

Why Has Regional Income Convergence in the U.S. Declined?

Peter Ganong and Daniel Shoag*

First Draft: July 2012, This Draft: March 2013

Abstract

The past thirty years have seen a dramatic decrease in the rate of income convergence across states and in population flows to wealthy places. These changes coincide with (1) an increase in housing prices in productive areas, (2) a divergence in the skill-specific real returns to living in productive places, (3) a redirection of low-skilled migration and (4) diminished human capital convergence due to migration. We develop a model where falling housing supply elasticity and endogenous labor mobility generates these patterns. Using a new panel measure of housing supply regulations, we demonstrate the importance of this channel. Income convergence continues in less-regulated places, while it has stopped in more-regulated places.

JEL Codes: E24, J23, J24, R14, R23, R52

Keywords: Convergence, Regulation, Land Use, Migration, Housing Prices

*Email: ganong@fas.harvard.edu (Harvard University) and dan_shoag@hks.harvard.edu (Harvard Kennedy School). We would like to thank Marios Angeletos, Robert Barro, George Borjas, Gary Chamberlain, Raj Chetty, Gabe Chodorow-Reich, Bob Ellickson, Emmanuel Farhi, Bill Fischel, Dan Fetter, Edward Glaeser, Claudia Goldin, Joe Gyourko, Larry Katz, and seminar participants at Harvard, Tel Aviv, Bar Ilan, Dartmouth and the NBER Summer Institute for their valuable feedback. Shelby Lin provided outstanding research assistance. We thank Erik Hurst for alerting us to the end of regional convergence and spurring us to work on an explanation. Peter Ganong gratefully acknowledges residence at MDRC when working on this project and funding from the Joint Center for Housing Studies and the NBER Pre-Doctoral Fellowship in Aging and Health. Daniel Shoag and Peter Ganong gratefully acknowledges support from the Taubman Center on State and Local Government. Animations illustrating the changes in income convergence, directed migration, and regulations can be found at <http://www.people.fas.harvard.edu/~ganong/motion.html>.

1 Introduction

The convergence in per-capita incomes across U.S. states from 1880 to 1980 is one of the most striking relationships in macroeconomics. During this period, incomes across states converged, on average, at a rate of 1.8% a year.¹ During the last thirty years this relationship has weakened considerably, as observed at the metro-area level in Berry and Glaeser (2005).² The top two panels of Figure 1 plot the relationship between income growth and initial income for the periods 1940-1960 and 1990-2010. For example, the average per-capita income in Connecticut was 4.37 times larger than the average per capita income in Mississippi in 1940. By 1960, that ratio had fallen to 2.28, and it fell again to 1.76 by 1980. Today, the residents of Connecticut still have average incomes 1.77 times the average in Mississippi.

The change in this relationship can be seen in the bottom panel of Figure 1. In this panel, we plot the convergence relationship (change in log income on initial log income) for rolling twenty-year windows. This figure shows that the convergence relationship was quite strong through 1980, but in the last thirty years, this pattern has largely disappeared. Similarly, the standard deviation of log income across states fell through 1980, a phenomenon known as sigma convergence, and then held steady afterward.³

During the period of strong convergence before 1980, population flowed from poor to rich states, and changes in population were well-predicted by initial income. Figure 2 plots “directed migration”: the relationship between the twenty year changes in log population and initial log income per-capita for the period 1940-1960. Over those decades, initial income per capita was associated with significantly higher population growth rates. This phenomenon was driven both by Western states like California and Nevada as well as Eastern states like Maryland, Connecticut, and Delaware.

The right panel of Figure 2 plots the same relationship for the period 1990-2010. As is evident in the figure, population no longer flows to richer states. In the bottom panel of Figure 2, we plot the extent of directed migration for rolling twenty-year windows.⁴ Directed

¹See Barro and Sala-i Martin (1992), Barro and Sala-i Martin (1991), and Blanchard and Katz (1992) for classic references. This relationship is shown in Appendix Figure 1. Throughout the analysis, we restrict our attention to the 48 contiguous US states. Our results are not sensitive to this restriction.

²See also chapter 2 of Crain (2003) and Figure 6 of DiCecio and Gascon (2008).

³See Appendix Table 1. This measure demonstrates that the estimated decline in convergence rates is not due to a reduction in the variance of initial incomes relative to a stationary shock process. The strong rate of convergence in the past as well as the decline today do not appear to be driven by changes in measurement error. When we use the Census measure of state income to instrument for BEA income, or vice-versa, we find similar results. The decline also occurs at the Labor Market Area level, using data from Haines (2010) and U.S. Census Bureau (2012).

⁴In Appendix Table 2, we report estimates using net migration (rather than population change), as well as estimates at the Labor Market Area level. These show a similar decline in the rate of directed migration.

migration was quite strong in the period prior to 1980, but in the last thirty years, this pattern has largely disappeared.⁵ As an example, from 2000 to 2010, Utah's population grew by 24%, whereas Massachusetts's population grew by just 3%. This occurred even though per capita incomes in Massachusetts were 55% higher in 2000.

What explains the decline in income convergence and directed migration? In this paper, we present evidence that new housing supply constraints in some productive places helps to explain these changes. Supply constraints limit migration and bias its composition toward high-skilled workers, both of which serve to slow income convergence. We test this theory using a new, panel measure of housing supply constraints. Today, productive places with unconstrained housing supply continue to experience net migration and slower than average income growth; productive places with supply constraints have almost zero net migration and relatively fast income growth.

We build a model in which each state buys labor from its residents and uses a linear technology to produce a differentiated good. Consumers have Dixit-Stiglitz preferences and regions engage in monopolistic competition, so demand for each region's good is downward-sloping. When the population of a city rises, the price of its good must fall for the goods market to clear, and hence wages must fall for the local labor market to clear. This provides a micro-foundation for downward-sloping labor demand. When housing supply is completely elastic, workers of all skill types gradually move to the productive locale, generating convergence. Low-skilled workers are more sensitive to changes in housing prices, and as housing supply becomes constrained, low-skill workers stop moving to the productive locale, leading to a decline in convergence.

This explanation is consistent with stylized facts about housing prices, the returns to migration, and migration flows in the U.S. Over the past fifty years, differences in incomes across states have been increasingly capitalized into housing prices (Van Nieuwerburgh and Weill (2010), Glaeser *et al.* (2005b) Gyourko *et al.* (2012)). Coincident with this increase, we show that (1) the returns to living in productive places net of housing costs have fallen dramatically for low-skilled workers but have remained large for high-skilled workers, (2) high-skilled workers continue to move to areas with high nominal income, and (3) low-skilled

⁵A number of papers have noted a recent decline in gross interstate migration rates (Kaplan and Schulhofer-Wohl (2012), and Molloy *et al.* (2011)). Kaplan and Schulhofer-Wohl (2012) show that gross migration rates remain much larger than net migration rates despite recent trends. Therefore, the overall decline in gross migration does not explain the decline in net migration documented in this study. There is also little evidence that weather-related amenities can explain the changes in migration patterns documented here. Research by Glaeser and Tobio (2007) suggests that population growth in the South since 1980 is driven by low housing prices rather than good weather. Though average January temperature is predictive of population growth, it is not correlated with high housing prices. Moreover, the relationship between temperature and population growth has remained stable or declined in the post-war period.

workers are now moving to areas with *low* nominal income but high income net of housing costs.

To better understand the effects of housing price increases, we construct a new measure of regulation as an instrument for housing prices. Our measure is a scaled count of the number of decisions for each state that mention “land use”, as tracked through an online database of state appeals court records. Changes in the regulation measure are strongly predictive of changes in housing prices, and we validate this measure of regulation using the existing cross-sectional survey data. To the best of our knowledge, this is the first panel measure of land use regulations, and the first direct evidence linking national price increases to changes in regulations.

Using differential regulation patterns at the state level over time, we provide evidence for both the case where directed migration drives convergence and the case where housing supply limits migration and convergence. Tight land use regulations weaken the historic link between high incomes and new housing permits. Instead, income differences across places are capitalized into prices. With constrained housing supply, net migration by all types from poor to rich places is replaced by high-skill workers moving in and low-skill workers moving out. Places with high lagged income and supply constraints show diminished human capital convergence and no net migration. Finally, income convergence persists among places unconstrained by these regulations, but is diminished in areas with supply constraints.

In this paper, we highlight a single channel – labor mobility – which can help explain both income convergence through 1980 and its subsequent disappearance from 1980 to 2010. In doing so, we link two previously separate literatures. First, past literature on convergence has focused on the role of capital, racial discrimination, or sectoral reallocations.⁶ We build on an older tradition of work by economic historians (Easterlin (1958) and Williamson (1965)) as formalized by Braun (1993), in which directed migration drives convergence. Meanwhile, the literature on recent trends in regional inequality has emphasized skill-biased technological change (SBTC), and its regional variants (Autor and Dorn (2013), Moretti (2012), Diamond (2012)). Of course, there were many changes over this time period, and multiple factors likely contributed to the decline in income convergence. The housing supply channel we describe, while distinct, is complementary to the SBTC channel. Our paper highlights one factor that significantly contributed to slowing convergence and altered migration patterns, and demonstrates its relevance empirically using a new measure of changing regulations.

The remainder of the paper proceeds as follows. In Section 2, we develop a formal

⁶See Barro and Sala-i Martin (1992), Caselli and Coleman (2001), Michaels *et al.* (2012), and Hsieh *et al.* (2012).

model to explore the role of labor migration and housing supply in convergence. Section 3 demonstrates that this model is consistent with several stylized facts about migration and housing prices. Section 4 introduces a new measure of land use regulation and directly assesses its impact on convergence, Section 5 considers alternative explanations, provides a calibration, and discusses how our supply constraints explanation complements the existing literature, and Section 6 concludes.

2 Model of Regional Migration, Housing Prices, and Convergence

Before exploring the impact of housing market changes empirically, we first develop a model of regional economies to structure our study of the relationship between directed migration, housing markets, and income convergence.

Our model considers two locations within a national market: “Productiveville” and “Reservationville.” We first describe Productiveville, whose local economy uses a linear technology in labor to produce a differentiated good for national consumption as in Armington (1969). Consumers have Dixit-Stiglitz preferences over regional varieties, meaning that national demand for this good slopes downward.⁷ Workers in Productiveville have heterogeneous skill types and choose labor supply based on local wages. Low-skill workers spend more of their income on housing. We solve for the general equilibrium in this economy, where product, land, and labor markets clear.

Next, we consider the interregional allocation of labor. Once we allow migration, labor inflows into Productiveville drive down the average skill level and the average wage. These declines lead to regional convergence. We prove a proposition explaining how an increase in housing prices in Productiveville alters these patterns and differentially deters low-skilled workers from migrating there. If the elasticity of land supply falls, housing prices rise, and migration flows become smaller and biased towards high-skilled workers. These changes lead to the end of convergence.

Our model builds on the existing regional economics literature. In particular, it is indebted to Blanchard and Katz (1992), which assumes that regional labor demand slopes downward and that labor migration is costly. Similarly, Braun (1993) solves a dynamic

⁷For example, suppose that Nevada produces gambling. Doubling the number of workers in Nevada would enable a doubling in the number of casinos. But because national demand for this product slopes downward, Nevada would have to cut product prices (and wages) in order to use all available labor.

model with costly labor migration from unproductive to productive places. In his model, migrants generate congestion in public goods that drive down regional incomes. We extend his model, adding heterogeneous skill types and micro-founding congestion through the land sector.

Finally, while we use this model for calibration in Section 5.1.1, our empirical analysis in Sections 3 and 4 does not rely upon the functional forms in the model. Rather, our interpretation of the data relies on three substantive assumptions embedded in our model:

1. Regional labor demand slopes downward, meaning that regional income rises less than one for one with increases in population. A few examples from the economic history literature help illustrate this concept. First, Acemoglu *et al.* (2004) study show that states which had more mobilization of men in World War II also had increased female labor force participation. After the war, both males and females in these places earned considerably lower wages. Second, Hornbeck (2012) studies the impact of a major negative permanent productivity shock, the Dust Bowl. Like Blanchard and Katz (1992), he finds that out-migration is the primary factor adjustment which allowed wages to recover to some extent. Third, Margo (1997) studies the impact of a positive productivity shock: the Gold Rush. At first, wages soared, but as people migrated in to California, wages declined. Many other papers find evidence for downward-sloping labor demand.⁸ However, the extent to which a labor inflow of one skill/ethnicity group affects other skill/ethnicity groups is an open debate.⁹ Some researchers have argued that additional population in a region might even raise the returns to labor through agglomeration spillovers (Combes *et al.* (2007)). In our view, over what domains the labor demand curve slopes downward is a vital and open research question. Here, we present an interpretation of the new facts we document that is consistent with downward-sloping labor demand.
2. Land is an inferior good within a city, meaning that low-skill workers spend a disproportionate share of income on land. The study which best matches this theoretical concept is Glaeser *et al.* (2008), which shows convincing evidence that the income elasticity of plot size is far below 1. Other evidence consistent with this assumption is

⁸See Iyer *et al.* (2011), Boustan *et al.* (2010), Cortes (2008), Angrist (1996), and Borjas (2003). Hamermesh (1987) surveys 28 published estimates of the aggregate long-run labor-demand elasticity from 1960 to 1980, and concludes that it fell in the range of -0.15 to -0.50. Card (2009) estimates the effect of changes in the relative supply of college-educated workers on the college-high school wage gap and reports elasticities between -0.26 and -0.42.

⁹See Card (1990) and Borjas (2003) for opposing views.

that after a negative productivity shock to a city, the share of low-skilled workers rises (Glaeser and Gyourko (2005) and Notowidigdo (2011)). In Appendix Figure 2, we plot the share of income spent on housing by income level, using education to instrument for income (Ruggles *et al.* (2010)). Low-skill workers spend 26% of their income on housing, while high-skill workers spend 16%, which is almost half of what low-skill workers spend.¹⁰ To be clear, our theoretical results require non-homothetic preferences but are *not* dependent on a particular income-elasticity of demand.¹¹

3. Labor migration is costly. Social ties, place-specific human capital, and a preference to live in one's own state all may make migration costly, and lead individuals not to migrate, even when large wage gains are available. For our context, differences in income across regions of the US are large and highly persistent, and there is slow, sustained migration from poor to rich places. Reconciling these time scales requires sizeable migration costs. Large costs are consistent with the recent literature on this subject. Kennan and Walker (2011), for example, estimate that the average mover faces moving costs in excess of \$300,000.

2.1 Within-state equilibrium

Here, we show the assumptions and key results from our model. A full derivation is given in Appendix A.1.

2.1.1 Individual Decisions: Labor Supply, Goods Demand, and Land Demand

For a representative agent of skill type ψ_k in state s , utility is defined by

$$\begin{aligned} \arg \max_{\{x_{ss'k}, l_{sk}\}} U &= C - \frac{l_{sk}^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}} \\ \text{subject to } \int p_{s'} x_{ss'k} ds' &= \pi + w_s \psi_k l_{sk} - r_s \\ C &= \left(\int x_{ss'k}^\rho ds' \right)^{\frac{1}{\rho}} \end{aligned}$$

¹⁰In this paper, because of data constraints, we study housing prices, which combine land prices with the value of unit amenities built on that land. Albouy and Ehrlich (2012) demonstrate the tight link between these two prices. It seems likely that high-skill workers spend more on unit amenities than low-skill workers. If this is the case, our estimates will understate the extent to which land is an inferior good. We conduct this analysis with MSA fixed effects, because the within-city Engel curve is the relevant statistic for our model. This is important because skilled workers tend to live in higher income MSAs, where housing prices are higher. This compositional effect biases the results of a national regression.

¹¹We prove our propositions under alternative non-homothetic preferences in Appendix A.7.

The wage rate w_s is the cost of one unit of effective labor, l_{sk} indexes effort, r_s is the rental price of land in state s , π are lump-sum federal transfers of profits, and $s' \in [0, 1]$ indexes other states. The first order condition for labor supply, with P as the standard index of variety prices, is

$$l_{sk}^{Supply} = (\psi_k \frac{w_s}{P})^\epsilon \quad (1)$$

With a statewide wage of w_s , a worker with numeraire skill ($\psi = 1$) will supply $l_1 = w_s^\epsilon$ units of labor. In comparison, a worker with skill ψ_k will have an effective labor supply of $\psi_k^{1+\epsilon} l_1$. Consumption of each regional good is given by $x_{ss'k} = p_{s'}^{-\sigma} \frac{(\pi + w_s \psi_k l_{sk} - r_s)}{P^{1-\sigma}}$, where $\sigma = \frac{1}{1-\rho}$.

We assume that every worker inelastically demands one plot of land, and that this land yields no consumption benefit. In Appendix A.7, we develop an extension of the model where consumers have endogenous, non-homothetic demand for housing that is increasing in income.

2.1.2 State Decisions: Labor Demand, Goods Supply and Land Supply

State-level production is given by the linear production function $Y_s = A_s \times L_s$, where L_s is effective labor supply. We view this functional form as a reduced form representation of a more complicated production process. The first term A_s can encompass capital differences, natural advantages, institutional strengths, different sectoral compositions, amenities, and agglomeration benefits. Firms in the state supply their good monopolistically, taking national demand curves and prices in other states as given. They solve

$$\begin{aligned} & \max_{p_s, L_s, X_s} p_s X_s(p_s) - w_s L_s \quad \text{subject to} \\ \text{(i)} \quad & X_s^{Demand} = p_s^{-\sigma} \underbrace{\frac{\int \int (\pi + \psi_k^{1+\epsilon} w_{s'}^{1+\epsilon} P^{-\epsilon}) ds' dk}{P^{1-\sigma}}}_{\text{global real income: } Y_g} \\ \text{(ii)} \quad & L_s^{Demand} = X_s^{Supply} / A \end{aligned}$$

where $\mu_{s'k}$ is the number of type k workers in place s' . The firm's optimal price is given by the first order condition $p_s = \frac{\sigma}{\sigma-1} \frac{w_s}{A_s}$. Given national demand, to meet production with this pricing schedule, labor demand in the state is given by

$$L_s^{Demand}(w_s) = w_s^{-\sigma} A_s^{\sigma-1} Y_g$$

Every state owns a continuum of plots of land, with heterogeneous costs to occupying each plot. The lowest cost plots are occupied first. Defining the number of plots needed as

H , the cost to occupying the marginal plot is $c(H) = H^{1/\beta}$, which implies that

$$H^{Supply}(r_s) = r_s^\beta$$

State-level production and land profits are returned to consumers as a nationwide lump-sum rebate π .¹²

2.1.3 Labor and Land Market-Clearing

An equilibrium within state s is a set of prices $\{w_s, r_s, p_s\}$ and allocations $\{\ell_{sk}, N_s, q_s\}$ such that individuals and firms are optimizing and the goods, land and labor markets clear.¹³ Define aggregate labor supply as individual labor supply times the share of individuals μ_{sk} in the regional market multiplied by the productivity per unit of labor supplied ψ_k .

$$L_s^{Supply}(w_s) = \left(\frac{w_s}{P}\right)^\epsilon \underbrace{\int \psi_k^{1+\epsilon} \mu_{sk} dk}_{\text{effective region pop: } \gamma_s} \quad (2)$$

Labor: Before, we derived labor supply and demand schedules as a function of individual supply and firm production. Now, we impose the labor market-clearing condition.

$$\begin{aligned} L_s^{Demand}(w) &= L_s^{supply}(w) \\ \Rightarrow w_s &= A_s^{\frac{\sigma-1}{\epsilon+\sigma}} \gamma_s^{\frac{-1}{\epsilon+\sigma}} Y^{\frac{1}{\epsilon+\sigma}} P^{\frac{\epsilon}{\epsilon+\sigma}} \end{aligned}$$

Land: In state s , there are $N_s = \int \mu_{sk} dk$ consumers. Recall that each consumer rents one plot of land, so $H_s^{Demand}(r) = N_s$. Using the land market-clearing condition, we set:

$$\begin{aligned} H_s^{Demand}(r) &= H_s^{Supply}(r) \\ \Rightarrow r_s &= N_s^{1/\beta} \end{aligned}$$

2.1.4 Indirect Utility and Comparative Statics

Now we can express the indirect utility for worker of skill ψ_k in a state s as a function of three state-specific parameters: productivity A_s , effective labor γ_s , and population N_s . The

¹²A model with durable housing stock would need to address the possibility of forward-looking behavior in housing markets. In the interests of simplicity, this channels is ruled out here.

¹³Note that we already used the goods market clearing condition in solving the producer's problem.

expression for indirect utility is:

$$v(A_s, \gamma_s, N_s, \psi_k) = \underbrace{\psi_k^{1+\epsilon} \left(A_s^{\frac{\sigma-1}{\epsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma}{\epsilon+\sigma}} Y_g^{\frac{1}{\epsilon+\sigma}} \gamma_s^{\frac{-1}{\epsilon+\sigma}} P^{\frac{\epsilon}{\epsilon+\sigma}} \right)^{1+\epsilon}}_{\text{income - effort cost}} \frac{P^{-(1+\epsilon)}}{1+\epsilon} \underbrace{- (\pi - N_s^{1/\beta}) P^{-1}}_{\text{profits-land price}} \quad (3)$$

The proof is given in Appendix A.2. One particularly important term for our subsequent analysis is γ_s , the region's "effective" population. We can decompose γ as

$$\gamma_s = \int \psi_k^{1+\epsilon} \mu_{sk} dk = \underbrace{\left(\frac{\int \psi_k^{1+\epsilon} \mu_{sk} dk}{\int \mu_{sk} dk} \right)}_{\text{average human capital}} \times \underbrace{\int \mu_{sk} dk}_{N_s \text{ population}}$$

One comparative static of interest is a population inflow that does not change the skill mix of the state. A population inflow pushes down wages and per capita income because of downward-sloping labor demand.

$$\varepsilon_{\text{population}}^{\text{per cap income}} = \frac{-(1+\epsilon)}{\sigma+\epsilon} < 0 \quad (4)$$

Another comparative static of interest a change in the state's average skill level, holding constant the total population. An increase in the state's skill level will mechanically raise per capita income.

$$\varepsilon_{\text{avg human capital}}^{\text{per cap income}} = \frac{\sigma-1}{\sigma+\epsilon} > 0 \quad (5)$$

These two elasticities summarize the impact of migration on income per capita and are formally derived in Appendix A.3. We return to them in the context of the calibration in Section 5.1.1.

2.2 Cross-state migration and dynamics

We add a time dimension to the model, in order to consider migration to Productiveville from a less productive place, Reservationville. We do not allow saving, so the model in the previous section describes the within-state equilibrium in each period. In Proposition 1, we characterize the impact of migration at date t on convergence, and in Proposition 2, we consider convergence dynamics over time.

Each period, workers draw stochastic i.i.d. moving costs with realization λ and then decide whether to move today or whether to stay with the possibility of moving in the future. We assume that moving cost λ is scaled by $\psi_k^{1+\varepsilon}$ such that it is proportional to time costs.¹⁴ Using the expression for indirect utility in equation 3, an agent's rule is to move if

$$\underbrace{\Lambda(A_P, \gamma_P, N_P, \psi_k)}_{\text{Value Function in Productiveville}} \geq \underbrace{\Lambda(A_R, \gamma_R, N_R, \psi_k)}_{\text{Value Function in Reservationville}} + \underbrace{\lambda \psi_k^{1+\varepsilon}}_{\text{Moving Cost}} \quad (6)$$

The cutoff λ_k^* is implicitly defined when equation 6 holds with equality.

Proposition 1

If at date t (1) income in Productiveville is sufficiently high ($v_P > v_R$),
 (2) land supply in the Productiveville is perfectly elastic ($\beta \rightarrow \infty$), and
 (3) average effective labor per capita is at least as high in Productiveville as in Reservationville ($\gamma_P/N_P \geq \gamma_R/N_R$),
 then migration generates per capita income convergence in that period.

Proof See Appendix A.4.

Next, we explore the implications of this model for convergence dynamics over time. We show that for some distribution of moving costs and parameters, forward-looking migration can generate constant rates of convergence. In this way, the model can match the observed sustained rates of income convergence before 1980.

Proposition 2

Assume ($\gamma_P/N_P = \gamma_R/N_R$). Then for some values of $\sigma, \varepsilon, A, \mu_{p,t_0}$, and μ_{R,t_0} , there is a distribution of moving costs $f(\lambda)$ such that the rate of convergence $\frac{d \ln(y_{Pt}) - d \ln(y_{Rt})}{\ln(y_{Pt-1}) - \ln(y_{R,t-1})}$, and the rate of directed migration $\frac{d \ln(N_{Pt}) - d \ln(N_{Rt})}{\ln(y_{Pt-1}) - \ln(y_{R,t-1})}$, are constant for all $t > 0$.

Proof See Appendix A.5.

¹⁴Conceptually, this assumption makes sense if the primary costs of moving are the search costs for a new job and the lost social networks for finding jobs in one's home state, both of which are proportional to time, not dollars.

The evidence presented in the introduction is consistent with Propositions 1 and 2: population flows from poor to rich places (“directed migration”) coincided with income convergence, and the end of income convergence coincided with the end of directed migration. Moreover, the rate of income convergence was relatively stable over a long period of time before 1980.

2.3 Migration by Skill Type

Next, we consider the impact of land supply elasticity on migration flows. High land prices make Productiveville particularly unattractive to low-skill workers. To show this mechanism transparently, we simplify the model in two ways. First, we abstract from the income determination in Reservationville and treat it simply as a location offering an agent of skill type k a fixed flow utility $\psi^{1+\epsilon}\Omega + 1 + \pi$. Second, we characterize migration flows in a static setting.

Define the flow gain from moving from this new, fixed flow utility location when land supply is completely elastic as:

$$\Delta(k) = \psi_k^{1+\epsilon} \left(A_{\text{Productiveville}}^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \gamma_{\text{Productiveville}}^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \varphi - \Omega \right)$$

where $\varphi = \left(\left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma}{\epsilon+\sigma}} Y_g^{\frac{1}{\epsilon+\sigma}} P_{\epsilon+\sigma}^{\frac{\epsilon}{\epsilon+\sigma}} \right)^{1+\epsilon} \frac{P^{-\epsilon}}{1+\epsilon}$ is a national constant that does not depend on state s .¹⁵ A potential migrant from group k adheres to the following decision rule:

$$\text{Move if } \underbrace{\Delta(k)}_{\Delta \text{Nominal Income}} - \underbrace{(N_{\text{Productiveville}}^{1/\beta} - 1)}_{\Delta \text{Housing Price}} > \lambda_k^* \quad (7)$$

Proposition 3

If Productiveville is attractive in nominal income terms ($\Delta(k) > 0$), then we can characterize migration flows to Productiveville as a function of the land supply elasticity β , and the normalized gain from moving $\tilde{\Delta}(k) = \log(\Delta(k) + 1) / \log(N_{\text{Productiveville}})$.

¹⁵Note that the π terms have cancelled in this expression.

	Land Supply Elasticity	Migration Flows
$Mig(\beta)$	$\beta \rightarrow \infty$	Mig indep of skill
	$\beta > \tilde{\Delta}(1)$	In-Mig by all types
	$\beta \in [\tilde{\Delta}(k-1), \tilde{\Delta}(k)]$	$\begin{cases} \text{In-Mig by skills} \geq k \\ \text{Out-Mig by skills} < k \end{cases}$
	$\beta < \tilde{\Delta}(K)$	Out-Mig by all types

Proof See Appendix A.6.

This proposition shows how different values of the land supply elasticity affect migration by skill type. When the land supply elasticity is high and land prices are low, all skill types migrate to Productiveville because of its higher productivity. When the land supply elasticity falls, land prices rise. As land prices rise, since land is a larger share of low-skilled consumption, low-skilled types are differentially discouraged from moving to Productiveville.¹⁶ Indeed, for some values of land prices, the low-skilled types move out while the high-skilled types continue to migrate to Productiveville.

3 Motivating Facts on Housing Prices and Migration

In this section, we highlight four stylized facts on the evolution of the flows of and returns to migration in the U.S. These facts motivated the model laid out in the previous section and its emphasis on the elasticity of housing supply. These facts are:

1. The differences in housing prices between rich and poor states have grown relative to the differences in incomes.
2. The “real” returns to living in productive places used to be large for all workers. They have since fallen differentially for low-skill workers.

¹⁶In a more complicated model with home ownership, low-skilled Productiveville residents would benefit from housing price appreciation. However, because their real income in the Productiveville (income - user cost of housing) would be low relative to their potential real income in other places, they would sell their houses and move to lower cost places.

