Economic Geography & Deportation*

Adam Harris
Class of 2017
Yale University

Advisor: Michael Peters
Department of Economics
Yale University

April 3, 2017

Abstract

I study the long-run economic effects of a policy of deporting undocumented immigrants. To evaluate such a policy, I develop a dynamic spatial growth model featuring migration, endogenous innovation, and economic geography. Using this framework, I simulate various deportation policy scenarios for the United States and quantify gains and losses in welfare and real wages, both in the aggregate and across states. From these results, I conclude that a policy of deporting undocumented immigrants leads to small increases in welfare and real wages of US workers but that these gains do not persist into the long run.

1 Introduction

In 2016, the political landscapes of the United States and European Union were reshaped by a surge of opposition to liberal immigration policies and calls for harsh enforcement of existing immigration laws. Most notably, a wave of support for more restrictive immigration policies was a driving force behind the victories of both the "Brexit" campaign in the United Kingdom and the Presidential campaign of Donald J. Trump in the United States. Although support for policies limiting immigration may derive from concerns related to culture and identity, both campaigns more often emphasized the negative economic effects of allowing large-scale immigration. For instance, in June 2016, former Conservative Party leader and

*I offer my sincerest thanks to Michael Peters for his excellent advising. His generosity in agreeing to serve as my thesis advisor while on leave and his insightful guidance have made this paper possible. I also thank others members of the Yale Economics Department who have guided me and helped me to develop economic intuition and research skills. In particular, I thank Fabian Eckert, Tony Smith, and Joseph Altonji.

prominent "leave" campaigner Iain Duncan Smith predicted that the recent uptick in immigration to Britain would result in "a 10% fall in wages." Vice Presidential candidate Mike Pence expressed similar concerns in August 2016, arguing that a "flood of illegal immigration has contributed mightily to depressing wages in this country and denying jobs and opportunities to Americans."

In supposing that large-scale immigration results in lower wages, Smith and Pence intuitively appeal to a simple static model of labor markets, in which the arrival of new immigrants shifts right the labor supply curve and lowers the equilibrium wage rate. In this model, deporting undocumented immigrants would be represented by the opposite shift and would thus increase the equilibrium wage. However, such a model is overly simplistic, and fails to acknowledge two features of labor markets and firm behavior that are critical in determining how a policy of deporting undocumented immigrants might affect economic outcomes. First, the simple static model on which Pence's argument implicitly relies assumes that the labor supplied by undocumented immigrants and by US citizens are perfect substitutes. This strong assumption is highly significant, as it implies that deportation would lead to reduced labor market competition and, thus, wage gains for US citizens. A model which assumed imperfect substitutibility between native and immigrant labor, on the other hand, would likely predict such effects to be much smaller. Second, the simple static model fails to account for a longer-run dynamic effect of deportation, which derives from the fact that the innovation decisions of firms respond to local labor supply. In a model featuring this linkage, the deportation of undocumented workers would reduce the incentives of US firms to innovate, resulting in lower rates of US productivity growth. The first issue has been explored extensively in the labor economics literature, with various approaches and findings summarized by Dustmann, Schönberg, and Stuhler (2016) [1]. In this paper, I focus on the second issue, developing a dynamic spatial growth model in which innovation responds to the distribution of labor across space. I use this model as a lens through which to study the long-run effects of deporting undocumented immigrants from the United States.

Over the past 25 years, the macroeconomic literature on growth has emphasized the link between ideas and growth, a connection highlighted most clearly by Romer (1990) [2]. In this class of idea-based growth models, a larger population spurs more rapid innovation. While the rate of new ideas per person might be the same in a large population as in a small one, the fact that ideas are non-rivalrous means that it is the quantity of new ideas rather than ideas per capita that matters for economic growth. To the extent that innovations are not perfectly mobile across space, the spatial distribution of population is then significant in

determining local innovation and thus, in the long run, local wages. Therefore, evaluating the long-run effects of deportation requires a model that features this relationship between population and innovation.

Desmet, Nagy, & Rossi-Hansberg (2017) [3] (hereafter referred to as DNRH) present a dynamic spatial growth model which incorporates geography and, in the spirit of Romer, features a relationship between population distribution and innovation. This work furthers the spatial economics literature in two significant ways. First, the DNRH model features a rich and realistic geography, which includes local productivity that evolves endogenously, realistic trade costs, endogenous amenities, spatial migration frictions, and land. Second, by imposing a restriction on the structure of migration costs (Assumption 1 in DNRH), the authors reduce the agent's migration decision, which would otherwise be a complicated dynamic problem, to a simple static decision. This is achieved by assuming that migration frictions take the form of region-specific entry costs and exit benefits. An agent pays a permanent multiplicative utility cost $m_1(r)$ to enter region r and receives a permanent multiplicative utility benefit $m_2(r)$ when leaving region r, with $m_1(r)$ and $m_2(r)$ being reciprocals. Each period, agents choose locations so as to maximize lifetime utility. However, the fact that the entry cost that an agent pays when entering a region is reimbursed upon leaving the region implies that she need not consider how her migration decision this period will affect the migration problem she faces in future periods. Thus, the agent's decision of where to locate reduces to a static problem. Furthermore, the choice of r_1, the agent's location next period, is independent of r_0, her location this period.

