Online Appendix for "Why Has Regional Income Convergence in the U.S. Declined?"

A Extended Model and Calibration, Proofs of Propositions 1 and 2

In this section we extend the model to allow for calibration and the simulation of shocks. Specifically, we add Stone-Geary preferences over housing and non-housing consumption to match the extent of non-homotheticity of housing demand observed in Online Appendix Figure B.1. We also add imperfectly substitutable skill types to to match the fact that no city arrives at a corner solution in skill composition. We report the results of a baseline calibration and then report the sensitivity to the model's two key assumptions. The goal of this exercise is to (1) show that the magnitude of the implied housing supply elasticity change was large enough to quantitatively account for the change in migration and convergence patterns and (2) to demonstrate the importance of the model's substantive assumptions (described at the start of Section 3) in generating the results.

A.1 Extended Model With Imperfectly Substitutable Skill Types

Individual Decisions: Goods Demand and Indirect Utility There are n_{jkt} agents are endowed as either skilled or unskilled in production $k \in \{s, u\}$, and have utility in state $j \in \{N, S\}$ at every date t of

$$\max_{\{c_{jkt}, h_{jkt}\}} \sum_t e^{-rt} ln(u_{jkt})$$

$$\text{where } u_{jkt} = c_{jkt}^{\beta} \left(h_{jkt} - \bar{H}\right)^{1-\beta} - \frac{\ell_{jkt}^{1+1/\varepsilon}}{1+1/\varepsilon}$$

$$\text{subject to } c_{jkt} + p_{jt} h_{jkt} = w_{jkt} l_{jkt} + \pi_t \tag{10}$$

The first order condition on labor supply implies:

$$l_{jkt} = \left(\beta^{\beta} \left(\frac{1-\beta}{p_{jt}}\right)^{1-\beta} w_{jkt}\right)^{\epsilon} \tag{11}$$

Workers' preferences take the Stone-Geary functional form with a baseline housing requirement $\bar{H}$ that is common for both high-skill and low-skill workers. This functional form generates non-homothetic housing demand.[33] To keep things simple, we assume inelastic labor supply and abstract from intertemporal markets by imposing a static budget constraint. Workers receive the local wage w_{jkt} for their skill type k and the price of housing relative to

[33] See Mulligan [2002] and Kongsamut et al. [2001] for other examples of papers using Stone-Geary preferences.

tradables is p_{jt}. Profits from both the housing sector and the tradable sector in North and South (π_t) are rebated lump-sum nationally. We can therefore write each agent's indirect utility as a function of the wage, price and preference parameters:

$$v_{jkt}(w_{jkt}, p_{jt}) = ln\left(\left(w_{jkt} + \pi_t - p_{jt}\bar{H}\right)\beta^{\beta}\left(\frac{1-\beta}{p_{jt}}\right)^{1-\beta}\right)$$

Labor Market Next, we turn to the production of tradables. State-level production is given by

$$Y_{jt} \quad = \quad A_j\left((l_{jut}n_{jut})^{\rho} + \theta(l_{jst}n_{jst})^{\rho}\right)^{\frac{1-\alpha}{\rho}}$$

where n_{jk} is the number of people of type k residing in state j.[34] We normalize $A_S = 1$ throughout, and assume $A_N > 1$. This term can encompass capital differences, natural advantages, institutional strengths, different sectoral compositions, amenities, and agglomeration benefits. Assuming labor earns its marginal product, we have:

$$w_{jut} = A_j\left(1-\alpha\right)\left(n_{jut}^{\rho} + \theta n_{jst}^{\rho}\right)^{\frac{1-\alpha-\rho}{\rho}}\left(n_{jut}\right)^{\rho-1} \tag{12}$$

$$w_{jst} = A_j\left(1-\alpha\right)\theta\left(n_{jut}^{\rho} + \theta n_{jst}^{\rho}\right)^{\frac{1-\alpha-\rho}{\rho}}\left(\theta n_{jst}\right)^{\rho-1} \tag{13}$$

Equilibrium in each these markets is given by the wage such that $l_{jkt}^{demand}(w_{jkt}, n_{jut}, n_{jst}) = l_{jkt}^{supply}(w_{jkt}, n_{jut}, n_{jst})$.

Housing Market Define the quantity of housing in place j at time t as H_{jt}. Every state is endowed with a housing supply at time zero equal to the demand of the initial population. Regulations can only affect new construction. Because they are designed to minimize the amount of cumulative development, we model them as imposing a convex cost as a function of the existing housing stock, where η, the measure of regulatory constraints, governs the elasticity of supply in growing regions. The marginal cost per unit of construction is

$$c(H_{jt}, H_{jt-1}) = \begin{cases} 0 & H_{jt} < H_{jt-1} \\ H_{jt}^{1/\eta} & H_{jt} \geq H_{jt-1} \end{cases}$$

All housing has a fixed maintenance cost to be habitable which we normalize to 1. So long as a city is growing, the price of all housing is equal to marginal cost of construction plus maintenance, so prices are:

$$p_{jt} = \begin{cases} 1 & if\ H_{jt} \leq H_{jt-1} \\ 1 + H_{jt}^{1/\eta} & if\ H_{jt} > H_{jt-1} \end{cases} \tag{14}$$

Regulations affect the dynamics of the system only in places where the population would otherwise be increasing. Demand for housing for each individual is equal to $\bar{H}+(1-\beta)\left(\frac{w_{jkt}+\pi_t}{p_{jt}}\right)$,

[34] This widely used form of imperfect substitution ensures an interior solution for skill ratios in equilibrium.

and therefore aggregate demand is

$$H_{jt} = n_{jut}\left(\bar{H} + (1-\beta)\left(\frac{w_{jut}+\pi_t}{p_{jt}}\right)\right) + n_{jst}\left(\bar{H} + (1-\beta)\left(\frac{w_{jst}+\pi_t}{p_{jt}}\right)\right) \tag{15}$$

We model regulations as affecting the elasticity of supply rather as a direct cost shock. This choice is motivated by empirical evidence that regulations affect the relationship between income and prices and not merely the price itself (see Figure 6 and Table 2). This choice is also consistent with the existing empirical work on regulations and housing (Saiz [2010]and Saks [2008]), and the dominant interpretation in the legal literature (Ellickson [1977]).