3. High-skilled workers continue to move to areas with high nominal income. Low-skilled workers are now moving to areas with lower nominal income, but high real income net of housing costs.
4. Migration used to contribute to human capital levels converging across states, but this process has slowed substantially.

These facts suggest that changes in the housing markets and the elasticity of housing supply are important factors.

3.1 Income Differences and Housing Prices

In the last fifty years, there has been a shift in the relationship between prices and incomes across states. Figure 3 plots the relationship between log income and log housing prices in the 1960 and 2010 Census/ACS. The slope of this relationship more than doubles over this period. In 1960, housing prices were 1 log point higher in a state with 1 log point higher income. In 2010, housing prices are 2 log points higher in a state whose income is 1 log point higher. The difference in prices between the rich areas and the poor areas has grown wider relative to the differences in income. The change in this relationship is not specific to the year 2010 or to the recent recession. In the bottom panel of Figure 3, we plot this relationship in the decennial census with the years 1960 and 2010 highlighted. Although there is a spike in 1990 due to the housing bubble in the prior decade, there is a clear upward trend in the slope.

3.2 Returns to Living in Rich Areas

We test for changing returns by examining the relationship between unconditional average income in a state and skill-group income net of housing prices (Ruggles *et al.* (2010)). The baseline results of this specification are shown in Figure 4. With i indexing households and s indexing state of residence, we show the regression:

$$\underbrace{Y_{is} - P_{is}}_{\text{Real Income}} = \alpha + \beta_{\text{lowskill}} \underbrace{Y_s}_{\text{Nominal Income}} + \beta_{\text{highskill}} Y_s \times S_{is} + \eta S_{is} + \gamma X_{is} + \varepsilon_{is}$$

where Y_{is} is household wage income, P_{is} is a measure of housing costs defined as 12 times the monthly rent or 5% of house value for homeowners, and S_{is} is the share of the household

that is high-skilled, and Y_s is the mean nominal wage income in the state.¹⁷ The left-hand side of the equation is a household-level measure of real income. This measure is regressed upon the state-level mean household income and the interaction of nominal income with the share of high-skilled workers in the household. In 1940, both lines have a strong positive slope, which shows that both groups gain from living in richer areas. In 2010, the slope for low-skilled workers is flatter. They no longer benefit as much when living in the richer states.

Appendix Table 3 shows the evolution of $\beta_{highskill}$ and $\beta_{lowskill}$ decade by decade. These coefficients measure the returns to an agent of low-skilled or high-skilled types to living in a state that is one dollar richer. For example, $\beta_{lowskill}$ is 0.88 in 1940, meaning that for low-skill workers, real income was \$0.88 higher in states with \$1.00 higher nominal income. $\beta_{lowskill}$ shows a secular decline from 1970 forward. By 2010, the real income of low-skilled households is only \$0.36 higher for each \$1 increase in nominal income. States with high nominal income no longer appear to offer substantially higher real income to low-skilled households.

The decade-specific coefficients on $\beta_{highskill}$ show a different pattern. In 1940 and 1960, high-skill and low-skill households had similar returns to migrating. Over time, though, $\beta_{highskill}$ becomes increasingly important, and is 0.61 by 2010. In other words, real income is three times more responsive to nominal income differences by state for high-skilled households than for low-skilled households. As the price-income ratio rose, the returns to living in high income areas for low-skilled households have fallen dramatically when housing prices rose, even as they have remained stable or grown for high-skilled households. These patterns could reflect the differential impact of an increase in housing prices, the regional impact of SBTC, or some combination of the two; additional analysis provides evidence of an independent role for housing prices.¹⁸

¹⁷Income net of housing cost is a household-level variable, while education is an individual-level variable. We conduct our analysis at the household level, measuring household skill using labor force participants ages 25-65. A person is defined as high-skilled if he or she has 12+ years of education in 1940, and 16+ years or a BA thereafter. The household covariates X_{is} are the size of the household, the fraction of household members in the labor force who are white, the fraction who are black, the fraction who are male, and a quadratic in the average age of the adult household members in the workforce.

¹⁸Appendix Table 3 reports the results of two robustness checks. First, to reduce the bias arising from the endogeneity of state of residence, we also provide instrumental variables estimates using the mean income level of the household workers' state of birth as an instrument. To be precise, we estimate $Y_{is} - P_{is} = \alpha + \beta_{lowskill}\hat{Y}_s + \beta_{highskill}\hat{Y}_s \times S_{is} + \gamma X_{is} + \varepsilon_{is}$, using $Y_{s,birth}$ and $Y_{s,birth} \times S_{is}$ as instruments for the two endogenous variables \hat{Y}_s and $\hat{Y}_s \times S_{is}$. Second, we demonstrate that housing costs have differentially changed housing prices in high nominal income places for low-skilled workers.

3.3 Migration Choices by Skill

Next, we examine the extent to which people moved from low to high income places. We estimate income in both nominal terms, and using the real income measure developed above. We estimate net migration using the Census question “where did you live 5 years ago?”, which was first asked in 1940 and last asked in 2000. We use the most detailed geographies available: State Economic Areas in 1940 (467 regions) and migration PUMAs in 2000 (1,020 regions).

In Figure 5, we examine migration patterns from 1935 to 1940. As is evident from the graphs, both high-skilled and low-skilled adults moved to places with higher nominal income.¹⁹ The same relationship holds true for real income. These results are similar to work by Borjas (2001), who finds that immigrants move to places which offer them the highest wages. In Figure 6, we examine migration patterns from 1995 to 2000. Although high-skilled adults are still moving to high unconditional nominal income locations, low-skilled adults are actually weakly migrating *away* from these locations.²⁰ This finding sharply contrasts with the results from the earlier period in which there was directed migration for both groups to high nominal income areas.

It is an apparent puzzle that low-skilled households would be moving away from productive places. However, this seeming contradiction disappears when we adjust income to reflect the group-specific means net of housing prices. Low real wages in high nominal income areas have made these areas prohibitively costly for low-skilled workers. Changes in observed migration patterns are consistent with the changes in the returns to migration shown above.²¹

¹⁹Migration and education are person-level variables, while income net of housing cost is a household-level variable. We conduct our analysis at the individual level, merging on area-by-skill measures of real income. To construct area-by-skill measures, we define households as high-skilled if the adult labor force participants in said household are high-skilled, and as low-skilled if none of them are high-skilled. For ease of analysis, we drop mixed households, which account for roughly 15% of the sample in both 1940 and 2000. Then, we compute $Y - P$ for each household and take the mean by area-skill group. The modest non-linearity amongst high-income places apparent in the 1940 results is due to Chicago and New York, both of which are very large cities that were hit hard by the Great Depression and failed to attract as many migrants as predicted.

²⁰Young *et al.* (2008) similarly show that from 2000 to 2006, low-income people migrated out from New Jersey, while high-income people migrated in.

²¹The specifications shown in Figures 5 and 6 involve some choices about how to parameterize housing costs and which migrants to study. Appendix Table 4 reports four robustness checks: doubling housing costs for the real income measure, excluding migrants within-state, using only whites, and using a place of birth migration measure. In 1940, all slopes are positive, and most are statistically significant. In 2000, all slopes are positive and statistically significant for high-skilled workers. For low-skilled workers, the coefficients broadly fit the patterns in Figure 6, although only sometimes are statistically significant.

3.4 Human Capital Flows

We now examine the effect of migration flows on aggregate human capital levels. We present evidence that the transition from directed migration to skill sorting appears to have substantially weakened human capital convergence due to migration. We follow the growth-accounting literature (e.g. Denison (1962), Goldin and Katz (2001)) in using a Mincer-style return to schooling to estimate a human capital index in the IPUMS Census files. Let k index human capital levels (education interacted with race/ethnicity), and let i index individuals. Within a year t , we estimate the returns to human capital using the specification

$$\log Inc_{ik} = \alpha_k + X_{ik}\beta + \varepsilon_{ik}$$

where Inc_{ik} is an individual's annual income, and X_{ik} includes demographic covariates and state fixed effects.²² To give the reader a sense of differences in human capital over time, Appendix Table 5 reports income and wage premia by year and education group. Define predicted income as $\widehat{Inc}_k = \exp(\hat{\alpha}_k)$ and $Share_{ks}$ as the share of people in human capital group k living in state s . Under the assumption of a fixed national return to schooling, a state's skill mix and these coefficients can be used to estimate its human capital. A state-level index is

$$\text{Human Capital}_s = \sum_k \widehat{Inc}_k \times Share_{ks}$$

Our research design exploits the fact that the Census asks people about both their state of residence and their state of birth. The state of birth question provides us with a no-migration counterfactual for the state-level distribution of human capital.²³ We can then compute the change in the human capital index due to migration as

$$\Delta HC_s = \sum_k \underbrace{\widehat{Inc}_k Share_{ks, \text{residence}}}_{\text{Realized Human Capital Allocation}} - \underbrace{\widehat{Inc}_k Share_{ks, \text{birth}}}_{\text{No-Migration Counterfactual}}$$

²²Skill level k is defined as the interaction of completed schooling levels (0 or NA, Elementary, Middle, Some HS, HS, Some College, College+), a dummy for black and a dummy for Hispanic. X_{ik} includes other demographic covariates (dummy for gender and a dummy for foreign born) and a fixed effect by state of residence. Admittedly, this framework does not allow for selection among migrants along other unobservables. This model of fixed national return to schooling is conservative, though, given the substantial literature showing that the South had inferior schooling quality conditional on years attained (e.g. Card and Krueger (1992)). Thus this measures is, if anything, likely to underestimate the human capital dispersion across states.

²³We limit the sample to people aged 25 and above so that we can measure completed education. To the extent that people migrate before age 25 (or their parents move them somewhere else), we may pick up migration flows from more than twenty years ago.

Next, we take the baseline measure of what human capital would have been in the absence of migration ($HC_{s,birth}$) and examine its relationship with how much migration changed the skill composition of the state (ΔHC_s). Specifically, we regress

$$\Delta HC_s = \alpha + \beta HC_{s,birth} + \varepsilon_s$$

Figure 7 shows the results of this regression for different years in the U.S. Census. We estimate a slope of $\hat{\beta} = -0.42$ in the 1960 Census. Of the human capital dispersion by state of birth, migration of low human capital workers to high human capital places was sufficient to eliminate 42% of the disparities in human capital. By 2010, migration would have eliminated only 12% of the remaining disparity.

4 A Panel Measure of Housing Regulations

These stylized facts suggest that changes in housing prices were an important contributor to changing migration and convergence patterns. The model formalized this idea and highlighted the importance of changes in the elasticity of housing supply. In this section, we explore the role of housing prices and the elasticity of housing supply directly. We introduce a new measure of housing regulations, corresponding to β in the model. Empirical work has shown tight links between prices and measures of land use regulation in the cross-section.²⁴ Propositions 1 and 2 showed that migration could drive convergence. Proposition 3 of our model showed that migration patterns depended crucially on β , the land supply elasticity. Here, we introduce a new panel measure for housing price regulations based on state appeals court records and validate it against the existing cross-sectional regulation measures. This measure is, to the best of our knowledge, the first panel of housing supply regulations in the United States. It offers the first direct statistical evidence that regulation increases covary directly with price increases.²⁵ We use this measure to test for our entire causal chain by showing that housing supply constraints reduce permits for new construction, raise prices, lower net migration, slow human capital convergence and slow income convergence. Finally, we use this measure to show that convergence should slow the most in places with housing

²⁴Examples include Glaeser *et al.* (2005a), Katz and Rosen (1987), Pollakowski and Wachter (1990), Quigley and Raphael (2005), and Rothwell (2012) using US data. See Brueckner and Sridhar (2012) for work on building restrictions in India.

²⁵We think that aggregate changes in density in the U.S. are unlikely to explain shifts in housing supply. Because the heterogeneity in densities across places greatly exceeds the aggregate change in density, it seems unlikely that substantial parts of America have reached a maximum building density. See Appendix Figure 3 for details.

supply constraints.

4.1 Measure Construction and Reliability

Our measure of land use regulations is based upon the number of state appellate court cases containing the phrase “land use” over time.²⁶ The phrase “land use” appears 42 times in the seminal case *Mount Laurel* decision issued by the New Jersey Supreme Court in 1975. Municipalities use a wide variety of tactics for restricting new construction, but these rules are often controversial and any such rule, regardless of its exact institutional origin, will likely be tested in court. This makes court decisions an omnibus measure which captures many different channels of restrictions on new construction. We searched the state appellate court records for each state-year using an online legal database. Gennaioli and Shleifer (2007) describe how the common law gradually accumulates a set of precedents over time. To ensure that our regulation measure captures the stock of restrictions rather than the flow of new laws, we set:

$$Regulation_{s,t} = \frac{\sum_{k=1920}^t LandUseCases_{sk}}{\sum_{k=1920}^t TotalCases_{sk}}$$

We begin our measure in 1940, and initialize it by using case counts from 1920-1940.²⁷ Appendix Table 6 shows selected values for each state. One immediate result from constructing this measure is that the land use restrictions have become increasingly common over the past fifty years. Figure 8 displays the national regulation measure over time, which rises strongly after 1960 and grows by a factor of four by 2010.²⁸

We test the reliability of this measure using cross-sectional land-use surveys. The first survey, from the American Institute of Planners in 1975, asked 21 land use-related questions of planning officials in each state (The American Institute of Planners (1976)).²⁹ To build a summary measure, we add up the total number of yes answers to the 21 questions for each

²⁶We show similar results for the phrase “zoning” in Appendix Table 7.

²⁷Note that because we measure land use relative to the total case load in the state, this variable is a measure of land use regulation relative to the total volume of legal activity. The states that have lots of regulation and legal activity on other fronts (e.g. New York) will not be marked as high regulation by this measure. Other states, that face little legal activity overall, such as Wyoming, may be classified as high regulation even though there are few constraints to new construction.

²⁸One of the first to document this increase was Ellickson (1977). Fischel (2004), Kahn (2011), and Schleicher (2012) explore the social, political, and technological reasons underlying this rise.

²⁹Saks (2008) also uses this survey as a measure of land use regulations.

state. As can be seen in Figure 8, the 1975 values of our measure are strongly correlated with this measure. Similarly, our measure is highly correlated with the 2005 Wharton Residential Land Use Regulation Index (WRLURI).³⁰ Finally, we compute residual Wharton regulation by fitting $WRLURI_{s,2005} = \alpha + \beta_1 API_{s,1975} + \varepsilon_s$ and then calculating $WRL\tilde{U}RI_{s,2005} = WRLURI_{s,2005} - \beta_1 API_{s,1975}$. We regress $WRL\tilde{U}RI_{s,2005}$ on our regulation measure and again find a strong positive correlation. This is further evidence that the measure truly captures the panel variation in regulations.

We examine the effect of the regulation measure on housing prices, with state and year fixed effects, in Appendix Figure 4. We find that that changes in regulation are highly correlated with changes in prices.³¹

4.2 Using Regulations to Test the Model

Having established that our regulation measure is a good proxy for housing supply elasticity, we test its direct effect on the convergence relationship. Before turning towards regressions, we first demonstrate the effect of land-use regulations on convergence graphically. Figure 9 shows differential convergence patterns among the high and low elasticity states. The convergence relationship within the low regulation, high elasticity states remains strong throughout the period. Conceptually, we can think of this group of states as reflecting Propositions 1 and 2, with within-group reallocations of people from low-income states to high-income states. In contrast, the convergence coefficients among states with low elasticities or tight regulations display a pronounced weakening over time. The patterns within this group reflect the skill-sorting of Proposition 3, where only high-skilled workers want to move to the high-income states.³²

We now turn towards regressions and explore the effect of regulations more rigorously on the entire convergence mechanisms described above. We divide our state-year observations into high and low regulation bins based on a fixed cutoff of the median value of the regulation index in 2005. This division allows us to exploit variation in regulation across years and

³⁰To construct state-level measures, we weighted the metro estimates in Gyourko *et al.* (2008) by 1960 population and imputed from neighbors where necessary.

³¹However, see Davidoff (2010) for a dissenting view about the impact of regulations on housing prices using cross-sectional data. Davidoff writes “Unfortunately, a panel of regulations is not available, so there is no way to determine if time series changes in regulations are associated with changes in supply.”

³²As a robustness check, we divide the states according to a measure of their housing supply elasticity based upon land availability and the WRLURI constructed by Saiz (2010). Note that unlike our regulation measure, the Saiz measure is fixed, so the groupings of states are stable. The results are shown in Appendix Figure 5.

states while still producing interpretable coefficients. In total, 12.5% of states fall in to the high regulation bin in 1970, and by 1990 this bin contains 40% of the states. We explore robustness to using the continuous measure and state-specific cutoffs in Appendix Table 7.

Our baseline specifications are of the following form:

$$Y_{s,t} = \alpha_t + \alpha_t \mathbf{1}[Reg_{s,t} > Reg_{Med,2005}] + \beta Inc_{s,t-20} + \beta_{\text{high reg}} Inc_{s,t-20} \times \mathbf{1}[Reg_{s,t} > Reg_{Med,2005}] + \varepsilon_{s,t} \quad (8)$$

The coefficients of interest, β and $\beta_{\text{high reg}}$, measure the effect of lagged income in low and high regulation state-years and are reported in Table 2.

The first left-hand side variable considered with this specification in Column 1 is housing permits issued relative to the housing stock. Absent land use restrictions, places with higher income will face greater demand for houses and will permit at a faster rate. Accordingly, the base coefficient on income is positive. The interaction term $\beta_{\text{high reg}}$ is negative: in the high-regulation regime there is no correlation between income and permits for new construction.

Column 2 analyzes housing prices. Whereas regulations weakened (or even reversed) the relationship between income and permitting, we expect regulations to steepen the relationship between income and prices. The uninteracted coefficient recovers the 1-to-1 relationship between log income and log prices found in the first panel of Figure 3. The interaction term indicates that the relationship between income and prices in high regulation state-years is 1.7. More of the income differences are capitalized in to prices in the high regulation regime.³³

Column 3 explores what effect regulations have on the link between income and population growth. In our model, states with high income per capita will draw migrants when the elasticity of housing supply is large. The uninteracted coefficient is 2.0, which is similar to the coefficients in the first panel of Appendix Table 1. When the elasticity of supply is low, the model predicts that housing prices make moving prohibitively costly and directed migration ceases. Consistent with the model, the interaction coefficient is large and negative (-2.8).

In our model, Proposition 1 showed that, when housing supply was perfectly elastic, migration would be skill-neutral and undo any initial human capital advantage held by productive places. We find in Column 4 that the uninteracted coefficient is sizeable and negative reflecting convergence in human capital in the unconstrained regime. The interaction coefficient is sizeable and positive, indicating that human capital convergence slows among high regulation observations.

³³Similarly, Saks (2008) finds that employment demand shocks are capitalized into prices rather than quantities in the high regulation regime.

Finally, Column 5 brings this analysis full circle by directly looking at the effect of high regulations on the convergence relationship. The uninteracted coefficient (-2.3) captures the strong convergence relationship that exists absent land use restrictions shown in the early years in Figure 1. However, the interaction coefficient is large and positive (3.2). This finding indicates that the degree of convergence among states in periods of high regulation is virtually non-existent.

Table 4 tightly links the theory from Section 2 to the observed data. The first row of coefficients describe a world where population flows to rich areas, human capital converges across places, and regional incomes converge quickly as in Propositions 1 and 2. The second row of coefficients is consistent with the low-elasticity regime described in Proposition 3, with increased capitalization and no net migration or convergence.

5 Robustness Checks and Alternative Explanations

Our analysis thus far has explored the role that housing regulations have played in changing skill-specific labor mobility and income convergence. Of course, other factors are likely to affect both patterns and it is important to consider whether these factors may be driving the results in the previous section. We divide our analysis of the interaction of these factors with the housing regulation channel into the following two pieces. First, we explore whether *observed* migration patterns can account for a quantitatively meaningful portion of convergence. We do this in a calibrated version of our model under different assumptions about the elasticity of local labor demand and labor supply. We appreciate that these are controversial parameters, and therefore feel that it is worthwhile to assess what conditions are needed for the theoretical forces in our model to meaningfully contribute to the changing trends documented above. We also explore the available evidence for other factors that contributed convergence before 1980, such as the accumulation of physical capital. In the following section, we explore other recent changes that contributed to the slowing of convergence and which might be correlated with increased regulation. In particular, we focus on the role of SBTC, and other unobserved shocks that may bias our estimates. We find that SBTC worked in tandem with housing supply constraints to lower the rate of convergence. Nevertheless, our evidence suggests that SBTC alone can not account for changes in regional migration and income growth, and that changes in housing markets are an important part of the story.

5.1 Migration's Role in Past Convergence

An influential literature has sought to explain income convergence in terms of physical capital convergence.³⁴ In the basic neo-classical model, labor is immobile and differences in income occur because of differences in physical capital per worker. The returns to capital are higher in capital-poor places than in capital-rich places, which drives faster growth and convergence.

Due to the lack of data, this mechanism has been difficult to assess empirically across U.S. states. In contrast, we show that the *observed* labor movements can account for a quantitatively meaningful component of income convergence trends in the context of our model under various parameter values. Additionally, we find that the data to evaluate the physical capital convergence hypothesis are too sparse to properly explore its role in income convergence. The data that do exist provide mixed support for the simple neo-classical model. Nevertheless, alternative models might give a larger role for physical capital, and additional data on this topic would be valuable.

5.1.1 Calibration – The Role of Changes in Labor

The model developed in Section 2 demonstrates two channels through which migration can affect income convergence. The model delivers a parsimonious, intuitive equation for the effect of migration on convergence. In terms of the model, a state's income per capita is given by $y_i = Y_i/N_i = \left(\pi + A^{\frac{(1+\epsilon)(\sigma-1)}{\epsilon+\sigma}} P^{-\epsilon} Y^{\frac{(1+\epsilon)}{\epsilon+\sigma}} \right) \left(\int \psi_k^{1+\epsilon} \mu_{ik} dk \right)^{\frac{\sigma-1}{\sigma+\epsilon}} \setminus \int \mu_{ik} dk$. Taking logs, a change in per capita income can be decomposed into changes in effective labor per capita and changes in population, with

$$d \log y_i = \underbrace{-\frac{(1+\epsilon)}{\sigma+\epsilon}}_{\epsilon_{\text{pop}}^{\text{per capita inc}}} \underbrace{d \log \left(\int \mu_{ik} dk \right)}_{\Delta \text{ population}} + \underbrace{\frac{\sigma-1}{\sigma+\epsilon}}_{\epsilon_{\text{avg skill level}}^{\text{per capita inc}}} \underbrace{d \log \left(\frac{\int \psi_k^{1+\epsilon} \mu_{ik} dk}{\int \mu_{ik} dk} \right)}_{\Delta \text{ effective labor per cap}}$$

The model delivers simple formulas for the effect of changes in effective labor per capita and population on per capita income ($\epsilon_{\text{pop}}^{\text{per capita inc}}$, $\epsilon_{\text{avg skill level}}^{\text{per capita inc}}$ from equations 4 and

³⁴Though the past literature has focused on capital convergence as a mechanism, Barro and Sala-i Martin (1992) were careful to note that income convergence has occurred more slowly in the data than would be predicted by their model with physical capital adjustment. We claim very slow adjustment of labor in response to regional income differentials. Since people usually have strong family and economic ties to where they grew up, as well as heterogeneous preferences over location, migration from unproductive to productive places could plausibly take a long time.

5). Note that for comparison, a neo-classical model that had exogenous labor supply and homogeneous labor would have had the smaller impact of $d \log y_i = -\frac{1}{\sigma} d \log \left(\int \mu_{ik} dk \right)$. Endogenous labor supply and human capital amplify the effects of migration. Converting to regression coefficients, we obtain:

$$\hat{\beta}_{\text{Convergence}} = \underbrace{\varepsilon_{\text{pop}}^{\text{per capita inc}}}_{\text{Directed Migration}} + \underbrace{\varepsilon_{\text{avg skill level}}^{\text{per capita inc}}}_{\text{Skill Convergence}} \quad (9)$$

Model Parameters The model allows us to convert estimates of direct migration (Figure 2) and human capital convergence (Figure 7) into estimates of income convergence (Figure 1). This mapping, however, requires us to take a stance on the elasticity of labor supply and the elasticity of substitution across regional goods, or implicitly, on the elasticity of local labor demand. These are controversial parameters, and so we consider a wide range of possible values in Appendix Table 9.³⁵ In order to make the calibration results easier to follow, we report our baseline estimates in Table 3, using $\sigma = \{2, 6\}$ ³⁶ and $\{\epsilon = 0, 0.6\}$.³⁷

Sample Moments Next, we turn to estimating the magnitudes of population flows and human capital flows.

Population Channel: We estimate all population flows in terms of 20-year windows, corresponding to the 20-year windows used for convergence in Figure 1. To estimate the effect of initial income on net migration, we regress

$$\log Pop_t - \log Pop_{t-20} = \alpha + \eta \log y_{t-20,s} + \varepsilon_s$$

³⁵Rothstein (2010) considers a calibration exercise that similarly requires assumptions on these two parameters, albeit in a very different context. We follow his lead and consider a wide range of parameter values, including labor supply elasticities $\epsilon = \{0, 0.3, 0.6, 1, 2\}$ and elasticities of substitution $\sigma = \{1, 2, 6, 10, \infty\}$. These values allow us to consider elasticities of labor demand ranging from 0 - 1.

³⁶Nakamura and Steinsson (2011) use an elasticity of substitution between domestic and out-of-state goods equal to 2, and Ruhl (2008) reports an elasticity with respect to long-run shocks equal to 6. This range, when paired with the values for the elasticity of labor supply above, yields elasticities of labor demand $\varepsilon_{\text{pop}}^{\text{per capita inc}}$ between -0.17 and -0.62. This range contains the empirical labor demand elasticity estimates of Rothstein (2008) and Acemoglu *et al.* (2004), and we consider a larger range of estimates in Appendix Table 9.

Additionally, in Appendix Figure 6, we plot the reduced form relationship between the twenty-year growth rates in state average income and population, controlling for state and year fixed effects. Though this relationship is obviously confounded towards zero by shocks that both increase income and draw population, the estimated slope (-0.20) is within our baseline calibration range.

³⁷Consider the following balanced growth path preferences, $\arg \max_{\ell_k} \ln(c) - \frac{\ell_k^{1+1/\epsilon}}{1+1/\epsilon}$, under the assumption that the elasticity of housing supply $\beta \rightarrow \infty$. Log utility implies static labor supply, or $\epsilon = 0$. Chetty (2012) characterizes $\varepsilon^{\text{Labor Supply}} = 0.6$ as the consensus estimate for the sum of extensive and intensive labor supply elasticities based on a meta-analysis of approximately thirty studies with quasi-experimental variation in tax rates.

The results are shown in Appendix Table 2. For the 20-year periods ending in 1950, 1960, 1970 and 1980, there was substantial net migration towards higher-income states. Averaging the coefficients from this period, we find that a doubling of initial income raised the annual net migration rate to a state by 1.26 percentage points.

Human Capital Channel: We estimate human capital flows by examining the change in human capital due to migration relative to baseline levels and its relation to initial income. Specifically, we regress

$$\log(\Delta HC_{ts} + HC_{t-20,s}) - \log(HC_{t-20,s}) = \alpha + \varphi \log y_{t-20,s} + \varepsilon_{st}$$

Details on the construction of this human capital measure are in Appendix B.1. The results are shown in the second row of Appendix Table 8. Human capital used to flow strongly from poor states to rich states, as estimated in the 1960 Census, but all of the estimates using the 2010 American Community Survey suggest that human capital levels are diverging slightly due to migration. Until now, we have worked with income as a measure of human capital. Because in our model labor supply is determined endogenously, in principle wage rates are a better analog for skill type. If we assume that reported annual hours is a measure of labor supply, we use a Mincerian regression to compute an hourly wage rate \widehat{wage}_k . In the model with endogenous labor supply, $\hat{inc}_k = \widehat{wage}_k^{1+\varepsilon}$. The results are shown in Appendix Table 8, rows (3) through (5). Higher labor supply elasticities exacerbate the impact of existing skill differences on the dispersion in the human capital stock. Therefore, convergence in skill levels then implies a greater degree of convergence in the human capital stock.³⁸

Finally we can use equation 9 to predict income convergence before 1980, $\hat{\beta}_{\text{Convergence}}^{pre1980}$, as well as the amount of income convergence predicted today, $\hat{\beta}_{\text{Convergence}}^{2010}$. Table 3 shows the results of this calibration, in which we estimate the change in the predicted convergence rate, $\hat{\beta}_{\text{Convergence}}^{2010} - \hat{\beta}_{\text{Convergence}}^{pre1980}$.³⁹ We find that for most parameter values, observed migration can account for much of convergence and that changes in convergence due to mi-

³⁸Supplemental analysis, not reported here, sheds more light on the mechanisms for human capital convergence. We find that the strong mid-century human capital convergence came from low-skill workers moving to high-skill states. Changes in black migration in particular are an important component of changes in human capital convergence; this dovetails with historical accounts of the Great Migration as a search for higher wages. Changes in foreign-born migration are less important. Finally, although black migration was relatively important for human capital convergence, we find that dropping black migration leaves our estimates of “directed migration” (the number of people moving from poor to rich places) relatively unaffected.