As DNRH perform an ambitious counterfactual exercise for a world comprising 64,800 regions, this assumption is critical to making it feasible to solve the model and perform counterfactual data exercises. This migration cost structure, while convenient for the exercise that the authors perform, makes it impossible to use the DNRH model to simulate realistic dynamic effects of deportation, represented by an exogenous change in the initial spatial distribution of population. The fact that each agent's location decision is independent of her current location implies that the population distribution for next period is independent of the current population distribution. This means that in a counterfactual scenario in which one exogenously changes the spatial population distribution, this change will not be persistent, and will thus have no significant long-run effects. While the DNRH model has the benefits of tractability and computational simplicity, it is not a suitable framework for analyzing the dynamic effects of deportation.

Allen & Donaldson (2017)[4] (hereafter referred to as AD) present a dynamic spatial model in which tractibility is achieved by assuming a generational structure in the migration problem rather than by imposing the restrictive assumption on migration costs employed by DNRH. AD assume that each agent is born in some region i, where her parent resides. The agent is not economically active during childhood, but upon reaching adulthood, she chooses a location j so as to maximize her own utility, but without regard for the lifetime utility of her offspring. This structure also results in a static migration decision and, thus, AD are able to achieve tractability in the migration problem while still allowing for flexible migration frictions and a realistic relation between next period's population distribution and the population distribution in the current period.

By combining the generational migration structure of AD with the innovation/growth structure from DNRH, I develop a model which offers an improved dynamic framework for simulating the effects of deportation. Using this new model, I perform experiments simulating the long-run economic effects of a policy of deporting undocumented immigrants from the United States. I find that while workers in every US state enjoy short-run benefits of reduced competition in local labor markets, these gains deteriorate over time. In the very long run, US workers see lower real wages under the deportation scenario as compared with the simulation without deportation.

This work extends a literature which employs quantitative models to study issues of economic geography. These quantitative models have their origin in Eaton & Kortum (2002) [5], work which reinvigorated traditional Ricardian models of international trade by introducing a framework that both generalized to a many-country setting and could be taken to data. More recently, others, most notably Allen & Arkolakis (2014) [6], have adapted this modeling technique to study the spatial distribution of economic activity across many regions. A complete taxonomy of this family of models is provided by Redding & Rossi-Hansberg (2016) [7]. Key features of these models include agglomeration and dispersion forces. A number of papers in the urban economics literature, including Ahlfeldt, Redding, Sturm & Wolf (2015) [8] and Greenstone, Hornbeck & Moretti (2010) [9], document substantial agglomeration externalities, and several others, including Ellison, Glaeser & Kerr (2010) [10], suggest mechanisms underlying the existence of such forces. The balance between these agglomeration forces and dispersion forces plays a central role in determining equilibrium properties and long-run behavior of quantitative spatial models.

The paper proceeds as follows. In Section 2, I present data on both bilateral migration

flows between the US and Mexico, as well as on undocumented immigration to the United States. I also discuss the recent history of US immigration policy, providing background for later evaluation of the economic implications of migration policy decisions. In Section 3, I outline the model and, in Section 4, analyze properties of its equilibrium system and balanced growth path. In Section 5, I outline algorithms for calibrating migration costs, as well as initial values of the state variables. In Section 6, I present and analyze results of both the baseline simulation and deportation policy experiments. In Section 7, I draw conclusions about the long-run effects of a policy of deporting undocumented immigrants from the US. Finally, in Section 8, I discuss the limitations of the computational experiments I perform and highlight opportunities for further theoretical work in this area of research.

2 US Immigration: Data and policy context

In this section, I present data on immigrants to the US—both legal and undocumented—to provide background for the US immigration policy debate.

The 2015 American Community Survey reported a United States' immigrant population of 43.3 million, representing 13.5% of all people living in the US. While, as Donald Trump frequently alluded to during the US Presidential campaign, this proportion is significantly higher than the rate typical for the US in the second half of the 20th century, it does not represent a departure from longer-run historical norms; in 1910, for example, 14.7% of the US population was foreign-born, and this proportion was even higher in the late 19th century (Migration Policy Institute). Though Trump differs from many mainstream Republicans in arguing that high levels of immigration, including legal immigration, are undesirable, policymakers on both sides of the aisle have long expressed concern about the level of undocumented immigration to the US. President Bill Clinton, for instance, declared in his 1996 book *Between Hope and History* that "We must not tolerate illegal immigration." He went on to say that his administration had "moved forcefully to protect American jobs...[and] removed 30,000 illegal workers from jobs across the country."

As of 2014, Pew Hispanic estimated the number of undocumented immigrants in the US at 11.1 million, with just over half of that number (5.85 million) being of Mexican origin.[11] While the relatively large undocumented population in the US is frequently cited to argue in favor of tighter US border security, more relevant figures that are often omitted from such discussions are net bilateral migration flows between the US and Mexico. During the late 1990s, estimated migration flows across the southern border of the US were quite one-sided,