Equilibrium Taking $\{n_{jut}, n_{jst}\}$ as given, prices $\{w_{jut}, w_{jst}, p_{jt}\}$ and allocations $\{c_{jkt}, l_{jkt}, H_{jkt}, \}$ that satisfy equations 10-15 constitute an equilibrium in the housing and labor markets. This equilibrium also allows us to write indirect utility as a function of the local population $(v_{jkt}(n_{jut}, n_{jst}))$.

Migration and Dynamics

Having characterized the equilibrium within a location, we turn to cross-location dynamics. Normalizing the national population of each skill type to 1, we define $\Delta_{kt} = v_{Nkt}(n_{Nut}, n_{Nst}) - v_{Slt}((1-n_{Nut}), (1-n_{Nst}))$ as the flow utility gains to living in the North. Note that when land supply is perfectly elastic ($\eta \to \infty$)and initial population allocations are balanced, Δ_{kt} does not depend on the skill type k.[35] We can now define the present discounted value of migrating from South to North as:

$$q_k\left(t\right) = \sum_{\tau=t}^{\infty} e^{-r\tau}\Delta_{k\tau} \tag{16}$$

These expressions depend upon exogenous parameters and shocks, as well as two state variables n_{Nut} and n_{Nst}.

Given these gains to migration, how many people migrate each period? We follow Braun [1993] in assuming that the migration rate is proportional to the present-discounted value of migrating:

$$\Delta ln(n_{Nkt}) - \Delta ln(n_{Skt}) = \psi q_k(t) \tag{17}$$

This equation holds exactly for i.i.d. migration cost draws from a specific distribution, or it can be viewed as a linear approximation of a more general class of processes.

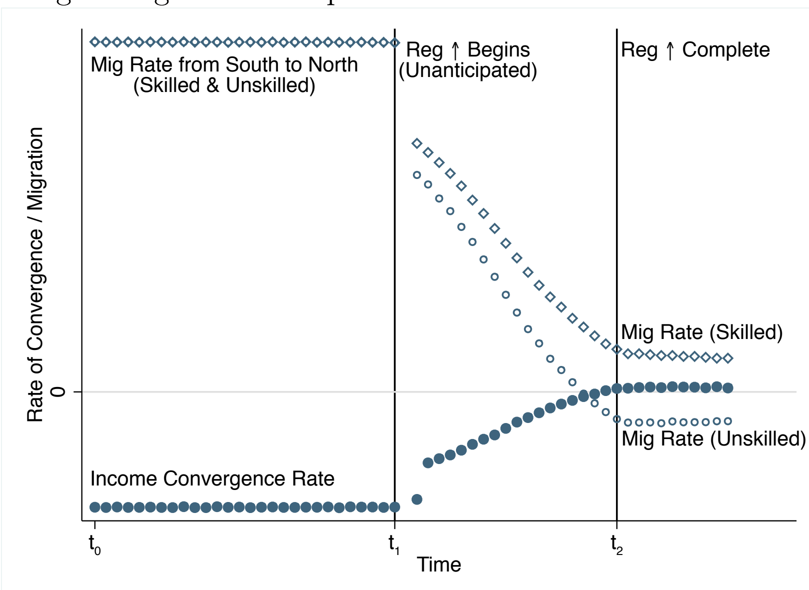
A.2 Simulation

The equations represented in 16 and 17 constitute a dynamic system in terms of two endogenous variables and exogenous shocks and parameters. To illustrate the dynamics of the system, we consider a numeric example. We plot the dynamics in a simulation where (I) the population of skilled and unskilled workers are evenly divided between North and South, (II) the housing supply in the North is completely elastic ($\eta \to 0$), and where (III)

[35]This holds under the normalization that $\bar{H} = \pi$.

the productivity parameter A_N is significantly greater than 1. Given these assumptions, the initial population in the South exceeds the steady-state population values.

The figure below illustrates the dynamics of the system from these conditions until time t_1.[36] When the housing supply in the North is completely elastic, the relative gains to migration are independent of skill type, and hence both high and low productivity workers migrate away from the South at the same constant rate. This directed migration makes labor more scarce in the South and more plentiful in the North, which yields a constant rate of convergence in per capita incomes between the regions. Additionally, if there were a larger fraction of unskilled workers in the South, then migration would have driven convergence by equating average human capital levels as well.



At date t_1, the elasticity of housing supply, η, begins to fall and reaches a new, permanently lower level at time t_2. This unanticipated shock increases housing prices in the growing North, and alters the value of living in the North in the future. Both skilled and unskilled migration rates fall, but they do not fall to the same degree. Skilled workers continue to find it worthwhile to move from South to North, but the increase in housing prices actually makes the North relatively unattractive to unskilled workers who begin to move in the opposite direction. The joint effect is that, by t_2, there is no more net migration from South to North and no further convergence in incomes per capita. Instead, migration flows lead to skill-sorting and segregation by skill type.