³⁹Changing the value of the labor supply elasticity has three different effects on the calibration. First, more responsive labor supply means that wages will change more in response to a population inflow, thereby amplifying $\varepsilon_{\text{pop}}^{\text{per capita inc}}$. Second, more responsive labor supply means that when a low-skilled person moves to a high-productivity place he or she supplies more labor. This effect dampens the human capital convergence channel $\varepsilon_{\text{avg skill level}}^{\text{per capita inc}}$. Finally, more responsive labor supply makes historic human capital

gration are of the same order of magnitude as the observed changes in the data.⁴⁰ Appendix Table 9 considers a wide range of parameter values and demonstrates that the calibrated impact is robust, though it does become less important as the elasticity of substitution across regional varieties grows arbitrarily large.

5.1.2 The Role of Capital

Past work, most notably Barro and Sala-i Martin (1992), has explored whether faster capital accumulation in poor states can explain regional convergence from 1880 to 1980 in a calibrated model. Empirical measures of the state-level capital stock are quite difficult to obtain, but three pieces of evidence suggest a more nuanced analysis of the role of capital in U.S. convergence is needed.⁴¹

First, we analyzed a panel state-level measure of capital using the Census of Wealth, which was constructed using local tax records and conducted on a roughly decennial basis until 1922 (Kuznets *et al.* (1964)). Because property taxes were the primary source of revenue for state and local governments at this time, we consider this measure to be fairly accurate. The data show mixed evidence for the capital-driven convergence hypothesis, as depicted in Appendix Figure 7. From 1880 to 1900 capital grew most quickly in *wealthy* states (which contradicts the neo-classical model), and from 1900 to 1920, capital grew most quickly in *poor* states (which is consistent with the neo-classical model). Income per capita was converging in both periods, and consistent with the explanation put forward in this paper, there was substantial net migration towards wealthy states in both periods as well.

Second, we plot a timeseries of regional bank lending rates from Landon-Lane and Rockoff (2007) in Appendix Figure 7. If the return on capital is falling in the amount of the capital stock, we should see high interest rates in poor areas and low interest rates in rich areas. In fact, we see high interest rates in the rich West, which is inconsistent with the neo-classical model. Additionally, these regional interest rates largely converged by the end of World War

differences play a larger role in historic income differences. Thus skill group changes have a larger impact on measured human capital.

⁴⁰Although our model matches the macro moments for pre-1980 income convergence and 2010 income convergence, the timeseries variation in our inputs does not exactly match the timeseries variation in the convergence rate. Directed migration falls a bit before income convergence falls, and human capital convergence persists longer than income convergence.

⁴¹Garofalo and Yamarik (2002) constructed state-level capital estimates by combining state-level industry employment composition with national industry-level capital-labor ratios. Although an important step forward, this measure characterizes the extent of convergence in industrial structure in the U.S., not the extent of capital convergence.

II. Under the neo-classical model, this finding would indicate that capital-labor ratios had converged as well.

Third, in some cases, capital may be a substitute rather than a complement for labor. Recent work by Hornbeck and Naidu (2012) suggests that an outflow of cheap black labor due to a flood in Mississippi in 1927 led to additional capital investment in the agricultural sector in subsequent years. Along similar lines, Lewis (2011) finds that increased low-skilled labor supply in the 1980s and 1990s was a substitute for capital investment in the manufacturing sector. In this case, capital investment would *amplify* the impact of changes in the labor supply rather than attenuate them. More research and more data are needed to seriously study the role of physical capital in U.S. income convergence, but the existing data are insufficient to draw strong conclusions about its role.

5.2 Other Factors Slowing Convergence

5.2.1 Different Steady States: Convergence Has Already Happened

Income gaps across states are smaller today than they were in the past. Perhaps differences in incomes today reflect steady-state differences due to permanent amenity differences. While possible, two pieces of evidence are inconsistent with this suggestion. First, a close examination of Figure 1 shows that from 1940 to 1960 there was within-group convergence among the rich states as well as among the poor states. The income differences between Connecticut and Illinois or Mississippi and Tennessee in 1940 are smaller than the differences between Connecticut and Mississippi in 1990, and yet there was substantial within-group convergence from 1940 to 1960 and much less from 1990 to 2010. Second, our analysis with the regulation measure (e.g. Figure 9) shows substantial within-group convergence in the low regulation group, suggesting that existing income differences today are sufficiently large and transitory as to make convergence possible.

5.2.2 Omitted Variable Bias and Reverse Causality

It is possible that the change in regulations coincided in space and time with another change that directly affected convergence patterns and that regulations have no direct causal effect on the convergence process. For example, high productivity areas might have experienced a negative demand shock for low-skilled workers that pushed them to emigrate and slowed convergence. If high-skilled workers are also more likely to endorse regulation, the relation-

ships demonstrated in Section 4 might be spurious. Similarly, if changes in regulations reflect attempts by forward-looking agents to capitalize on independently slowing convergence, then regulation might have no causal impact despite the evidence shown above. We test for these stories and the general possibility of omitted variable bias in a number of ways in Table 4, which explores cross-sectional convergence from 1955-1980 and 1980-2005. Columns 1 and 2 demonstrate (like Figure 9) that high and low-regulation states in 2005 had similar convergence rates before 1980, but that convergence slowed in high-regulation states after these restrictions were enacted. To address the reverse causality issue, Columns 3 and 4 demonstrate this relationship controlling for state level measures of industry and skill composition from Autor and Dorn (2013), and the coefficients are virtually unchanged.⁴² Of course, controlling for potentially confounding covariates does not address the possibility of reverse causality through unobserved channels. To investigate the plausibility of this hypothesis, we once again make use of the rich time-series variation in our regulation measure. Although regulation was low across the board in 1955, there is still cross-sectional variation in our measure for that year. This variation is predictive of subsequent increases in regulation, and the correlation between the measure in 1955 and 2005 is 0.45. Dividing the states at the median in 1955, we classify the states above the median in that year as having a latent tendency to regulate.

We explore the effect of this latent tendency to regulate on income convergence before and after 1980 in Columns 5 and 6. The table demonstrates that these states (like states with high regulation today) displayed similar convergence behavior before 1980. In the period after 1980, once these latent tendencies had been activated in the form of high regulations, these states experience a sizeable drop in their degree of income convergence.

We conduct a similar exercise in the last two columns by splitting states at the median based upon the geographic availability of develop-able land using data from Saiz (2010).⁴³ Again, the table demonstrates that states with low geographic land availability did not display different convergence behavior before 1980. In the period with tight building restrictions after 1980, however, these states also experience a reduction in their rates of income convergence. The similarity of these patterns is striking given that the correlation between our two predetermined elasticity measures – the high latent regulation dummy and low geographic land availability dummy – is only 0.33. While sizeable, these interaction coefficients are

⁴²Specifically, we control for their measures of the share of workers in routine occupations, the college to non-college population ratio, immigrants as a share of the non-college population ratio, manufacturing employment share, the initial unemployment rate, the female share, the share age 65+, and the share earning less than the 10 year ahead minimum wage. We aggregate their data to the state level via population weighting.

⁴³We population-weight metro areas using 1960 counts to derive state-wide averages.

smaller than the ones reported in Table 2. This could be consistent with pre-1980 measures of housing supply elasticity explaining some, but not all, of the decline in convergence.

This evidence suggests that housing supply restrictions played a crucial role in the end of convergence and rules out the possibility that regulations are merely a by-product of an alternate shock. Table 4 shows that if housing supply restrictions did not affect income convergence, then regulations must be correlated with a non-related convergence-ending shock, and this new shock must be correlated with states' geography and historical legal structures. Moreover, such an explanation would have to explain the fact that neither feature influenced convergence rates prior to the period of high land use regulation. Although it is possible to generate such an explanation, articulating such a story is sufficiently complicated that we feel the weight of the evidence supports a role for housing supply restrictions.

5.2.3 Skill Biased Technological Change

Conceptually, skill-biased technological change (SBTC) could slow the rate of convergence for several reasons.

Consider an increase in the skill premium. This change would have two effects on convergence rates. It mechanically widens the income gaps between richer, more educated states and poorer, less educated ones. Additionally, in our model it raises the returns to migration for high-skill workers living in low-income states. The change in the returns to migration is complementary to our supply constraints story – both forces serve to make migration to rich places more heavily weighted towards skilled workers. As for the magnitude of the mechanical effect, Autor *et al.* (2008) estimate that the college-high school premium rose from 0.40 in 1980 to 0.64 in 2000. The share of people with a BA in (henceforth “share BA”) in 1980 had a standard deviation of 3 percentage points across states, and the mechanical increase in the skill premium would have reduced the annual convergence rates by roughly 0.18. Moreover, this estimate of the mechanical effect of SBTC is overstated in the context of our framework, as the estimate incorporates the increasing concentration of high-skilled workers into productive places in the calibrated skill premium changes. Alternatively, we can calculate the mechanical contribution of SBTC using the more refined skill bins used in Section 3.4. Using this refinement, changing skill premiums would have reduced annual convergence rates after 1980 by 0.34. The observed change in annual convergence rates was 1.11, meaning that the mechanical effect of SBTC provides a significant but incomplete account for the change.⁴⁴

⁴⁴Though a national increase in the premium does not account for the entire change in convergence, there

Finally, it is possible that as skills have become more important, incomes of everyone in high share BA places would rise (Moretti (2012); Glaeser and Saiz (2004); Berry and Glaeser (2005)) due to agglomeration externalities. Under this theory, incomes would grow more quickly in these places, slowing convergence. One testable prediction which differentiates this story (a “demand shock” in productive areas story) from our housing supply constraints story comes from skill-specific migration patterns. A positive demand shock should *raise* in-migration rates for all workers. If this demand shock mostly affected skilled workers, then it should raise the migration rate for skilled workers. In contrast, a negative housing supply shock predicts sharply falling in-migration by low-skill workers and a steady or slightly declining in-migration for high-skill workers. Although information-economy cities such as San Francisco, Boston and New York offer high nominal wages to all workers (typically in the top quintile nationally), after adjusting for housing costs all three cities offer below average returns to low-skill workers (typically in the bottom decile). In Table 5, we examine the flows of low-skilled and high-skilled workers in 1980 and 2010 to high skilled states as measured in 1980. This period and independent variable were chosen to be consistent with the literature on skill agglomerations.

There has been a remarkable shift in the composition of migration to high share BA places. From 1980 to 2010, there was a large decrease in the in-migration rate of low-skilled workers to high share BA states, and no change or a small decline in the in-migration rate of high-skilled workers to high share BA states.⁴⁵ These results suggest that rising share BA in high share BA areas documented by other researchers may partially be the result of out-migration by low-skill workers and increased domestic human capital production, rather than increasing in-migration by high-skill workers. Overall, SBTC and its place-specific variants are complementary with the supply constraints story developed here. When supply is constrained, increases in demand for skilled labor serve to further slow convergence.

is evidence that the skill premium has risen most in places which have a high share of workers with a BA (Moretti (2012)). Such a correlation would raise the mechanical contribution of SBTC to the end of convergence, yet a model with SBTC alone could not reproduce the changing migration patterns documented below. Moreover, when performing a similar back of the envelope calculation to the national premium change considered above, again using the refinement in Section 3.4, we find that allowing the skill premium to vary by location has only a modest effect. The estimated mechanical decline in convergence rates changes from 0.34 to 0.38.

⁴⁵This is consistent with results reported earlier using the 5-year net migration measure. Figure 6 showed that workers with less than a BA are leaving high nominal income areas, and Figure 2 showed that population in high nominal income areas was growing more slowly than the national rate.

6 Conclusion

For more than 100 years, per-capita incomes across U.S. states were strongly converging and population flowed from poor to wealthy areas. In this paper, we claim that these two phenomena are related. By increasing the available labor in a region, migration drove down wages, reduced labor supply and induced convergence in human capital levels.

Over the past thirty years, both the flow of population to productive areas and income convergence have slowed considerably. We show that the end of directed population flows, and therefore the end of income convergence, can be explained by a change in the relationship between income and housing prices. Although housing prices have always been higher in richer states, housing prices now capitalize a far greater proportion of the income differences across states. In our model, as prices rise, the returns to living in productive areas fall for low-skilled households, and their migration patterns diverge from the migration patterns of the high-skilled households. The regional economy shifts from one in which labor markets clear through net migration to one in which labor markets clear through skill-sorting, which slows income convergence. We find patterns consistent with these predictions in the data.

To identify the effect of these price movements, we introduce a new panel instrument for housing supply. Prior work has noted that land-use regulations have become increasingly stringent over time, but panel measures of regulation were unavailable. We create a proxy for these measures based on the frequency of land-use cases in state appellate court records. First, we find that tighter regulations raise the extent to which income differences are capitalized into housing prices. Second, tighter regulations impede population flows to rich areas and weaken convergence in human capital. Finally, we find that tight regulations weaken convergence in per capita income. Indeed, though there has been a dramatic decline in income convergence nationally, places that remain unconstrained by land use regulation continue to converge at similar rates.

These findings have important implications not only for the literature on land-use and regional convergence, but also for the literature on inequality and segregation. A simple back of the envelope calculation shown in Appendix Table 10 finds that cross-state convergence accounted for approximately 30% of the drop in hourly wage inequality from 1940 to 1980 and that had convergence continued apace through 2010, the increase in hourly wage inequality from 1980 to 2010 would have been approximately 10% smaller. The U.S. is increasingly characterized by segregation along economic dimensions, with limited access for most workers to America’s most productive cities. We hope that this paper will highlight the role land use restrictions play in supporting this segregation.

References

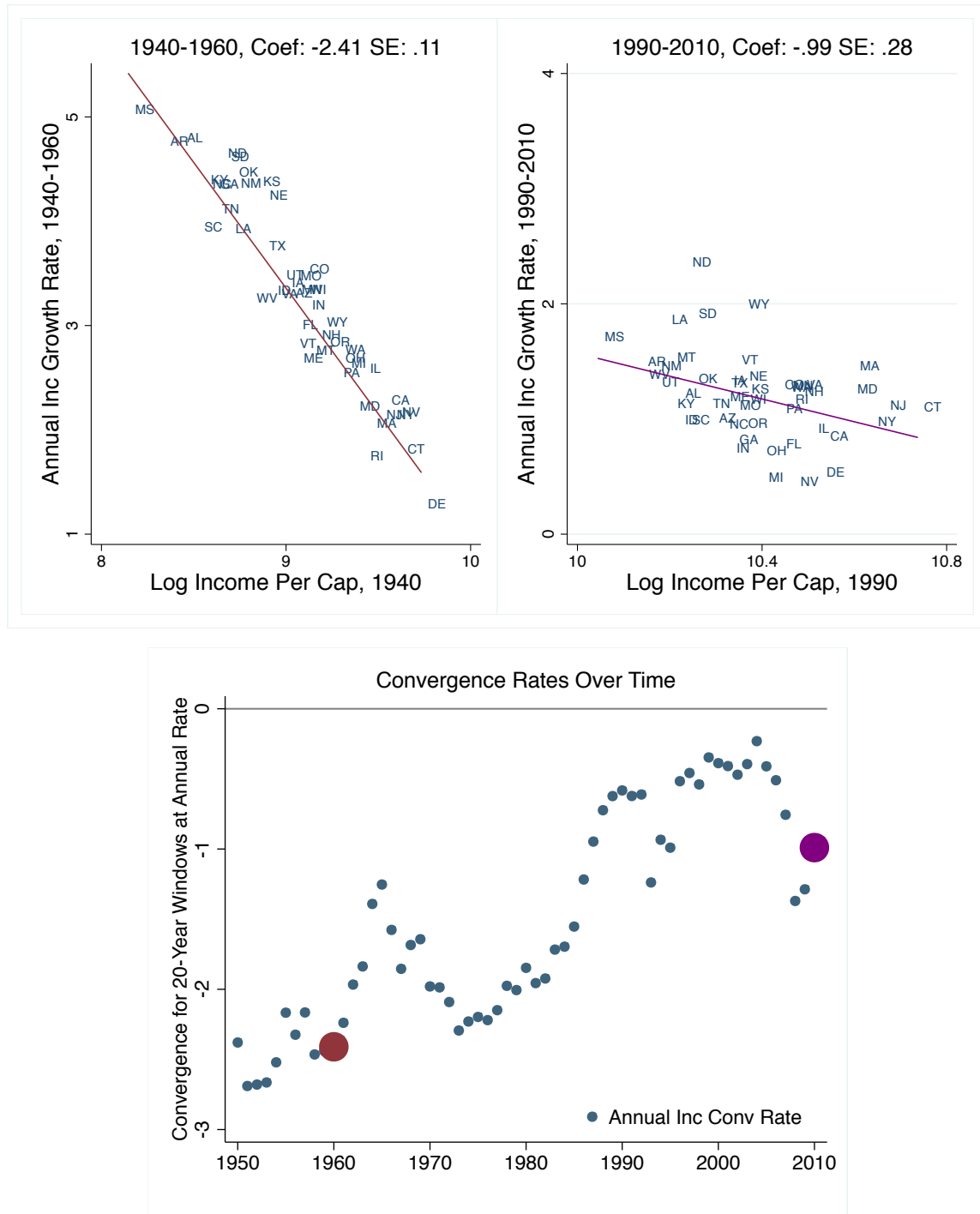
- Acemoglu, D., Autor, D. H. and Lyle, D. (2004) Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury, *The Journal of Political Economy*, **112**.
- Albouy, D. and Ehrlich, G. (2012) Metropolitan Land Values and Housing Productivity, *mimeo*.
- Angrist, J. (1996) Short-run demand for Palestinian labor, *Journal of Labor Economics*.
- Armington, P. S. (1969) A Theory of Demand for Products Distinguished by Place of Production, *Staff Papers- International Monetary Fund*, **16**, 159–178.
- Autor, D. H. and Dorn, D. (2013) The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market, *American Economic Review*.
- Autor, D. H., Katz, L. F. and Kearney, M. S. (2008) Trends in U.S. Wage inequality: Revising the Revisionists, *Review of Economics and Statistics*, **90**, 300–323.
- Barro, R. J. and Sala-i Martin, X. (1991) Convergence across States and Regions, *Brookings Papers on Economic Activity*, pp. 107–182.
- Barro, R. J. and Sala-i Martin, X. (1992) Convergence, *Journal of Political Economy*, **100**, 223–251.
- Berry, C. R. and Glaeser, E. (2005) The Divergence of Human Capital Levels across Cities, *Papers in Regional Science*, **84**, 407–444.
- Blanchard, O. J. and Katz, L. F. (1992) Regional Evolutions, *Brookings Papers of Economic Activity*, **23**, 1–76.
- Borjas, G. J. (2001) Does Immigration Grease the Wheels of the Labor Market?, *Brookings Papers on Economic Activity*, **1**, 69–133.
- Borjas, G. J. (2003) The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market, *Quarterly Journal of Economics*, **118**, 1335–1374.
- Boustan, L. P., Fishback, P. V. and Kantor, S. (2010) The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression, *Journal of Labor Economics*, **28**, 719–746.
- Braun, J. (1993) *Essays on economic growth and migration*, Ph.D. thesis, Harvard.
- Brueckner, J. K. and Sridhar, K. S. (2012) Measuring welfare gains from relaxation of land-use restrictions: The case of India’s building-height limits, *Regional Science and Urban Economics*, **42**, 1061–1067.
- Bureau of Economic Analysis (2012) State Personal Income, Tables SA1-SA3.
- Card, D. (1990) The impact of the Mariel boatlift on the Miami labor market, *Industrial and Labor Relations Review*, **43**, 245–257.
- Card, D. (2009) Immigration and Inequality, *American Economic Review: Papers and Proceedings*, **99**, 1–21.
- Card, D. and Krueger, A. B. (1992) Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States, *The Journal of Political Economy*, **100**, 1–40.
- Caselli, F. and Coleman, W. J. (2001) The U.S. Structural Transformation and Regional Convergence: A Reinterpretation, *Journal of Political Economy*, **109**, 584–616.
- Chetty, R. (2012) Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply, *Econometrica*.

- Combes, P.-P., Duranton, G., Gobillon, L. and Roux, S. (2007) Estimating Agglomeration Economies with History, Geography, and Worker Effects, in *Agglomeration Economics* (Ed.) E. L. E. L. Glaeser, The University of Chicago Press, February, pp. 15–66.
- Cortes, P. (2008) The Effect of Low-Skilled Immigration on U.S. Prices: Evidence from CPI Data, *Journal of Political Economy*, **116**.
- Crain, W. M. (2003) *Volatile States: Institutions, Policy, and the Performance of American State Economies*, University of Michigan Press, Ann Arbor, MI.
- Davidoff, T. (2010) What Explains Manhattan’s Declining Share of Residential Construction?, *Journal of Public Economics*, **94**, 508–514.
- Denison, E. F. (1962) *The Sources of Economic Growth in the United States and the Alternatives Before Us*, Committee for Economic Development, New York.
- Diamond, R. (2012) The Welfare Effects of Changes in Local Wages, Prices and Amenities Across US Cities, *Harvard Working Paper*.
- DiCecio, R. and Gascon, C. S. (2008) Income Convergence in the United States: A Tale of Migration and Urbanization, *Annals of Regional Science*, **45**, 365–377.
- Easterlin, R. A. (1958) Long Term Regional Income Changes: Some Suggested Factors, *Papers in Regional Science*, **4**, 313–325.
- Ellickson, R. C. (1977) Suburban Growth Controls: An Economic and Legal Analysis, *The Yale Law Journal*, **86**, 385–511.
- Fischel, W. (2004) An economic history of zoning and a cure for its exclusionary effects, *Urban Studies*, pp. 317–340.
- Garofalo, G. A. and Yamarik, S. (2002) Regional Convergence: Evidence from a New State-by-State Capital Series, *Review of Economics and Statistics*, **84**, 316–323.
- Gennaioli, N. and Shleifer, A. (2007) The Evolution of Common Law, *Journal of Political Economy*, **115**, 43–68.
- Glaeser, E. L. and Gyourko, J. (2005) Urban Decline and Durable Housing, *Journal of Political Economy*, **113**, 345–375.
- Glaeser, E. L., Gyourko, J. and Saks, R. (2005a) Why is Manhattan So Expensive? Regulation and the Rise in House Prices, *Journal of Law and Economics*, pp. 1–50.
- Glaeser, E. L., Gyourko, J. and Saks, R. E. (2005b) Why Have Housing Prices Gone Up?, *American Economic Review*, **95**, 329–333.
- Glaeser, E. L., Kahn, M. E. and Rappaport, J. (2008) Why do the Poor Live in Cities? The Role of Public Transportation, *Journal of Urban Economics*, **63**, 1–24.
- Glaeser, E. L. and Saiz, A. (2004) The Rise of the Skilled City, *Brookings-Wharton Papers on Urban Affairs*, pp. 47–105.
- Glaeser, E. L. and Tobio, K. (2007) The Rise of the Sunbelt, *NBER Working Paper 13071*.
- Goldin, C. and Katz, L. F. (2001) The Legacy of U. S. Educational Leadership: Notes on Distribution and Economic Growth in the 20th Century, *American Economic Review: Papers and Proceedings*, **91**, 18–23.

- Gyourko, J., Mayer, C. and Sinai, T. (2012) Superstar Cities, *mimeo*.
- Gyourko, J., Saiz, A. and Summers, A. (2008) A New Measure of the Local Regulatory Environment for Housing Markets: Regulatory Index, *Urban Studies*, **45**, 693–729.
- Haines, M. R. (2010) *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 [Computer file]. ICPSR02896-v3.*, Inter-university Consortium for Political and Social Research, Ann Arbor, MI.
- Hamermesh, D. (1987) The Demand for Labor in the Long Run, in *Handbook of Labor Economics*.
- Hornbeck, R. (2012) The Enduring Impact of the American Dust Bowl: Short and Long-run Adjustments to Environmental Catastrophe, *American Economic Review*, **Forthcomin**.
- Hornbeck, R. and Naidu, S. (2012) When the Levee Breaks : Labor Mobility and Economic Development in the American South, *Mimeo*.
- Hsieh, C. T., Hurst, E., Jones, C. I. and Klenow, P. J. (2012) The Allocation of Talent and U.S. Economic Growth, *mimeo*.
- Iyer, L., Meng, X., Qian, N. and Zhao, X. (2011) The General Equilibrium Effects of Chinese Urban Housing Reforms on the Wage Structure, *mimeo*.
- Kahn, M. E. (2011) Do Liberal Cities Limit New Housing Development? Evidence from California, *Journal of Urban Economics*, **69**, 223–228.
- Kaplan, G. and Schulhofer-Wohl, S. (2012) Understanding the Long-Run Decline in Interstate Migration, *mimeo*.
- Katz, L. and Rosen, K. T. (1987) The Interjurisdictional Effects of Growth Controls on Housing Prices, *The Journal of Law and Economics*, **30**, 149.
- Kennan, J. and Walker, J. (2011) The Effect of Expected Income on Individual Migration Decisions, *Econometrica*, **79**, 211–251.
- Kuznets, S., Miller, A. R. and Easterlin, R. A. (1964) *Population Redistribution and Economic Growth: Analyses of economic change*, American Philosophical Society, Philadelphia.
- Landon-Lane, J. and Rockoff, H. (2007) The origin and diffusion of shocks to regional interest rates in the United States, 1880 to 2002., *Explorations in Economic History*, **44**, 487–500.
- Lewis, E. (2011) Immigration, Skill Mix, and Capital-Skill Complementarity, *Quarterly Journal of Economics*, **126**, 1029–1069.
- Margo, R. A. (1997) Wages in California During the Gold Rush.
- Michaels, G., Rauch, F. and Redding, S. J. (2012) Urbanization and Structural Transformation, *Quarterly Journal of Economics*, **127**, 535–586.
- Molloy, R., Smith, C. L. and Wozniak, A. (2011) Internal Migration in the United States, *Journal of Economic Perspectives*, **25**, 1–24.
- Moretti, E. (2012) *The New Geography of Jobs*, Houghton Mifflin Harcourt, Boston.
- Nakamura, E. and Steinsson, J. (2011) Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions, *NBER Working Paper 17391*.
- Notowidigdo, M. J. (2011) The Incidence of Local Labor Demand Shocks, *NBER Working Paper 17167*.

