$\hat{\lambda}^F_{rt} = (1 - \pi) \lambda^F_{rt} \quad \text{(A23)}$$
$$\Delta\hat{n}_{rt} = \Delta n_{rt} - \pi \lambda^F_{rt} \quad \text{(A24)}$$

where $\lambda^F_{rt}$ and $\Delta n_{rt}$ are their true counterparts. And the equation for the (net) internal inflow is:

$$\lambda^I_{rt} = \delta_1^c \lambda^F_{rt} + \delta_2^c \Delta n_{rt} \quad \text{(A25)}$$
Based on these:

\[
\begin{align*}
\text{Var} \left( \hat{\lambda}_{rt} \right) &= (1 - \pi)^2 V^F \\
\text{Var} \left( \Delta \hat{n}_{rt} \right) &= V^n + \pi^2 V^F - 2\pi C^{Fn} \\
\text{Cov} \left( \Delta \hat{n}_{rt}, \hat{\lambda}_{rt} \right) &= (1 - \pi) \left( C^{Fn} - \pi V^F \right) \\
\text{Cov} \left( \hat{\lambda}_{rt}, \lambda_{rt} \right) &= (1 - \pi) \left( \delta_1^c V^F + \delta_2^c C^{Fn} \right) \\
\text{Cov} \left( \Delta \hat{n}_{rt}, \lambda_{rt} \right) &= (\delta_1^c - \pi \delta_2^c) C^{Fn} + \delta_2 V^n - \pi \delta_1^c V^F
\end{align*}
\]

where

\[
\begin{align*}
V^F &\equiv \text{Var} \left( \lambda_{rt} \right) \\
V^n &\equiv \text{Var} \left( \Delta n_{rt} \right) \\
C^{Fn} &\equiv \text{Cov} \left( \Delta n_{rt}, \lambda_{rt} \right)
\end{align*}
\]

Substituting (A26)-(A30) into equation (30) then gives:

\[
\begin{pmatrix}
\delta_1^c \\
\delta_2^c
\end{pmatrix} = 
\begin{pmatrix}
\text{Var} \left( \hat{\lambda}_{rt} \right) & \text{Cov} \left( \Delta \hat{n}_{rt}, \hat{\lambda}_{rt} \right) \\
\text{Cov} \left( \Delta \hat{n}_{rt}, \hat{\lambda}_{rt} \right) & \text{Var} \left( \Delta \hat{n}_{rt} \right)
\end{pmatrix}^{-1}
\begin{pmatrix}
\text{Cov} \left( \hat{\lambda}_{rt}, \lambda_{rt} \right) \\
\text{Cov} \left( \Delta \hat{n}_{rt}, \lambda_{rt} \right)
\end{pmatrix}
\]

E.2 Estimates of employment elasticity

In this section, I offer estimates of the employment elasticity \( \eta \) to local population. For simplicity, the model in the main text assumes that employment adjusts instantaneously: equation (11) contains no dynamic terms. But Amior and Manning (2018) do find evidence of such dynamics, and their Online Appendices show how one can derive an estimating equation of the following form:

\[
\Delta n_{rt} = \eta_h + \eta_1 \Delta l_{rt} + \eta_2 (n_{rt-1} - l_{rt-1}) + \eta_3 b_{rt} + A_r \eta_{At} + \varepsilon_{rt}
\]
and (net) internal inflows:

\[
\Delta n_{rt} = \eta_t + \eta_1 F \lambda_{rt}^F + \eta_1 I \lambda_{rt}^I + \eta_2 (n_{rt-1} - l_{rt-1}) + \eta_3 b_{rt} + A_r \eta_{At} + \varepsilon_{rt} \tag{A36}
\]

Clearly, OLS estimates of (A35) and (A36) cannot be interpreted causally. Omitted demand shocks in the errors will generate confounding positive correlation between employment on the left hand side and population on the right. The natural instrument for population growth \(\Delta l_{rt}\) is the enclave shift-share. However, as column 1 of Table A3 shows, this has no power: this reflects the one-for-one crowd-out identified in the main text. In column 2, like Beaudry, Green and Sand (2018) and Amior and Manning (2018), I use maximum January temperature as an instrument: Rappaport (2007) shows that Americans have been moving to places with milder winters. (In these specifications, I exclude January temperature and its interactions with year effects from the \(A_r\) amenity vector on the right hand side.) This has a strong positive effect on population.

To identify the impact of \(\lambda_{rt}^F\) and \(\lambda_{rt}^I\) separately in (A36), I use both the enclave shift-share and January temperature as instruments. As expected, the shift-share has a large positive effect on \(\lambda_{rt}^F\) (column 3 of Table A3) and a large negative effect on \(\lambda_{rt}^I\) (column 4); whereas January temperature has a large effect on \(\lambda_{rt}^I\), but matters little for \(\lambda_{rt}^F\). As before, I instrument the lagged employment rate with the lagged Bartik: the final three columns show a large positive effect.

I present OLS and IV estimates of (A35) and (A36) in Table A4. The OLS elasticity to \(\Delta l_{rt}\) is essentially 1, and the effects of \(\lambda_{rt}^F\) and \(\lambda_{rt}^I\) are 0.87 and 0.99 respectively. Once I apply the instrument, I estimate smaller effects - as one might expect. Column 3, which instruments \(\Delta l_{rt}\) with the enclave shift-share, has insufficient power to identify anything - for the reasons explained above. Once I use the January temperature instrument, I identify an elasticity to \(\Delta l_{rt}\) of 0.74.\(^{26}\) And in column 4, the effects of \(\lambda_{rt}^F\) and \(\lambda_{rt}^I\) are now 0.61 and 0.78 respectively.

### E.3 Derivation of true employment elasticity \(\eta\) and undercoverage bias \(\pi\)

What do these estimates imply about the undercoverage bias, \(\pi\)? Given I am studying crowd-out in response to immigration, I choose to identify \(\eta\) with the elasticity of employment to \(\lambda_{rt}^F\), i.e. \(\eta_1 F\) in (A36).

\(^{26}\)Amior and Manning (2018) estimate a slightly larger \(\eta_1\) of 0.79. This can be attributed to two differences: they use one more decade of data (their sample includes the 1950s), and they do not adjust employment for demographic composition.
It should be emphasized that any undercoverage will bias my estimate of \( \eta_F \) upwards. To derive an expression for this bias, I follow identical steps to those of Section E.2. Abstracting from contemporaneous shocks and initial conditions in (A36), the true model for employment can be written as:

\[
\Delta n_{rt} = \eta_F \lambda^F_{rt} + \eta_I \lambda^I_{rt} \tag{A37}
\]