A.3 Calibration

We set θ, the premium for skilled versus unskilled workers, equal to 1.7. This is representative of the BA/non-BA relative wages in data, holding race and gender constant. We set the

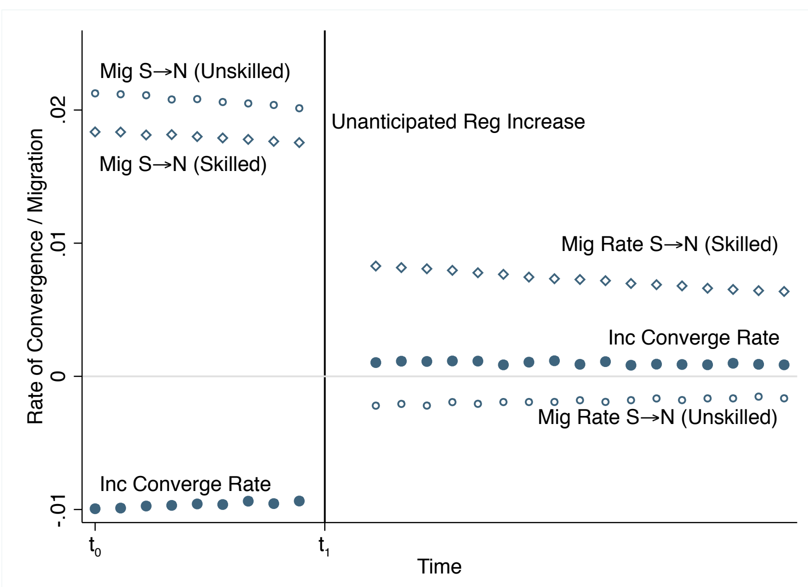
[36]This graph is meant to illustrate the model's dynamics. To do this, we set $\theta = 1.7$, $\alpha = 0.33$, ρ=0.9, β=0.25, H=0.25, $A_n = 2$, $\psi = 0.005$, and r=0.05. We then simulated a falling housing supply elasticity by having $1/\eta$ ascend from a value near zero to 0.25.

elasticity of subsitution between skilled and unskilled workers, ρ, equal to 0.6 as in Card [2009]. The initial share of skilled workers living in the North is set to 0.69, and the initial share of unskilled workers is set to 0.63. This matches the population distribution in 1950, when splitting states in to "North" and "South" at the median based on per capita incomes. The total population of each skill type is normalized to one.

We use the two parameters of the utility function, $\bar{H}$ and β, to match the Engel-curve for housing estimated in Section 3. This entails setting $\beta = .06$ and $\bar{H} = .25$ in Online Appendix. This parameter choice means that we can analyze whether the nonhomotheticity we observe for housing within labor markets is large enough to generate the changes we see in migration for the observed change in housing prices. The discount rate r , treating each period as one year, and the labor share of production $(1 - \alpha)$ are set to 0.05 and 0.65 as in much of the literature. The elasticity of labor supply ϵ is set to 0.6 as in Chetty [2013]. We set A, the relative productivity parameter, equal to 1.8. This is consistent with a fraction of 85% of the population residing in the North in the steady state given equalized skill distributions.

Finally, we are left to calibrate the moving cost parameter ψ, the elasticity parameter η, and the size of the elasticity shock. We initially set η equal to 0.4, which generates roughly a 1 to 1 relationship between log prices and log per capita income, matching the relationship in the data for 1950 and 1960 as reported in Figure 2. The parameter ψ is set equal to .002 to match the speed of directed migration observed prior to the explosion of land use regulations.

We simulate a shock that lowers η to 0.4 to 0.135 after 10 periods. This drop is calibrated to match the change in the log price to log income ratio, which in the data (Figure 3) rises to 2 from 1. The dynamics of the system to this shock displayed below.



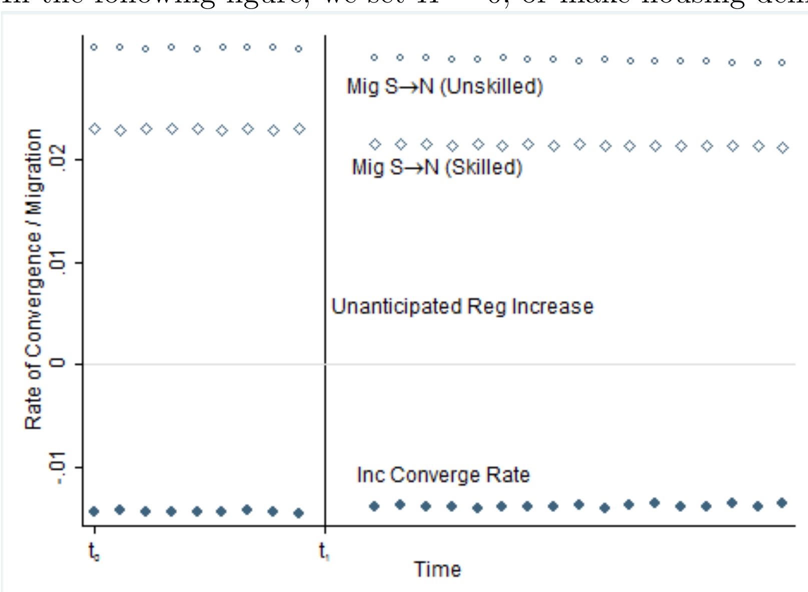
The figure shows that, before the shock, total directed migration averaged slightly less than 2% per year as in the data. Both skilled and unskilled workers migrate from South to North, with unskilled workers actually moving at a slightly faster rate due to initial skill imbalances. The convergence rate before the shock is slightly less than 1% per year. The rate in the data is closer to 2% per year, meaning that under this calibration, the migration mechanism can account for roughly 50% of convergence prior to the regulatory shock.