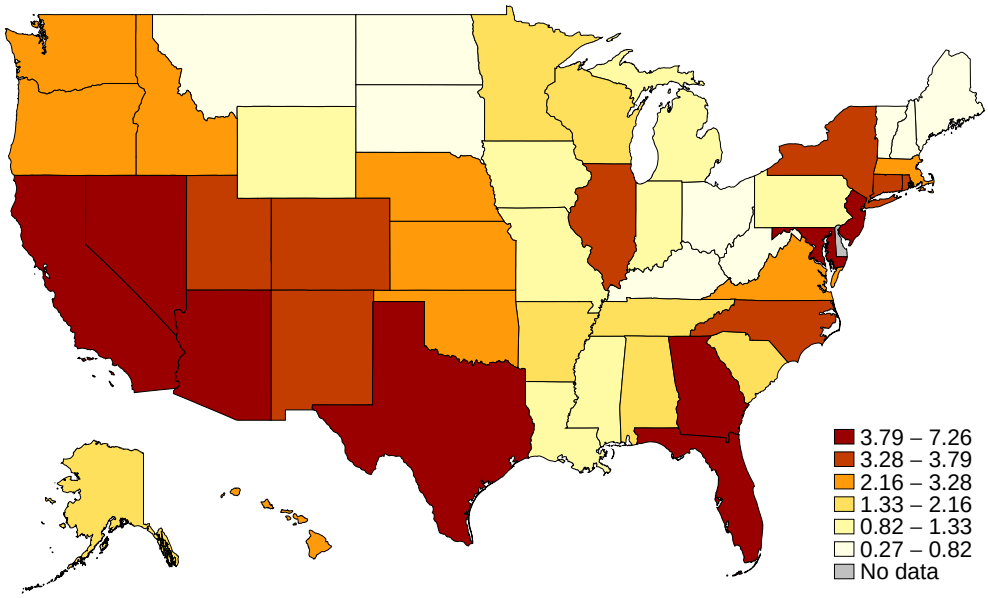
with approximately 2.94 million immigrants entering the US from Mexico and only 670,000 migrating from Mexico to the US between 1995 and 2000. However, the direction of these net flows have since reversed. Between 2009 and 2014, about 870,000 immigrants moved from Mexico to the US and about 1 million migrated from the US to Mexico. The 2014 Mexican National Survey of Demographic Dynamics from which these estimates were calculated also offers insight into the factors underlying this trend. While deportation did account for about 14% of the 1 million migrants from the US to Mexico in the 2009-2014 period, most left the US voluntarily, with 61% citing family reunification as the reason for migration. [12]

The fact that bilateral net migration flows from Mexico to the US are now negative leads one to question whether Trump's campaign promise of "an impenetrable physical wall on the southern border" would help achieve his goal of reducing the number of undocumented immigrants living in the US. However, another policy pledge which Trump made as President-elect, the deportation of two to three million undocumented immigrants, would undoubtedly have a larger effect. Although Trump's tone in discussing undocumented immigrants and deportation has been markedly more aggressive than that of his predecessor, this suggested policy would not be a violent departure from that carried out by the Obama Administration. According to the Department of Homeland Security's "Yearbook of Immigration Statistics," more than 3.1 million compulsory removals[1] of undocumented immigrants were conducted during the fiscal years 2008-2015.[14]

While immigration enforcement is (at least in theory) a federal issue and most immigration policy decisions are made at the national level, a policy of deporting undocumented immigrants has significant local implications. Undocumented immigrants living in the US are overwhelmingly concentrated in the Southwest, the Southeast, and few metropolitan areas of the Northeast and Midwest (see Figure 1). This highly uneven geographic distribution means that US immigration policy has significant differential implications across the 50 states. This makes evaluating the economic effects of a policy of deportation an interesting economic geography problem. In what follows, I will develop a quantitative spatial model and use it to evaluate effects both aggregate and local.

[1]Note that this figure does not represent the number of undocumented immigrants living in the United States who were deported. This includes people apprehended while attempting to cross the border or shortly after crossing the border. According to the *Los Angeles Times*[13], this category accounted for nearly two-thirds of the deportations recorded in 2013.

Figure 1: Undocumented immigrants as a percent of state population [15]



[2]

3 The Model

In this section, I outline the quantitative spatial model which I will use as a framework for evaluating the economic effects of deportation. I begin by characterizing economic geography in the model. Second, I outline and solve the consumption and migration decisions faced by agents. Third, I define the evolution of technology and outline and solve the problem faced by firms. Finally, I show that a special case of my model is isomorphic to that of Allen & Donaldson.

3.1 Geography

The world comprises N spatially distinct regions denoted $i, j = 1, 2, \ldots, N$. Economic geography in this model is defined by land, local amenities, migration costs, and trade costs.

Each region i has measure of land $H_i > 0$. The exogenous local amenity associated with living in region i at time t is $\bar{u}_i > 0$. Let $D \in \mathbb{R}^{N \times N}$ be a matrix of exogenous iceberg trade costs, with $D_{i,j} \geq 1$ representing the cost of transporting goods from region i to region j. Finally, let $M \in \mathbb{R}^{N \times N}$ be the matrix of exogenous iceberg migration costs, with $M_{i,j} \geq 1$

[2]Figures 1-4 and 11-12 were constructed using the MapTile package for Stata. I acknowledge Michael Stepner of MIT for building this package and for providing map data files for the US to the public.

representing the cost (in terms of utility) of moving from region i to region j.

For the purpose of my deportation experiments, I set $N = 51$, with regions representing the 50 United States and Mexico. In order to make estimation feasible (see Section 5), I impose the following structure for bilateral migration costs:

$$M_{i,j} = \begin{cases} 1 & i = j \\ \delta_1 & \text{if } i, j \text{ are both in US} \\ \delta_2 & \text{if } i \text{ is in US and } j = \text{Mexico.} \\ \delta_3 & \text{if } j \text{ is in US and } i = \text{Mexico.} \end{cases}$$

where $\delta_1, \delta_2, \delta_3 \geq 1$.

3.2 Agent's problem

A measure of adult agents resides in each region, and the total population of adult agents across all regions is $\bar{L}$. Some agents are citizens or legal residents while others are undocumented immigrants. L_{it}^C denotes the citizens and legal residents per unit land, while L_{it}^U denotes the undocumented immigrants per unit land residing in region i in period t. In this model, the two groups do not differ in preferences or in labor productivity. The total population density of region i is $L_{it} = L_{it}^C + L_{it}^U$.