But given we only observe \( \Delta \hat{n}_{rt} \) and \( \lambda^F_{rt} \) with error, I estimate:

\[
\Delta \hat{n}_{rt} = \hat{\eta}_F \hat{\lambda}^F_{rt} + \hat{\eta}_I \hat{\lambda}^I_{rt} \tag{A38}
\]

where \( \Delta \hat{n}_{rt} \) and \( \hat{\lambda}^F_{rt} \) are defined as in (A23) and (A24). The coefficients \( \hat{\eta}_F \) and \( \hat{\eta}_I \) will identify:

\[
\begin{pmatrix}
\hat{\eta}_F \\
\hat{\eta}_I
\end{pmatrix}
= \begin{pmatrix}
\text{Var} \left( \lambda^F_{rt} \right) & \text{Cov} \left( \lambda^I_{rt}, \lambda^F_{rt} \right) \\
\text{Cov} \left( \lambda^I_{rt}, \lambda^F_{rt} \right) & \text{Var} \left( \lambda^I_{rt} \right)
\end{pmatrix}^{-1}
\begin{pmatrix}
\text{Cov} \left( \hat{\lambda}^F_{rt}, \Delta \hat{n}_{rt} \right) \\
\text{Cov} \left( \hat{\lambda}^I_{rt}, \Delta \hat{n}_{rt} \right)
\end{pmatrix}
\tag{A39}
\]

\[
= \begin{pmatrix}
(1 - \pi)^2 V^F & (1 - \pi) C^{FI} \\
(1 - \pi) C^{FI} & V^I
\end{pmatrix}^{-1}
\begin{pmatrix}
(1 - \pi) \left( \hat{\eta}_F V^F + \hat{\eta}_I C^{FI} \right) \\
\hat{\eta}_F C^{FI} + \hat{\eta}_I V^I
\end{pmatrix}
\]

where \( V^F \equiv \text{Var} \left( \lambda^F_{rt} \right) \), \( V^I \equiv \text{Var} \left( \lambda^I_{rt} \right) \) and \( C^{FI} \equiv \text{Cov} \left( \lambda^F_{rt}, \lambda^I_{rt} \right) \). It follows from (A39) that \( \hat{\eta}_I \) identifies \( \eta_I \), while \( \hat{\eta}_F \) is upward biased:

\[
\frac{\hat{\eta}_I - \eta_I}{\hat{\eta}_F} = \pi \tag{A40}
\]

Equation (A40) describes a negative relationship between the true \( \eta_F \) and \( \pi \), for a given \( \hat{\eta}_F \) estimate - which I set as 0.61, based on column 4 of Table A4. Together with the positive (\( \eta, \pi \)) relationship described in Figure 1 in the main text (and identifying \( \eta \) with \( \eta_F \)), this allows me to identify \( \eta \) and \( \pi \): these take values of 0.44 and 0.27 respectively.

For completeness, I also consider what happens if I calibrate the employment elasticity \( \eta \) using the estimated response to \( \lambda^I_{rt} \), i.e. \( \eta_I \) in (A36). This is equal to 0.78 and is (in principle) uncontaminated by undercoverage bias. Figure 1 would then imply a \( \pi \) of 0.44. Though a \( \pi \) of 0.27 (my preferred estimate) seems plausible in the context of the evidence described in Section 6.1, 0.44 does appear too large.
F  Cohort effects in within-area estimates

The difference between the pooled cross-section and longitudinal estimates in Section 7.2 is suggestive of large cohort effects, though perhaps not conclusively so. For example, it could be that the disparity is driven by events in the initial five years of each decade (excluded from the longitudinal sample).

In Table A5, I test for cohort effects more explicitly by exploiting information in the census on natives’ state of birth. As a reference, the first three columns present again the CZ-level pooled cross-section estimates of $\delta_{w1}$ in equation (35), identical to Table 8 in the main text. Columns 4-6 then offer state-level estimates of $\delta_{w1}$, again using pooled cross-sections. My sample consists of 49 geographical units (the 48 states of the continental US plus the District of Columbia) and three decadal observations (over 1970-2000, for comparability with the estimates in Table 8). As with the CZ estimates, the first stage (using the education-specific enclave instrument $m_{sr}$) has substantial power for all education delineations. And the IV estimates of the native-only response (column 5) look very similar to the comparable estimates for CZs (column 3).

Recall the dependent variable in column 5, $\lambda^{I,N}_{gr}$, is the contribution of natives to group $g$ population growth among state $r$ residents. In column 6, I now replace this with $\lambda^{I,N}_{BPgr}$: the contribution of natives to group $g$ population growth among those born (rather than residing) in state $r$. The column 6 estimates will now describe the contribution of cohort effects to state $r$’s education composition (though given that a third of Americans live outside their birth state, it should understate any such effects). Remarkably, these numbers are all larger than the state of residence effects in column 5 - and for the first two delineations, substantially so.

In other words, foreign inflows to a given state exert a larger impact on the education composition of natives born in that state (i.e. the pure birth cohort effect) than on those residing in it (which accounts for both birth cohorts and mobility). This suggests any contribution of internal mobility to the $\delta_{w1}$ estimate in column 5 is more than fully offset by cohort effects.

G  Reconciliation with Card (2001)

The seminal reference in the crowd-out literature is Card (2001). He offers within-area estimates of crowd-out, i.e. $\delta_{w1}$ in (35), but exploiting longitudinal residential information in the US census (respondents were asked where they lived five years previously: see Section 7.2). This approach should address concerns about cohort effects, but he still
estimates a positive value for $\delta^w_w$ - with each new immigrant to an area-skill cell attracting (on net) 0.25 additional residents. This appears to conflict with my own longitudinal estimates in column 5 of Table 8 in the main text. In this appendix, I attempt to reconcile my results with his. The divergence of our estimates is mostly explained by the delineation of skill groups (Card uses six imputed occupation groups) and choice of right hand side controls. The sample of geographical areas also plays a role.

I begin my attempting to replicate Card’s results. In line with his paper, I study variation across the 175 largest MSAs in the 5% census extract of 1990. The sample is restricted to individuals aged 16 to 68 with more than one year of potential experience. In constructing his sample, Card uses all foreign-born individuals in the census extract and a 25% random sample of the native-born. I instead use the full sample of natives, and this may (at least partly) account for some small discrepancies between his estimates and my replication. Card delineates six skill groups by probabilistically assigning individuals into broad occupation categories (laborers and low skilled services; operative and craft; clerical; sales; managers; professional and technical), conditional on their education and demographic characteristics. This assignment is based on predictions from a multinomial logit model, estimated separately for native men, native women, migrant men and migrant women; and I follow the procedure set out in his appendix. This approach offers the advantage of accounting for any occupational downgrading of migrants (see e.g. Dustmann, Schoenberg and Stuhler, 2016).

Card estimates a specification very similar to (35), except he uses first order approximations of $\lambda^I_{gr}$ and $\lambda^F_{gr}$. Specifically:

$$
\left( L_{gr,1990} - L_{gr,1990}^F \right) - L_{gr,1985} = \delta^w_0 + \delta^w_1 \frac{L_{gr,1990}^F}{L_{gr,1985}} + X_{gr} \delta^w_X + d_g + d_r + \varepsilon_{gr}
$$

where $L_{gr,1990}$ is the population of skill group $g$ in area $r$ in the census year (1990); $L_{gr,1985}$ is the local population five years previously, based on responses to the 1990 census; and $L_{gr,1990}^F$ is the number of immigrants in the skill-area cell in 1990 who were living abroad in 1985. Thus, the dependent variable is the contribution of natives and earlier (pre-1985) migrants to population growth (net of emigrants from the US, who do not appear in the sample), and the regressor $\frac{L_{gr,1990}^F}{L_{gr,1985}}$ is the contribution of immigration to that growth. To be more precise, Card actually uses the total (within-cell) population growth $\frac{L_{gr,1990} - L_{gr,1985}}{L_{gr,1985}}$ as the dependent variable, but this is a cosmetic difference: it simply adds a value of 1 to the $\delta^w_1$ coefficient (see Peri and Sparber, 2011). $X_{gr}$ is a vector of mean characteristics of individuals in the $(g, r)$ cell: these consist of mean age, mean age squared, mean years

---

27The 1990 census microdata includes sub-state geographical identifiers known as Public Use Microdata Areas (PUMAs), and a concordance between PUMAs and MSAs can be found at: https://usa.ipums.org/usa/volii/puma.shtml. A number of PUMAs straddle MSA boundaries; and following Card, I allocate the population of a given PUMA to an MSA if at least half that PUMA’s population resides in the MSA.
of schooling and fraction black, separately for both natives and migrants in the cell, and (for migrants only) mean years in the US. Finally, $d_g$ and $d_r$ are full sets of skill group and area fixed effects respectively.