When a shock calibrated to match changing price ratios hits, both directed migration and income convergence cease as in the data. The rate of income convergence falls roughly 1%, similar to the change in the rate of beta-convergence reported in Figure 1. Thus, while the migration channel can only account for half of the level of convergence, changes in migration can account for roughly 100% of the change. The cessation of total directed migration masks different trends for skilled and unskilled workers. Skilled workers continue to move from South to North at a reduced, but still significant rate. Unskilled migration, which had previously exceeded skilled migration, stops completely. Thus net migration has turned into skill-sorting across locations as in the data.

In the text, we described the model as depending on two substantive assumptions: non-homothetic housing demand and downward sloping labor demand curves. In these next two exercises, we demonstrate that by re-running our simulations while turning off each channel.

Homothetic Housing Demand

In the following figure, we set $\bar{H} = 0$, or make housing demand homothetic.

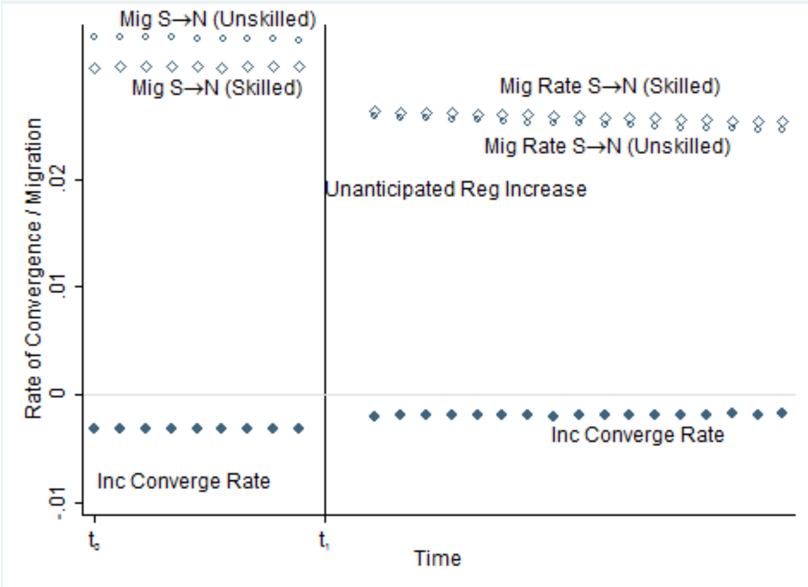


The model's initial dynamics look similar, though the utility gains from moving to the North are now slightly larger for both skilled and unskilled workers. As a result, the directed migration rates are higher for both groups and convergence is slightly stronger. The major difference occurs after the elasticity shock. When there is no longer a large, unavoidable housing cost the utility costs of an elasticity shock are smaller. While both directed migration rates and convergence rates fall, the total impact is much smaller. Further, though direct migration rates fall– the change is symmetic for both unskilled and skilled workers. As a result, there is no switch to skill sorting following the shock.

Non-Downward Sloping Demand

Alternately, we can consider the case when decreasing returns to scale set in very slowly.

To do this, we set $\alpha = .05$. The otherwise identically calibrated model now produces the results below.



The degree of convergence is significantly smaller than before, as it now occurs primarily via the human capital channel. Directed migration rates are larger, as there is less of a dampening effect on wages. Qualitatively, the elasticity shock has a similar impact to the baseline model, but quantitatively the change in convergence rates is far smaller due to the absence of the downward-sloping labor demand channel.

A.4 Proof of Proposition 1

Because $\lim_{\xi\to\infty} N^{1/\xi} = 1$, all terms in equation (6) are proportional to $\psi_k^{1+\epsilon}$, so $x_k^* = x^* \forall k$. In this case, the share of people leaving the South for the North is $F(x^*)$ for all skill types, where F is the CDF for the distribution of moving costs. Under assumptions (1) and (2) all groups find it worthwhile to move to the North, and under assumption (3) these flows are of a lower average skill than the average skill level in North. Using equation (4), we see that per capita incomes fall in the North.

A.5 Proof of Proposition 2

Define the gain to moving when housing costs are perfectly elastic as

$$\Delta_N^k = \psi_k^{1+\varepsilon}\left(\frac{1}{1+\varepsilon}((1-\alpha)A\gamma^{-\alpha})^{\frac{1+\varepsilon}{1+\alpha\varepsilon}} - \Omega\right)$$

A potential migrant from group k follows decision rule

$$\text{Move if } \Delta_N^k - (N_N^{1/\xi} - 1) > x_k^* \psi_k^{1+\varepsilon}\Omega$$

Given the cutoff migration rule $\frac{\Delta_N^k - N^{\frac{1}{\xi}} + 1}{\psi_k^{1+\varepsilon}\Omega} = x_k^*$, we can derive the following relationships between changes in housing supply elasticity ξ and (I) directed migration (II) skill sorting and (III) convergence.

(I) A fall in ξ reduces direct migration. Migration is defined as $DirectMig = \sum_k F(x_k^*)\mu_k^S$. Taking the above expression for x_k^*, we have $\frac{dx_k^*}{d\xi} = \frac{-1}{\psi_k^{1+\varepsilon}\Omega} * \frac{N^{\frac{1}{\xi}} log(N)}{\xi^2}$. This implies that $\frac{d}{d\xi}DirectMig = \sum_k F'(x_k^*)\frac{-1}{\psi_k^{1+\varepsilon}\Omega} * \frac{N^{\frac{1}{\xi}} log(N)}{\xi^2}\mu_k^s d\xi>0$, because F' is always positive by assumption as a CDF. In other words, a decrease in ξ lowers x_k^* or reduces migration.