In each period, adult agents derive utility from consumption of a variety of goods.

Preferences Agents consume varieties $\omega \in [0, 1]$. Adult consumers have CES preferences over this continuum of varieties with elasticity of substitution σ. For convenience, I set $\rho = \frac{\sigma-1}{\sigma}$. Solving the utility maximization problem, the utility from consumption of an agent living in region i at time t is

$$\left[\int_0^1 (c_{it}^\omega)^\rho \, d\omega\right]^{\frac{1}{\rho}} = \frac{y_{it}}{\left[\int_0^1 (p_{it}^\omega)^{1-\sigma} \, d\omega\right]^{\frac{1}{1-\sigma}}} = \frac{y_{it}}{P_{it}}$$

where y_{it} is the income of an agent living in region i and P_{it} is the CES price index.

Welfare is given by

$$W_{it} = \left(\bar{u}_i L_{it}^{-\lambda}\right) \frac{y_{it}}{P_{it}} \tag{1}$$

Note that in addition to $\bar{u}_i$, the exogenous component of the local amenity introduced above, there is also an endogenous component. L_{it} denotes the population density of region i at time t, so $L_{it}^{-\lambda}$ with $\lambda > 0$ represents a dispersion force; all else equal, agents prefer to reside in less densely populated regions.

Adult agents also face a migration decision, in which they observe the wages, prices, and amenities offered by each region and choose one in which to reside. Migration decisions in this model follow the generational structure of Allen-Donaldson:

Timing The following steps summarize the lifetime of an agent who is an adult in period t:

1. In period $t-1$, the agent is born in region i, the region where her parent lives. She is not economically active during childhood.

2. At the start of period t, the agent reaches adulthood and chooses j, the region in which she will live, consume, and work as an adult. This decision is represented by the maximization problem presented in (2). Note that in choosing j, the agent seeks to maximize her own welfare, but does not consider the lifetime welfare of her child. This assumption is crucial for delivering tractability. If this assumption were relaxed, the migration decision would be a complicated dynamic problem and solving would require considering agents' intertemporal preferences and expectations of future states of the economy.

3. In period t, the agent gives birth to a child in region j.

Migration problem An agent k born in i in period $t-1$ chooses her period t location by solving the problem

$$\max_j \frac{W_{jt}}{M_{ij}} \times \epsilon_j^k \tag{2}$$

where ϵ_j^k is agent k's idiosyncratic taste shock drawn from a Fréchet distribution. Assume $\epsilon_j^k \sim \text{Fréchet}(1/\Omega)$ is i.i.d. across individuals and locations with $\Omega > 0$ exogenous. Then, migration shares follow from the max-stable property of the Fréchet distribution, along with

the Law of Large Numbers. Recall that L_{it} denotes the population *density* of region i at time t. Let ℓ_{ijt} denote the proportion of people born in i in $t-1$ who move to j in t:

$$\ell_{ijt} = \left(\frac{W_{jt}/M_{ij}}{\Pi_{it}} \right)^{1/\Omega} \tag{3}$$

where

$$\Pi_{it} = \left(\sum_{k=1}^{N} \left(W_{kt}/M_{ik} \right)^{1/\Omega} \right)^{\Omega} \tag{4}$$

Thus, the period t population of j is given by

$$H_j L_{jt} = \sum_i H_i L_{i,t-1} \left(\frac{W_{jt}/M_{ij}}{\Pi_{it}} \right)^{1/\Omega} \tag{5}$$

$$= W_{jt} \sum_i H_i L_{i,t-1} \left(M_{ij} \Pi_{it} \right)^{-1/\Omega} \tag{6}$$

Note that this equation, which defines the model's population dynamics, differs from the corresponding equation in DNRH (equation (7) in DNRH) in one important respect. In the latter, the vector L_t is independent of the vector L_{t-1}. In DNRH, this path independence comes from the fact that, as a result of the structure imposed on migration costs, the fraction of people who move from region i to region j does not depend on i. That is, in the DNRH model, an equal proportion of people living in each region will move to j. My model (like the AD model), on the other hand, features path dependence; the spatial allocation of labor next period depends directly on the spatial allocation of labor this period. This allows for a more realistic characterization of population dynamics and, in particular, will deliver more reasonable predictions for how these dynamics respond to an exogenous change in the population distribution, such as the deportation of undocumented immigrants.

3.3 Technology and innovation

In this subsection, I outline the structure of technology and present the problem faced by firms.[3]

[3] As the supply side of my model follows closely DNRH, the material in this section has in very similar form been used in my Senior Project in Applied Mathematics, titled "Exploring the economic effects of mass deportation in a spatial growth model". That project used the DNRH model to perform counterfactual experiments and, thus, required outlining the DNRH model in detail. I include the material here for completeness.