The instrument for $\frac{L_{gr,1990}}{L_{gr,1985}}$ is a first order approximation of (37) in the main text, specifically $\sum_o \phi_{or,1985} L_{os,1990}$, where $\phi_{or,1985}$ is the share of origin $o$ migrants who lived in area $r$ in 1985, and $L_{os,1990}$ is the number of new origin $o$ migrants who arrived in the US between 1985 and 1990. I use the same 17 origin country groups as Card.

In his baseline OLS specification (with 175 MSAs and observations weighted by cell population), Card estimates $\delta_{1w}$ as 0.25, with a standard error of 0.04. Card’s IV estimate is also 0.25, but with a standard error of 0.05. I record these estimates in column 1 of Table A6.

I attempt to replicate these estimates in column 2 and achieve similar numbers for Card’s six-group occupation scheme (bottom row). In the remaining rows, I re-estimate the model for the four education delineations from Table 8 in the main text: (i) college graduates / non-graduates; (ii) at least one year of college / no college; (iii) high school dropouts / all others; (iv) four groups: dropouts, high school graduates, some college and college graduates. In the fifth row, I also study a classification with two imputed occupation groups: all those two-digit occupations with less than 40% college share in 1990, versus all those with more than 40%. I assign individuals probabilistically to these groups using the same multinomial logit procedure (conditioning on the same demographic characteristics) as for Card’s six group delineation in the final row. Looking at column 2, it appears that the choice of skill delineation makes no significant difference to the estimates. In column 3, I cluster errors by state: the standard errors are now larger, but the difference is not dramatic.

Much of the action comes in column 4, when I exclude the mean demographic controls in $X_{gr}$ from the right hand side. Almost all the estimates of $\delta_{1w}$ are now negative, and they are statistically significant for the college graduate, college and two-group occupation schemes, with IV coefficients of -2.14, -0.45 and -0.47 respectively. The controls in $X_{gr}$ absorb much of the (within-area) variation in the migration shock: a regression of the enclave instrument on the $d_g$ and $d_r$ fixed effects yields an R squared of 0.858, and including the controls raises the R squared to 0.928. Of course, these controls may be picking up important skill-specific shocks which I have neglected: the purpose of this exercise is merely to understand how our results can be reconciled.

---

28 Using his population growth dependent variable, this comes out as 1.25 - from which I subtract 1. See the final column of Table 4 of Card (2001).

29 The occupational distribution in college share is strongly bipolar, and 40% is the natural dividing line.
Column 5 extends the geographical sample to all identifiable MSAs (raising the total from 175 to 320), and column 6 extends it to cover 49 additional regions consisting of the non-metro areas in each state\(^{30}\) (so 369 areas in total). The latter modification ensures the area sample is comprehensive of the US, similarly to the CZs I use in the main text. There may be good reason to exclude non-metro areas; but again, the purpose of this exercise is merely to reconcile our results. These sample extensions make the coefficients larger (more negative) for all skill delineations, and the IV estimates in column 6 are now statistically significant for all but the four-group education delineation.

In the final column, I replace the left and right hand side variables with \(\log \left( \frac{L_{gr, 1990} - L_{F, 1990}}{L_{gr, 1985}} \right)\) and \(\log \left( \frac{L_{gr, 1985} + L_{F, 1990}}{L_{gr, 1985}} \right)\) respectively, using the functional form I apply in this paper: see equations (22) and (23). This makes a negligible difference to the results. The final column can now be compared to my longitudinal estimates in the main text (column 5 of Table 8): the results look similar. Just as with the education groups, moving to a finer occupation classification (i.e. from the penultimate to the final row) yields a smaller \(\delta_1^{w}\) estimate: the discussion in Section 7.2 offers an intuition for this result.

\(^{30}\)Based on the allocation procedure described above, all of New Jersey is already classified as part of an MSA. The “49 additional regions” cover the remaining 49 states.
### Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Means by decade</th>
<th>Percentiles</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1960s</td>
<td>1970s</td>
</tr>
<tr>
<td>Foreign inflow: $\lambda_{rt}$</td>
<td>0.020</td>
<td>0.033</td>
</tr>
<tr>
<td>Enclave shift-share, $m_{rt}$</td>
<td>0.016</td>
<td>0.025</td>
</tr>
<tr>
<td>Emp rate (end of decade)</td>
<td>0.624</td>
<td>0.659</td>
</tr>
<tr>
<td>Bartik shift-share, $b_{rt}$</td>
<td>0.173</td>
<td>0.227</td>
</tr>
</tbody>
</table>

This table reports descriptive statistics for key variables of interest: population-weighted means by decade, and percentiles of the full distribution. Employment rates are adjusted for local demographic composition.

### Table 2: First stage for crowding out estimates

<table>
<thead>
<tr>
<th></th>
<th>Foreign inflow: $\lambda_{rt}$</th>
<th>$\Delta$ log emp</th>
<th>Lagged log emp rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Current enclave, $m_{rt}$</td>
<td>0.919*** (0.084)</td>
<td>1.229*** (0.119)</td>
<td>1.173*** (0.105)</td>
</tr>
<tr>
<td>Lagged enclave, $m_{rt-1}$</td>
<td>-0.399*** (0.056)</td>
<td>-0.377*** (0.053)</td>
<td>-0.424*** (0.095)</td>
</tr>
<tr>
<td>Current Bartik, $b_{rt}$</td>
<td>0.092*** (0.025)</td>
<td>0.078*** (0.026)</td>
<td>0.121*** (0.034)</td>
</tr>
<tr>
<td>Lagged Bartik, $b_{rt-1}$</td>
<td>0.063*** (0.019)</td>
<td>0.064*** (0.019)</td>
<td>0.160*** (0.028)</td>
</tr>
<tr>
<td>Amenity×yr controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Geography</td>
<td>CZ</td>
<td>CZ</td>
<td>Yes</td>
</tr>
<tr>
<td>Year sample</td>
<td>60-10</td>
<td>60-10</td>
<td>70-10</td>
</tr>
<tr>
<td>Observations</td>
<td>3,610</td>
<td>3,610</td>
<td>2,888</td>
</tr>
</tbody>
</table>

This table reports first stage estimates corresponding to the crowding out specifications in Table 3. All specifications control for year effects and the amenity variables (interacted with year effects). Errors are clustered by state, 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 3: Estimates of crowding out across CZs and states