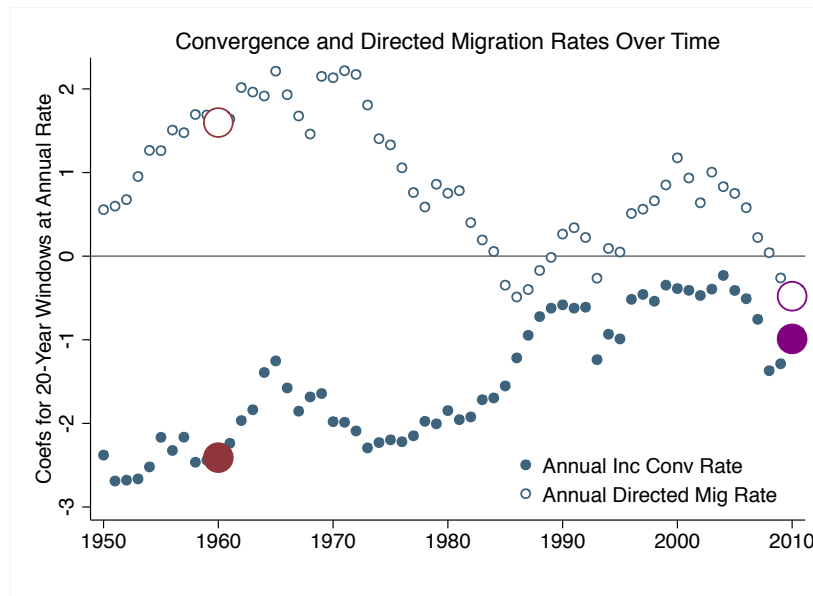
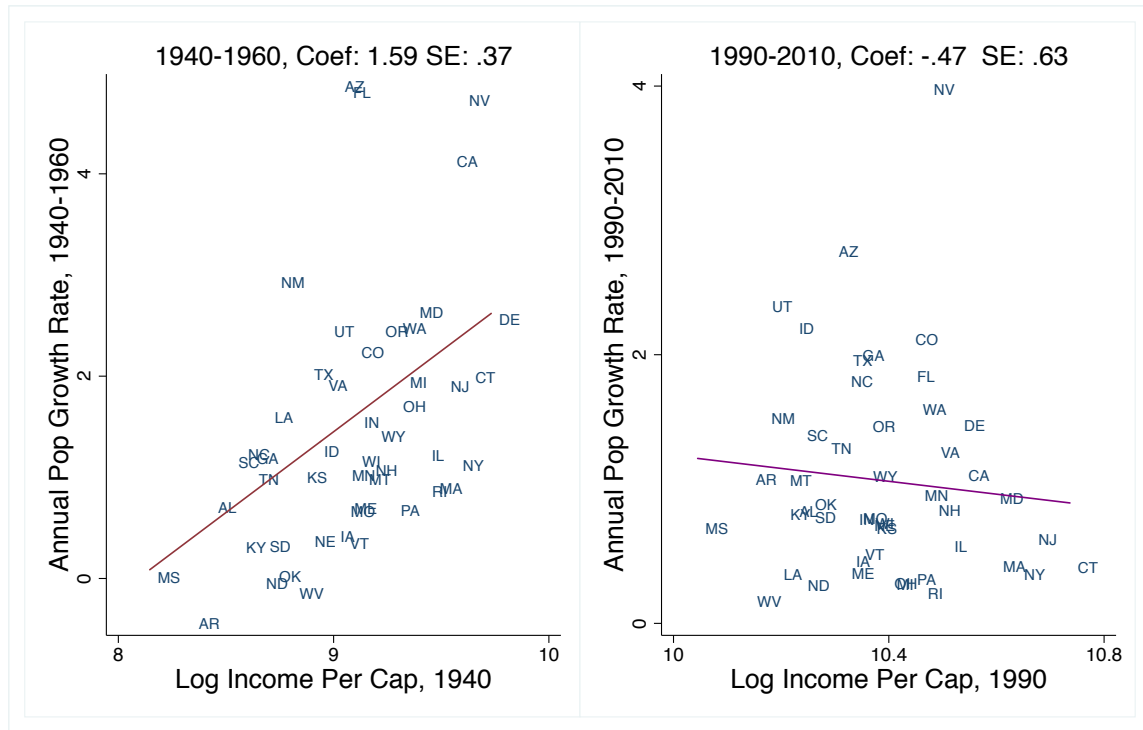
- Pollakowski, H. O. and Wachter, S. M. (1990) The Effects of Land-Use Constraints on Housing Prices, *Land Economics*, **66**.
- Quigley, J. and Raphael, S. (2005) Regulation and the High Cost of Housing in California, *American Economic Review*, **95**, 323–328.
- Rothstein, J. (2008) The Unintended Consequences of Encouraging Work: Tax Incidence and the EITC, *mimeo*.
- Rothstein, J. (2010) Is the EITC as Good as an NIT? Conditional Cash Transfers and Tax Incidence, *American Economic Journal: Economic Policy*, pp. 177–208.
- Rothwell, J. (2012) Housing Costs, Zoning, and Access to High-Scoring Schools, *Metropolitan Policy Program at Brookings*.
- Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B. and Sobek, M. (2010) Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database].
- Ruhl, K. J. (2008) The International Elasticity Puzzle, *Mimeo*.
- Saiz, A. (2010) The Geographic Determinants of Housing Supply, *Quarterly Journal of Economics*, **125**, 1253–1296.
- Saks, R. E. (2008) Job Creation and Housing Construction: Constraints on Metropolitan Area Employment Growth, *Journal of Urban Economics*, **64**, 178–195.
- Schleicher, D. (2012) City Unplanning, *mimeo*.
- The American Institute of Planners (1976) *Survey of State Land Use Planning Activity*.
- U.S. Census Bureau (2012) USACounties.
- Van Nieuwerburgh, S. and Weill, P.-O. (2010) Why Has House Price Dispersion Gone Up?, *Review of Economic Studies*, **77**, 1567–1606.
- Williamson, J. G. (1965) Regional Inequality and the Process of National Development: A Description of the Patterns, *Economic Development and Cultural Change*, **13**, 1– 84.
- Young, C., Varner, C. and Massey, D. S. (2008) Trends in New Jersey Migration: Housing, Employment, and Taxation, *Policy Research Insitute at the Woodrow Wilson School*.

FIGURE 1
The End of Income Convergence



Notes: The y-axis in the first two panels is the annual growth rate of income per capita. The third panel plots coefficients from 20-year rolling windows. The larger red and purple dots correspond to the coefficients from the top two panels. Income data from the Bureau of Economic Analysis (2012). Alaska, Hawaii, and DC are omitted here, and in all subsequent figures and tables.

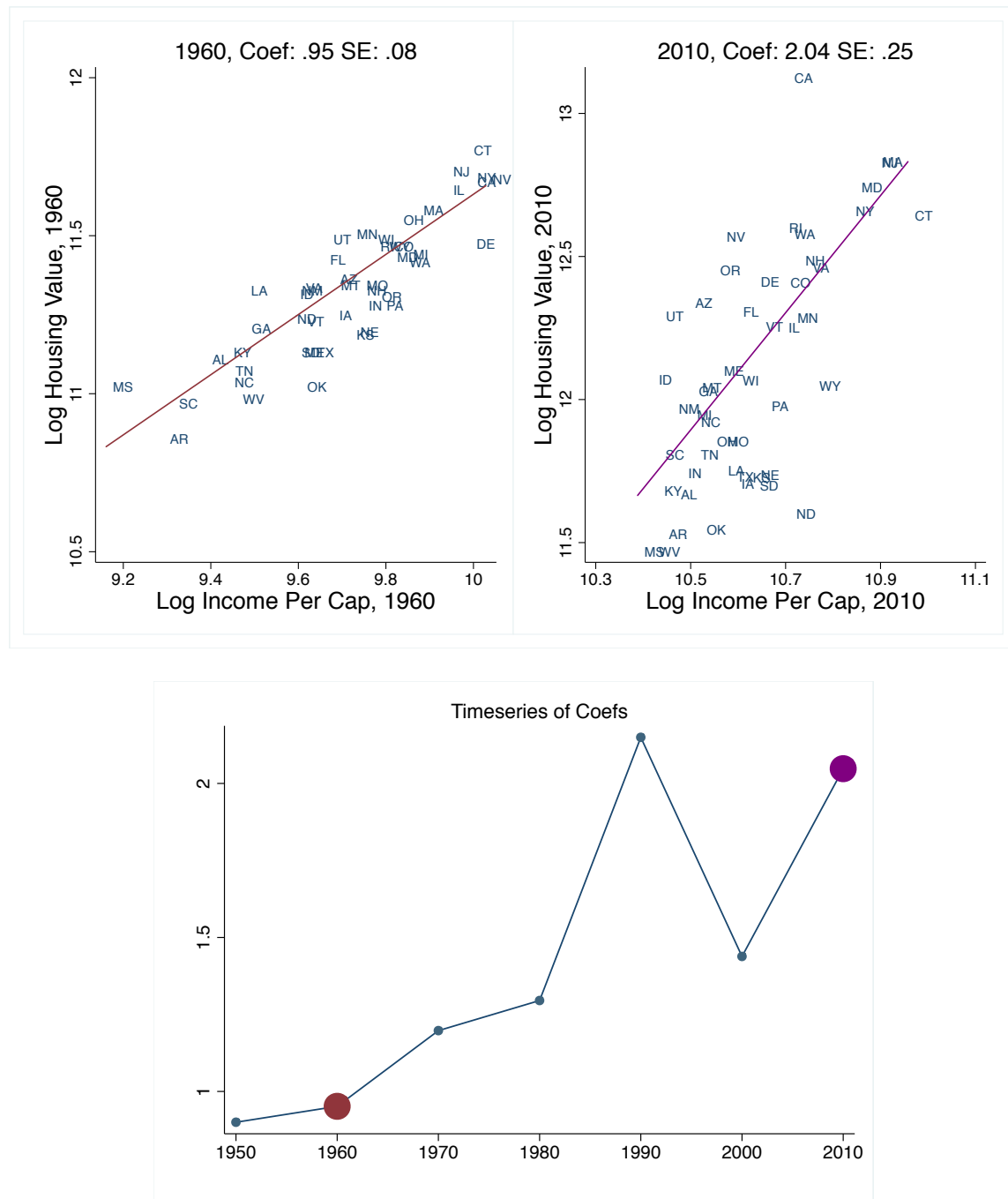
FIGURE 2
The End of Directed Migration



Notes: The y-axis in the first two panels is the annual growth rate of log population. The third panel plots coefficients from 20-year rolling windows for population changes and income changes. The larger red and purple dots correspond to the coefficients from the top two panels.

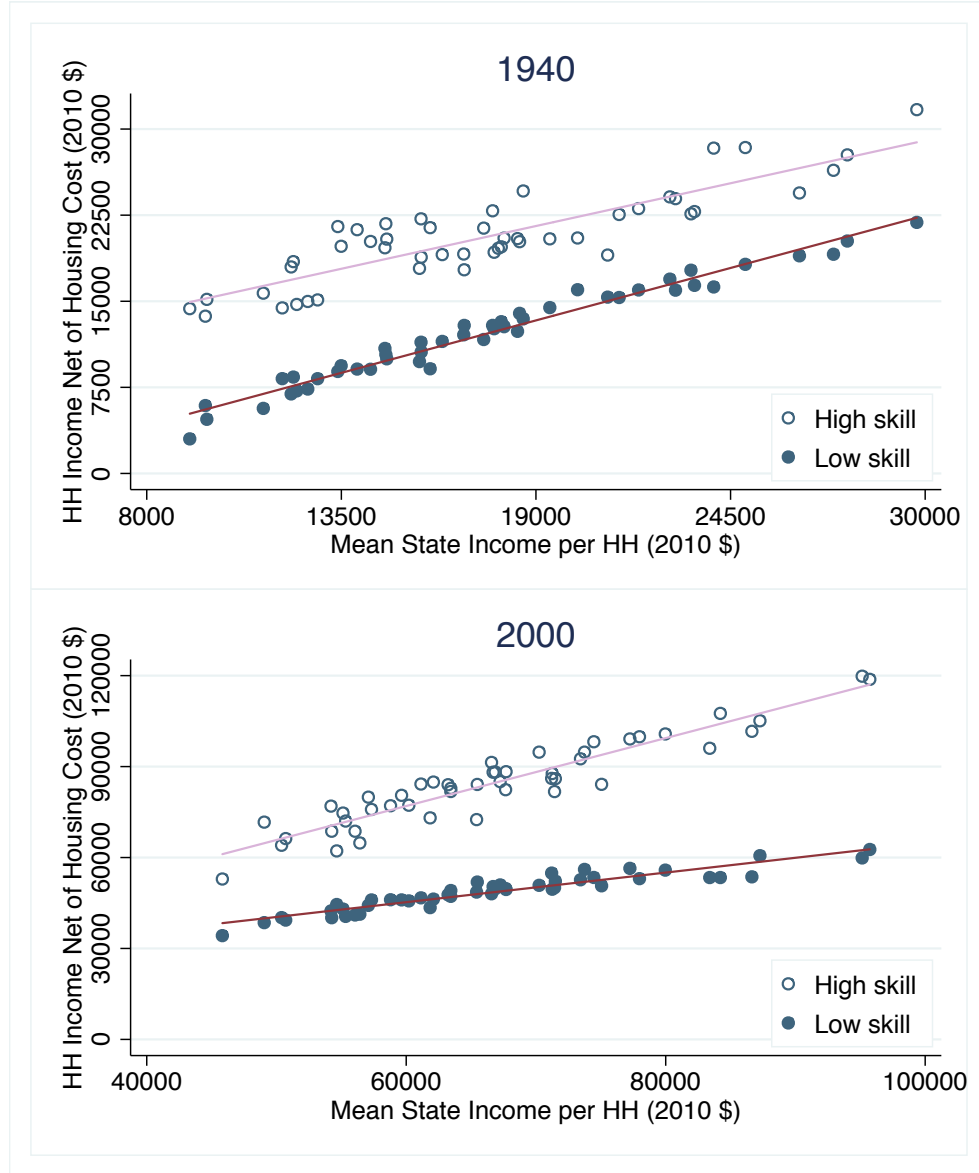
FIGURE 3

Rising Prices in High Income States



Notes: The first two panels regress median housing value on income per capita at the state level. The third panel plots coefficients from 20-year rolling windows. The larger red and purple dots correspond to the coefficients from the first two panels.

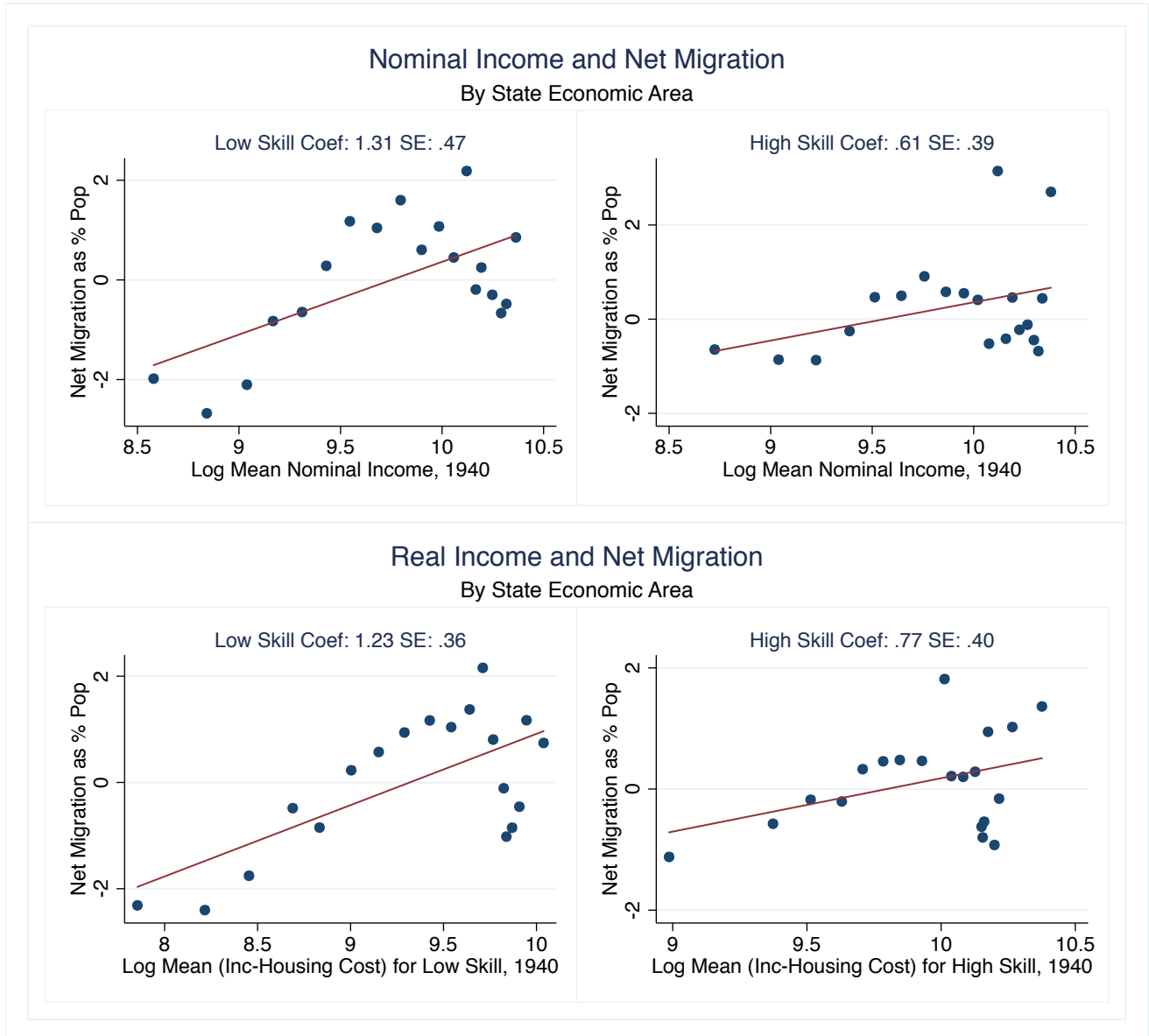
FIGURE 4
Returns to Migration: Skill-Specific Income Net of
Housing Cost



Notes: This figure plots the relationship between unconditional mean household income and mean skill-specific income net of housing costs. The x-axis plots state-wide average household wage income for households with at least one labor force participant aged 25-65. The y-axis plots the mean household wage income net of housing cost for high- and low-skilled households after controlling for household demographics (see Section 3.1 for details). Housing costs are defined as 5% of house value for homeowners and 12X monthly rent for renters. High-skilled households are defined as households in which all adult workers have 12+ years of education in 1940 (a bachelor's degree in 2000) and low-skilled households are defined as households in which no worker adult worker has this level of education. The roughly 15% of mixed skill-type households in each year are dropped from the construction of the dependent variable, but not from the computation of unconditional state average income.

FIGURE 5

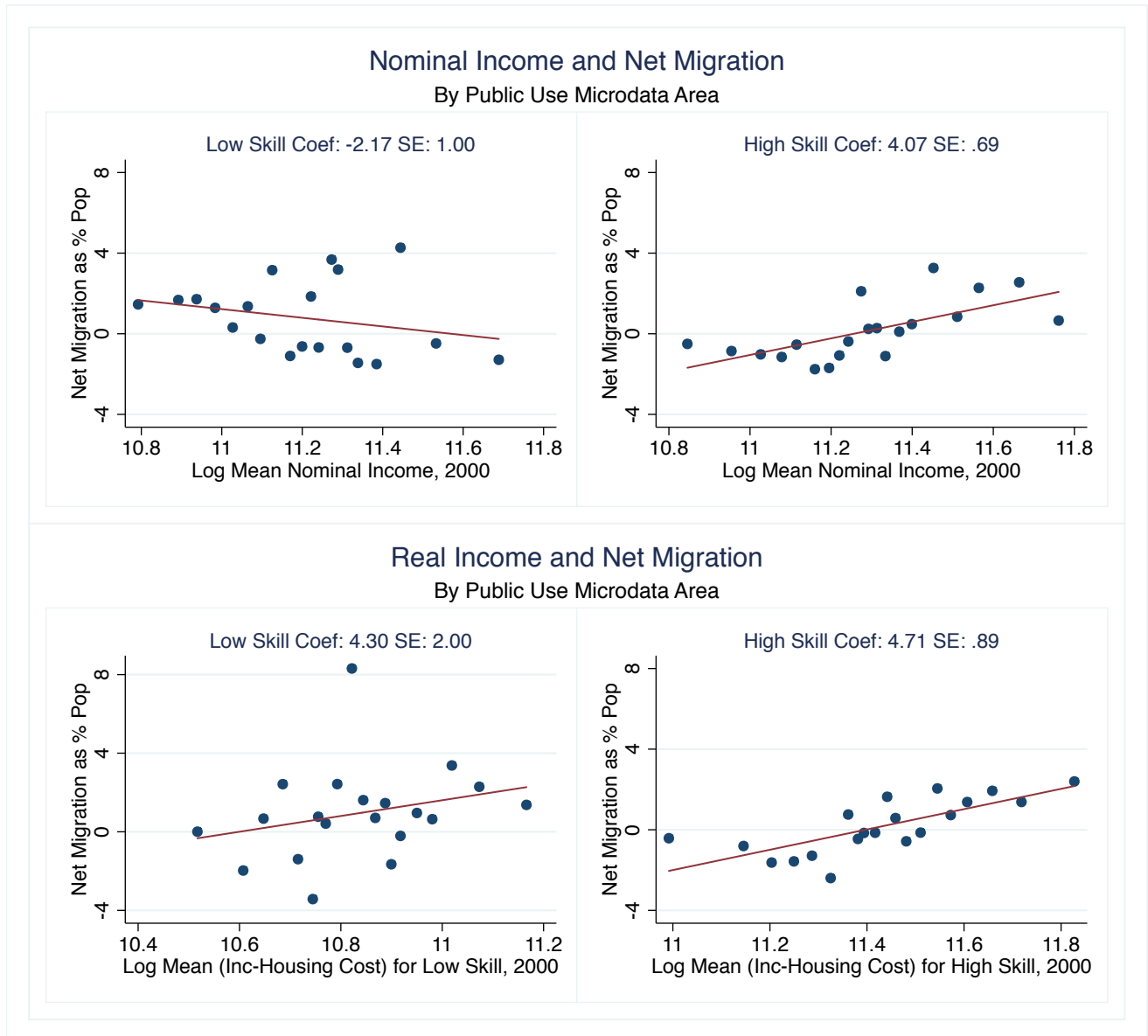
Migration Flows by Skill Group: Nominal vs. Real Income, 1940



Notes: These panels plot net migration as a fraction of the population ages 25-65 for 466 State Economic Areas (SEA) in the 1940 IPUMS Census extract. Each panel stratifies the SEAs into 20 quantiles by income, weighting each SEA by its population, and then computes the mean net migration within each quantile. The two top panels plot net migration as a function of the log household wage income in the SEA, for individuals with less than 12 years of education (left) and those with 12+ years (right). The two bottom panels plot the migration rates for these skill groups against the log skill-group mean value of household wage income net of housing costs. Housing costs are defined as 5% of house value for homeowners and 12X monthly rent for renters. All x-axis variables are computed for non-migrating households with at least one labor force participant aged 25-65.

FIGURE 6

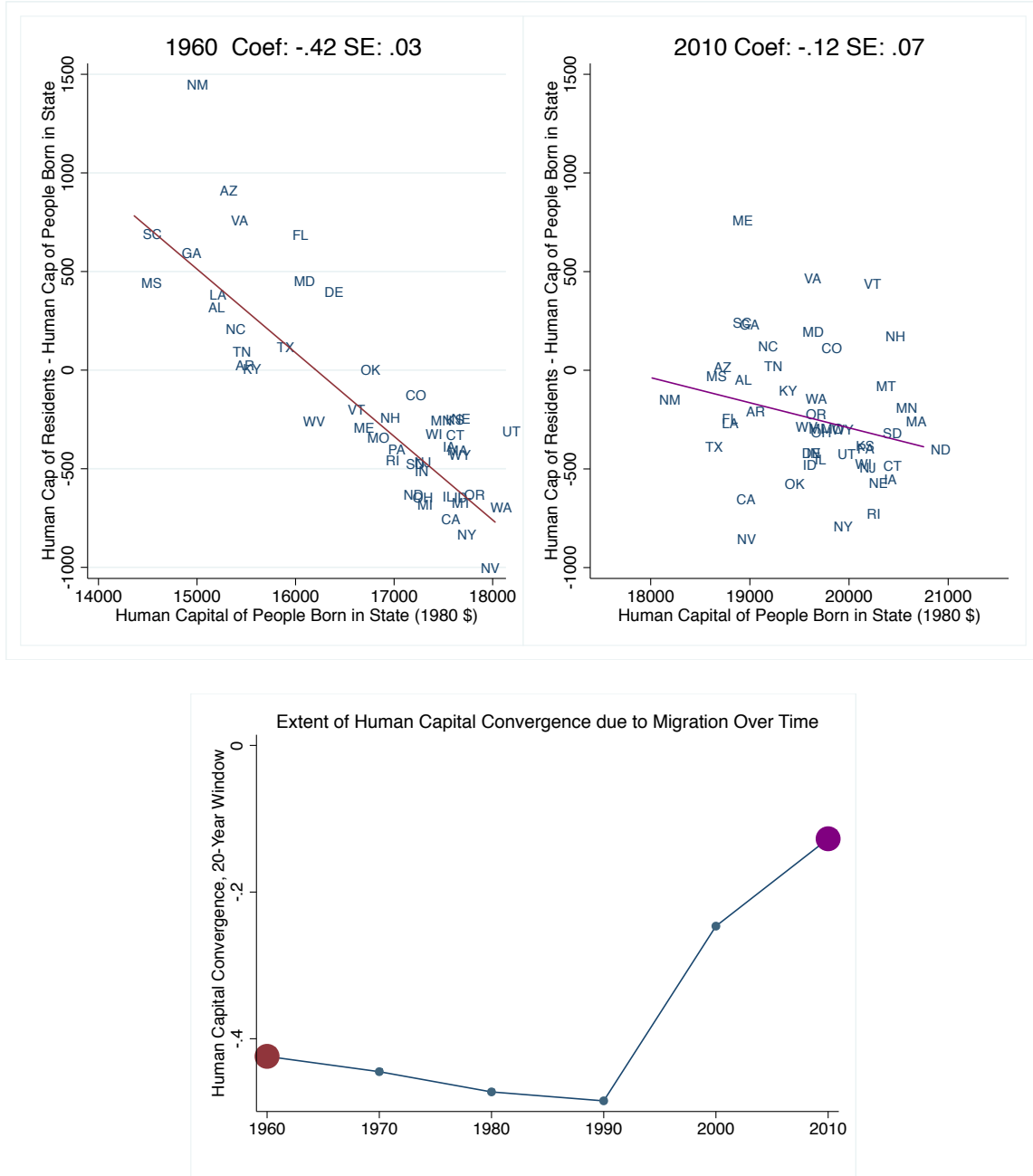
Migration Flows by Skill Group: Nominal vs. Real Income, 2000



Notes: These panels plot net migration as a fraction of the population ages 25-65 for 1,020 3-digit Public Use Microdata Area (PUMA) in the 2000 IPUMS 5% Census extract. Each panel stratifies the PUMAs into 20 quantiles by income, weighting each PUMA by its population, and then computes the mean net migration within each quantile. The two top panels plot migration rates as a function of log household wage income in the PUMA, for individuals with less than 4 years of college (left) and with 4+ years (right). The two bottom panels plot the migration rates for these skill groups against the skill-group mean value of household wage income net of housing costs. Housing costs are defined as 5% of house value for homeowners and 12X monthly rent for renters. All x-axis variables are computed for non-migrating households with at least one labor force participant aged 25-65.

FIGURE 7

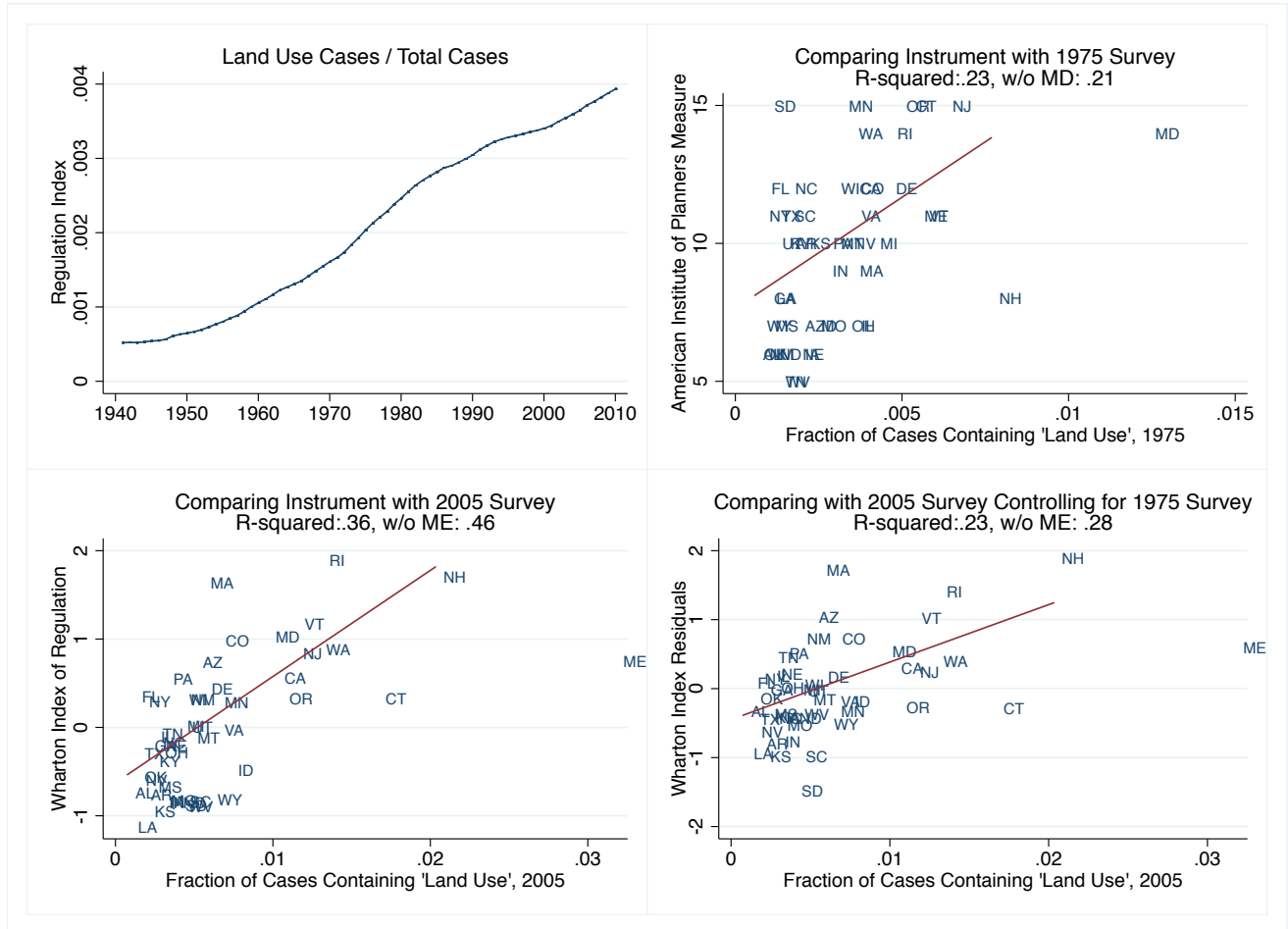
The Decline of Human Capital Convergence



Notes: Human capital index is estimated by regressing $\log Inc_{ik} = \alpha_k + X_{ik}\beta + \varepsilon_{ik}$, and then constructing $Human\ Capital_s = \sum_k \exp(\hat{\alpha}_k) \times Share_{ks}$. We separately estimate the human capital index by state of residence and by state of birth, to develop a no-migration counterfactual. The top panels show figures from a regression of $HumanCap_{s,res} - HumanCap_{s,birth} = \alpha + \beta HumanCap_{s,birth} + \varepsilon_s$ in 1960 and 2010. The skill premium ($\{\alpha_k\}$) is estimated in the 1980 Census, and these estimates are applied to skill shares in different Census years. See the data appendix for details. The bottom panel plots a timeseries of coefficients. The larger red and purple dots correspond to the coefficients from the first two panels.