As in DNRH, a firm produces a good $\omega \in [0,1]$ according to the following production function

$$q_{it}^{\omega} = \left(\phi_{it}^{\omega}\right)^{\gamma_1} \left(L_{it}^{\omega}\right)^{\mu} z_{it}^{\omega} \tag{7}$$

where q_{it}^{ω} is the period-t output per unit land of a firm located in region i that produces good ω. ϕ_{it}^{ω} is a level of "innovation," chosen by the firm, which in part determines its productivity. Innovation is costly, and the firm must employ $\nu \left(\phi_{it}^{\omega}\right)^{\xi}$ extra units of labor in order to enjoy this innovation. L_{it}^{ω} is the amount of labor (per unit land) employed by the firm for production (i.e., not for innovation). z_{it}^{ω} is the realization of a random variable that exogenously shifts the firm's productivity. Draws are i.i.d. across goods and periods, and are distributed Fréchet, with CDF $F(z,i) = \exp\left(-\tau_{it} \left(L_{it}\right)^{\alpha} z^{-\theta}\right)$, where $\alpha \geq 0$ and $\theta > 0$ are exogenous parameters, and τ_{it} evolves according to an endogenous dynamic process.

In particular, τ_{it} is a state variable representing the non-random component of productivity, and its evolution depends on the innovation decisions of firms in region i. In addition, technology in this model diffuses across space, so τ_{it} also depends indirectly on the innovation decisions of firms in other regions.

$$\tau_{it} = \left(\phi_{t-1,i}\right)^{\theta\gamma_1} \left[\sum_{j=1}^{N} \frac{1}{N} \tau_{t-1,j}\right]^{1-\gamma_2} \left(\tau_{t-1,i}\right)^{\gamma_2} \tag{8}$$

where $\gamma_1, \gamma_2 \in [0,1]$.

Following DNRH, the productivity shock z_{it}^{ω} is identical across firms producing a particular good ω in a particular region i. Thus, within a region i, firms producing ω have identical productivity and face identical wage rates, rental rates, and trade costs. Bertrand competition among identical firms implies that, in equilibrium, these firms will set identical prices, resulting in perfect local competition. Under perfect competition, firms in region i will bid up the rental rate to the point at which all firms earn zero profits.

In DNRH, firms choose innovation so as to maximize the present discounted value of profits. However, the fact that, in my model, all agents are short-lived and there are no capital markets suggests that firms too should be short-lived. I assume that firms choose the levels of production and innovation so as to maximize current-period profits.[4]

[4]Note that the assumption that firms are short-lived is made simply for consistency, and that this assumption does not substantively alter the behavior of firms in this model. Although firms in DNRH seek

Below, I present this static profit-maximization problem (9) and its two first-order conditions, (10) and (11). First, I define notation for prices: Let $p_t^\omega(i,i)$ be the price of a unit of ω produced in region i and sold in region i. Prices will vary across space as a result of transportation costs: The price of a unit of ω produced in i and sold in j is $p_t^\omega(j,i) = D_{i,j} p_t^\omega(i,i)$ where $D_{i,j} \geq 1$ is the iceberg trade cost of transporting goods from i to j.

The firm seeks to maximize period-t profits, which I write as revenue less costs of factor inputs. Let R_{it} and w_{it} denote, respectively, the rental rate of land and the wage rate in region i at time t.

$$\max_{L_{it}^\omega, \phi_{it}^\omega} \quad p_t^\omega(i,i)\left(\phi_{it}^\omega\right)^{\gamma_1}\left(L_{it}^\omega\right)^\mu z_{it}^\omega - w_{it}L_{it}^\omega - w_{it}\nu\left(\phi_{it}^\omega\right)^\xi - R_{it} \tag{9}$$

The first-order condition with respect to labor, L_{it}^ω:

$$\mu p_t^\omega(i,i)\left(\phi_{it}^\omega\right)^{\gamma_1}\left(L_{it}^\omega\right)^{\mu-1} z_{it}^\omega = w_{it} \tag{10}$$

And with respect to innovation, ϕ_{it}^ω:

$$\gamma_1 p_t^\omega(i,i)\left(\phi_{it}^\omega\right)^{\gamma_1-1}\left(L_{it}^\omega\right)^\mu z_{it}^\omega = \xi w_{it}\nu\left(\phi_{it}^\omega\right)^{\xi-1} \tag{11}$$

Solving (10) and (11), I find the total number of people employed by firms producing variety ω in region i. Note that this is the sum of production workers, L_{it}^ω, and innovation workers, $\nu\left(\phi_{it}^\omega\right)^\xi$. First, dividing the left-hand side of (10) by the left-hand side of (11) and dividing the right-hand side of (10) by the right-hand side of (11) gives

$$\frac{\mu\phi_{it}^\omega}{\gamma_1 L_{it}^\omega} = \frac{1}{\xi\nu\left(\phi_{it}^\omega\right)^{\xi-1}} \quad \text{or, equivalently,} \quad \frac{L_{it}^\omega}{\mu} = \frac{\xi\nu\left(\phi_{it}^\omega\right)^\xi}{\gamma_1} \tag{12}$$

It follows that total employment (per unit land) by firms producing good ω in region i is

$$\bar{L}_{it}^\omega \equiv L_{it}^\omega + \nu\left(\phi_{it}^\omega\right)^\xi = \frac{L_{it}^\omega}{\mu}\left[\mu + \frac{\gamma_1}{\xi}\right] \tag{13}$$

to maximize the present value of profits for all periods, the competitive structure of local land market results in the firm earning zero profits in all future periods. Thus, maximizing the present value of profits for all periods is equivalent to maximizing static profits.