<table>
<thead>
<tr>
<th></th>
<th>Conditional</th>
<th></th>
<th></th>
<th></th>
<th>Unconditional</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\lambda_{rt}$</td>
<td>$\lambda_{rt}$</td>
<td>$\lambda_{rt}$</td>
<td>$\lambda_{rt}$</td>
<td></td>
<td>$\lambda_{rt}$</td>
<td>$\lambda_{rt}$</td>
<td>$\lambda_{rt}$</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>Foreign inflow: $\lambda_{Frt}$</td>
<td>-0.883***</td>
<td>-0.913***</td>
<td>-0.761***</td>
<td>-1.096***</td>
<td>-1.099***</td>
<td>-0.787***</td>
<td>-0.235</td>
<td>-0.949**</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.065)</td>
<td>(0.200)</td>
<td>(0.130)</td>
<td>(0.153)</td>
<td>(0.167)</td>
<td>(0.228)</td>
<td>(0.379)</td>
</tr>
<tr>
<td>$\Delta$ log emp</td>
<td>0.882***</td>
<td>0.743***</td>
<td>0.520***</td>
<td>0.831***</td>
<td>0.833***</td>
<td>0.534***</td>
<td>0.071</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.043)</td>
<td>(0.072)</td>
<td>(0.207)</td>
<td>(0.221)</td>
<td>(0.119)</td>
<td>(0.168)</td>
<td>(0.239)</td>
</tr>
<tr>
<td>Lagged log emp rate</td>
<td>0.251***</td>
<td>0.556***</td>
<td>0.646***</td>
<td>0.677***</td>
<td>0.679***</td>
<td>0.534***</td>
<td>0.136</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.105)</td>
<td>(0.109)</td>
<td>(0.099)</td>
<td>(0.096)</td>
<td>(0.119)</td>
<td>(0.168)</td>
<td>(0.239)</td>
</tr>
<tr>
<td>Current Bartik, $b_{rt}$</td>
<td>0.290***</td>
<td>0.907***</td>
<td>0.290***</td>
<td>0.907***</td>
<td>0.300***</td>
<td>0.071</td>
<td>0.136</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.103)</td>
<td>(0.060)</td>
<td>(0.103)</td>
<td>(0.105)</td>
<td>(0.167)</td>
<td>(0.233)</td>
<td>(0.233)</td>
</tr>
<tr>
<td>Lagged Bartik, $b_{rt-1}$</td>
<td>-0.388***</td>
<td>-0.984***</td>
<td>-0.469**</td>
<td>-0.469**</td>
<td>-0.469**</td>
<td>0.136</td>
<td>0.239</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.161)</td>
<td>(0.124)</td>
<td>(0.161)</td>
<td>(0.124)</td>
<td>(0.167)</td>
<td>(0.233)</td>
<td>(0.233)</td>
<td>(0.233)</td>
</tr>
</tbody>
</table>

Specification: OLS IV OLS IV IV IV IV IV IV
Instruments: $m_{rt}, b_{rt}, b_{rt-1}$ $m_{rt}, b_{rt-1}, m_{rt}, b_{rt-1}, m_{rt}$ $m_{rt}$ $m_{rt}$ $m_{rt}$ $m_{rt}$ $m_{rt}$
$F$-stat for $\lambda_{Frt}$: 93.68 126.47 106.79 124.92 295.91
$F$-stat for $\Delta n_{rt}$: 56.93 34.70 31.00 - - - -
$F$-stat for $n_{rt-1} - l_{rt-1}$: - - - - - - -
Amenity×yr controls: Yes Yes Yes Yes Yes Yes Yes Yes
Geography: CZ CZ CZ CZ CZ CZ State
Year sample: 60-10 60-10 60-10 60-10 60-10 60-10 60-10
Observations: 3,610 3,610 3,610 3,610 3,610 2,888 245

Columns 1-2 report OLS and IV estimates of the conditional crowding out specification (across CZs), (19), and columns 3-7 report the unconditional specification, (20). There are (up to) three endogenous variables: the foreign inflow, $\lambda_{Frt}$, the log employment change, and the lagged log employment rate. The corresponding instruments are the enclave shift-share $m_{rt}$ and the current and lagged Bartiks. Column 7 replaces the dependent variable with its lag, so it omits the initial decade. Column 8 replicates column 6 for state-level data. I report Sanderson-Windmeijer F-statistics which account for multiple endogenous variables. Errors are clustered by state, and robust standard errors are in parentheses. Observations are weighted by lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.
Table 4: Robustness of unconditional IV crowd-out to controls and decade

<table>
<thead>
<tr>
<th></th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Controlling for year effects only</td>
<td>0.273</td>
<td>-0.726</td>
<td>-0.041</td>
<td>-0.943***</td>
<td>-0.538**</td>
<td>-0.526**</td>
</tr>
<tr>
<td></td>
<td>(0.944)</td>
<td>(0.635)</td>
<td>(0.250)</td>
<td>(0.225)</td>
<td>(0.252)</td>
<td>(0.246)</td>
</tr>
<tr>
<td>+ Current Bartik</td>
<td>-0.745</td>
<td>-0.268</td>
<td>-0.455</td>
<td>-0.921***</td>
<td>-0.572**</td>
<td>-0.689***</td>
</tr>
<tr>
<td></td>
<td>(1.134)</td>
<td>(0.466)</td>
<td>(0.350)</td>
<td>(0.260)</td>
<td>(0.251)</td>
<td>(0.217)</td>
</tr>
<tr>
<td>+ Lagged log ER (instrumented)</td>
<td>-0.709</td>
<td>-0.238</td>
<td>-0.744*</td>
<td>-0.327</td>
<td>-0.564**</td>
<td>-0.753***</td>
</tr>
<tr>
<td></td>
<td>(1.139)</td>
<td>(0.318)</td>
<td>(0.441)</td>
<td>(0.421)</td>
<td>(0.246)</td>
<td>(0.239)</td>
</tr>
<tr>
<td>+ Climate controls</td>
<td>-1.967**</td>
<td>-2.088***</td>
<td>-0.973***</td>
<td>-1.343***</td>
<td>-0.845***</td>
<td>-1.396***</td>
</tr>
<tr>
<td></td>
<td>(0.908)</td>
<td>(0.467)</td>
<td>(0.302)</td>
<td>(0.256)</td>
<td>(0.180)</td>
<td>(0.192)</td>
</tr>
<tr>
<td>+ Coastline dummy</td>
<td>-2.032**</td>
<td>-2.087***</td>
<td>-0.865**</td>
<td>-1.119***</td>
<td>-0.637***</td>
<td>-1.263***</td>
</tr>
<tr>
<td></td>
<td>(0.947)</td>
<td>(0.473)</td>
<td>(0.350)</td>
<td>(0.251)</td>
<td>(0.189)</td>
<td>(0.228)</td>
</tr>
<tr>
<td>+ Log pop density 1900</td>
<td>-1.657***</td>
<td>-1.797***</td>
<td>-0.726***</td>
<td>-1.100***</td>
<td>-0.558***</td>
<td>-1.107***</td>
</tr>
<tr>
<td></td>
<td>(0.610)</td>
<td>(0.220)</td>
<td>(0.201)</td>
<td>(0.276)</td>
<td>(0.215)</td>
<td>(0.256)</td>
</tr>
<tr>
<td>+ Log distance to closest CZ</td>
<td>-1.626***</td>
<td>-1.917***</td>
<td>-0.877***</td>
<td>-1.203***</td>
<td>-0.638***</td>
<td>-1.137***</td>
</tr>
<tr>
<td></td>
<td>(0.634)</td>
<td>(0.197)</td>
<td>(0.188)</td>
<td>(0.298)</td>
<td>(0.236)</td>
<td>(0.251)</td>
</tr>
<tr>
<td>+ Amenity×yr effects (i.e. all controls)</td>
<td>-1.626***</td>
<td>-1.917***</td>
<td>-0.877***</td>
<td>-1.203***</td>
<td>-0.638***</td>
<td>-1.096***</td>
</tr>
<tr>
<td></td>
<td>(0.634)</td>
<td>(0.197)</td>
<td>(0.188)</td>
<td>(0.298)</td>
<td>(0.236)</td>
<td>(0.130)</td>
</tr>
<tr>
<td>Native contribution only (all controls)</td>
<td>-0.926*</td>
<td>-1.873***</td>
<td>-0.751***</td>
<td>-0.870***</td>
<td>0.101</td>
<td>-0.715***</td>
</tr>
<tr>
<td></td>
<td>(0.476)</td>
<td>(0.190)</td>
<td>(0.176)</td>
<td>(0.302)</td>
<td>(0.194)</td>
<td>(0.127)</td>
</tr>
<tr>
<td>Observations</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>3,610</td>
</tr>
</tbody>
</table>