FIGURE 8

Regulation Measure: Timeseries and Validity



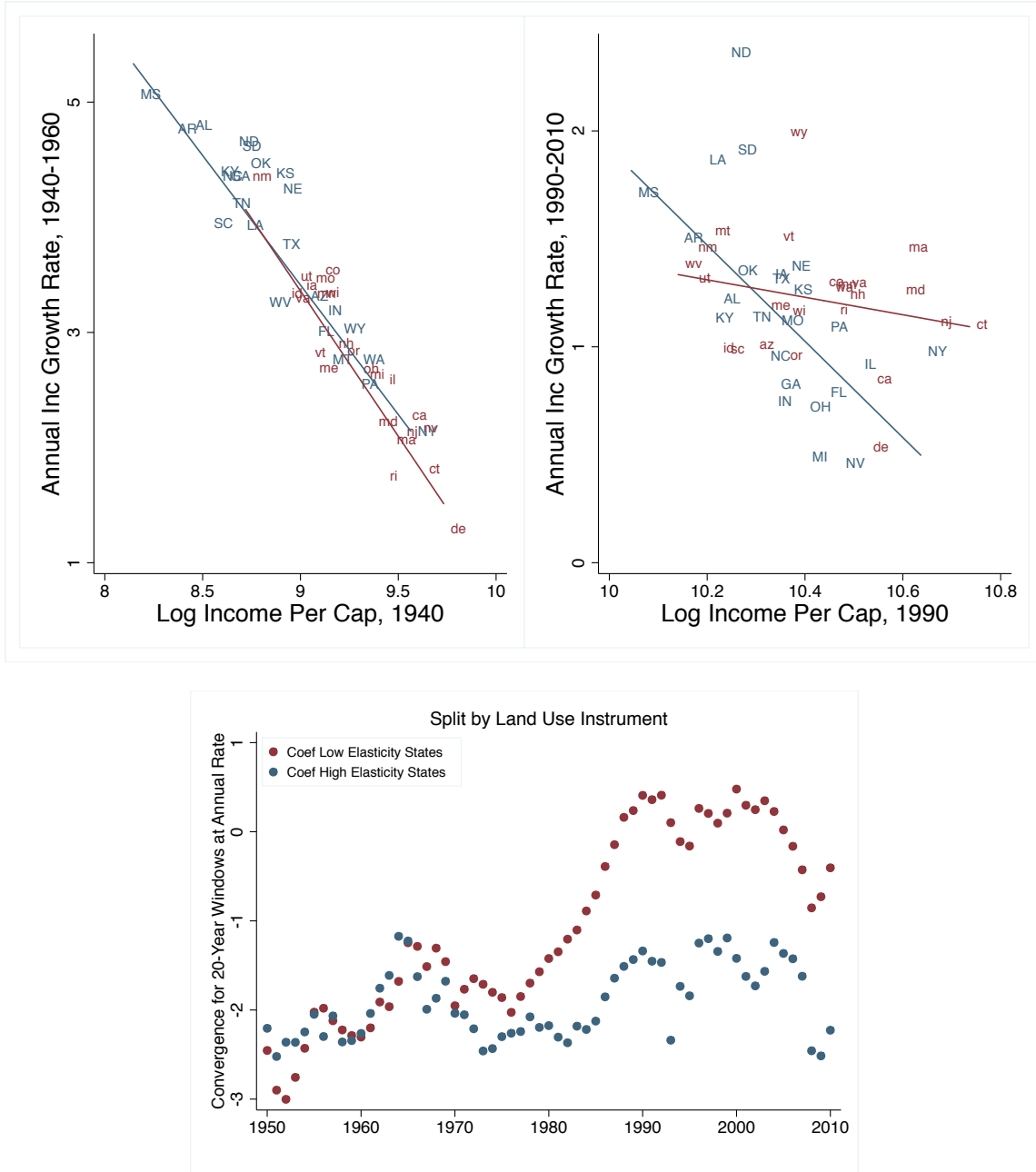
Notes: The top left panel plots the cumulative frequency of cases containing the phrase "land use" in the state appeals court databases as a fraction of total cases to date. The years 1920-1940 are included in the totals to initialize the series but are not plotted.

The top right panel plots the relationship between the 1975 values of the regulation measure introduced in the text and the sum of affirmative answers to the regulation questions asked in the 1975 American Institute of Planners Survey of State Land Use Planning Activities. The fit line excludes MD.

The lower left panel plots the relationship between the 2005 values of the regulation measure introduced in the text and the 2005 Wharton Residential Land Use Regulatory Index. The fit line excludes ME.

The lower right panel plots on the y-axis 2005 Wharton residuals, conditional on the 1975 survey measure. In all three panels, the court-based regulation measure is predictive of the other regulation measures.

FIGURE 9
Income Convergence by Housing Supply Elasticity



Notes: The top panels reconstruct Figure 1, by relabeling states according to their regulation levels. Blue states (upper case) have above median housing supply elasticity using our regulation measure and red states (lower case) have below median elasticity. Regulation is measured in 1941 for left panel and in 2005 for right panel.

The bottom panel depicts the coefficients from running: $\Delta Inc_{s,t} = \alpha_t + \beta Inc_{s,t-20} + \varepsilon_{s,t}$ over rolling twenty year windows. The regressions are estimated separately for two equally sized groups of states, split along their measure of land use regulations from the legal database. The groups are split at the median every year, so the composition of the groups changes over time.

TABLE 1
Summary Statistics

	1940		1960		1980		2000	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Personal Income Per Capita (\$000, 2012 \$)	8.83	3.18	16.34	3.15	26.63	3.63	38.41	5.95
Population (Million)	2.73	2.69	3.72	3.80	4.69	4.76	5.83	6.26
Human Capital Level Per Capita (1 indicates everyone is non-black non-Hispanic with HS degree)	0.79	0.08	0.93	0.06	1.09	0.04	1.27	0.05
Median House Price (\$000, 2012 \$)	39.7	15.4	85.2	18.6	129.4	32.1	152.3	44.5
Regulation Measure (share of cases relating to land use*100)	0.07	0.04	0.14	0.09	0.35	0.26	0.53	0.47
Fraction Age 25-65 with								
12+ Years of Education	0.26	0.44	0.45	0.5	0.73	0.44	0.88	0.33
16+ Years of Education	0.05	0.21	0.08	0.28	0.18	0.38	0.27	0.44

Sources: IPUMS Census extract, BEA Income estimates, Price index from Lindert and Sutch (2006) and an online database of state appellate court documents.

Notes: n=48 states, excluding Alaska, Hawaii, and DC. Dollar amounts are in real 2012 dollars. Human Capital Level is measured using education, race, and age. See Section 3.3 for details on Human Capital Level construction. The regulation measure for 1941 is reported in place of the unavailable 1940 statistics.

TABLE 2
Impacts of Regulation on Permits, Prices, Migration, and Convergence

	Annual Construction Permits _t % of Housing Stock (1)	Log House Price _{t-20} (2)	ΔLog Population _{t, t-20} Annual Rate in % (3)	Δ Log Human Capital _{t,t-20} Annual Rate in % (4)	Δ Log Income Per Cap _{t,t-20} Annual Rate in % (5)
<u>Land Use Reg Measure</u> 1(landuse _{it} >landuse ^{median} ₂₀₀₅)					
Log Inc Per Cap _{t-20}	1.865 (1.667)	1.026*** (0.0904)	2.042** (0.915)	-0.239** (0.0958)	-2.306*** (0.209)
Log Inc Per Cap _{t-20}	-2.423 (1.790)	0.667*** (0.185)	-2.812** (1.087)	0.297** (0.144)	3.199*** (0.382)
*High Reg					
Year*High Reg FEs	Y	Y	Y	Y	Y
R ²	0.182	0.859	0.129	0.837	0.782
N	1,536	336	2,976	288	2,976

Notes: The table reports the coefficients β_1 and β_2 from regressions of the form: $\Delta \ln y_{it} = \alpha_t + \alpha_t I(\text{reg} > x) + \beta_1 \ln y_{it-1} + \beta_2 \ln y_{it-1} \times I(\text{reg} > x) + \varepsilon_{it}$. The construction of the land use regulation measure is described in the text, and the high regulation cutoff value is set to the measure's 2005 median value. The dependent variables are new housing permits from the Census Bureau, the median log housing price from the IPUMS Census extracts, population change, the change in log human capital due to migration, and the change in log per-capita income. Standard errors clustered by state. *** p<0.01, ** p<0.05, * p<0.1

TABLE 3
Calibration: Role of Population and Human Capital in Income Convergence

<u>Inputs from Model and Literature</u>					
Elasticity of Substitution for Regional Goods σ				{2, 6}	
Elasticity of labor supply ε				{ 0, 0.6}	
(1) elasticity of income wrt $\Delta\text{pop} = -(1+\varepsilon)/(\sigma+\varepsilon)$ from Equation 2					
(3) elasticity of income wrt $\Delta\text{HumanCapital} = (\sigma-1)/(\sigma+\varepsilon)$ from Equation 3					
<u>Inputs from Data</u>		1950-1980		2010	Difference
(2) avg directed mig coef (Table 2)		1.26		-0.48	-1.74
(4) avg sort coef (Appendix Table 7)					
$\varepsilon = 0$		-0.15		0.03	0.19
$\varepsilon = 0.6$		-0.24		0.08	0.33
		Population Channel		Human Capital Channel	
		Change in		Change in Human	
		Directed		Capital	
<u>Calculation</u>	$\varepsilon^{\text{inc}}_{\Delta\text{pop}}$	Migration		$\varepsilon^{\text{inc}}_{\Delta\text{HumanCap}}$	Convergence
		Annual Rate in %		Annual Rate in %	
Elasticities					Change in Predicted
ε : Labor Supply					Income Convergence
σ : Subs for Goods					Annual Rate in %
		(1)	* (2)	+	(3) * (4) = (5)
ε	σ				
0	2	-0.50	-1.74	1.00	0.19
0	6	-0.17	-1.74	1.00	0.19
0.6	2	-0.62	-1.74	0.77	0.33
0.6	6	-0.24	-1.74	0.91	0.33
Average Observed Convergence Coef ($\beta^{\Delta\text{inc}}_{\text{laginc}}$)					
Observed Convergence Rate, 20-Year Windows Ending In 1950-1980					-2.10
Observed Convergence Rate, 20-Year Window Ending In 2010					-0.99
Change in Convergence					1.11

Notes: This table estimates the role of migration in income convergence. Migration drives convergence through population flows from poor to rich states (measured in Table 2) and human capital flows from rich to poor states (measured in Table 3). The effect of these changes on income per capita is calibrated using the model (equation 7) and different assumptions on the elasticity of labor supply and demand. We consider four scenarios: balanced growth preferences and a labor supply elasticity of 0.6, and elasticities of substitution across regional goods equal to 2 and 6. See the text for additional details.

TABLE 4
Latent Tendency to Regulate, Geographic Land Availability, and Convergence

Year	Log Income Per Cap _t - Log Income Per Cap _{t-25} (Annual Rate in %)							
	1980 (1)	2005 (2)	1980 (3)	2005 (4)	1980 (5)	2005 (6)	1980 (7)	2005 (8)
Log Inc Per Cap _{t-25}	-1.836*** (0.0955)	-1.609*** (0.368)	-1.959*** (0.264)	-3.287*** (0.366)	-1.712*** (0.190)	-1.368*** (0.390)	-1.833*** (0.231)	-1.355*** (0.480)
Log Inc Per Cap _{t-25} *	0.410 (0.258)	1.769*** (0.587)	0.529* (0.283)	1.948*** (0.422)	0.212 (0.295)	1.165* (0.602)	0.262 (0.249)	1.105* (0.626)
Low Elasticity Measure								
Controls	Low Elas. Dummy		Low Elas. Dummy Autor-Dorn Skill Measures		Low Elas. Dummy		Low Elas. Dummy	
R ²	0.738	0.249	0.787	0.720	0.730	0.190	0.737	0.163
N	48	48	48	48	48	48	48	48
Low Elasticity Measure	Above Median Regulation 2005		Above Median Regulation 2005		Above Median Regulation 1955		Low Geographic Land Availability	

Notes: The table reports the coefficients β_1 and β_2 from regressions of the form

$$\Delta \ln y_{it,t-25} = \alpha_1 + \alpha_2 I(\text{reg} > x) + \beta_1 \ln y_{it-25} + \beta_2 \ln y_{it-25} \times I(\text{reg} > x) + \varepsilon_i$$

for the periods 1955-1980 and 1980-2005. In all columns the states are divided into equally-sized high and low elasticity groups. The first two columns divide states based on the value of their regulation instrument in 2005. The second two columns repeat this exercise with controls from Autor and Dorn (AER, Forthcoming) for the share of workers in routine occupations, the college to non-college population ratio, immigrants as a share of the non-college population ratio, manufacturing employment share, the initial unemployment rate, the female share, the share age 65+, and the share earning less than the 10 year ahead minimum wage. We aggregate their data to the state level via population weighting. In the third set of columns, we divide states based on the value of their regulation instrument in 1955 ("latent regulation"). The final two columns divide states based on the population-weighted land availability constructed from Saiz (2010). The table reports the annualized rate over 25 years. 2005 is chosen as a cutoff for the high-regulation period to avoid confounding from the 2007-2009 recession. The correlation between the high regulation dummy in 2005 and the high regulation dummy in 1955 is .58. The correlation between both regulation dummies and the low geographic availability dummy is .33. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

TABLE 5
Migration By Skill Group and Share BA

Panel A: Total Migration (Extensive + Intensive Margin)

	# Residents - # Born in State as % of Total State Pop			
	Low-Skill (1)	High-Skill (2)	Total Mig (2) + (1)	Difference (2)-(1)
<i>1980 Census, n=48</i>				
Share BA, 1980	2.624*** (0.479)	0.762*** (0.131)	3.386*** (0.550)	-1.862***
<i>2010 American Community Survey, n=48</i>				
Share BA, 1980	0.490** (0.235)	0.614*** (0.138)	1.104*** (0.354)	0.124
Coef 2010 - Coef 1980	-2.134***	-0.148	-2.282***	

Panel B: Choice of Destination | Decision to Leave Birth State (Intensive Margin)

	# Migrants to state j from state of birth j' - Pop j / (Pop National - Pop j')		
	Low-Skill (1)	High-Skill (2)	Difference (2)-(1)
<i>1980 Census, n=2256</i>			
Share BA, 1980	0.116* (0.0608)	0.173*** (0.0460)	0.057**
<i>2010 American Community Survey, n=2256</i>			
Share BA, 1980	-0.0297 (0.0400)	0.129*** (0.0326)	0.149***
Coef 2010 - Coef 1980	-0.136**	-0.044	

Notes: This table examines differences by skill group and over time in migration to high BA states.

Panel A measures net migration of 25-44 year olds relative to state of birth as a share of the state's total population. There is one observation per state, and robust SE are in parentheses. This measure is attractive because it captures both the decision to migrate and the choice of destination, but it is sensitive to differential trends in domestic BA production.

Panel B corrects for this issue and focuses on choice of destination among those who choose to migrate within the 48 continental states. Each observation is a state of origin by state of destination pair. We examine whether people who migrate are disproportionately attracted to states with high share BA. We normalize each observation by subtracting the ratio of the population of the destination state to the population of all states (dropping the population of the state of origin). Observations are weighted by the total number of migrants from the origin state, and the standard errors are clustered by destination.

Share BA is calculated using people ages 25-65. Low-skill is defined as having less than a BA. High skill is defined as having a BA or higher. *** p<0.01, ** p<0.05, * p<0.1

A Theory Appendix

A.1 General Equilibrium Model With Downward-Sloping Product Demand

We characterize a general equilibrium model for a state with a labor market and a goods market. Later, we describe this state as “Reservationville”, because it serves as the reservation locale in our model. We also describe another state which has the same labor and goods market, along with a land market. The addition of the land market is mathematically straightforward, and we omit it here.

A.1.1 Individual Decisions: Labor Supply and Product Demand

Individuals in the region “home s ” of skill type k consume a basket of differentiated goods. Each region $s' \in [0, 1]$ produces one good $x_{s'}$. For each unit of labor they supply, l_{sk} , they offer the market $\psi_k l_{sk}$ units of “effective labor.” Individuals solve the following problem, taking the local price for effective labor w_s and the national price for products $\{p_{s'}\}$ as exogenous

$$\begin{aligned}
 U_{sk} &= \left(\int x_{ss'k}^\rho ds' \right)^{\frac{1}{\rho}} - \frac{l_{sk}^{1+\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}} + \lambda \left(\pi + w_s \psi_k l_{sk} - \int p_{s'} x_{ss'k} ds' \right) \\
 \frac{\partial U}{\partial x_s} &= \left(\left(\int x_{ss'k}^\rho ds' \right)^{\frac{1-\rho}{\rho}} \right) x_{ss'k}^{\rho-1} - \lambda p_{s'} = 0 \quad \forall s' \\
 \frac{\partial U}{\partial l} &= -l_{sk}^{1/\epsilon} + \lambda w_s \psi_k = 0 \\
 \frac{\left(\left(\int x_{ss'k}^\rho ds' \right)^{\frac{1-\rho}{\rho}} \right) x_{ss'k}^{\rho-1}}{p_{s'}} &= \frac{l_{sk}^{1/\epsilon}}{w_s \psi_k} \Rightarrow \\
 l_{sk}^{supply}(w) &= \psi_k^\epsilon w_s^\epsilon \left(\underbrace{\frac{x_{ss'k}^{\rho-1}}{p_{s'}} \left(\left(\int x_{ss'k}^\rho ds' \right)^{\frac{1-\rho}{\rho}} \right)}_{\xi_{sk}: \text{marginal utility of consumption}} \right)^\epsilon
 \end{aligned} \tag{10}$$

Equation (10) holds for all goods $s' \in [0, 1]$. We assume that labor supply of different types aggregates linearly, so

$$L_s^{supply}(w) = \int l_{sk}^{supply}(w) \mu_{sk} dk \tag{11}$$

We now apply the standard Dixit-Stiglitz solution techniques to derive the demand for any individual good s' in terms of its own price $p_{s'}$, individual labor income $\psi_k w_s l_{sk}$ and the aggregate price index P . Define $\sigma = \frac{1}{1-\rho}$. The first order conditions for $\{\frac{dU}{dx_{s'}}\}$ imply that an individual's consumption of two goods s' and s'' must have the following ratio:

$$\begin{aligned}
 \frac{x_{ss''k}}{x_{ss'k}} &= \left(\frac{p_{s''}}{p_{s'}} \right)^{-\sigma} \Rightarrow x_{ss''k} = x_{ss'k} \left(\frac{p_{s''}}{p_{s'}} \right)^{-\sigma} \\
 p_{s''} x_{ss''k} &= p_{s'} x_{ss'k} \left(\frac{p_{s''}}{p_{s'}} \right)^{-\sigma} \text{ for any } s'' \\
 \text{Integrating } \int p_{s''} x_{ss''k} ds'' &= \int p_{s'} x_{ss'k} \left(\frac{p_{s''}}{p_{s'}} \right)^{-\sigma} ds'' \\
 \pi + w_s \psi_k l_{sk} &= x_{ss'k} p_{s'}^\sigma \underbrace{\int p_{s''}^{1-\sigma} ds''}_{P^{1-\sigma}}
 \end{aligned}$$

$$x_{ss'k} = p_{s'}^{-\sigma} \frac{(\pi + w_s l_{sk} \psi_k)}{P^{1-\sigma}} \quad (12)$$

Recall that l_{sk} is actually $l_{sk}^{supply}(w)$ from equation (10) which governed labor supply. We now substitute in for the labor supply elasticity above, to write an individual's demand for good $x_{s'}$ as:

$$x_{ss'k}^{Demand}(p, w, \xi, P) = \frac{p_{s'}^{-\sigma} (\pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} \xi_{sk})}{P^{1-\sigma}} \quad (13)$$

where ξ_{sk} is a scaling of household marginal utility. To derive this scaling recall from equation (10) that

$$\xi_{sk} = \left(\frac{x_{ss'k}^{\rho-1}}{p_{s'}} \left(\left(\int x_{ss'k}^\rho ds' \right)^{\frac{1-\rho}{\rho}} \right) \right) \quad \forall sk$$

Recall equation (12), that demand by consumer sk for good s' is $x_{ss'k} = p_{s'}^{-\sigma} \frac{(\pi + w_s l_{sk} \psi_k)}{P^{1-\sigma}}$. Plugging the demand equation into the marginal utility expression gives

$$\begin{aligned} \xi_{sk} &= \left(\frac{p_{s'}^{-\sigma(\rho-1)}}{p_{s'}} \left(\frac{\pi + w_s l_{sk} \psi_k}{P^{1-\sigma}} \right)^{\rho-1} \left(\left(\int \left(p_{s'}^{-\sigma} \frac{(\pi + w_s l_{sk} \psi_k)}{P^{1-\sigma}} ds' \right)^\rho \right)^{\frac{1-\rho}{\rho}} \right) \right) \\ &= \left(\frac{p_{s'}^{-\sigma(\rho-1)}}{p_{s'}} \left(\frac{\pi + w_s l_{sk} \psi_k}{P^{1-\sigma}} \right)^{\rho-1} \left(\frac{\pi + w_s l_{sk} \psi_k}{P^{1-\sigma}} \right)^{1-\rho} \left(\int p_{s'}^{-\rho\sigma} ds' \right)^{\frac{1-\rho}{\rho}} \right) \\ &= \left(\left(\int p_{s'}^{-\frac{\rho}{1-\rho}} ds' \right)^{\frac{1-\rho}{\rho}} \right) \\ &= P^{-1} \end{aligned}$$

where P is the standard Dixit-Stiglitz price aggregator. This shows that ξ_{sk} is the same for any sk . Intuitively, labor supply levels are determined by the real wage level, so $\xi_{sk} = P^{-1}$ yields labor supply condition:

$$l_{sk}^{supply}(w) = \left(\frac{\psi_k w_s}{P} \right)^\epsilon \quad (14)$$

as well as consumer demand:

$$x_{ss'k}^{Demand}(p, w, \xi, P) = \frac{p_{s'}^{-\sigma} (\pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} P^{-\epsilon})}{P^{1-\sigma}} \quad \forall s, s' \quad (15)$$

Finally, we can integrate over all the individuals $s'k$ to calculate an aggregate demand curve for good s :

$$X_s^{demand} = \underbrace{\frac{\int \int (\pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} P^{-\epsilon}) ds' dk}{P^{1-\sigma}}}_{\text{national real income: } Y_g} p_s^{-\sigma}$$

A.1.2 Firm Decisions: Product Supply and Labor Demand

We assume that each region has a single firm s , which takes the national demand curve and local wages as exogenous. Firms rebate profits π nationally and equally across types. The firm serves the national market, in which it faces aggregate demand $X_s(p_s)$ but hires L_s labor locally at wage w_s .

$$\max_{p_s, L_s, X_s} p_s X_s(p_s) - w_s L_s \quad \text{subject to}$$

- (i) $X_s^{Demand} = p_s^{-\sigma} Y_g$
- (ii) $L_s^{Demand} = X_s^{Supply} / A$

$$\begin{aligned}
&\Longleftrightarrow \max_{p_s} p_s^{1-\sigma} Y_g - \frac{w_s}{A} p_s^{-\sigma} Y_g \\
&\Rightarrow \text{Firm FOC: } p_s = \frac{\sigma}{\sigma-1} \frac{w_s}{A}
\end{aligned} \tag{16}$$

Note that firm optimization

A.1.3 Product Market Equilibrium and Labor Demand

Note that product market equilibrium $X_s^{Demand} = X_s^{Supply}$ is already guaranteed by firm optimization as a monopolist. Combining constraint (i) from the firm's problem with firm price equation (16)

$$X_s^{demand} = Y_g \left(\frac{\sigma}{\sigma-1} \frac{w_s}{A} \right)^{-\sigma}$$

Inverting the production function $X = AL$ as $L = X/A$ gives a company's labor demand as a function of wages and downward-sloping demand for their good.

$$L^{Demand}(w_s) = A^{\sigma-1} \left(\frac{\sigma}{\sigma-1} \right)^{-\sigma} w_s^{-\sigma} Y_g \tag{17}$$

A.1.4 Labor Market Equilibrium

Recall that aggregate labor supply is given by the individual labor supply decision, equation (14), times the share of individuals μ_{sk} in the regional market multiplied by the productivity per unit of labor supplied ψ_k .