The zero-profit condition allows me to write the equilibrium rental rate as

$$R_{it} = p_t^{\omega}(i,i)\left(\phi_{it}^{\omega}\right)^{\gamma_1}\left(L_{it}^{\omega}\right)^{\mu} z_{it}^{\omega} - w_{it}L_{it}^{\omega} - w_{it}\nu\left(\phi_{it}^{\omega}\right)^{\xi} \tag{14}$$

From (10) and (12),

$$p_t^{\omega}(i,i)\left(\phi_{it}^{\omega}\right)^{\gamma_1}\left(L_{it}^{\omega}\right)^{\mu} z_{it}^{\omega} = \frac{w_{it}L_{it}^{\omega}}{\mu} = w_{it}\frac{\xi\nu\left(\phi_{it}^{\omega}\right)^{\xi}}{\gamma_1} \tag{15}$$

Then, substituting into (14) gives

$$R_{it} = \left[\frac{\xi(1-\mu)}{\gamma_1} - 1\right] w_{it}\nu\left(\phi_{it}^{\omega}\right)^{\xi} \tag{16}$$

Under perfect competition, firms take the rental rate of land, R_{it}, as well as the wage rate, w_{it}, as given. Thus, one can solve (16) for ϕ_{it}^{ω}. Then, the optimal choice of L_{it}^{ω} follows from (12), and total employment by firms producing ω in region i follows from (13). It is important to note that the idiosyncratic productivity shock z_{it}^{ω} does not appear in (12), (13), or (16). Thus, the firm's innovation and employment decisions are independent of this shock. As this shock is the only feature differentiating firms that produce different varieties, it follows that ϕ_{it}^{ω} and L_{it}^{ω} are identical across ω. All firms in region i make the same production and innovation decisions irrespective of the goods they produce.

From these conclusions, it is straightforward to derive the input unit cost in location i at time t

$$\mathrm{mc}_{it} \equiv \left[\frac{1}{\mu}\right]^{\mu}\left[\frac{\nu\xi}{\gamma_1}\right]^{1-\mu}\left[\frac{\gamma_1 R_{it}}{w_{it}\nu\left(\xi(1-\mu)-\gamma_1\right)}\right]^{(1-\mu)-\frac{\gamma_1}{\xi}} w_{it} \tag{17}$$

$$= \left[\frac{1}{\mu}\right]^{\mu}\left[\frac{\nu\xi}{\gamma_1}\right]^{1-\mu}\phi_{it}^{\xi(1-\mu)-\gamma_1} w_{it} \tag{18}$$

and the price of a unit of good ω produced at i and sold at i

$$p_{it}(i,i) = \frac{\mathrm{mc}_{it}}{z_{it}^{\omega}} \tag{19}$$

And, since trade costs are passed through completely to consumers, the price of a unit of good ω produced at i and sold at j is

$$p_t(j,i) = D_{i,j}\frac{\mathrm{mc}_{it}}{z_{it}^{\omega}} \tag{20}$$

Exports Consumers in region j will buy good ω from whichever region offers the lowest price. Thus, the proportion of goods produced in i that are sold in j is

$$\pi_t(j,i) = \Pr\left(p_t(j,i) \leq p_t(j,k) \text{ for all } k\right) \tag{21}$$

$$= \frac{\tau_{it}\left(L_{it}\right)^{\alpha}\left[\mathrm{mc}_{it}D_{j,i}\right]^{-\theta}}{\sum_{k\in\mathcal{R}}\tau_{kt}\left(L_{kt}\right)^{\alpha}\left[\mathrm{mc}_{kt}D_{j,k}\right]^{-\theta}} \tag{22}$$

For a formal derivation of (22), see Appendix B of DNRH. This tractable expression for export shares in this model results from the max-stable property of the Fréchet distribution. The feature of Fréchet-distributed productivity shocks follows the example of Eaton & Kortum (2002)[5].

The CES price index for region i can be written as

$$P_{it} = \kappa_1\left[\sum_{j=1}^{N}\tau_{jt}L_{jt}^{\alpha}\left(\mathrm{mc}_{jt}D_{i,j}\right)^{-\theta}\right]^{-\frac{1}{\theta}} \tag{23}$$

where $\kappa_1 = \Gamma\left(\frac{1-\sigma}{\sigma}+1\right)^{\frac{1}{1-\sigma}}$. The derivation uses a change of variable; since the probability distribution of prices is known, it is convenient to integrate over prices rather than over goods. This derivation is explained fully in Eaton & Kortum (2002)[5].

Trade balance As is common in spatial models of trade, I impose trade balance. For all $i = 1, 2, \ldots, N$,

$$w_{it}H_iL_{it} = \sum_{j=1}^{N}\pi_t(j,i)w_{jt}H_jL_{jt}$$

By substituting in (22), I derive

$$\begin{aligned} w_{it}H_i\left(L_{it}\right)^{1-\alpha}\frac{1}{\tau_{it}}\left(\mathrm{mc}_{it}\right)^{\theta} &= \sum_{j=1}^{N}\frac{\left(D_{ji}\right)^{-\theta}}{\sum_{k\in\mathcal{R}}\tau_{kt}\left(\bar{L}_{kt}\right)^{\alpha}\left[\mathrm{mc}_{kt}D_{j,k}\right]^{-\theta}}w_{jt}H_jL_{jt} \\ &= \sum_{j=1}^{N}\left(D_{j,i}\right)^{-\theta}\left(\frac{P_{jt}}{\kappa_1}\right)^{\theta}w_{jt}H_jL_{jt} \end{aligned}$$

which can be rearranged as follows:

$$w_{it}^{1+\theta} H_i L_{it}^{1-\alpha} \frac{1}{\tau_{it}} \left(\left[\frac{1}{\mu} \right]^{\mu} \left[\frac{\nu \xi}{\gamma_1} \right]^{1-\mu} \right)^{\theta} (\phi_{it})^{\theta(1-\mu)\xi - \theta\gamma_1} = \sum_{j=1}^{N} (D_{j,i})^{-\theta} \left(\frac{P_{jt}}{\kappa_1} \right)^{\theta} w_{jt} H_j L_{jt} \tag{24}$$