(II) A fall in ξdisproportionately impacts low skilled. To show this note $\frac{d^2 x_k^*}{d\xi d\psi} = (1+\epsilon)N^{\frac{1}{\xi}}log(N)\psi_k^{-2-\varepsilon}\Omega^{-1}\xi^{-2} > 0$. So the reduction in x is smaller in absolute value for greater values of ψ. So a given decrease in ξ reduces migration by more for lower skills. Thus a fall in ξ will lead migration to become more skill biased.

(III) A fall in ξ reduces convergence. Per capita income in the North can be written:

$$\frac{\sum_k w_k \ell_k(\mu_k^N + F(x_k^*)\mu_k^S)}{\sum_k \mu_k^N + F(x_k^*)\mu_k^S} = \frac{\sum \psi_k^{1+\varepsilon}\left((1-\alpha)A\underbrace{(\sum_k(\mu_k + F(x_k^*)\mu_k^S)\psi_k^{1+\varepsilon})^{-\alpha}}_{\gamma\text{: effective labor}}\right)^{\frac{1+\varepsilon}{1+\alpha\varepsilon}}(\mu_k^N + F(x_k^*)\mu_k^S)}{\sum_k \mu_k^N + F(x_k^*)\mu_k^S}$$

We are abstracting from productivity in the South, and so convergence is equivalent to the impact on per capita income in the North. The sign of the derivative $\frac{d}{d\xi}(\frac{\sum_k w_k\ell_k(\mu_k^N+F(x_k^*)\mu_k^S)}{\sum_k \mu_k^N+F(x_k^*)\mu_k^S})$ is determined by the following condition:

$$\underbrace{\frac{(1-\alpha)\sum_k \psi_k^{1+\varepsilon}\mu_k^S F'(x_k^*)\frac{dx_k^*}{d\xi}}{\sum_k \mu_k^S F'(x_k^*)\frac{dx_k^*}{d\xi}}}_{Marginal-Impact} \lesseqgtr \underbrace{\frac{\sum_k \psi_k^{1+\varepsilon}(\mu_k^N + F(x_k^*)\mu_k^S}{\sum_k \mu_k^N + F(x_k^*)\mu_k^S}}_{Average-Skill}$$

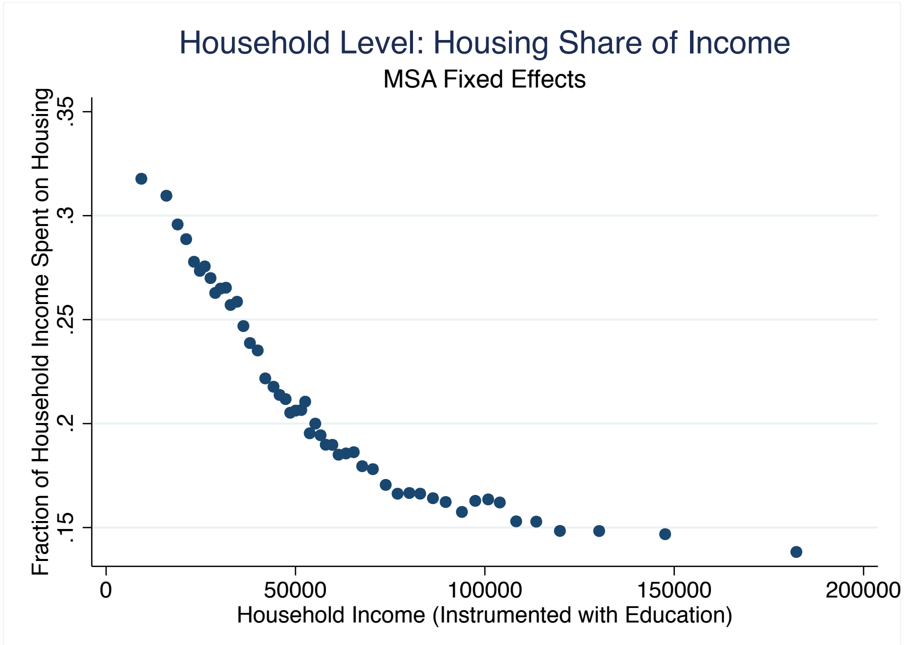
This is intuitive. Incomes rise if the marginal immigrant is more skilled than the average resident in the North, after adjusting for downward sloping demand $1-\alpha$. We can simplify this condition further by plugging in $\frac{dx_k^*}{d\xi}$ and observing that, at perfectly elastic housing supply $F(x_k^*)$ and $F'(x_k^*)$is the same for all k. Given our maintained assumption that the North is higher skill, on average, than the South we can derive a sufficient condition by setting μ_k^Nequal to zero. This allows us to rewrite the condition as

$$\underbrace{(1-\alpha)\frac{\sum_k \mu_k^S}{\sum_k \mu_k^S \frac{1}{\psi_k^{1+\varepsilon}}}}_{Marginal-Impact} \lesseqgtr \underbrace{\frac{\sum_k \psi_k^{1+\varepsilon}\mu_k^S}{\sum_k \mu_k^S}}_{Average-Skill}$$

By Jensen's inequality, the left-hand side of this equation is less than the right hand side. As a result, the derivative of income with respect to the housing supply elasticity is negative , and an increase in ξ reduces per capita income in the North. Thus a fall in ξ will reduce convergence.