Now, for state s , we equate aggregate labor supply, equation (11) and aggregate labor demand, equation (17). Also define $\gamma_s = \int \psi_k^{1+\varepsilon} \mu_{sk} dk$ as the amount of "effective skill" in a city.

$$\begin{aligned}
L_s^{Demand}(w) &= L_s^{supply}(w) \\
A^{\sigma-1} \left(\frac{\sigma}{\sigma-1} \right)^{-\sigma} w_s^{-\sigma} Y_g &= \int l_{sk}^{supply}(w) \mu_{sk} dk \\
&= \int \psi_k \left(\psi_k \frac{w_s}{P} \right)^\varepsilon \mu_{sk} dk \\
&= w_s^\varepsilon P^{-\varepsilon} \int \psi_k^{1+\varepsilon} \mu_{sk} dk \\
\Rightarrow w_s^{\varepsilon+\sigma} &= A^{\sigma-1} \left(\frac{\sigma}{\sigma-1} \right)^{-\sigma} Y_g \gamma_s^{-1} P^\varepsilon \\
w_s^{market-clearing} &= A^{\frac{\sigma-1}{\varepsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma}{\varepsilon+\sigma}} Y_g^{\frac{1}{\varepsilon+\sigma}} \gamma_s^{\frac{-1}{\varepsilon+\sigma}} P^{\frac{\varepsilon}{\varepsilon+\sigma}}
\end{aligned}$$

A.2 Indirect Utility in Productiveville

We have the utility function:

$$U_{sk} = \left(\int x_{ss'k}^\rho ds' \right)^{\frac{1}{\rho}} - \frac{l_{sk}^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}}$$

Recall from equations (14) and (15) that

$$\begin{aligned}
l_{sk}^{supply}(w) &= \left(\frac{\psi_k w_s}{P} \right)^\varepsilon \\
x_{ss'k}^{Demand}(p, w, \xi, P) &= \frac{p_{s'}^{-\sigma} (\pi + \psi_k^{1+\varepsilon} w_s^{1+\varepsilon} P^{-\varepsilon} - r_s)}{P^{1-\sigma}}
\end{aligned}$$

where r_i is the rental cost of housing in Productiveville. We can rewrite utility as:

$$\begin{aligned}
V &= \left(\int \left(\frac{p_{s'}^{-\sigma} (\pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} P^{-\epsilon} - r_s)}{P^{1-\sigma}} \right)^{\rho} ds' \right)^{\frac{1}{\rho}} - \frac{(\psi^{\epsilon} (\frac{w_s}{P})^{\epsilon})^{\frac{1+\epsilon}{\epsilon}}}{1 + \frac{1}{\epsilon}} \\
&= (\pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} P^{-\epsilon} - r_s) \left(\int \left(\frac{p_{s'}^{-\sigma}}{P^{1-\sigma}} \right)^{\rho} ds' \right)^{\frac{1}{\rho}} - \frac{(\psi^{\epsilon} (\frac{w_s}{P})^{\epsilon})^{\frac{1+\epsilon}{\epsilon}}}{1 + \frac{1}{\epsilon}} \\
&= (\pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} P^{-\epsilon} - r_s) P^{-1} - \frac{(\psi^{\epsilon} (\frac{w_s}{P})^{\epsilon})^{\frac{1+\epsilon}{\epsilon}}}{1 + \frac{1}{\epsilon}} \\
&= \left(\psi_k^{1+\epsilon} w_s^{1+\epsilon} \frac{P^{-\epsilon}}{1 + \epsilon} + (\pi - r_s) \right) P^{-1}
\end{aligned}$$

From before we had, $w_s^{market-clearing} = A^{\frac{\sigma-1}{\epsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma}{\epsilon+\sigma}} Y_g^{\frac{1}{\epsilon+\sigma}} \gamma_s^{\frac{-1}{\epsilon+\sigma}} P^{\frac{\epsilon}{\epsilon+\sigma}}$, so

$$V = P^{-1} \left(\psi_k^{1+\epsilon} \left(A^{\frac{\sigma-1}{\epsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma}{\epsilon+\sigma}} Y_g^{\frac{1}{\epsilon+\sigma}} \gamma_s^{\frac{-1}{\epsilon+\sigma}} P^{\frac{\epsilon}{\epsilon+\sigma}} \right)^{1+\epsilon} \frac{P^{-\epsilon}}{1 + \epsilon} + (\pi - r_s) \right)$$

A.3 Regional Income Per Capita

With the market-clearing wage, we can go back to the individual labor supply condition, equation (14), to solve for per capita income for type sk

$$\begin{aligned}
y_{sk} &= \pi + w_s \psi_k l_{sk} = \pi + \psi_k^{1+\epsilon} w_s^{1+\epsilon} P^{-\epsilon} \\
&= \pi + \psi_k^{1+\epsilon} A^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \gamma_s^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma(1+\epsilon)}{\sigma+\epsilon}} Y_g^{\frac{1+\epsilon}{\epsilon+\sigma}} P^{\frac{\epsilon(1+\epsilon)}{\epsilon+\sigma}} P^{\frac{-\epsilon(\epsilon+\sigma)}{\epsilon+\sigma}} \\
&= \pi + \psi_k^{1+\epsilon} A^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \gamma_s^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma(1+\epsilon)}{\sigma+\epsilon}} Y_g^{\frac{1+\epsilon}{\epsilon+\sigma}} P^{\frac{\epsilon(1-\sigma)}{\epsilon+\sigma}}
\end{aligned}$$

Integrating this over different skill groups, we have per capita labor income in a place as

$$\begin{aligned}
Y_s &= \int y_{sk} \mu_{sk} dk = \int \pi \mu_{sk} dk + \int \left(\psi_k^{1+\epsilon} A^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \gamma_s^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma(1+\epsilon)}{\sigma+\epsilon}} Y_g^{\frac{1+\epsilon}{\epsilon+\sigma}} P^{\frac{\epsilon(1-\sigma)}{\epsilon+\sigma}} \right) \mu_{sk} dk \\
&= \int \pi \mu_{sk} dk + \left(A^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma(1+\epsilon)}{\sigma+\epsilon}} Y_g^{\frac{1+\epsilon}{\epsilon+\sigma}} P^{\frac{\epsilon(1-\sigma)}{\epsilon+\sigma}} \right) \gamma_s^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \underbrace{\left(\int \psi_k^{1+\epsilon} \mu_{ik} dk \right)}_{\gamma_s} dk
\end{aligned}$$

Note that

$$\gamma_s^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \gamma_s = \gamma_s^{\frac{-(1+\epsilon)+\sigma+\epsilon}{\sigma+\epsilon}} = \gamma_s^{\frac{\sigma-1}{\sigma+\epsilon}}$$

and define

$$\kappa = \left(A^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma(1+\epsilon)}{\sigma+\epsilon}} Y_g^{\frac{1+\epsilon}{\epsilon+\sigma}} P^{\frac{\epsilon(1-\sigma)}{\epsilon+\sigma}} \right)$$

Then, normalizing $\pi = 0$, we have

$$Y_s = \kappa \left(\int \psi_k^{1+\epsilon} \mu_{sk} dk \right)^{\frac{\sigma-1}{\sigma+\epsilon}}$$

$$= \kappa \left(\frac{\int \psi_k^{1+\epsilon} \mu_{sk} dk}{\int \mu_{sk} dk} \right)^{\frac{\sigma-1}{\sigma+\epsilon}} \left(\int \mu_{sk} dk \right)^{\frac{\sigma-1}{\sigma+\epsilon}}$$

Per capita income is equal to

$$\begin{aligned} \frac{Y_i}{\int \mu_{sk} dk} &= \kappa \left(\frac{\int \psi_k^{1+\epsilon} \mu_{sk} dk}{\int \mu_{sk} dk} \right)^{\frac{\sigma-1}{\sigma+\epsilon}} \left(\int \mu_{sk} dk \right)^{\frac{\sigma-1}{\sigma+\epsilon}} \left(\int \mu_{sk} dk \right)^{-1} \\ &= \kappa \left(\frac{\int \psi_k^{1+\epsilon} \mu_{sk} dk}{\int \mu_{sk} dk} \right)^{\frac{\sigma-1}{\sigma+\epsilon}} \left(\int \mu_{sk} dk \right)^{\frac{-(1+\epsilon)}{\sigma+\epsilon}} \\ \Rightarrow \log \frac{Y_i}{\int \mu_{ik}} &= \log \kappa + \underbrace{\frac{\sigma-1}{\sigma+\epsilon} \log \left(\frac{\int \psi_k^{1+\epsilon} \mu_{sk} dk}{\int \mu_{sk} dk} \right)}_{\substack{\text{per cap income} \\ \varepsilon_{\text{population}}}} - \underbrace{\frac{(1+\epsilon)}{\sigma+\epsilon} \log \left(\int \mu_{sk} dk \right)}_{\substack{\text{per cap income} \\ \varepsilon_{\text{avg human capital}}}} \end{aligned}$$

A.4 Proof of Proposition 1

Using equations (4) and (5), we see that $\varepsilon_{\text{pop}}^{\text{per capita inc}} < 0$ and $\varepsilon_{\text{avg skill level}}^{\text{per capita inc}} > 0$. Thus we need only show $\lambda_k^* = \lambda^* \forall k$ to verify the proposition. We note that because housing supply is completely elastic ($\lim_{\beta \rightarrow \infty} N^{1/\beta} = 1$), $\Lambda_{\text{productiveville}}$ is proportional to $\psi_k^{1+\epsilon}$, just as moving costs are. Thus the proposition is proved if $\Lambda_{\text{Reservationville},k}$ is proportional to $\psi_k^{1+\epsilon}$.

To ease the math, define \tilde{v}_t as the excess flow utility earned by a person with $\psi = 1$ in Productiveville relative to Reservationville at time t . Further define the excess value function to living in Productiveville $\tilde{\Lambda}_{p,T} = \sum_{t=T}^{\infty} \beta^t \tilde{v}_t$. By our assumptions on the production process, $\tilde{\Lambda}_{p,k,T} = \psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T}$.

With these definitions in hand, agents in Reservationville draw stochastic moving costs $\lambda \sim F$ proportional to $\psi_k^{1+\epsilon}$ and decide whether or not to migrate. Migrating with lower moving costs is strictly better, generating a cutoff value for each type λ_k^* below which they move and above which they don't. λ_k^* is defined by an indifference condition on moving:

$$\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \psi_k^{1+\epsilon} \lambda_k^* = \beta \Lambda_{\text{Reservationville},k,T}$$

. Note that:

$$\begin{aligned} \Lambda_{\text{Reservationville},k,T} &= \int_0^{F^{-1}\left(\frac{\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \beta \Lambda_{Rk}}{\psi_k^{1+\epsilon}}\right)} \left(\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \lambda \psi_k^{1+\epsilon} \right) f(\lambda) d\lambda + \int_{F^{-1}\left(\frac{\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \beta \Lambda_{Rk}}{\psi_k^{1+\epsilon}}\right)}^1 \beta \Lambda_{R,T} f(\lambda) d\lambda \\ \Lambda_{\text{Reservationville},k,T} &= \frac{\psi_k^{1+\epsilon} F^{-1}\left(\frac{\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \beta \Lambda_{Rk}}{\psi_k^{1+\epsilon}}\right) \left(\tilde{\Lambda}_{p,T} - .5 F^{-1}\left(\frac{\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \beta \Lambda_{Rk}}{\psi_k^{1+\epsilon}}\right) \right)}{\left(1 - \beta + \beta F^{-1}\left(\frac{\psi_k^{1+\epsilon} \tilde{\Lambda}_{p,T} - \beta \Lambda_{Rk}}{\psi_k^{1+\epsilon}}\right) \right)} \end{aligned}$$

Finally, we complete the proof by guessing and verifying. If we assume $\Lambda_{\text{Reservationville},T} = \psi_k^{1+\epsilon} \zeta$, then plugging above, we get

$$\Lambda_{\text{Reservationville},T,k} = \psi_k^{1+\epsilon} \underbrace{\frac{F^{-1}\left(\tilde{\Lambda}_{p,T} - \beta \zeta\right) \left(\tilde{\Lambda}_{p,T} - .5 F^{-1}\left(\tilde{\Lambda}_{p,T} - \beta \zeta\right) \right)}{\left(1 - \beta + \beta F^{-1}\left(\tilde{\Lambda}_{p,T} - \beta \zeta\right) \right)}}_{\zeta}$$

A.5 Proof of Proposition 2

A.5.1 The Path of Income and Population Over Time

Define $\tau = \left(A^{\frac{(\sigma-1)(1+\epsilon)}{\epsilon+\sigma}} \left(\frac{\sigma}{\sigma-1} \right)^{\frac{-\sigma(1+\epsilon)}{\sigma+\epsilon}} Y_g^{\frac{1+\epsilon}{\epsilon+\sigma}} P^{\frac{\epsilon(1-\sigma)}{\epsilon+\sigma}} \right)$, a national constant. Per capita income in an area is a function of this constant and the local population:

$$\frac{Y_s}{\int \mu_{sk}} = \tau \left(\int \mu_{sk} dk \right)^{\frac{-(1+\epsilon)}{\sigma+\epsilon}}$$

Define $\epsilon = \frac{(1+\epsilon)}{\sigma+\epsilon}$ as the elasticity of both per capita income and indirect utility with respect to population. Further, let A be the ratio of relative productivity in N relative to S . Here we use notation N for the Northern rich region and S for the Southern poor region; this is interchangeable with Productiveville and Reservationville. We then have per capita incomes:

$$y_{Nt} = A\tau (\mu_{Nt})^{-\epsilon} \text{ and } y_{St} = \tau (\mu_{St})^{-\epsilon}$$

Let x be the share of people leaving place S for place N . The gap in per capita income growth rates between North and South is

$$\begin{aligned} \Delta \ln(y) &= \ln(y_{Nt}) - \ln(y_{St}) \\ &= -\epsilon \left(\ln \left(\frac{\mu_{Nt}}{\mu_{Nt-1}} \right) \right) - \epsilon \left(\ln \left(\frac{\mu_{St}}{\mu_{St-1}} \right) \right) \\ &= -\epsilon \left(\ln \left(\frac{\mu_{Nt-1} + x\mu_{St-1}}{\mu_{Nt-1}} \right) - \ln \left(\frac{(1-x)\mu_{St-1}}{\mu_{St-1}} \right) \right) \\ &= -\epsilon \left(\ln \left(1 + \frac{x\mu_{St-1}}{\mu_{Nt-1}} \right) - \ln(1-x) \right) \end{aligned}$$

The convergence rate is the gap in per capita growth rates divided by the gap in levels. We set this to a negative constant κ .

$$\frac{\Delta \ln(y)}{\Delta \ln(y_{t-1})} = \frac{-\epsilon \left(\ln \left(1 + \frac{x\mu_{St-1}}{\mu_{Nt-1}} \right) - \ln(1-x) \right)}{\ln(A) - \epsilon \ln \left(\frac{\mu_{Nt-1}}{\mu_{St-1}} \right)} = \underbrace{\kappa}_{\text{convergence}}$$

Given this constant, we can solve for x :

$$\begin{aligned} -\epsilon \left(\ln \left(\frac{1 + \frac{x\mu_{St-1}}{\mu_{Nt-1}}}{1-x} \right) \right) &= \kappa \left(\ln \left(A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon \right) \right) \\ e^{-\epsilon} \left(\frac{1 + \frac{x\mu_{St-1}}{\mu_{Nt-1}}}{1-x} \right) &= e^{\kappa} A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon \\ x \frac{\mu_{St-1}}{\mu_{Nt-1}} &= e^{\kappa+\epsilon} A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon (1-x) - 1 \\ x \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} + e^{\kappa+\epsilon} A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon \right) &= e^{\kappa+\epsilon} A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon - 1 \\ x &= \frac{e^{\kappa+\epsilon} A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon - 1}{\left(\frac{\mu_{St-1}}{\mu_{Nt-1}} + e^{\kappa+\epsilon} A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon \right)} \end{aligned}$$

Because $A \left(\frac{\mu_{St-1}}{\mu_{Nt-1}} \right)^\epsilon = \frac{Y_{Nt-1}}{Y_{St-1}}$, we can rewrite this as

$$x = \frac{e^{\kappa+\epsilon} \left(\frac{Y_{Nt-1}}{Y_{St-1}} \right) - 1}{e^{\kappa+\epsilon} \left(\frac{Y_{Nt-1}}{Y_{St-1}} \right) + A^{\frac{-1}{\epsilon}} \left(\frac{Y_{Nt-1}}{Y_{St-1}} \right)^{\frac{1}{\epsilon}}}$$

Finally, define Y_{N0} as income in the North and Y_{S0} as income in the South at $t = t_0$. Then

$$x_t^* = x = \frac{e^{\kappa+\epsilon} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right) - 1}{e^{\kappa+\epsilon} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right) + A^{\frac{-1}{\epsilon}} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)^{\frac{1}{\epsilon}}}$$

We need optimal migration from the South to produce this fraction of the Southern population moving North for each time t . Below, we derive conditions under which this fraction is declining over time. It is intuitive that the share of the Southern population moving would fall over time, because as migration rates should fall as the benefit to moving falls. Still, the ratio between the amount of directed migration and the initial income gap will be constant, so that income convergence continues at constant rate.

A.5.2 Individual Migration Decisions

Consider an agent in the South deciding whether to move to the North today or stay in the South, with the possibility of moving in the future, valued at V_{T+1}^* . This agent discounts the future at rate ρ . In each period, agents draw i.i.d. moving costs $\lambda \sim F$. Define $\lambda_T^* = F^{-1}(x_T^*)$. The agent will move if

$$\begin{aligned} \text{Gain to Moving at } T - \text{Flow Cost} &> \text{Benefit to Moving Later} \\ \sum_{t=T}^{\infty} e^{-\rho t} \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} - \lambda_T &> e^{-\rho T} V_{T+1}^* \end{aligned}$$

At x_T^* , the agent is indifferent between moving and staying. This implies that

$$\lambda_T^* \equiv F^{-1}(x_T^*) = \underbrace{\sum_{t=T}^{\infty} e^{-\rho t} \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)}}_{\text{Gains to moving at } T} - \underbrace{e^{-\rho T} V_{T+1}^*}_{\text{Benefit to Staying}}$$

The benefit to waiting is that expected future migration costs are lower. We know at each period how likely it is that the agent would choose to move in all future periods. So we can integrate up the value the agent gets from eventually winding up in Productiveville. The difference between that and the value of moving today is the expected savings in moving costs. This defines the distribution of moving costs for the part of the distribution hit covered by the sequence $\{x_t^*\}_{t=0}^{\infty}$.

$$\begin{aligned} \underbrace{\sum_{t=T}^{\infty} \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{-(\rho+\kappa)t+\kappa t_0}}_{\text{Excess Gains to moving at } T} - \underbrace{\sum_{t=T+1}^{\infty} \left(\prod_{j=T}^t (1-x_j^*) \right) x_t^* \sum_{t_2=t}^{\infty} \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{-(\rho+\kappa)t_2+\kappa t_0}}_{\text{Excess Gains From Moving Eventually}} &= \underbrace{F^{-1}(x_T^*) - E[\text{Future MC} | T]}_{\text{Cost to moving now - Eventual Moving Cost}} \\ &= F^{-1}(x_T^*) - \sum_{t=T+1}^{\infty} \prod_{j=T}^t (1-x_j^*) \int_0^{F^{-1}(x_t^*)} \lambda \cdot f(\lambda) \cdot d\lambda \end{aligned}$$

A.5.3 Finding An Interior Solution

To finish the proof, we need to show that $\frac{dx_t^*}{dt} < 0$ for $t > 0$. Because income gaps between North and South are falling, this implies that we need the fraction of Southern residents leaving each period to be declining. This ensures that the dynamic problem described above has an interior solution. Recall from the previous section that

$$\begin{aligned}
 x_t^* &= \frac{e^{\kappa+\epsilon} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right) - 1}{e^{\kappa+\epsilon} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right) + A^{-\frac{1}{\epsilon}} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)^{\frac{1}{\epsilon}}} \\
 \frac{dx_t^*}{dt} &= \underbrace{\frac{e^{\kappa+\epsilon} \left(\frac{Y_{N0}}{Y_{S0}} \right) \kappa e^{\kappa(t-t_0)}}{e^{\kappa+\epsilon} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right) + A^{-\frac{1}{\epsilon}} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)^{\frac{1}{\epsilon}}}}_{P1} \\
 &\quad - \underbrace{\frac{x_t^*}{\left(e^{\kappa+\epsilon} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right) + A^{-\frac{1}{\epsilon}} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)^{\frac{1}{\epsilon}} \right)}}_{P2} \times \\
 &\quad \underbrace{\left(e^{\kappa+\epsilon} \left(\frac{Y_{N0}}{Y_{S0}} \right) \kappa e^{\kappa(t-t_0)} + \frac{A^{-\frac{1}{\epsilon}} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)^{\frac{1}{\epsilon}} \left(\frac{Y_{N0}}{Y_{S0}} \right) \kappa e^{\kappa(t-t_0)}}{\epsilon \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)} \right)}_{P3}
 \end{aligned}$$

$$\frac{dx_t^*}{dt} < 0 \iff P1 < P2 \times P3$$

Cancelling terms, we can rewrite that as:

$$\kappa e^{\kappa+\epsilon} < \kappa x_t^* \times \underbrace{\left(e^{\kappa+\epsilon} + \frac{A^{-\frac{1}{\epsilon}} \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)^{\frac{1}{\epsilon}}}{\epsilon \left(1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)} \right)} \right)}$$

Define $\eta = 1 + \left(\frac{Y_{N0}}{Y_{S0}} \right) e^{\kappa(t-t_0)}$.

$$\begin{aligned}
 \kappa e^{\kappa+\epsilon} &< \kappa x_t^* \times \left(e^{\kappa+\epsilon} + \frac{A^{-\frac{1}{\epsilon}} \eta^{\frac{1-\epsilon}{\epsilon}}}{\epsilon} \right) \\
 e^{\kappa+\epsilon} &> \frac{e^{\kappa+\epsilon} \eta - 1}{e^{\kappa+\epsilon} \eta + A^{-\frac{1}{\epsilon}} \eta^{1/\epsilon}} \times \left(e^{\kappa+\epsilon} + \frac{A^{-\frac{1}{\epsilon}}}{\epsilon} (\eta)^{\frac{1-\epsilon}{\epsilon}} \right) \\
 \Rightarrow (e^{\kappa+\epsilon})^2 \eta + e^{\kappa+\epsilon} A^{-\frac{1}{\epsilon}} \eta^{1/\epsilon} &> (e^{\kappa+\epsilon})^2 \eta - e^{\kappa+\epsilon} + e^{\kappa+\epsilon} \eta \frac{A^{-\frac{1}{\epsilon}}}{\epsilon} (\eta)^{\frac{1-\epsilon}{\epsilon}} - \frac{A^{-\frac{1}{\epsilon}}}{\epsilon} (\eta)^{\frac{1-\epsilon}{\epsilon}} \\
 \Rightarrow e^{\kappa+\epsilon} \left(A^{-\frac{1}{\epsilon}} \eta^{1/\epsilon} + 1 - \frac{A^{-\frac{1}{\epsilon}}}{\epsilon} (\eta)^{\frac{1}{\epsilon}} \right) &> -\frac{A^{-\frac{1}{\epsilon}}}{\epsilon} (\eta)^{\frac{1-\epsilon}{\epsilon}} \\
 \Rightarrow e^{\kappa+\epsilon} &> -\frac{A^{-\frac{1}{\epsilon}}}{\epsilon} (\eta)^{\frac{1-\epsilon}{\epsilon}} \setminus \left(1 - \frac{\epsilon - 1}{\epsilon} A^{-\frac{1}{\epsilon}} (\eta)^{\frac{1}{\epsilon}} \right) = \eta^{\frac{1-\epsilon}{\epsilon}} \setminus \left((1-\epsilon) \eta^{\frac{1}{\epsilon}} - \epsilon A^{\frac{1}{\epsilon}} \right)
 \end{aligned}$$

Plugging back in for η gives

$$e^{\kappa+\epsilon} > \frac{\left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right) e^{\kappa(t-t_0)}\right)^{\frac{1-\epsilon}{\epsilon}}}{\left((1-\epsilon) \left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right) e^{\kappa(t-t_0)}\right)^{\frac{1}{\epsilon}} - \epsilon A^{\frac{1}{\epsilon}}\right)}$$

We need this to be true at both $t = t_0$ and $t = \infty$. At $t = t_0$, $e^{\kappa(t-t_0)} = e^0 = 1$, and at $t = \infty$, $e^{-\infty} = 0$. This gives us the conditions:

$$\begin{aligned} e^{\kappa+\epsilon} &> \max\left\{\frac{\left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)\right)^{\frac{1-\epsilon}{\epsilon}}}{\left((1-\epsilon) \left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)\right)^{\frac{1}{\epsilon}} - \epsilon A^{\frac{1}{\epsilon}}\right)}, \frac{1}{(1-\epsilon - \epsilon A^{1/\epsilon})}\right\} \\ &= \frac{\left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)\right)^{\frac{1-\epsilon}{\epsilon}}}{\left((1-\epsilon) \left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)\right)^{\frac{1}{\epsilon}} - \epsilon A^{\frac{1}{\epsilon}}\right)} \\ \kappa &> \log \frac{\left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)\right)^{\frac{1-\epsilon}{\epsilon}}}{\left((1-\epsilon) \left(1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)\right)^{\frac{1}{\epsilon}} - \epsilon A^{\frac{1}{\epsilon}}\right)} - \epsilon \end{aligned}$$

So as long as this combination of ϵ , A , and $1 + \left(\frac{Y_{N0}}{Y_{S0}}\right)$ are sufficiently small, then there exists some moving cost distribution F such that convergence occurs at a constant rate.

A.6 Proof of Proposition 3

We again assume that $\lambda \geq 0$, $\Delta(k) > 0$ for all k , and that $N_P > 1$. The final assumption is without loss of generality and simply implies that an increase in β raises prices in the state ($\frac{\partial p_P}{\partial \beta} > 0$). ($N_P^{1/\beta} - 1$) is continuous and strictly monotonic in β and $\lim_{\beta \rightarrow 0} (N_P^{1/\beta} - 1) \rightarrow \infty$. By Bolzano's theorem a strictly monotonic and continuous function will intersect a fixed value at most once on the half-closed interval bounded by 0. The indirect utility excluding housing in Productiville, $\Delta(k)$, is by definition continuous and strictly monotonic in ψ . As a result, there is a unique β_k^* for which the left-hand side of equation 7 equal zero for each skill type, and this unique value β_k^* is decreasing in k . For $\beta > \beta_k^*$, the left-hand side of equation 7 is strictly positive implying in-migration for skill type k . For $\beta < \beta_k^*$, the left-hand side of equation 7 is strictly negative implying out-migration for skill type k . The first case, where $\beta \rightarrow \infty$, is proven in proposition 1.

A.7 Endogenous Housing Demand

In Section 2, we derived a model in which all individuals had the same demand for one plot of land, which meant that housing demand was completely inelastic. That assumption is convenient for exposition, but not necessary for our results. Here we derive the two propositions in the text while allowing for endogenous, non-homothetic housing demand.

A.7.1 Individual Decisions: Labor Supply and Housing Demand

We assume quasi Stone-Geary preferences, with a minimum housing requirement of H per worker. Workers solve the following problem:

$$\arg \max_{c, h, L} U = \beta \ln(h - H) + (1 - \beta) \ln c - \frac{L^{1+\frac{1}{\epsilon}}}{1 + \frac{1}{\epsilon}} + \lambda [w\psi L - c - ph]$$

The first order conditions are

$$U_c : \frac{1-\beta}{c} = \lambda$$

$$U_h : \frac{\beta}{h-H} = \lambda p$$

$$U_L : L^{1/\epsilon} = \lambda w \psi$$

which implies the following ratios between labor, consumption, and housing:

$$L = c^{-\epsilon} (1-\beta)^\epsilon w^\epsilon \psi^\epsilon$$

$$p(h-H) = \frac{\beta c}{(1-\beta)}$$

We can use these ratios along with the household budget constraint to derive expressions for household allocations in terms of prices and parameters:

$$c^{1+\epsilon} + c^\epsilon p H (1-\beta) - (1-\beta)^{1+\epsilon} w^{1+\epsilon} \psi^{1+\epsilon} = 0$$

We now simplify the analysis by assuming $\epsilon = 1$, we can apply the quadratic formula to the previous equation:

$$c = \frac{-pH(1-\beta) + (1-\beta)\sqrt{p^2 H^2 + 4w^2 \psi^2}}{\underbrace{2}_{\Lambda}}$$

$$h = \frac{\beta \Lambda}{(1-\beta)p} + H$$

$$L = \frac{\psi w (1-\beta)}{\Lambda}$$

A.7.2 Indirect Utility

Given these allocations, household utility can be expressed in terms of parameters and prices:

$$V(p, H, w, \psi, \beta) = \beta \ln \left(\frac{\beta \Lambda}{(1-\beta)p} \right) + (1-\beta) \ln \Lambda - \frac{1}{2} \left(\frac{\psi w (1-\beta)}{\Lambda} \right)^2$$

Note that if $H = 0$, $p = 1$, $\Lambda = (1-\beta)w\psi$. As a result, we have a simple expression:

$$V(1, 0, w, \psi, \beta) = \beta \ln(\beta w \psi) + (1-\beta) \ln((1-\beta)w\psi) - \frac{1}{2}$$

$$= \ln(w) + \ln(\psi) + \underbrace{\beta \ln \beta + (1-\beta) \ln(1-\beta)}_{\text{Constant}} - \frac{1}{2} \quad (18)$$

An individual's indirect utility, and therefore the attractiveness of migration, increases in the city wage w . Further if we now assume that the utility in Reservationville is proportional to $\ln(\psi)$, and that moving costs are additive, migration is once again independent of skill type. Because we did not change the firm side of the economy, equation (14) is sufficient for establishing the results derived in Proposition 1.