Rental income As in DNRH, I assume local land ownership. That is, total rent payments in region i are distributed equally among current residents of i, so the total income of an agent living in i is

$$\begin{aligned} y_{it} &= w_{it} + \frac{R_{it}}{L_{it}} \\ &= w_{it} \left[1 + \left(\frac{\xi (1-\mu)}{\gamma_1} - 1 \right) \nu \frac{\phi_{it}^{\xi}}{L_{it}} \right] \end{aligned} \tag{25}$$

I recall that the first order conditions of the static profit maximization problem of a firm producing good ω, along with the zero-profit condition, give

$$\frac{(\phi_{it}^{\omega})^{1/\xi}}{L_{it}^{\omega}} = \frac{\gamma_1}{\mu \nu \xi}$$

where L_{it}^{ω} is the amount of *production* labor per unit land employed by the firm. Then, using the fact that the *total* labor (in both production and innovation) employed by the firm is $\bar{L_{it}}^{\omega} = \frac{L_{it}^{\omega}}{\mu} \left[\mu + \frac{\gamma_1}{\xi} \right]$, along with the fact that the employment and innovation decisions of the firm are independent of ω,

$$\frac{\phi_{it}^{\xi}}{L_{it}} = \frac{\gamma_1}{\nu} \times \frac{1}{\gamma_1 + \mu\xi} \tag{26}$$

Substituting into (25), I find that an agent's total income is a constant multiple of her wage:

$$y_{it} = \frac{w_{it}}{\mu} \tag{27}$$

Then, the unit cost expression simplifies to

$$\mathrm{mc}_{it} = \kappa_2^{-1/\theta} L_{it}^{(1-\mu) - \frac{\gamma_1}{\xi}} w_{it}$$

where $\kappa_2 = \left(\mu^{-\mu} \left(\frac{\nu}{\gamma_1} \right)^{\frac{\gamma_1}{\xi}} \left(\frac{1}{\gamma_1 + \mu\xi} \right)^{(1-\mu) - \frac{\gamma_1}{\xi}} \right)^{-\theta}$.

3.4 Isomorphism with Allen-Donaldson model

In the previous subsection, I outlined the technology and innovation side of my model which follows DNRH and features endogenous productivity growth. This is one respect in which my model differs from that of Allen-Donaldson, which has exogenous productivity. Another difference is that my model, unlike that of AD, also features land as a factor of production. I show in Appendix A that, considering a single period t, a special case of my model in which all regions have an equal measure of land is isomorphic to the Allen-Donaldson model. The two models are linked by Table 2, which is also presented in Appendix A.

4 Equilibrium

In order to use this model for calibrations, simulations, and policy experiments studying the effects of deporting undocumented immigrants, I must first define what I mean by an equilibrium and determine whether an equilibrium exists. If so, I must find a method for solving for equilibrium and determine whether the resulting solution is unique.

I begin by defining two terms. The first, *temporary equilibrium*, refers to an equilibrium for the economy at a particular point in time. The second, *dynamic equilibrium*, refers to a sequence of temporary equilibria representing the path of the economy over time.

Definition 1 Given vectors $\tau_t, L_{t-1} \in \mathbb{R}^N_{>0}$, a *temporary equilibrium* is a set of vectors $\{w_t, L_t, W_t, \Pi_t\}$ which solves the following system of equations:

1. From equation (4):

$$\Pi_{it}^{1/\Omega} = \sum_{j=1}^{N} \left(W_{jt}/M_{ij}\right)^{1/\Omega}$$

2. From equation (6):

$$L_{it} W_{it}^{-1/\Omega} = \sum_{j=1}^{N} \frac{H_j}{H_i} M_{j,i}^{-1/\Omega} \Pi_{jt}^{-1/\Omega} L_{j,t-1}$$

3. From equations (23) and (1): Price equations

$$w_{it}^{-\theta} L_{it}^{\lambda\theta} W_{it}^{\theta} = \kappa_2 \kappa_1^{-\theta} \sum_{j=1}^{N} \bar{u}_i^{\theta} D_{ij}^{-\theta} \tau_{jt} w_{jt}^{-\theta} L_{jt}^{\alpha-\theta(1-\mu)+\frac{\theta\gamma_1}{\xi}}$$

4. From equation (24) and (1): Trade balance condition

$$w_{it}^{1+\theta} L_{it}^{1-\alpha+\theta(1-\mu)-\frac{\theta\gamma_1}{\xi}} = \kappa_1^{-\theta}\kappa_2\mu^{-\theta}\sum_{j=1}^{N}\frac{H_j}{H_i}\tau_{it}\bar{u}_j^{\theta}\left(D_{j,i}\right)^{-\theta} w_{jt}^{1+\theta} L_{jt}^{1-\lambda\theta} W_{it}^{-\theta}$$

Note that I have used (26) to substitute in for ϕ_{it} in each of the above equations.

Definition 2 Given vectors $L_0, \tau_0 \in \mathbb{R}^N_{>0}$, a *dynamic equilibrium* is a path of temporary equilibria and technology vectors. That is, a dynamic equilibrium comprises a set of vectors $\{w_t, L_t, W_t, \Pi_t, \tau_t\}_{t=1}^T$ such that, for each $t = 1, \ldots, T$, the following two conditions are satisfied:

1. The vectors τ_t, τ_{t-1}, and L_{t-1} satisfy the technology evolution equations (8) and (26).
2. Given vectors τ_t and L_{t-1}, the set of vectors $\{w_t, L_t, W_t, \Pi_t\}$ constitute a temporary equilibrium.