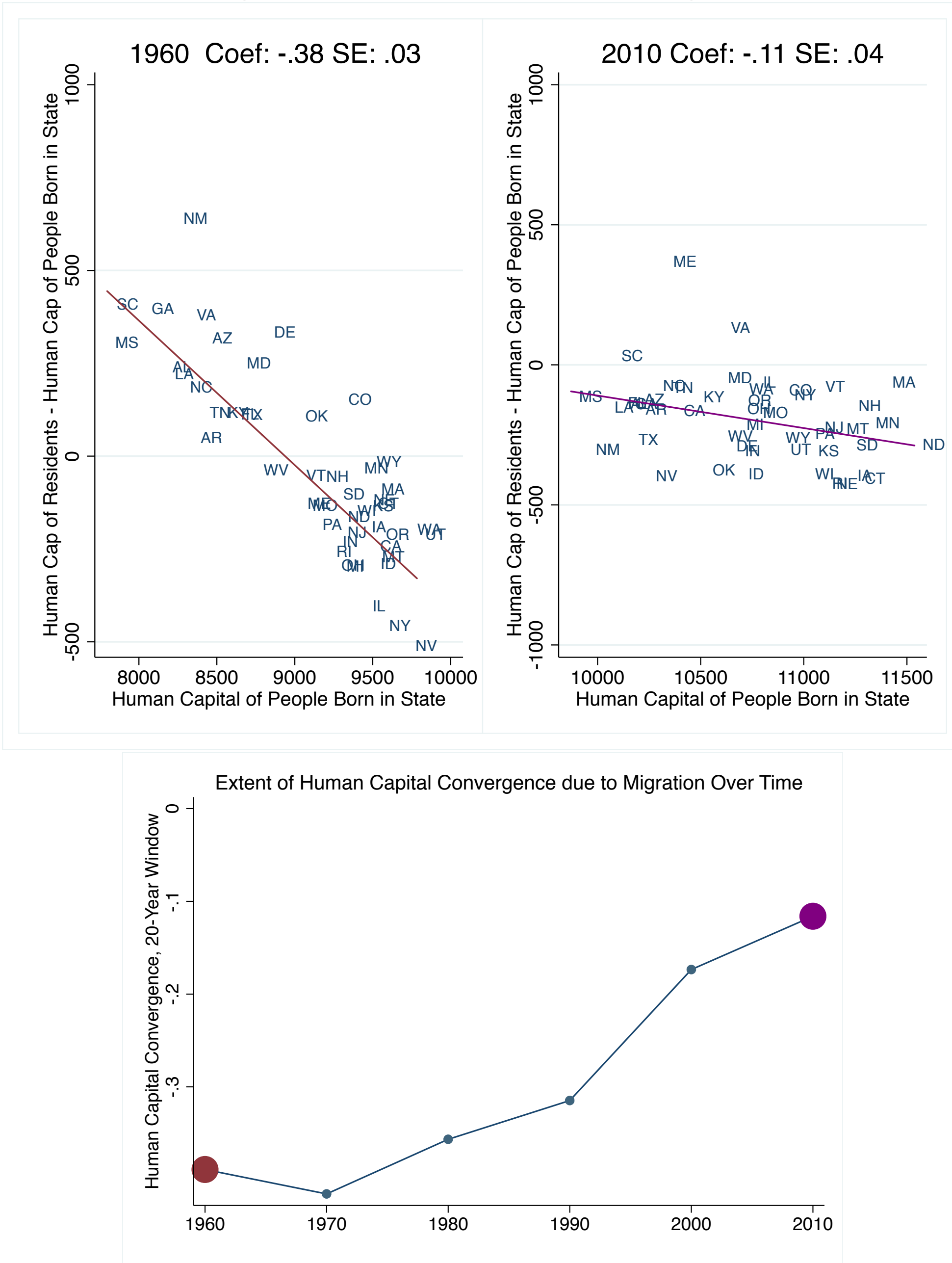
B Appendix Figures

Figure B.1: Housing Is An Inferior Good



Note: This figure plots the relationship between the share of household income spent on housing and average household income in the 2010 ACS, conditional on MSA-level fixed effects. Because annual income is volatile, we instrument for it using education levels. Specifically, we construct predicted income for each household by summing the average wages associated with the detailed education level of all the household's prime age members (25-65). We divide the sample into 50 bins based on household predicted income and plot the average housing share for each bin, controlling for 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.

Figure B.2: Human Capital Convergence



Notes: Human capital index is estimated by regressing $\log Inc_{ik} = \alpha_k + X_{ik}\beta + \varepsilon_{ik}$ in the 1980a Census, where α_k is a set of seven education indicators, and then constructing Human Capital$_j = \sum_k exp(\hat{\alpha}_k) \times Share_{kj}$. 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 $HumanCap_{j,res} - HumanCap_{j,birth} = \alpha + \beta HumanCap_{j,birth} + \varepsilon_j$ in 1960 and 2010. More detail is provided on the construction of this index in Section 4.2. The bottom panel plots a time-series of coefficients. The larger red and purple dots correspond to the coefficients from the first two panels.

C Why Did Land Use Regulations Change?

Since Ellickson [1977]'s seminal article, it has been widely accepted that municipalities' land use restrictions serve to raise property values for incumbent homeowners.[37] In this section, we examine the institutional and demographic factors which may have led such regulations to become more widespread and more effective in constraining supply across an entire region.

Many land use scholars point to a landmark shift toward new stringencies in regulations in the 1960's and 1970's. Fischel [2004] argues that in the wake of racial desegregation, land use restrictions allowed suburban residents to keep out minorities using elevated housing prices, and that environmentalism provided a sanitized language for this ideology. He writes "I submit that neighbour empowerment and double-veto systems, in conjunction with local application of environmental laws, changed metropolitan development patterns after 1970." In a book on land use regulation, Garrett [1987] writes

> A changing public attitude toward growth and development within many local communities emerged in the early 1960s. Two factors were simultaneously responsible for this change. First, there was an increasing concern over environmental issues, and it was apparent that certain types of economic development were detrimental to the environment. Second, economic analysis began to demonstrate that all forms of economic development did not generate a positive fiscal impact in every community.