A.7.3 Comparative Statics for Housing Prices and Returns to Migration

To compute the effect of housing prices on migration incentives, we take the derivative of indirect utility with respect to housing prices.

$$\begin{aligned}
V &= \beta \ln \left(\frac{\beta \left[\frac{-pH + \sqrt{p^2 H^2 + 4w^2 \psi^2}}{2} \right]}{p} \right) + (1 - \beta) \ln \left(\frac{-pH(1 - \beta) + (1 - \beta) \sqrt{p^2 H^2 + 4w^2 \psi^2}}{2} \right) \\
&\quad - \frac{\left(\frac{2\psi w(1 - \beta)}{-pH(1 - \beta) + (1 - \beta) \sqrt{p^2 H^2 + 4w^2 \psi^2}} \right)^2}{2} \\
&= \beta \ln \left(\frac{-pH + \sqrt{p^2 H^2 + 4w^2 \psi^2}}{p} \right) + (1 - \beta) \ln \left(-pH + \sqrt{p^2 H^2 + 4w^2 \psi^2} \right) - 2 \left(\frac{\psi w}{-pH + \sqrt{p^2 H^2 + 4w^2 \psi^2}} \right)^2 \\
&\quad + \text{constant terms}
\end{aligned}$$

$$\frac{\partial V}{\partial p} = \frac{pH + \sqrt{p^2 H^2 + 4w^2 \psi^2}}{p \left(pH - \sqrt{p^2 H^2 + 4w^2 \psi^2} \right)} < 0 \quad (19)$$

Equation (19) demonstrates that higher housing prices make moving to the Productiveville less attractive for all types.

We explore how this effect varies by skill type by taking the cross-partial of this derivative with respect to an individual's skill type ψ . That expression yields:

$$\frac{\partial V}{\partial p \partial \psi} = \frac{8\psi H w^2}{\sqrt{p^2 H^2 + 4w^2 \psi^2} \left(pH - \sqrt{p^2 H^2 + 4w^2 \psi^2} \right)^2} > 0 \quad (20)$$

Equations (19) and (20) establish that higher prices reduce utility, and hence the incentive to migrate. This mechanism has less of an impact on the utility of higher skilled agents, which replicates Proposition 2. Intuitively, the result emerges because the minimum housing requirement causes an individual's housing share of expenditure to fall with income. Therefore, higher housing prices induce a larger utility loss for poorer agents.

B Data Appendix

B.1 Human Capital Convergence

Calibrating the extent of human capital convergence across multiple years for all workers is more involved than the reduced forms for people aged 25-44 that are discussed in the text. For each year t , we estimate the returns to human capital using the specification

$$\log WageIncome_{ik} = \alpha_k + X_{ik}\beta + \varepsilon_{ik}$$

For all human capital analyses (Appendix Tables 5, 8, and 10), we winsorize income or wages at the 1st and 99th percentile. Skill level k is defined empirically as the interaction of completed schooling levels (0 or NA, Elementary, Middle, Some HS, HS, Some College, College+), a dummy for black and a dummy for Hispanic. We also run the same regression with $\log \frac{WageIncome_{ik}}{AnnualHours_{ik}}$ as the dependent variable.

We made every effort to code annual hours consistently across years. Because hours last week are reported in intervals in 1960 and 1970, we code each observation using the sample mean of hours within that interval from other years. In 2000 and 2010, the hours data come from the usual number of hours per week.

For people aged 25-44, we computed $\widehat{Inc}_k = \exp(\alpha_k)$ and

$$\Delta HC_s^{25-44} = \sum_k \widehat{Inc}_k Share_{25-44, ks, residence} - \widehat{Inc}_k Share_{25-44, ks, birth}$$

For the calibration, we are interested in the migration flows of older workers, not just those aged 25-44. For people aged 45-64, we add superscript t to index time for when the cohort was observed, we measure the impact of migration in the last 20 years as:

$$\Delta HC_{st}^{45-64} = \sum_k \widehat{Inc}_k (Share_{45-64, ks, residence}^t - Share_{25-44, ks, residence}^{t-20})$$

Specifically, we compare the human capital level of people residing in state s of ages 45-64 today to the human capital level of people residing in state s of ages 25-44 in the Census 20 years ago. To ensure that we are picking up changes in the human capital mix rather than changes in the human capital premia, we use premia \widehat{Inc}_k calculated for year t . For example, for the human capital change from 1960-1980, we would calculate \widehat{Inc}_k from the 1980 Census. Then, letting s_{st}^{25-44} index the share of people aged 25-44 in state s and $\Delta s_{st}^{25-44} = s_{st, res}^{25-44} - s_{st, birth}^{25-44}$, we can compute the total change in human capital due to migration as

$$\Delta HC_{st} = \underbrace{s_{st}^{25-44} \Delta HC_{st}^{25-44}}_{\text{Young Cohort}} + \underbrace{s_{st}^{45-64} \Delta HC_{st}^{45-64}}_{\text{Old Cohort}} + \underbrace{\Delta s_{st}^{25-44} (HC_{st-1}^{25-44} - HC_{st-1}^{45-64})}_{\text{Cohort Share Change} \times \text{Cohort Diff}}$$

We also conduct robustness checks where we characterize the overall change in human capital (including within-state accumulation) as

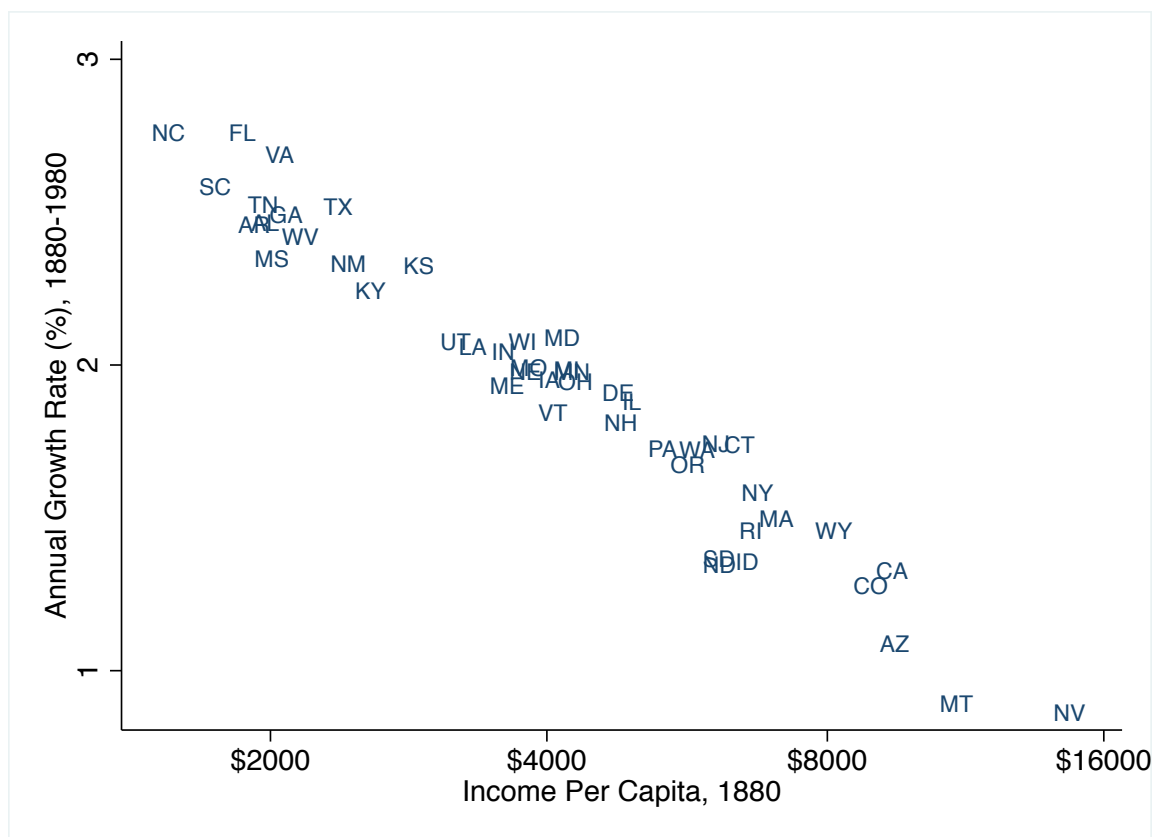
$$\log(HC_t) - \log(HC_{t-20})$$

B.2 Labor Market Area Analysis

For the analysis in Appendix Tables 1 and 2, we construct a panel of income and population at the Labor Market Area (LMA) level. LMAs are linked by intercounty commuting flows and partition the United States (Tolbert and Sizer, 1996). LMA population is constructed simply by adding the population of constituent counties. LMA income is estimated as the population-weighted average of county-level income. The income series uses median family income from 1950-2000 from Haines (2010) and USACounties (2012). In 1940 and 2010, the series is unavailable. In 1940, we use pay per manufacturing worker from Haines (2010). Pay per manufacturing worker which had a correlation of 0.77 with median family income in 1950, a year when both series were available. In 2010, we use median household income from USACounties (2012), which had a correlation of 0.98 with median family income in 2000, a year when both series were available.

APPENDIX FIGURE 1

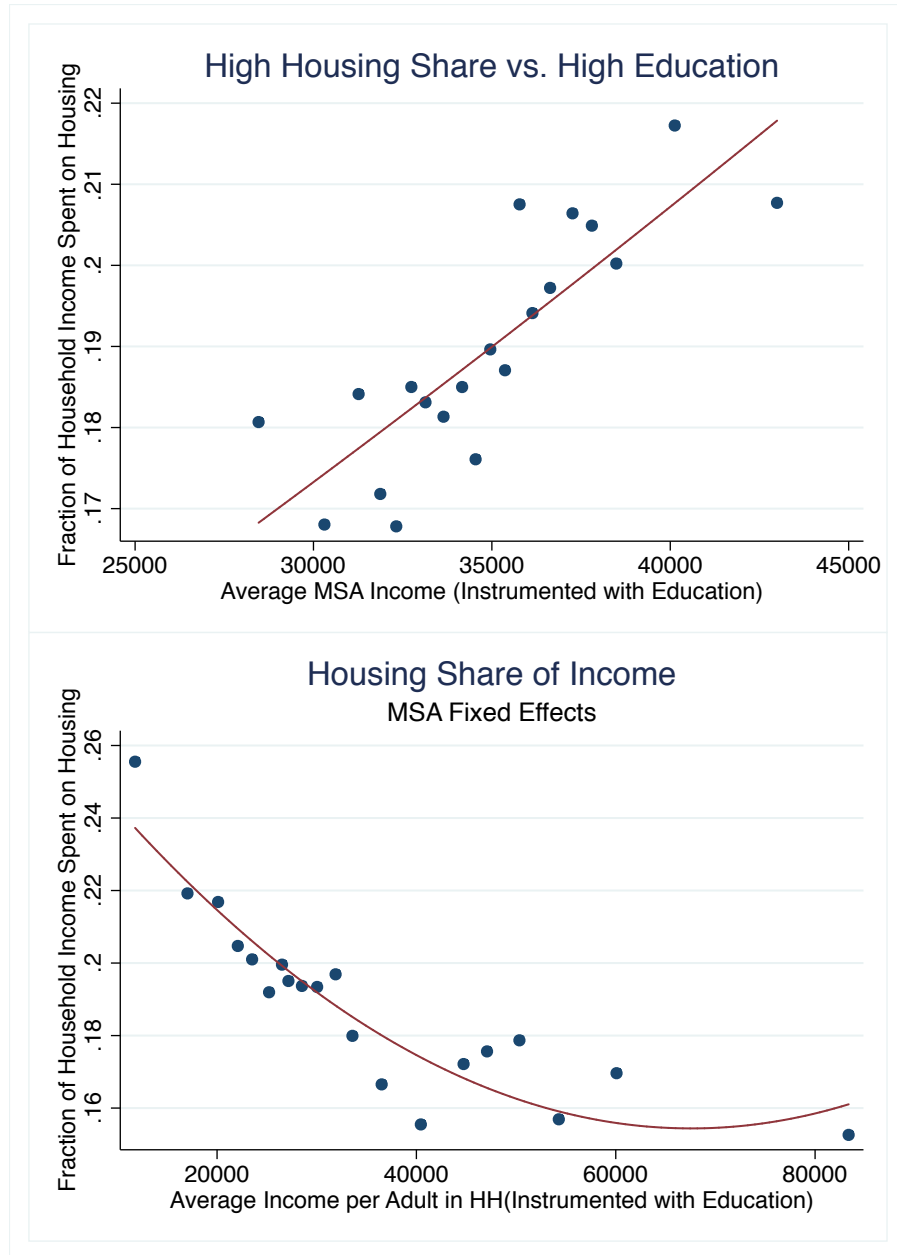
Income Convergence in the United States



Notes: Income data from Kuznets *et al.* (1964), and the Bureau of Economic Analysis (2012). All dollars are in real 2012 terms. Oklahoma was a federal territory at the time of the 1880 Census and so no separate income measurement is available.

APPENDIX FIGURE 2

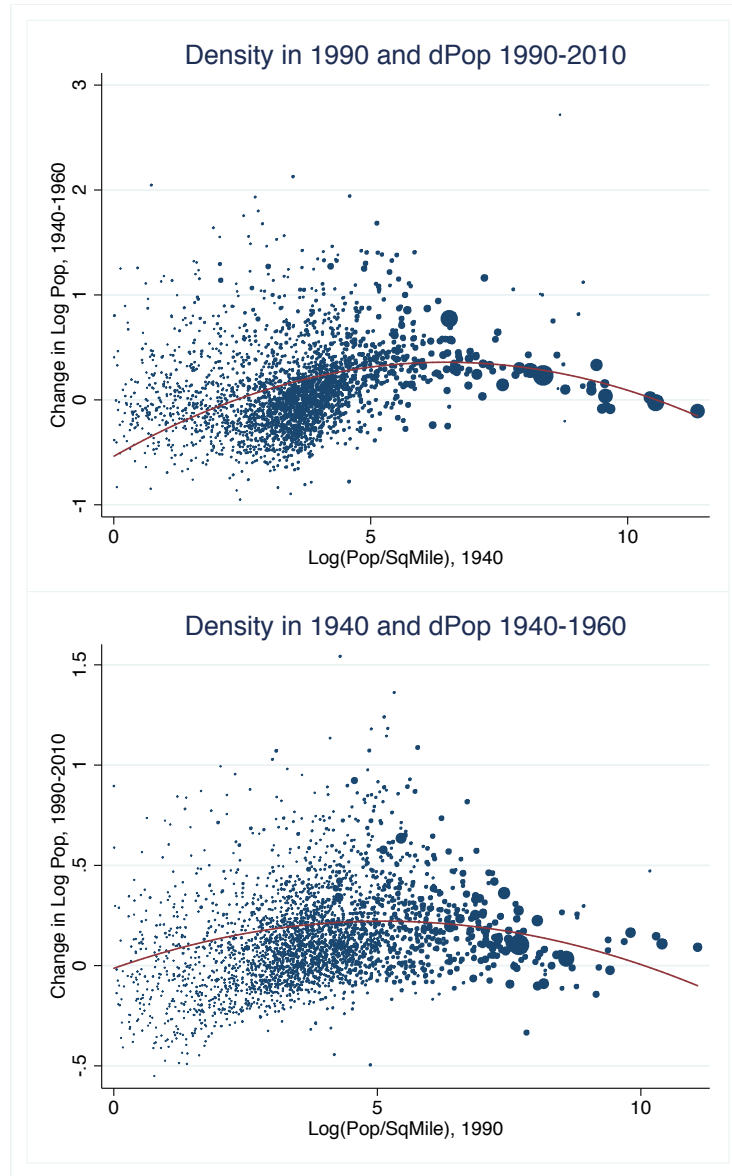
Engel Curve for Housing



Note: This figure plots the relationship between the share of household income spent on housing and average income in the 2010 ACS. Annual income is volatile, meaning that an apparently non-homothetic cross-sectional relationship between housing share and annual income might not reflect the true relationship between housing share and permanent income. To address this issue, we instrument for an individual's income using her education level. We then compute average income as the average of predicted income for adults in the household. The top panel stratifies MSAs into 20 quantiles by MSA average income, and then computes the housing share within each quantile. The bottom panel stratifies households into 20 quantiles by household average income net of an MSA fixed effect, and then computes the housing share within each quantile, conditional on MSA fixed effects. Housing expenditure is computed as twelve times monthly rents or 5% of housing costs. Housing shares above 100% and below zero are excluded.

APPENDIX FIGURE 3

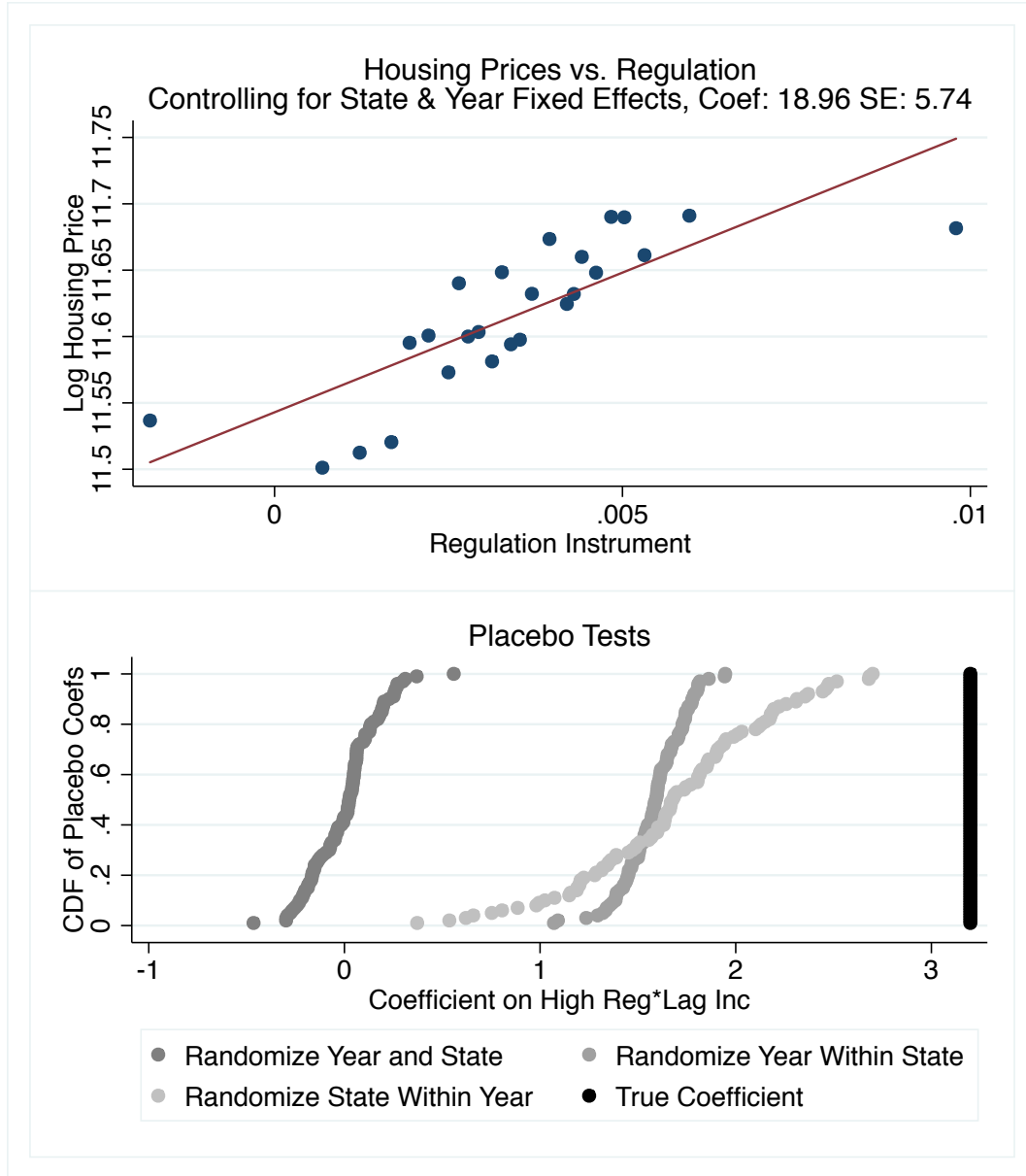
Effect of County Density On Population Growth



Notes: This figure plots the correlation of log county-level densities on with subsequent 20-year log population growth for 1940 and 1990. Dot size is proportional to lagged population. The quadratic fit line is constructed weighting each observation by its lagged population. In both panels, high-density places have lower growth than the national average. In 1940, the population-weighted mean log density was 5.51, and in 1990, the population-weighted mean density was 6.05. The heterogeneity in county densities is enormous compared to this difference in means. In addition, had the densest counties in America reached a “maximum density” between 1940 and 1990, then we would expect to see a sharp drop in population growth there after 1990. No such relation is evident in the data.

APPENDIX FIGURE 4

Regulation Measure: Relevance and Placebo Tests

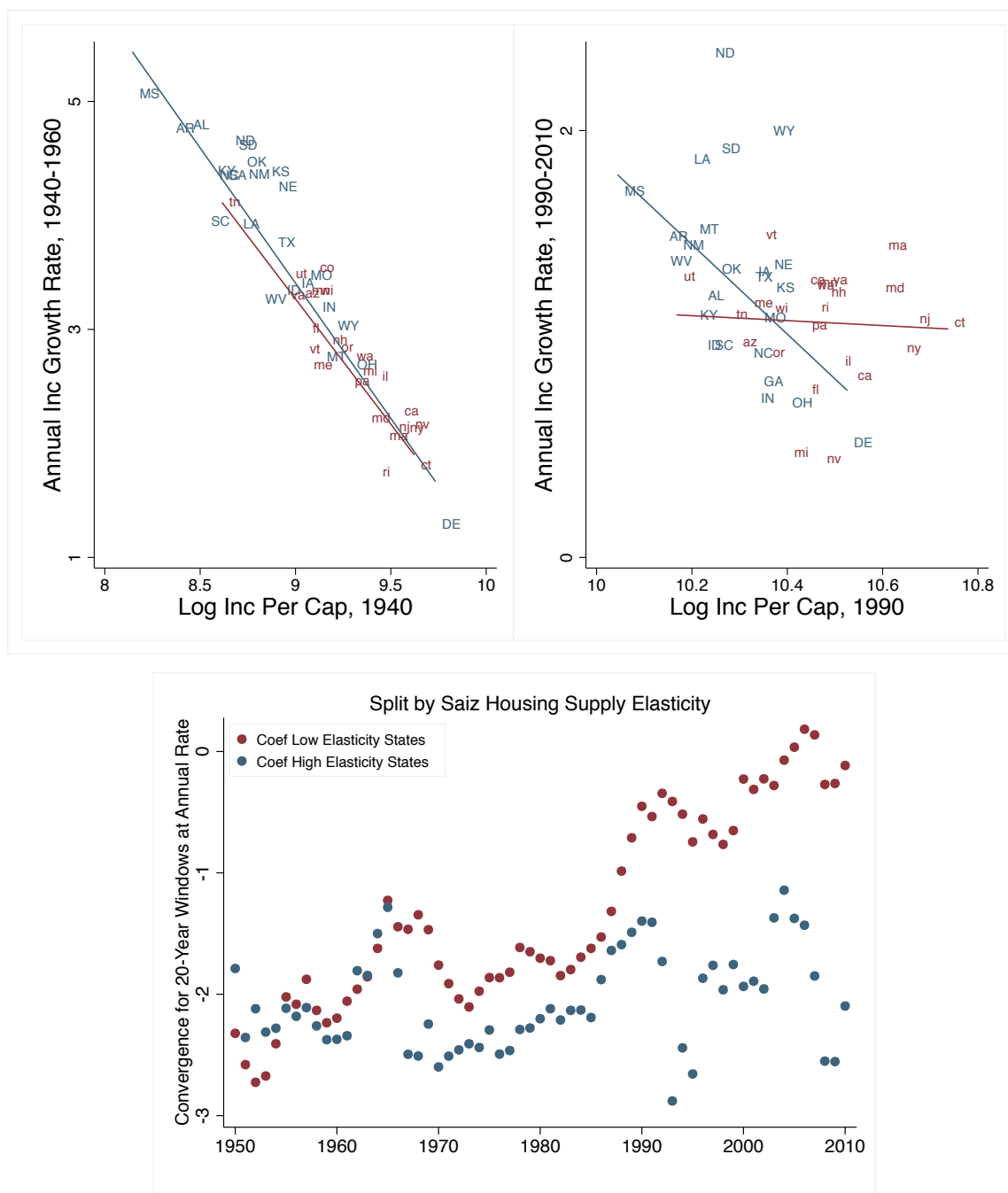


Notes: The first panel depicts the relationship between the regulation measure described in the text and log housing prices after controlling for state and year fixed effects. The coefficient and standard error for this regression are presented below the title. The data are divided in to 20 equally sized bins for the presentation. $n=288$, one observation for year-decade from 1960 forward. Standard errors clustered by state.

The second panel plots the cumulative distribution functions of the coefficients estimated in three sets of placebo monte-carlo experiments. The CDFs plot the coefficients β from the regression $\Delta \log Inc_{it} = \alpha + \beta \mathbf{1}(West_{it} > q_{50}^{West}) \times \log Inc_{i,t-20} + \eta \mathbf{1}(West_{it} > q_{50}^{West}) + \gamma \log Inc_{i,t-20} + \varepsilon_{it}$, where q_{50}^{West} is the median value of the regulations in 2005. The monte carlo experiments reassign the actual regulation values either (a) across both states and time (b) across years within a state, preserving the fixed cross-sectional variation, and (c) across states within a year, preserving the fixed time-series variation. The coefficient from the same regression using the true regulation series is plotted as a vertical line in black.

APPENDIX FIGURE 5

Income Convergence by Housing Supply Elasticity (Saiz)

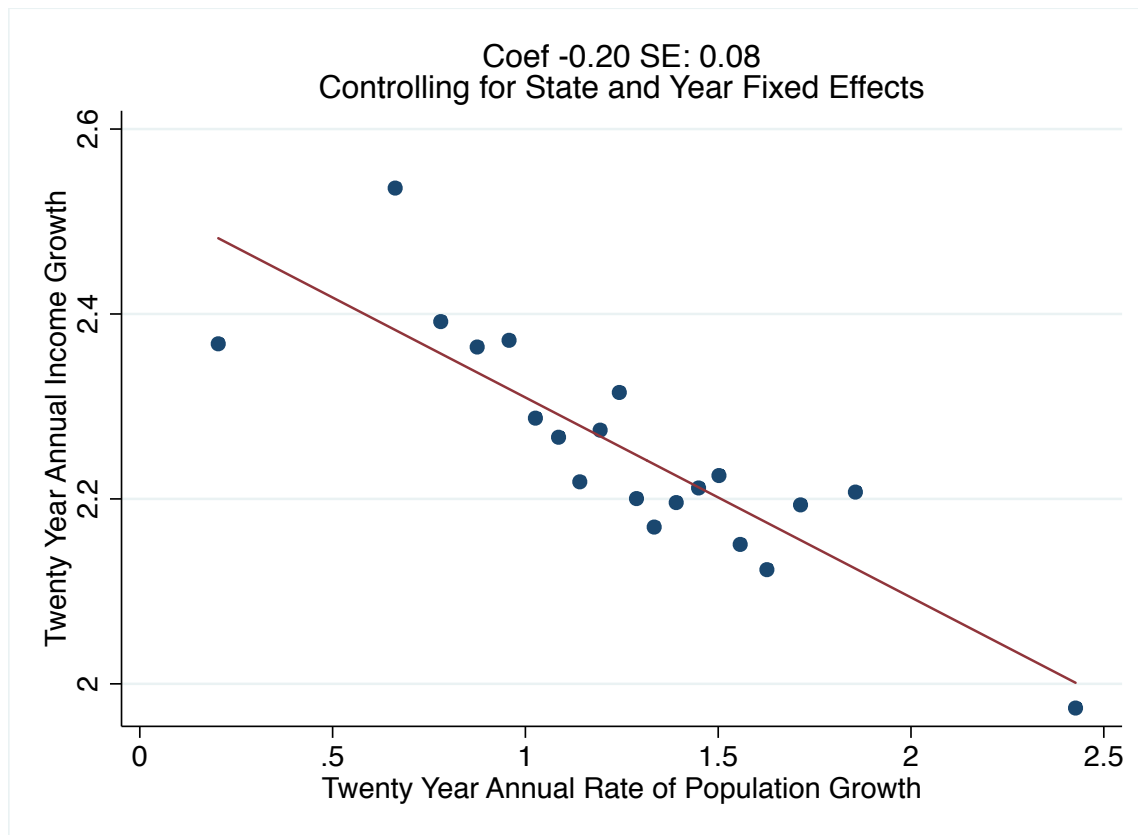


Notes: The top panels reconstruct Figure 2, relabeling states according to their housing supply elasticities levels, split along their measure of housing supply elasticity in Saiz (2010). We weight the time-invariant MSA-level measures from Saiz by population to produce state-level estimates and impute a value for Arkansas based on neighboring states. Blue states (upper case) have above median housing supply elasticity and red states (lower case) have below median elasticity.

The bottom panel depicts the coefficients β from running: $\Delta Inc_{s,t} = \alpha_t + \beta Inc_{s,t-20} + \varepsilon_{s,t}$ over rolling twenty year windows. The regressions are estimated separately for two equally sized groups of states.