This table tests the robustness of my IV estimates of $\delta_u^i$ (in column 4 of Table 3) to the choice of controls and decadal sample. The first five columns estimate $\delta_u^i$ separately by decade, and column 6 pools all decades. Moving down the rows of the table, I show how my $\delta_u^i$ estimate changes as progressively more controls are included. All specifications include the foreign inflow $\lambda_F$ (instrumented with the enclave shift-share, $m_t$) and year effects. The second row controls additionally for a current Bartik, $b_{rt}$; the third row includes the (endogenous) lagged employment rate (together with its lagged Bartik instrument, $b_{r, t-1}$); and the various amenities are then progressively added - until the penultimate row, which includes the full set of controls I use in Table 3. The final row replaces the dependent variable with the contribution of natives alone, $\lambda_{t, N}$, using the full set of controls. Errors are clustered by state, 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 5: IV effects of foreign inflows on local labor market outcomes

<table>
<thead>
<tr>
<th></th>
<th>Employment rates</th>
<th>Wages</th>
<th>Housing costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Native</td>
<td>Migrant</td>
<td>Native</td>
</tr>
<tr>
<td></td>
<td>Current (1)</td>
<td>Current (2)</td>
<td>Current (3)</td>
</tr>
<tr>
<td><strong>Foreign inflow:</strong> λ^*_{grt}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.210***</td>
<td>-0.190**</td>
<td>-0.350***</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.092)</td>
<td>(0.075)</td>
</tr>
<tr>
<td><strong>Lagged log emp rate</strong></td>
<td>-0.411***</td>
<td>-0.414***</td>
<td>-0.469**</td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.091)</td>
<td>(0.204)</td>
</tr>
<tr>
<td><strong>Current Bartik</strong></td>
<td>0.259***</td>
<td>0.255***</td>
<td>0.333***</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.034)</td>
<td>(0.038)</td>
</tr>
<tr>
<td><strong>Lagged Bartik</strong></td>
<td>-0.144***</td>
<td>0.069*</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.036)</td>
<td>(0.079)</td>
</tr>
<tr>
<td><strong>Instruments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amenity×yr controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year sample</td>
<td>60-10</td>
<td>60-10</td>
<td>60-10</td>
</tr>
</tbody>
</table>

This table reports estimates of (20), but replacing the dependent variable with various local outcomes: changes in the log (composition-adjusted) employment rate and mean residualized wage, separately for natives and migrants, and residualized housing rents and prices. See notes under Table 3 for further details about the specification, and see Table 2 for the first stage. The observation count is a little smaller in column 5: I am unable to compute composition-adjusted migrant employment rates for 11 small CZs in the 1960s. Errors are clustered by state, 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 6: IV effects of foreign inflows by education

|                                        | Δ log pop | Foreign contrib: λ^*_{grt} | Residual contrib: λ^*_{grt} | Employment rates | Wages | Housing costs |
|                                        | (1)       | (2)                         | (3)                         | (4)               | (5)               | (6)              | (7)   | (8)   | (9) |
| Coll grads                             | -0.261*   | 0.816***                    | -0.977***                   | -0.042*           | -0.192***         | 0.185**          | 0.319** | 0.259** | 0.625** |
|                                        | (0.159)   | (0.041)                     | (0.184)                     | (0.025)           | (0.057)           | (0.080)          | (0.157) | (0.109) | (0.278) |
| Non-grads                              | -0.145    | 1.033***                    | -1.274***                   | -0.270***         | -0.258***         | -0.030           | -0.088 | 0.020  | 0.321  |
|                                        | (0.135)   | (0.011)                     | (0.136)                     | (0.069)           | (0.066)           | (0.130)          | (0.084) | (0.109) | (0.291) |

This table reports IV estimates of δ^*_{grt} in (21), estimated for various outcomes separately for college graduates and non-graduates. All specifications include 3,610 observations (722 CZs over five decadal periods), with the exception of column 5 (whose samples are 2,693 and 3,590 respectively), for the same reasons as discussed under Table 5. The right hand side of the estimating equation is identical to that of column 2 (Panel A) of Table 3; see table notes for details. Errors are clustered by state, 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 7: Conditional and unconditional crowd-out by decade

<table>
<thead>
<tr>
<th></th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
<th>2000s</th>
<th>All years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Unconditional crowd-out: $\hat{\delta}_u^1$</td>
<td>-1.626**</td>
<td>-1.917***</td>
<td>-0.877***</td>
<td>-1.203***</td>
<td>-0.638***</td>
<td>-1.096**</td>
</tr>
<tr>
<td></td>
<td>(0.634)</td>
<td>(0.197)</td>
<td>(0.188)</td>
<td>(0.298)</td>
<td>(0.236)</td>
<td>(0.130)</td>
</tr>
<tr>
<td>Conditional crowd-out: $\hat{\delta}_c^1$</td>
<td>-0.631**</td>
<td>-1.158***</td>
<td>-0.689***</td>
<td>23.645</td>
<td>-0.807***</td>
<td>-0.913***</td>
</tr>
<tr>
<td></td>
<td>(0.261)</td>
<td>(0.128)</td>
<td>(0.244)</td>
<td>(1886.724)</td>
<td>(0.199)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>Observations</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>722</td>
<td>3,610</td>
</tr>
</tbody>
</table>

The first row offers estimates of unconditional crowding out, i.e. equation (20), separately by decade. (This is identical to the penultimate row of Table 4.) The second row replicates this exercise for conditional crowding out, i.e. equation (19). Errors are clustered by state, 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 8: Within-CZ IV estimates of $\delta_w^1$

<table>
<thead>
<tr>
<th></th>
<th>Pooled cross-sections</th>
<th>Longitudinal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First stage coefficient</td>
<td>Full residual contrib: $\lambda_{yrt}$ only: $\lambda_{yrt}^I$</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>CG / &lt; CG</td>
<td>0.539***</td>
<td>1.502***</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.295)</td>
</tr>
<tr>
<td>Coll / &lt; Coll</td>
<td>0.662***</td>
<td>1.040***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>HSD / &gt; HSD</td>
<td>0.785***</td>
<td>0.980***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.088)</td>
</tr>
<tr>
<td>4 edu groups</td>
<td>0.744***</td>
<td>1.330***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.095)</td>
</tr>
</tbody>
</table>

This table reports within-area estimates of $\delta_w^1$, based on equation (35). Columns 1-3 are based on pooled decadal cross-sections between 1970 and 2000, and columns 4-6 exploit longitudinal information on changes in residence over 1975-1980, 1985-1990 and 1995-2000. Columns 1 and 4 report the first stage coefficients on the education-specific enclave shift-share, $m_{yrt}$. And the remaining columns report IV estimates of $\delta_w^1$, both for the full residual contribution (natives and old migrants) and for natives only. The four rows offer estimates for different education-based skill delineations: (i) college graduates / non-graduates, (ii) at least one year of college / no college, (iii) high school dropouts / all others, and (iv) four groups: high school dropouts, high school graduates, some college and college graduates. All specifications control for both CZ-year and education-year interacted fixed effects. Errors are clustered by state, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged cell-specific population share. *** p<0.01, ** p<0.05, * p<0.1.
Table A1: Contribution of inflows and outflows to crowding out across CZs