Along similar lines, the American Land Planning Law textbook (Taylor and Williams [2009]) write that, after a period in the 1900's during which courts typically held the application of restrictions to particular tracts of land to be invalid, the courts "went to the other extreme, tending to uphold anything for which there was anything to be said." Our statistical regulation measure is broadly consistent with this argument, although the change in the intellectual climate described above somewhat preceded the run-up in our measure – the flow of new land use cases rose sharply from 1970 to 1990.

Because land use rules are administered at the local level, there are no seminal Supreme Court cases which marked this new era of jurisprudence. Among state cases, scholars typically cite *Mount Laurel vs. National Association for the Advancement of Colored Persons* (NAACP) as among the most important. Philadelphia suburb Mount Laurel, at the time composed primarily of single family houses, adopted rules which required that developers of multi-family units provide in leases that (1) no school-age children may occupy a one-bedroom unit and (2) no more than two children may occupy a two-bedroom unit. In addition, should a development have more than 0.3 children per unit on average, the developers were required to pay any additional tuition costs. The NAACP sued, and in 1975, the New Jersey Supreme Court ruled in its favor, finding that each community had to provide its "fair share" of "low- and moderate-income housing."

[37] Blanchflower and Oswald [2013] demonstrate the link between homeownership and land use regulation empirically.

While the NAACP won the case, Mount Laurel and like-minded suburbs won the war. Mount Laurel's new planning ordinance rezoned only 20 of its 14,300 acres, choosing locations such that "the new zones had serious physical difficulties and restrictions created by the ordinance that rendered their actual development for low-cost housing virtually impossible" (Garrett [1987]). In 1977, the state Supreme Court issued a new ruling in the *Oakwood at Madison* decision, which substantially rolled back its prior decision, finding instead that that courts were not competent to determine what constituted a "fair share". These cases led to the "Mount Laurel Doctrine," wherein judges began to play a continuing role in monitoring local zoning policies, but the sea change had already occurred in New Jersey. From 1970 to 2010, its urban population grew at an annual rate of 0.4%, less than half the national average for this period.[38]

New state and regional environmental restrictions on land use, detailed in a White House report titled "The Quiet Revolution in Land Use Control", added another constraint on new construction. These restricions played a crucial role in preventing construction on a metro-wide level, an argument highlighted by Ellickson [1977]. In a Tiebout model where consumers choose locations, if some municipalities restrict construction as Mount Laurel did, and other places respond by issuing more permits, then the aggregate impact on new units and average prices could be zero. For example, in the East Bay region in California, while many municipalities restricted construction, the coastal city of Emeryville adopted developer-friendly policies, yielding much higher-density units. In 1969, the California Legislature gave the San Francisco Bay Conservation and Development Commission the power to require permits from anyone seeking to develop land along the shoreline (Bosselman and Callies [1971]). The Commission then blocked a plan by Emeryville to fill the Bay and construct large developments there.[39] The East Bay has remained an attractive place to live, but with no municipality willing to allow new construction, housing prices across the East Bay have soared in recent years.

Local variation in regulations is not randomly assigned; it is the product of substantial work by local governments and regulatory bodies. There is some recent work on the political economy of the regulations. Kahn [2011] shows that in California, cities which vote Democratic tend to issue fewer housing permits. Hilber and Robert-Nicoud [2013] and Schleicher [2013] develop political economy stories where changes in the share of developed land, and in the structure of city politics, respectively, cause changes in land use policies.

D Data from State Supreme Court and Appellate Court Cases

Westlaw is a online database of court documents. In March of 2012, we accessed Westlaw's database of reported state supreme and appellate court decisions (at 1.next.westlaw.com). For each year and state from 1940 to 2011 we recorded the total number of cases in the

[38]Urban population is defined as population living in a Primary Metropolitan Statistical Area.

[39]A change in town leadership in the election of 1987 also led to a slowdown in new development. Nevertheless, Emeryville today still has some of the highest-density construction in the East Bay and this new regional authority further limited Tiebout competition.

database, the total number of cases in the database containing the string "land use", and the total number of cases containing the term "zoning". For example, in 1976 there were 1,303 documents in the database in California of which 31 contained the string "land use". These counts were then used to construct proxies for the severity of land use regulation as described in the text.

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

$\Delta \ln var_{it}$ (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. LMA population is constructed 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).

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
Δ Log Pop_{it} (Annual Rate in %) = $\alpha+\beta_t \ln y_{it-1}+\varepsilon_{it}$							
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 $ObservedPop_t$ - (Pop_{t-10} + $Births_{t,t-10}$ + $Deaths_{t,t-10}$). 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 (net $migration_{t,t-20}$ + pop_{t-20}) - log(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. LMA population is constructed 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).