APPENDIX FIGURE 6

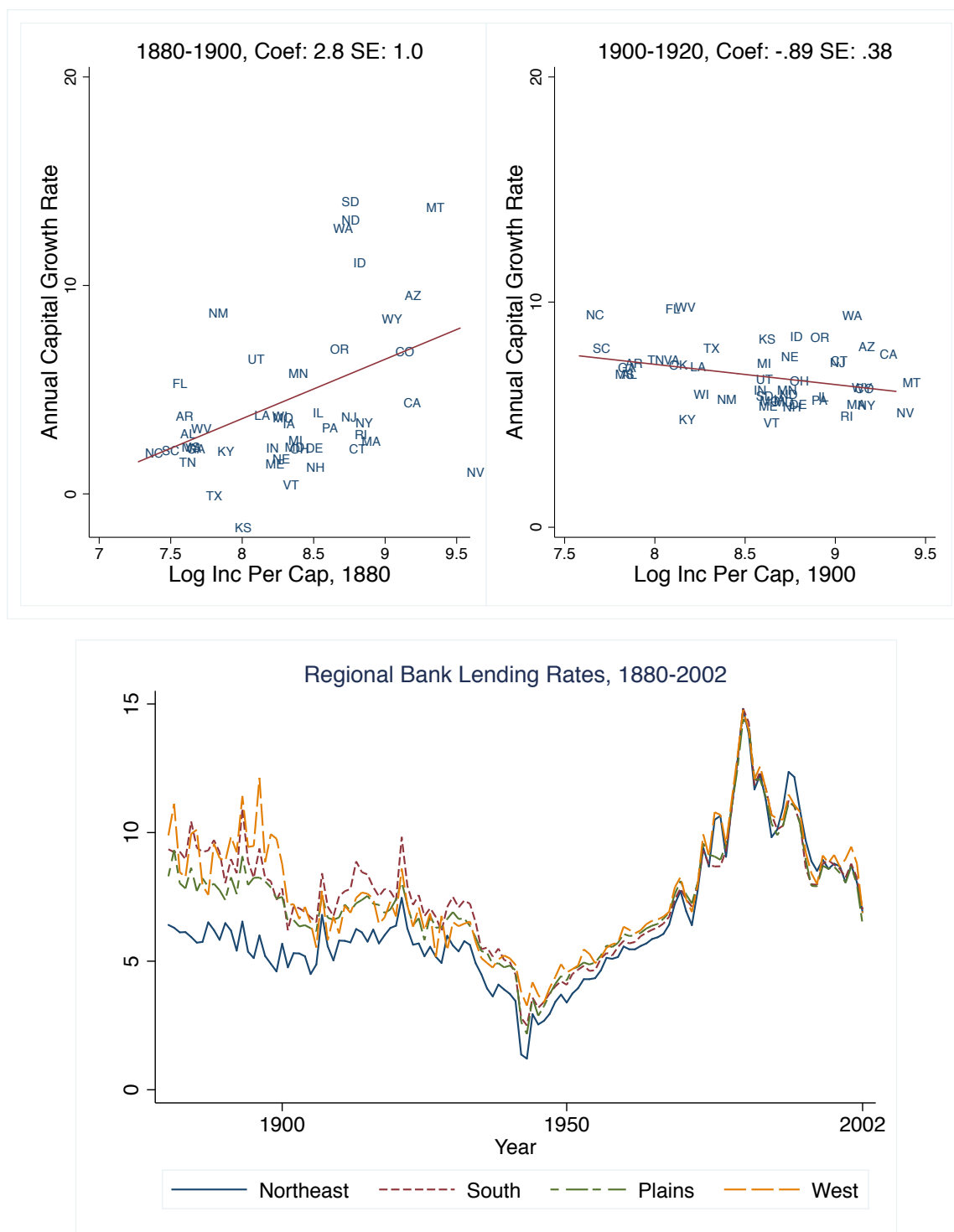
Reduced Form Changes in Income vs. Changes in Population



Note: This figure plots the relationship between twenty year annual growth rates in state income per capita against twenty year growth rates in state population, controlling for state and year fixed effects. State-year bins are divided into twenty quantiles based on population growth rates and average annual growth rates are calculated for each bin, conditional on state and year fixed effects.

APPENDIX FIGURE 7

Capital Convergence, 1880-1920, and Interest Rates, 1880-2002



Notes: The first panel shows capital convergence (or lack thereof) between 1880 and 1920 using estimates come from Kuznets *et al.* (1964). The original source for the capital estimates is the Census of Wealth.

Data series were assembled by Landon-Lane and Rockoff (2007) from bank reports.

APPENDIX TABLE 1
 σ Convergence, IV Estimates of Convergence and Labor Market Area Convergence

Panel A: Cross-Sectional Standard Deviation of Income							
	1950	1960	1970	1980	1990	2000	2010
BEA Log Inc Per Cap	0.236	0.199	0.155	0.137	0.150	0.150	0.138
Panel B: Additional Convergence Regressions							
$\Delta \ln y_{it}$ (Annual Rate in %) = $\alpha + \beta_t \ln y_{it-1} + \varepsilon_{it}$							
20 year period ending in...							
OLS BEA	1950	1960	1970	1980	1990	2000	2010
Coefficient	-2.38	-2.41	-1.98	-1.85	-0.58	-0.39	-0.99
Standard Error	0.16	0.11	0.16	0.15	0.31	0.46	0.29
OLS Census							
Coefficient	--	-1.82	-2.33	-2.42	-0.36	-0.26	-1.33
Standard Error	--	0.13	0.16	0.12	0.33	0.50	0.32
IV BEA with Census							
Coefficient	--	-2.46	-1.65	-1.59	-0.37	-0.22	-1.23
Standard Error	--	0.12	0.22	0.25	0.32	0.46	0.42
IV Census with BEA							
Coefficient	--	-1.81	-2.42	-2.37	-0.48	-0.27	-0.84
Standard Error	--	0.12	0.18	0.14	0.38	0.59	0.27
Panel C: Convergence at Labor Market Area Level							
$\Delta \ln \text{var}_{it}$ (Annual Rate in %) = $\alpha + \beta_t \ln y_{it-1} + \varepsilon_{it}$							
20 year period ending in...							
Income Convergence	1950	1960	1970	1980	1990	2000	2010
Coefficient	--	-0.97	-1.69	-2.13	-0.21	0.23	-0.26
Standard Error	--	0.19	0.10	0.13	0.18	0.26	0.16

Notes: Panel A. This panel reports the standard deviation of log income per capita across states. This corresponds to the σ convergence concept in Barro and Sala-i-Martin (1992).

Panel B. Table 2 calculates convergence coefficients using data on personal income from the BEA. That specification is biased in the presence of classical measurement error. We address the bias issue by instrumenting for the BEA measure using an alternative Census measure and vice versa. The Census measure is log wage income per capita for all earners, except in 1950 where it is only household heads. The first stage F-statistics range from 189 to 739. Classical measurement error is not an issue in these IV regressions, and the convergence coefficients display a similar time-series pattern.

Panel C. This panel replicates the "OLS Census" specification from this table at the Labor Market Area (LMA) level, with each LMA weighted by its population. LMAs are 382 groups of counties which partition the United States. See Appendix B.2 for details on construction of LMA sample.

APPENDIX TABLE 2
Directed Migration From Poor to Rich States and Labor Market Areas

	20 year period ending in...						
	1950	1960	1970	1980	1990	2000	2010
$\Delta \text{Log Pop}_{it}$ (Annual Rate in %) = $\alpha + \beta_i \ln y_{it-1} + \varepsilon_{it}$							
Baseline, State-Level							
Coefficient	0.56	1.60	2.13	0.75	0.26	1.18	-0.48
Standard Error	0.27	0.37	0.60	0.78	1.03	1.05	0.64
Y: Net Migration (Birth-Death Method)							
Coefficient	1.16	2.68	2.92	1.14	0.78	1.06	-0.49
Standard Error	0.19	0.36	0.59	0.77	0.97	1.02	0.58
Y: Net Migration (Survival Ratio Method)							
Coefficient	1.29	2.04	2.20	0.67	0.05	--	--
Standard Error	0.23	0.35	0.58	0.77	0.92	--	--
Baseline, Labor Market Area Level							
Coefficient	--	1.82	1.73	-0.02	-0.88	0.17	0.13
Standard Error	--	0.31	0.26	0.32	0.42	0.41	0.25

Sources: BEA Income estimates, Ferrie (2003) and Fishback et al. (2006)

Notes: Robust standard errors are shown below coefficients. Birth-death method uses state-level vital statistics data to calculate net migration as $\text{ObservedPop}_t - (\text{Pop}_{t-10} + \text{Births}_{t,t-10} + \text{Deaths}_{t,t-10})$. Survival ratio method computes counterfactual population by applying national mortality tables by age, sex, and race to the age-sex-race Census counts from 10 years prior. The dependent variable for the last two rows is $\log(\text{net migration}_{t,t-20} + \text{pop}_{t-20}) \cdot \log(\text{pop}_{t-20})$. Both published series end in 1990, and we use vital statistics to construct the birth-death measure through 2010.

The first three rows show state-level analysis. The fourth shows results at the Labor Market Area (LMA) level, with each LMA weighted by its population. LMAs are 382 groups of counties which partition the United States. See Appendix B.2 for details on construction of LMA sample.

APPENDIX TABLE 3
Returns to Living in a High Income State by Skill

	1940	1960	1970	1980	1990	2000	2010
<u>Panel A: Returns to Migration (OLS)</u>							
	Income Net of Housing Costs						
Average State Income	0.880*** (0.020)	0.736*** (0.026)	0.786*** (0.042)	0.726*** (0.077)	0.657*** (0.035)	0.539*** (0.035)	0.356*** (0.046)
Average State Income X HH Skill	-0.180** (0.066)	0.133 (0.077)	0.090 (0.066)	0.040 (0.116)	0.227** (0.071)	0.614*** (0.092)	0.610*** (0.075)
N	255,391	306,576	339,412	2,116,772	2,924,925	3,142,015	694,985
<u>Panel B: Returns to Migration (IV for State of Residence with State of Birth)</u>							
	Income Net of Housing Costs						
Average State Income	0.932*** (0.029)	0.776*** (0.038)	0.859*** (0.055)	0.772*** (0.093)	0.667*** (0.036)	0.488*** (0.035)	0.258*** (0.051)
Average State Income X HH Skill	-0.212*** (0.057)	-0.036 (0.100)	-0.083 (0.097)	-0.353** (0.137)	0.222** (0.086)	0.708*** (0.122)	0.614*** (0.123)
N	255,391	306,576	339,412	2,116,772	2,924,925	3,142,015	694,985
<u>Panel C: Differential Impacts of Housing Costs in High-Income States (OLS)</u>							
	Log Housing Costs						
Log Average State Income	1.203*** (0.113)	1.116*** (0.104)	1.492*** (0.172)	1.645*** (0.436)	2.765*** (0.326)	2.357*** (0.314)	2.612*** (0.363)
Log Average State Income X HH Skill	0.531*** (0.120)	-0.241** (0.080)	-0.206 (0.131)	-0.382 (0.294)	-0.379* (0.183)	-0.761*** (0.200)	-0.749*** (0.213)
N	269,172	324,895	358,036	2,104,099	2,857,523	3,159,098	707,898

Notes: All standard errors are clustered by state. *** p<0.01, ** p<0.05, * p<0.1

Panel A. This panel reports the coefficients β_1 and β_2 from the regression $Y_i - P_i = \alpha + \gamma \text{Skill}_i + \beta_1 Y + \beta_2 Y * \text{Skill}_i + \theta X_i + \varepsilon_i$, where Y_i and P_i measure household wage income and housing costs respectively, Y measures average state income and X_i are household covariates. Household Skill_i is the fraction of household adults in the workforce who are skilled, defined as 12+ years of education in 1940 and 16+ years thereafter. Household covariates are the size of the household, the fraction of adult workers who are black, white, and male, and a quadratic in the average age of adult household workers. Housing costs P_i are defined as 5% of house value or 12 times monthly rent for renters. 1950 is omitted since income data are available only for household heads.

Panel B. The IV regressions replicate panel A, but instrument for average state income and its interaction with household skill using the average income of the state of birth of adult household workers. The first stage F-statistics in these regressions exceed 80.

Panel C. This panel reports the coefficients β_1 and β_2 from the regression $\log(P_i) = \alpha + \gamma \text{Skill}_i + \beta_1 \log(Y) + \beta_2 \log(Y) * \text{Skill}_i + \theta X_i + \varepsilon$.

APPENDIX TABLE 4
Migration Flows by Skill Group: Nominal vs. Real Income

	Dep Var: 5-Year Net Migration as Share of Total Pop				
	Baseline (1)	Double Housing Cost (2)	Exclude In-State Mig (3)	Only Whites (4)	Mig Measure Birth State (5)
<u>Panel A: Low-Skill People, 1940</u>					
Log Nominal Income	1.313*** (0.470)	-- --	1.049** (0.438)	1.007** (0.443)	1.086** (0.443)

Log Group-Specific Income Net of Housing	1.236*** (0.364)	1.109*** (0.274)	1.017*** (0.350)	0.980*** (0.352)	0.995*** (0.338)
<u>Panel B: High-Skill People, 1940</u>					
Log Nominal Income	0.611 (0.392)	-- --	0.617 (0.419)	0.585 (0.387)	0.475 (0.411)

Log Group-Specific Income Net of Housing	0.773* (0.400)	0.899** (0.337)	0.905* (0.462)	0.821* (0.415)	0.701 (0.513)
<u>Panel C: Low-Skill People, 2000</u>					
Log Nominal Income	-2.173** (1.006)	-- --	-2.456*** (0.792)	-2.377*** (0.757)	0.281 (8.453)

Log Group-Specific Income Net of Housing	4.309** (2.007)	6.042*** (2.140)	-0.357 (1.167)	1.725 (1.418)	-11.99 (11.51)
<u>Panel D: High-Skill People, 2000</u>					
Log Nominal Income	4.077*** (0.694)	-- --	1.786*** (0.611)	2.894*** (0.649)	19.32*** (5.373)

Log Group-Specific Income Net of Housing	4.715*** (0.894)	3.634*** (1.280)	1.937*** (0.701)	3.593*** (0.874)	14.06*** (4.567)

Note: Each cell represents the results from a different regression. The table regresses 5 year net-migration rates on average income and skill-specific income net of housing. Low-skill is defined as having less than 12 years of education in 1940 and less than a BA in 2000. In 1940, the unit of observation is State Economic Area, with n=455 to 466, depending on specification. In 2000, the unit of observation is three-digit Public Use Microdata Areas, with n=1,020. The baseline case reproduces the results in Figures 5 and 6. The second column shows the effect of doubling the housing costs described in the text to control for non-housing price differences across places. The third column excludes intra-state migrants in calculating net-migration rates. The fourth column excludes non-white migrants in calculating net-migration rates. The final measure calculates migrants as the number of residents residing outside their state of birth. Additional details are presented in the text. Standard errors clustered by state. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX TABLE 5
Income and Wage Premia by Human Capital Level

	Elem	Middle	<HS	HS	<BA	BA+
Panel A: Annual Income						
1940	0.38	0.56	0.77	1.00	1.17	1.63
1950	0.50	0.66	0.81	1.00	1.07	1.35
1960	0.44	0.65	0.83	1.00	1.13	1.55
1970	0.52	0.68	0.82	1.00	1.14	1.60
1980	0.59	0.68	0.79	1.00	1.15	1.55
1990	0.60	0.67	0.73	1.00	1.25	1.86
2000	0.66	0.68	0.72	1.00	1.25	1.91
2010	0.64	0.68	0.70	1.00	1.28	2.17
Panel B: Hourly Earnings						
1940	0.53	0.69	0.86	1.00	1.19	1.61
1950	0.67	0.79	0.90	1.00	1.13	1.34
1960	0.60	0.76	0.90	1.00	1.14	1.48
1970	0.65	0.76	0.88	1.00	1.16	1.59
1980	0.69	0.76	0.87	1.00	1.13	1.48
1990	0.68	0.74	0.84	1.00	1.19	1.70
2000	0.70	0.74	0.81	1.00	1.20	1.76
2010	0.69	0.74	0.79	1.00	1.22	1.89

Notes: Returns to education are expressed relative to workers with 12 years of education (1940-1980) and a high school degree (1990-2010). Income sample is all people ages 25-64, except 1950, when only incomes of household heads were recorded. Earnings sample is all people with positive wage income. Income and earnings are winsorized at 1st and 99th percentile to minimize the influence of outliers.

Panel A: Returns to education in terms of annual wage income are calculated using a Mincerian regression. The specification generally follows Delong, Goldin and Katz (2003), with dummies for black, Hispanic, and foreign-born, a quartic in experience interacted with female, and state dummies.

Panel B: Mincerian regression uses annual wage income divided by annual hours.

APPENDIX TABLE 6

Cumulative Fraction of "Land Use" Cases To Date

State	1950	1980	2010
AL	0.025	0.071	0.077
AR	0.045	0.161	0.174
AZ	0.032	0.261	0.542
CA	0.088	0.520	1.007
CO	0.128	0.422	0.568
CT	0.110	0.518	1.739
DE	0.105	0.370	0.679
FL	0.053	0.102	0.114
GA	0.024	0.108	0.205
IA	0.072	0.221	0.311
ID	0.179	0.295	0.885
IL	0.096	0.373	0.232
IN	0.042	0.305	0.290
KS	0.074	0.231	0.205
KY	0.065	0.148	0.258
LA	0.086	0.098	0.101
MA	0.084	0.449	0.629
MD	0.152	1.344	1.013
ME	0.156	0.738	3.388
MI	0.091	0.450	0.420
MN	0.095	0.478	0.761
MO	0.084	0.269	0.304
MS	0.030	0.092	0.245
MT	0.097	0.347	0.521
NC	0.073	0.171	0.286
ND	0.035	0.260	0.384
NE	0.121	0.185	0.307
NH	0.196	0.943	2.407
NJ	0.192	0.765	1.345
NM	0.135	0.157	0.430
NV	0.343	0.393	0.145
NY	0.025	0.094	0.180
OH	0.102	0.348	0.260
OK	0.032	0.080	0.153
OR	0.089	0.891	1.079
PA	0.056	0.308	0.331
RI	0.197	0.526	1.360
SC	0.092	0.173	0.446
SD	0.026	0.177	0.451
TN	0.065	0.186	0.288
TX	0.075	0.116	0.141
UT	0.128	0.182	0.539
VA	0.130	0.535	0.660
VT	0.062	0.594	1.292
WA	0.050	0.503	1.373
WI	0.114	0.317	0.425
WV	0.076	0.166	0.518
WY	0.000	0.207	0.676

Data from an online database of state appellate court documents. Cumulative count begins in 1920 and runs to date reported. The table is expressed as 100 times the share of land use cases to date.

APPENDIX TABLE 7
Impacts of Alternate Regulation Measures on Permits, Prices, Migration, and Convergence

	Annual Construction Permits _t % of Housing Stock (1)	Log House Price _{t-20} (2)	ΔLog Population _{t, t-20} Annual Rate in % (3)	Δ Log Human Capital _{t,t-20} Annual Rate in % (4)	Δ Log Income Per Cap _{t,t-20} Annual Rate in % (5)
<u>Zoning Reg Measure 1(zoning_{it}>zoning^{median}₂₀₀₅)</u>					
Log Inc Per Cap _{t-20}	2.945 (1.957)	0.993*** (0.0815)	2.613** (1.046)	-0.391*** (0.0882)	-2.144*** (0.223)
Log Inc Per Cap _{t-20} *	-4.230**	0.669***	-3.322**	0.337*	2.929***
"Zoning" High Reg	(2.071)	(0.216)	(1.330)	(0.172)	(0.410)
N	1,488	288	2,928	240	2,928
R ²	0.216	0.859	0.161	0.895	0.793
<u>Continuous Reg Measure</u>					
Log Inc Per Cap _{t-20}	1.926 (1.484)	1.015*** (0.123)	1.720** (0.812)	-0.289** (0.114)	-2.168*** (0.267)
Log Inc Per Cap _{t-20} *	-216.9	56.46*	-176.3*	26.64	294.2***
Continuous Reg	(129.5)	(31.52)	(89.39)	(17.62)	(58.78)
R ²	0.194	0.845	0.127	0.843	0.798
<u>State Specific Reg Measure 1(landuse_{it}>3*landuse_{i,1941})</u>					
Log Inc Per Cap _{t-20}	4.591 (3.243)	0.920*** (0.0862)	2.223*** (0.783)	-0.359** (0.134)	-2.800*** (0.251)
Log Inc Per Cap _{t-20} *	-4.266	0.623***	-1.150	0.403**	2.292***
State Spec High Reg	(3.223)	(0.158)	(0.689)	(0.162)	(0.537)
N	1,536	336	2,976	288	2,976
R ²	0.251	0.824	0.147	0.830	0.773
Year*High Reg FEs	Y	Y	Y	Y	Y

Notes: The table reports the coefficients β_1 and β_2 from regressions of the form:

$$\Delta \ln y_{it} = \alpha_t + \alpha_t I(\text{reg} > x) + \beta_1 \ln y_{it-1} + \beta_2 \ln y_{it-1} \times I(\text{reg} > x) + \varepsilon_{it}$$

This table uses three alternate land use regulation measures. The first panel uses appellate court case counts for the term "zoning", and the 'high regulation' cutoff value is set to the measure's 2005 median value. The second panel defines uses the continuous land use regulation measure, interacted both with the year dummies and with initial income. The third panel defines a state as high regulation when its "land use" rolling document count to date reaches three times its own 1941 value. The dependent variables are new housing permits from the Census Bureau, the median log housing price from the IPUMS Census extracts, population change, the change in log human capital due to migration, and the change in log per-capita income. Standard errors clustered by state. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX TABLE 8
Human Capital Convergence

	1960	1980	1990	2000	2010
<i>Convergence in State Human Capital Levels</i>					
(1) Y: 20-year Change in Human Capital due to Migration (Annual Rate in %)					
X: Human Capital t-20	-0.30	-0.21	-0.19	-0.15	-0.07
<i>Convergence in Human Capital Levels of High- and Low-Income States</i>					
Y: $\log(20\text{-year Change in Human Capital due to Migration} + \text{Human Capital } t-20) - \log(\text{Human Capital } t-20)$ (Annual Rate in %)					
X: Log Inc Per Cap t-20					
<i>Assumptions for Human Capital Measurement</i>					
(2) Annual income	-0.25	-0.13	-0.05	-0.13	0.05
(3) Hourly wage, impute hours with $\epsilon = 0$	-0.19	-0.12	-0.03	-0.08	0.03
(4) Hourly wage, impute hours with $\epsilon = 0.6$	-0.30	-0.19	-0.05	-0.10	0.08
(5) Hourly wage, impute hours with $\epsilon = 2.6$	-0.58	-0.44	-0.15	-0.09	0.33
<i>Alternate Samples</i>					
(6) 20-year Change in Human Capital	-0.21	-0.26	-0.18	-0.37	-0.07

Notes: Each cell is a coefficient from a cross-sectional regression of change in human capital on lagged human capital or income. See Appendix Table 2 for income and wage premia by year and skill group and the Data Appendix for details on index construction. Data on the 1950-1970 are missing because the Census only asked education and income of household heads in 1950.

(1) Human capital index for state i at date t calculated as $HC_{it} = \sum_j inc_{jt} * share_{ijt}$. Change in Human Capital due to Migration is pop-weighted sum of

HC by state of residence minus HC by state of birth (ages 25-44) plus

HC by state of residence (ages 45-64) minus HC by state of residence 20 years ago (for people then ages 25-44). Groups j include education, black and Hispanic dummies and inc_{jt} is calculated using a Mincerian regression.

(2) Replicates specification (1), with y variable as $\log(\Delta \text{HumanCapital} + \text{HumanCapital}_{t-20}) - \log(\text{HumanCapital}_{t-20})$ and x variable as Log Inc Per Cap_{t-20}.

(3, 4, 5) Replicates row 2, with alternative labor supply assumptions. $HC_{it} = \sum_j wage_{jt}^{(1+\epsilon)} * share_{ijt}$, where wage is calculated as annual wage income/annual hours and ϵ is parameterized to reflect labor supply differences by skill level from the model (Section 2.1.1).

(6) $\log(\text{HumanCapital}_t) - \log(\text{HumanCapital}_{t-20})$. Unlike the measures developed above, which focuses exclusively on changes due to migration, this measure includes changes in human capital accumulation by nonmovers.

APPENDIX TABLE 9
Calibration: Role of Population and Human Capital in Income Convergence

<u>Inputs from Model</u>					
elasticity of income wrt $\Delta\text{pop} = -(1+\epsilon)/(\sigma+\epsilon)$ from Equation 2					
elasticity of income wrt $\Delta\text{HumanCapital} = (\sigma-1)/(\sigma+\epsilon)$ from Equation 3					
<u>Inputs from Data</u>					
	1950-1980	2010	Difference		
avg directed mig coef (Table 2)	1.26	-0.48	-1.74		
avg sort coef on income (Table 3)	-0.19	0.05	0.24		
Observed Change in Convergence Rates	1.11				
<u>Calibrated Change Due to Observed Changes in Migration</u>					
ϵ/σ	$\sigma=$ 1	$\sigma=$ 2	$\sigma=$ 6	$\sigma=$ 10	$\sigma=$ ∞
$\epsilon=$ 0	1.980	1.111	0.531	0.415	0.241
$\epsilon=$ 0.3	1.925	1.193	0.589	0.454	0.241
$\epsilon=$ 0.6	1.890	1.256	0.641	0.490	0.241
$\epsilon=$ 1	1.860	1.320	0.704	0.536	0.241
$\epsilon=$ 2	1.819	1.425	0.833	0.636	0.241

Notes: This table estimates the role of migration in income convergence. Migration drives convergence through population flows from poor to rich states (measured in Table 2) and human capital flows from rich to poor states (measured in Table 3). The effect of these changes on income per capita is calibrated using the model (equation 7) and different assumptions on the elasticity of labor supply and demand. We consider a wide range of parameter values. The results above demonstrate that the *observed* changes in migration can account for a significant portion of the change in convergence, though this effect becomes weaker as the elasticity of substitution across regional varieties grows large.

APPENDIX TABLE 10
Inequality Impacts of Convergence and its Demise

Panel A: Inequality Counterfactual without Convergence (1940-1980)		
	Std Dev of Log Hourly Earnings -- Full-time Males	
	State ^a	Total ^b
1940	0.300	0.781
1950	0.227	0.672
1960	0.183	0.580
1970	0.147	0.600
1980	0.106	0.618
σ Convergence (1940-1980) ^c	65%	
1980 No Convergence Counterfactual: SD $[Y + Y_{state1940}*(1-0.35)]^d$		0.674
Δ Inequality		
1980 Observed - 1940 Observed		-0.163
1980 No Convergence Counterfactual - 1940 Observed		-0.107
Share of Δ Inequality Accounted for By Convergence		34%
Panel B: Inequality Counterfactual if Convergence Continued (1980-2010)		
	Std Dev of Log Hourly Earnings -- Full-time Males	
	State	Total
1980	0.106	0.618
1990	0.125	0.622
2000	0.098	0.643
2010	0.115	0.678
2010 Convergence Counterfactual: SD $[Y - Y_{state1980}*(1-0.35)]^e$		0.674
Δ Inequality		
2010 Observed - 1980 Observed		0.060
2010 Convergence Counterfactual - 1980 Observed		0.056
Share of Δ Inequality Accounted for By End of Convergence		8%

Sample uses hourly earnings for men ages 18-65 with nonallocated positive earnings, who worked at least 40 weeks last year and at least 30 hours per week in the Census. Sample is winsorized at the 1st and 99th percentile in order to limit the influence of outliers.

a. Population-weighted standard deviation of mean state-by-year log hourly earnings.

b. Standard deviation of log hourly earnings. Conceptually, this measure includes both state-level and residual variation in earnings.

c. σ Convergence = $1 - SD_{state1980} / SD_{state1940}$. Note that this measure uses hourly earnings, and is different from the measure of σ Convergence developed in Appendix Table 1, which uses per capita income.

d. Rather than using observed state income in 1980, we predict state income using 1940 state income and the observed convergence rate of 65% to calculate $Y_{state1980hat} = 0.35 * Y_{state1940}$. We characterize the counterfactual distribution of earnings in the absence of state income convergence as $Y + Y_{state1940} - Y_{state1980hat}$.

e. Method follows note (d), except that we calculate the counterfactual with convergence as $Y - Y_{state1980} + Y_{state2010hat}$.