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th></th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Net flow</td>
<td>Inflow</td>
<td>Outflow</td>
</tr>
<tr>
<td>$\lambda_{t5}$</td>
<td>$\lambda_{t5}$</td>
<td>$\lambda_{t5}$</td>
<td>$\lambda_{t5}$</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td></td>
</tr>
<tr>
<td>Foreign inflow: $\lambda_{F}^{t5}$</td>
<td>-0.500</td>
<td>-0.296</td>
<td>0.147</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
<td>(0.385)</td>
<td>(0.152)</td>
</tr>
<tr>
<td>Log ER lagged 10 yrs</td>
<td>0.199***</td>
<td>0.162***</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.051)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Current decadal Bartik</td>
<td>0.286**</td>
<td>0.400***</td>
<td>0.161***</td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.115)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>SW F-stat for foreign inflow</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SW F-stat for lagged ER</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Observations</td>
<td>2,166</td>
<td>2,166</td>
<td>2,166</td>
</tr>
</tbody>
</table>

This table offers OLS and IV estimates of the 5-year unconditional crowding out effect, based on equation (A21), and disaggregates these into the (approximate) contributions from internal inflows and outflows. Variable definitions and data sources are given in Section C. The flow data covers the intervals 1965-70, 1975-80, 1985-90 and 1995-2000. The 5-year foreign inflow is instrumented with a 5-year enclave shift-share in the IV specification, based on settlement patterns five years previously. The log employment rate, lagged ten years (e.g. measured at 1960 for the 1965-70 flow interval), is instrumented using a lagged decadal Bartik. I also control for a current decadal Bartik, year effects and the amenity variables (interacted with year effects) described in Section 3. Errors are clustered by state, and robust standard errors are reported in parentheses. Each observation is weighted by the 5-year lagged local population share. *** p<0.01, ** p<0.05, * p<0.1.
Table A2: Robustness of crowd-out estimates to sample and specification

<table>
<thead>
<tr>
<th>Sample Basic</th>
<th>Large CZs</th>
<th>Low m$_{rt}$</th>
<th>3,610</th>
<th>985</th>
<th>3,544</th>
<th>3,610</th>
<th>3,610</th>
<th>3,610</th>
<th>3,610</th>
</tr>
</thead>
</table>

Panel A: Weighted estimates

<table>
<thead>
<tr>
<th>$\delta_{it}$</th>
<th>$\lambda_{it}$</th>
<th>$\Delta ln(y_{it})$</th>
<th>$\Delta ln(y_{it-1})$</th>
<th>$\Delta ln(y_{it-1})$</th>
<th>$\Delta ln(y_{it-2})$</th>
<th>$\Delta ln(y_{it-2})$</th>
<th>$\Delta ln(y_{it-3})$</th>
<th>$\Delta ln(y_{it-3})$</th>
<th>$\Delta ln(y_{it-4})$</th>
<th>$\Delta ln(y_{it-4})$</th>
<th>$\Delta ln(y_{it-5})$</th>
<th>$\Delta ln(y_{it-5})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{it}$</td>
<td>$\lambda_{it}$</td>
<td>$\Delta ln(y_{it})$</td>
<td>$\Delta ln(y_{it-1})$</td>
<td>$\Delta ln(y_{it-1})$</td>
<td>$\Delta ln(y_{it-2})$</td>
<td>$\Delta ln(y_{it-2})$</td>
<td>$\Delta ln(y_{it-3})$</td>
<td>$\Delta ln(y_{it-3})$</td>
<td>$\Delta ln(y_{it-4})$</td>
<td>$\Delta ln(y_{it-4})$</td>
<td>$\Delta ln(y_{it-5})$</td>
<td>$\Delta ln(y_{it-5})$</td>
</tr>
</tbody>
</table>

Panel B: Unweighted estimates

<table>
<thead>
<tr>
<th>$\delta_{it}$</th>
<th>$\lambda_{it}$</th>
<th>$\Delta ln(y_{it})$</th>
<th>$\Delta ln(y_{it-1})$</th>
<th>$\Delta ln(y_{it-1})$</th>
<th>$\Delta ln(y_{it-2})$</th>
<th>$\Delta ln(y_{it-2})$</th>
<th>$\Delta ln(y_{it-3})$</th>
<th>$\Delta ln(y_{it-3})$</th>
<th>$\Delta ln(y_{it-4})$</th>
<th>$\Delta ln(y_{it-4})$</th>
<th>$\Delta ln(y_{it-5})$</th>
<th>$\Delta ln(y_{it-5})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{it}$</td>
<td>$\lambda_{it}$</td>
<td>$\Delta ln(y_{it})$</td>
<td>$\Delta ln(y_{it-1})$</td>
<td>$\Delta ln(y_{it-1})$</td>
<td>$\Delta ln(y_{it-2})$</td>
<td>$\Delta ln(y_{it-2})$</td>
<td>$\Delta ln(y_{it-3})$</td>
<td>$\Delta ln(y_{it-3})$</td>
<td>$\Delta ln(y_{it-4})$</td>
<td>$\Delta ln(y_{it-4})$</td>
<td>$\Delta ln(y_{it-5})$</td>
<td>$\Delta ln(y_{it-5})$</td>
</tr>
</tbody>
</table>

This table offers alternative estimates of $\delta_{it}$ in equation (20), implementing different IV strategies and variable specifications. Panel A reports weighted estimates by the lagged population share, and Panel B reports unweighted estimates. Columns 1-7 in each panel do not control for CZ fixed effects, while columns 8-10 do. The dependent variable in each specification is reported in the field above the column number. The instruments I use in each specification are reported at the bottom of the table. The Sanderson-Windmeijer (2016) F-statistics account for multiple endogenous variables. Column 2 restricts the sample to CZs with at least 100,000 individuals aged 16-64 in 1960, and column 3 restricts it to observations with enclave shift-share $m_{rt}$ values below 0.1. All specifications control for the lagged employment rate (always instrumented with the lagged Bartik $b_{r-1}$), the current Bartik $b_{r}$, year effects and the amenity variables (interacted with year effects) described in Section 3. Errors are clustered by state, and robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Table A3: First stage for employment response

<table>
<thead>
<tr>
<th></th>
<th>Log pop change: $\Delta \ell_{rt}$</th>
<th>Foreign inflow: $\lambda^F_{rt}$</th>
<th>Residual contribution: $\lambda^I_{rt}$</th>
<th>Lagged log emp rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Current enclave $m_{rt}$</td>
<td>-0.073</td>
<td>0.922***</td>
<td>-1.030***</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>(0.139)</td>
<td>(0.087)</td>
<td>(0.174)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>Max Jan temp</td>
<td>0.333***</td>
<td>0.038**</td>
<td>0.314***</td>
<td>-0.141***</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.019)</td>
<td>(0.062)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Current Bartik $b_{rt}$</td>
<td>0.546***</td>
<td>0.545***</td>
<td>0.083***</td>
<td>0.477***</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.114)</td>
<td>(0.023)</td>
<td>(0.124)</td>
</tr>
<tr>
<td></td>
<td>0.286***</td>
<td>0.283***</td>
<td>0.069***</td>
<td>0.235***</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.067)</td>
<td>(0.020)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>SW F-test</td>
<td>0.26</td>
<td>75.88</td>
<td>121.32</td>
<td>72.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max temp Jan ×yr controls</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Year sample</td>
<td>60-10</td>
<td>60-10</td>
<td>60-10</td>
<td>60-10</td>
</tr>
<tr>
<td>Observations</td>
<td>3,610</td>
<td>3,610</td>
<td>3,610</td>
<td>3,610</td>
</tr>
</tbody>
</table>