APPENDIX TABLE 3

Returns to Living in a High Income State by Skill

	1940	1960	1970	1980	1990	2000	2010
Panel A. Returns to Migration (OLS)							
				Income Net of Housing Costs			
Average State Income	0.880***	0.736***	0.786***	0.726***	0.657***	0.539***	0.356***
	(0.020)	(0.026)	(0.042)	(0.077)	(0.035)	(0.035)	(0.046)
Average State Income X	-0.180**	0.133	0.090	0.040	0.227**	0.614***	0.610***
HH Skill	(0.066)	(0.077)	(0.066)	(0.116)	(0.071)	(0.092)	(0.075)
N	255,391	306,576	339,412	2,116,772	2,924,925	3,142,015	694,985
Panel B: Returns to Migration (IV for State of Residence with State of Birth)							
				Income Net of Housing Costs			
Average State Income	0.932***	0.776***	0.859***	0.772***	0.667***	0.488***	0.258***
	(0.029)	(0.038)	(0.055)	(0.093)	(0.036)	(0.035)	(0.051)
Average State Income X	-0.212***	-0.036	-0.083	-0.353**	0.222**	0.708***	0.614***
HH Skill	(0.057)	(0.100)	(0.097)	(0.137)	(0.086)	(0.122)	(0.123)
N	255,391	306,576	339,412	2,116,772	2,924,925	3,142,015	694,985

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 $Y_i - P_i = \alpha + \gamma Skill_i + \beta_1 Y + \beta_2 Y * Skill_i + \theta X_i + \varepsilon_i$, 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.

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)
Panel A: Low-Skill People, 1940					
Log Nominal Income	1.313***	--	1.049**	1.007**	1.086**
	(0.470)	--	(0.438)	(0.443)	(0.443)
Log Group-Specific Income Net of Housing	1.236***	1.109***	1.017***	0.980***	0.995***
	(0.364)	(0.274)	(0.350)	(0.352)	(0.338)
Panel B: High-Skill People, 1940					
Log Nominal Income	0.611	--	0.617	0.585	0.475
	(0.392)	--	(0.419)	(0.387)	(0.411)
Log Group-Specific Income Net of Housing	0.773*	0.899**	0.905*	0.821*	0.701
	(0.400)	(0.337)	(0.462)	(0.415)	(0.513)
Panel C: Low-Skill People, 2000					
Log Nominal Income	-2.173**	--	-2.456***	-2.377***	0.281
	(1.006)	--	(0.792)	(0.757)	(8.453)
Log Group-Specific Income Net of Housing	4.309**	6.042***	-0.357	1.725	-11.99
	(2.007)	(2.140)	(1.167)	(1.418)	(11.51)
Panel D: High-Skill People, 2000					
Log Nominal Income	4.077***	--	1.786***	2.894***	19.32***
	(0.694)	--	(0.611)	(0.649)	(5.373)
Log Group-Specific Income Net of Housing	4.715***	3.634***	1.937***	3.593***	14.06***
	(0.894)	(1.280)	(0.701)	(0.874)	(4.567)

Note: Each cell represents the results from a different regression. The table regresses 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 4 and 5. The second column shows the effect of doubling the housing cost measure 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. Standard errors clustered by state. *** p<0.01, ** p<0.05, * p<0.1

APPENDIX TABLE 5

Impacts of "Zoning" Regulation Measures on Permits, Prices, Migration, and Convergence

	Annual Construction $Permits_t$ % of Housing Stock	Log House Price $_{t-20}$	ΔLog $Population_{t, t-20}$ Annual Rate in %	Δ Log Human $Capital_{t,t-20}$ Annual Rate in %	Δ Log Income Per $Cap_{t,t-20}$ Annual Rate in %
	(1)	(2)	(3)	(4)	(5)
Zoning Reg Measure $1(zoning_{it} > zoning^{median}_{2005})$					
Log Inc Per Cap_{t-20}	5.955***	0.683***	2.507***	-0.0502***	-2.179***
	(2.165)	(0.114)	(0.690)	(0.0159)	(0.141)
Log Inc Per Cap_{t-20} * High Reg	-7.246***	1.032***	-3.646***	-0.00284	1.294***
	(2.456)	(0.255)	(1.064)	(0.0329)	(0.453)
N	1,536	384	2,448	288	2,448
R^2	0.293	0.886	0.158	0.325	0.818
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:

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

This table uses an alternate land use regulation measures of court case counts for the term "zoning". 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 6

Impact of Alternate Scaling of Land Use Variable on Directed Migration and Income Convergence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	ΔLog Population$_{t, t,t+20}$				Δ Log Income Per Cap$_{t,t+20}$			
Log Income	1.688**	2.274**	1.492	1.785**	-2.034***	-2.131***	-2.193***	-2.035***
	(0.637)	(0.860)	(1.043)	(0.702)	(0.102)	(0.0994)	(0.146)	(0.0995)
Log Income * Regulation Measure (Centiles of land use cases per capita, baseline measure)	-1.875***				1.304***			
	(0.608)				(0.393)			
Log Income * Regulation Measure (Centiles of land use cases per square mile)		-2.723***				1.287***		
		(0.950)				(0.317)		
Log Income * Regulation Measure (Centiles of land use cases per local govt)			-2.600**				1.721***	
			(1.168)				(0.418)	
Log Income * Regulation Measure (Centiles of land use cases as share of total cases)				-1.794***				1.153***
				(0.591)				(0.328)
Observations	2,448	2,448	1,872	2,448	2,448	2,448	1,872	2,448
R-squared	0.142	0.187	0.077	0.145	0.811	0.834	0.793	0.812

Notes: The table reports the coefficients β and β_{reg} from regressions of the form: $\ln y_{it} = \alpha_t + \alpha\, reg_{it} + \beta \ln y_{it} + \beta_{reg} \ln y_{it} reg_{it} + \varepsilon_{it}$. The regulation measure is centiles of: land use cases per capita, land use cases per square mile, land use cases per local government, and land use cases as a share of total cases in the appellate court data base . The centiles are scaled between zero and one. Data on local governments come from the Census of Local Governments, and is interpolated in off-years. Sample sizes are lower for those regressions, as the first data begin in 1952, as opposed to 1940. The dependent variables are the population change (directed migration) and the change in log per-capita income (convergence). The data are annual. Standard errors clustered by state. *** p<0.01, ** p<0.05, * p<0.1