This table reports first stage estimates corresponding to the employment growth specifications in Table A4. I report Sanderson-Windmeijer F-statistics which account for multiple endogenous variables. All specifications control for year effects and the amenity variables (interacted with year effects) described in Section 3. However, I only include the maximum January temperature control (and year interactions) in those specifications where it does not serve as an instrument. Errors are clustered by state, 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 A4: Estimates of employment response

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$\Delta$ log pop</td>
<td>1.025***</td>
<td>4.000</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(6.533)</td>
</tr>
<tr>
<td>Foreign contribution: $\lambda^F_{rt}$</td>
<td>0.870***</td>
<td>0.605***</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Residual contribution: $\lambda^I_{rt}$</td>
<td>0.994***</td>
<td>0.776***</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Lagged log emp rate</td>
<td>-0.218***</td>
<td>-2.757</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(5.197)</td>
</tr>
<tr>
<td>Current Bartik</td>
<td>0.143***</td>
<td>-1.714</td>
</tr>
<tr>
<td></td>
<td>(0.029)</td>
<td>(4.156)</td>
</tr>
<tr>
<td>Instruments</td>
<td>-</td>
<td>$m_{rt}$</td>
</tr>
<tr>
<td>Amenity ×yr controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Max temp Jan ×yr controls</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year sample</td>
<td>60-10</td>
<td>60-10</td>
</tr>
<tr>
<td>Observations</td>
<td>3,610</td>
<td>3,610</td>
</tr>
</tbody>
</table>

This table reports OLS and IV estimates for models of local employment growth, i.e. equations (A35) and (A36). All specifications control for year effects and the amenity variables (interacted with year effects) described in Section 3. However, I only include the maximum January temperature control (and year interactions) in those specifications where it does not serve as an instrument. Errors are clustered by state, 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 A5: Within-area IV estimates of $\delta^w_1$: Cohort effects

<table>
<thead>
<tr>
<th></th>
<th>Czs: Pooled cross-sections</th>
<th>States: Pooled cross-sections</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First stage coefficient</td>
<td>Full residual contrib: $\lambda_{grt}$</td>
<td>Natives only: $\lambda^N_{grt}$</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>CG / &lt; CG</td>
<td>0.530***</td>
<td>1.502***</td>
<td>1.638***</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.295)</td>
<td>(0.369)</td>
</tr>
<tr>
<td>Coll / &lt; Coll</td>
<td>0.662***</td>
<td>1.040***</td>
<td>1.046***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.132)</td>
<td>(0.162)</td>
</tr>
<tr>
<td>HSD / &gt; HSD</td>
<td>0.785***</td>
<td>0.980***</td>
<td>1.410***</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.088)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>4 edu groups</td>
<td>0.744***</td>
<td>1.330***</td>
<td>1.521***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.095)</td>
<td>(0.209)</td>
</tr>
</tbody>
</table>

This table explores the presence of cohort effects in the pooled cross-section IV estimates of $\delta^w_1$ in equation (35), using a range of education-based skill delineations. As a reference, the first three columns reproduce the CZ-level pooled cross-section estimates of $\delta^w_1$ from Table 8 in the main text, based on the three decadal periods between 1970 and 2000. Columns 4 reproduces the first stage estimates using state-level data (more specifically the 48 states of the continental US plus the District of Columbia). Column 5 estimates the IV effect of skill-specific foreign inflows $\hat{\lambda}_F$ on the native contribution to skill group $g$ population growth in state $r$: i.e. the state-level version of column 3. Column 6 replaces the dependent variable with the contribution of natives to group-specific population growth among those born (rather than residing) in state $r$. All specifications control for both area-year and skill-year interacted fixed effects. Errors are clustered by state, and robust standard errors are reported in parentheses. Each observation is weighted by the lagged cell-specific population share. *** p<0.01, ** p<0.05, * p<0.1.
Table A6: Reconciliation with within-area estimates from Card (2001)

<table>
<thead>
<tr>
<th></th>
<th>Card (2001): Replication</th>
<th>... with errors</th>
<th>... excluding demog</th>
<th>... with remaining MSAs</th>
<th>... with full area sample</th>
<th>... with alternative specification for flow variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>175 MSAs, weighted</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>Panel A: OLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG / &lt; CG</td>
<td>-0.153</td>
<td>-0.153</td>
<td>-1.948***</td>
<td>-3.270***</td>
<td>-3.595***</td>
<td>-3.878***</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.572)</td>
<td>(0.549)</td>
<td>(1.124)</td>
<td>(1.099)</td>
<td>(1.145)</td>
</tr>
<tr>
<td>Coll / &lt; Coll</td>
<td>0.212</td>
<td>0.212</td>
<td>-0.346***</td>
<td>-0.393***</td>
<td>-0.642***</td>
<td>-0.725***</td>
</tr>
<tr>
<td></td>
<td>(0.172)</td>
<td>(0.248)</td>
<td>(0.111)</td>
<td>(0.101)</td>
<td>(0.143)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>HSD / &gt; HSD</td>
<td>0.118</td>
<td>0.118</td>
<td>-0.084</td>
<td>-0.105</td>
<td>-0.259***</td>
<td>-0.291***</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(0.161)</td>
<td>(0.090)</td>
<td>(0.088)</td>
<td>(0.095)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>4 edu groups</td>
<td>0.162</td>
<td>0.162</td>
<td>-0.117</td>
<td>-0.191**</td>
<td>-0.349***</td>
<td>-0.409***</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.097)</td>
<td>(0.076)</td>
<td>(0.096)</td>
<td>(0.129)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>2 occup groups</td>
<td>0.106</td>
<td>0.106</td>
<td>-0.486***</td>
<td>-0.801***</td>
<td>-0.942***</td>
<td>-1.036***</td>
</tr>
<tr>
<td></td>
<td>(0.185)</td>
<td>(0.248)</td>
<td>(0.145)</td>
<td>(0.244)</td>
<td>(0.287)</td>
<td>(0.299)</td>
</tr>
<tr>
<td>6 occup groups</td>
<td>0.25***</td>
<td>0.214***</td>
<td>0.214**</td>
<td>-0.071</td>
<td>-0.181***</td>
<td>-0.230***</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.045)</td>
<td>(0.098)</td>
<td>(0.046)</td>
<td>(0.055)</td>
<td>(0.067)</td>
</tr>
</tbody>
</table>

Panel B: IV

<table>
<thead>
<tr>
<th></th>
<th>Card (2001): Replication</th>
<th>... with errors</th>
<th>... excluding demog</th>
<th>... with remaining MSAs</th>
<th>... with full area sample</th>
<th>... with alternative specification for flow variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>175 MSAs, weighted</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td>CG / &lt; CG</td>
<td>-0.563</td>
<td>0.563</td>
<td>-2.143***</td>
<td>-1.687*</td>
<td>-2.350**</td>
<td>-2.615***</td>
</tr>
<tr>
<td></td>
<td>(1.280)</td>
<td>(2.912)</td>
<td>(0.750)</td>
<td>(0.867)</td>
<td>(0.927)</td>
<td>(0.949)</td>
</tr>
<tr>
<td>Coll / &lt; Coll</td>
<td>0.389***</td>
<td>0.389</td>
<td>-0.449***</td>
<td>-0.499***</td>
<td>-0.747***</td>
<td>-0.846***</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
<td>(0.254)</td>
<td>(0.132)</td>
<td>(0.121)</td>
<td>(0.168)</td>
<td>(0.175)</td>
</tr>
<tr>
<td>HSD / &gt; HSD</td>
<td>0.297***</td>
<td>0.297***</td>
<td>-0.059</td>
<td>-0.096</td>
<td>-0.262***</td>
<td>-0.295***</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.133)</td>
<td>(0.073)</td>
<td>(0.071)</td>
<td>(0.080)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>4 edu groups</td>
<td>0.484***</td>
<td>0.484***</td>
<td>0.004</td>
<td>0.041</td>
<td>-0.136</td>
<td>-0.162*</td>
</tr>
<tr>
<td></td>
<td>(0.117)</td>
<td>(0.118)</td>
<td>(0.077)</td>
<td>(0.084)</td>
<td>(0.086)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>2 occup groups</td>
<td>0.244*</td>
<td>0.244</td>
<td>-0.469***</td>
<td>-0.653***</td>
<td>-0.806***</td>
<td>-0.885***</td>
</tr>
<tr>
<td></td>
<td>(0.139)</td>
<td>(0.298)</td>
<td>(0.109)</td>
<td>(0.109)</td>
<td>(0.138)</td>
<td>(0.141)</td>
</tr>
<tr>
<td>6 occup groups</td>
<td>0.25***</td>
<td>0.255***</td>
<td>0.255**</td>
<td>-0.054</td>
<td>-0.123**</td>
<td>-0.192***</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.045)</td>
<td>(0.115)</td>
<td>(0.043)</td>
<td>(0.044)</td>
<td>(0.051)</td>
</tr>
</tbody>
</table>

This table offers a reconciliation with Card’s (2001) within-area estimates of crowd-out, based on equation (A41). Card’s OLS and IV estimates of $\delta^*_w$ for his six-group imputed occupation scheme are presented in column 1. These are taken from Table 4 of his paper, based on the 175 largest MSAs of the 1990 census extract, with observations weighted by cell populations. (Card reports his estimates as the effect on aggregate population growth within the cell, but I subtract one from his numbers for comparability with my specification; see Peri and Sparber, 2011.) I attempt to replicate his results in column 2. In columns 3, I cluster standard errors by state. Column 4 excludes the demographic controls from the regression. Column 5 extends the geographical sample to all identifiable MSAs (raising the total to 320), and column 6 extends it to cover 49 additional regions consisting of the non-metro areas in each state (so 369 areas in total). Finally, column 7 replaces the left and right hand side variables with log \( \left( \frac{L_{g,1990} - L_{g,1985}}{L_{g,1985}} \right) \) and log \( \left( \frac{L_{g,1985} + L_{g,1990}}{L_{g,1985}} \right) \) respectively, applying the functional form I use in the main text. I present estimates for both Card’s six-group occupation scheme (bottom row) and the other skill delineations described in Appendix G. *** p<0.01, ** p<0.05, * p<0.1.
Figure 1: Implied relationship between $\pi$ and $\eta$

This figure plots the relationship between $\pi$ and $\eta$ implied by (34). I calibrate this equation using estimates from Table 3 (columns 2 and 4): specifically, $\hat{\delta}_1 = -0.913$, $\hat{\delta}_2 = 0.743$ and $\hat{\delta}_u = 1.096$.

Figure A1: Graphical illustration of crowding out estimates

This figure presents Frisch-Waugh type plots for the unconditional $\delta^u_t$ estimates in columns 3 and 4 of Table 3. See Appendix D.1 for details.
<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1677</td>
<td>Swati Dhingra, Silvana Tenreyro</td>
<td>The Rise of Agribusiness and the Distributional Consequences of Policies on Intermediated Trade</td>
</tr>
<tr>
<td>1676</td>
<td>Jeffrey Grogger, Ria Ivandic, Tom Kirchmaier</td>
<td>Comparing Conventional and Machine-Learning Approaches to Risk Assessment in Domestic Abuse Cases</td>
</tr>
<tr>
<td>1675</td>
<td>Mary Amiti, Stephen J. Redding, David E. Weinstein</td>
<td>Who’s Paying for the U.S. Tariffs? A Longer-Term Perspective</td>
</tr>
<tr>
<td>1674</td>
<td>Cheng Chen, Claudia Steinwender</td>
<td>Import Competition, Heterogeneous Preferences of Managers and Productivity</td>
</tr>
<tr>
<td>1673</td>
<td>Stefano Bolatto, Alireza Naghavi, Gianmarco Ottaviano, Katja Zajc Kejzar</td>
<td>Intellectual Property and the Organization of the Global Value Chain</td>
</tr>
<tr>
<td>1672</td>
<td>Timothy Besley, Isabel Roland, John Van Reenen</td>
<td>The Aggregate Consequences of Default Risk: Evidence from Firm-Level Data</td>
</tr>
<tr>
<td>1671</td>
<td>Jan-Emmanuel De Neve, Daisy Fancourt, Christian Krekel, Richard Layard</td>
<td>A Local Community Course That Raises Mental Wellbeing and Pro-Sociality</td>
</tr>
<tr>
<td>1670</td>
<td>Tommaso Sonno</td>
<td>Globalization and Conflicts: The Good, the Bad and the Ugly of Corporations in Africa</td>
</tr>
<tr>
<td>1669</td>
<td>Michael Amior</td>
<td>Immigration, Local Crowd-Out and Undercoverage Bias</td>
</tr>
<tr>
<td>Page</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>1668</td>
<td>Antoine Berthou, John Jong-Hyun Chung, Kalina Manova, Charlotte Sandoz Dit Bragard</td>
<td>Trade, Productivity and (Mis)allocation</td>
</tr>
<tr>
<td>1667</td>
<td>Holger Breinlich, Elsa Leromain, Dennis Novy, Thomas Sampson</td>
<td>Exchange Rates and Consumer Prices: Evidence from Brexit</td>
</tr>
<tr>
<td>1666</td>
<td>Fabrice Defever, Michele Imbruno, Richard Kneller</td>
<td>Trade Liberalization, Input Intermediaries and Firm Productivity: Evidence from China</td>
</tr>
<tr>
<td>1665</td>
<td>Philippe Aghion, Antonin Bergeaud, Richard Blundell, Rachel Griffith</td>
<td>The Innovation Premium to Soft Skills in Low-Skilled Occupations</td>
</tr>
<tr>
<td>1664</td>
<td>Filip Gesiarz, Jan-Emmanuel De Neve, Tali Sharot</td>
<td>The Motivational Cost of Inequality: Pay Gaps Reduce the Willingness to Pursue Rewards</td>
</tr>
<tr>
<td>1663</td>
<td>Felix Koenig</td>
<td>Technical Change and Superstar Effects: Evidence From the Roll-Out of Television</td>
</tr>
<tr>
<td>1662</td>
<td>Enrico Moretti, Claudia Steinwender, John Van Reenen</td>
<td>The Intellectual Spoils of War? Defense R&amp;D, Productivity and International Spillovers</td>
</tr>
<tr>
<td>1661</td>
<td>Decio Coviello, Andrea Ichino, Nicola Persico</td>
<td>Measuring the Gains from Labor Specialization</td>
</tr>
<tr>
<td>1660</td>
<td>Nicolás González-Pampillón</td>
<td>Spillover Effects from New Housing Supply</td>
</tr>
</tbody>
</table>