Having defined temporary and dynamic equilibria, it is clear that, given initial distributions of population and technology, solving for a dynamic equilibrium entails solving for a series of temporary equilibria and, at each step, updating the technology vector, τ_t. Since (8) and (26), the equations for updating technology, map strictly positive vectors τ_{t-1} and L_{t-1} to a strictly positive vector τ_t, the questions of whether a dynamic equilibrium exists and, if so, whether it is unique are equivalent to the questions of whether the temporary equilibrium, for each t, exists and is unique.

4.1 Temporary equilibrium properties

In this section, I show that I can rewrite the conditions defining a temporary equilibrium in a simpler form and, using this simplified system, find conditions for the existence and uniqueness of a temporary equilibrium.

My analysis of questions the existence and uniqueness of the temporary equilibrium will be simplified by making the following assumption regarding the structure of iceberg trade costs in the model and using the implications of this assumption to rewrite the system of equations defining a temporary equilibrium.

Assumption 1 *Let D, the matrix of bilateral trade costs, be symmetric, i.e. $D_{i,j} = D_{j,i}$ for all regions i, j.*

Theorem 2 from Allen, Arkolakis, & Takahashi (2014) [16] (hereafter referred to as AAT) states that in a general equilibrium model yielding gravity equation of the form $X_{ij} = K_{ij}\gamma_i\delta_j$ where K is a "quasi-symmetric" $N \times N$ matrix, assuming balanced trade is equivalent to assuming that the origin and destination shifters of a region i are equal up to scale. That is

$$\gamma_i = \kappa\delta_i$$

To give intuition, I have included the proof of the special case of AAT Theorem 2 for symmetric trade frictions in Appendix B.

In my model,

$$\begin{aligned} X_{ijt} &= \pi_t(j,i) w_{jt} H_j L_{jt} \\ &= \left(\kappa_1^{-\theta}\kappa_2\mu^{-\theta}\right) D_{ji}^{-\theta} \left[\tau_{it} w_{it}^{-\theta} L_{it}^{\alpha-\theta(1-\mu)+\frac{\gamma_1\theta}{\xi}}\right] \left[\bar{u}_j H_j w_{jt}^{1+\theta} L_{jt}^{1-\lambda\theta} W_{jt}^{-\theta}\right] \end{aligned}$$

Thus, the origin- and destination-specific components are

$$\begin{aligned} \gamma_i &= \tau_{it} w_{it}^{-\theta} L_{it}^{\alpha-\theta(1-\mu)+\frac{\gamma_1\theta}{\xi}} \\ \delta_i &= \bar{u}_i H_i w_{it}^{1+\theta} L_{it}^{1-\lambda\theta} W_{it}^{-\theta} \end{aligned}$$

Applying AAT Theorem 2,

$$\bar{u}_i H_i w_{it}^{1+\theta} L_{it}^{1-\lambda\theta} W_{it}^{-\theta} = \kappa\tau_{it} w_{it}^{-\theta} L_{it}^{\alpha-\theta(1-\mu)+\frac{\gamma_1\theta}{\xi}}$$

Solving this equation for the wage gives

$$w_{it} = \left(\kappa\left(\bar{u}_i H_i\right)^{-1} \tau_{it} L_{it}^{-1+\lambda\theta+\alpha-\theta(1-\mu)+\frac{\gamma_1\theta}{\xi}} W_{it}^{-\theta}\right)^{\frac{1}{1+2\theta}}$$

This gives an explicit expression for equilibrium wages in terms of fundamentals and other equilibrium objects. I can substitute this expression into the price equations, thus reducing the temporary equilibrium system to a system of $3N$ equations in $3N$ unknowns. Notice that κ cancels, so the price equations and trade balance condition in the temporary equilibrium

system are replaced by

$$L_{it}^{\tilde{\theta}\left(1-\lambda\theta-\alpha+\theta(1-\mu)-\frac{\gamma_1\theta}{\xi}\right)+\lambda\theta} W_{it}^{\theta\left(1+\tilde{\theta}\right)} \tag{28}$$
$$= \kappa_2\kappa_1^{-\theta}\sum_{j=1}^{N} D_{ji}^{-\theta}\left(\bar{u}_i^{\theta-\tilde{\theta}} H_i^{-\tilde{\theta}}\tau_{it}^{\tilde{\theta}}\right)\left(\tau_{jt}^{1-\tilde{\theta}}\bar{u}_j^{\tilde{\theta}} H_j^{\tilde{\theta}} L_{jt}^{\tilde{\theta}(1-\lambda\theta)+(1-\tilde{\theta})(\alpha-\theta(1-\mu)+\frac{\theta\gamma_1}{\xi})} W_{jt}^{\tilde{\theta}\theta}\right)$$

where $\tilde{\theta} = \frac{\theta}{1+2\theta}$.

Having rewritten and combined two of the conditions defining a temporary equilibrium, I am prepared to prove that there exists a unique, strictly positive temporary equilibrium. To do so, I will leverage a theorem from Allen, Arkolakis, & Li (2015) [17] (hereafter referred to as AAL). Theorem 1 from AAL gives conditions for the existence and uniqueness of solutions to a system of equations of the form

$$\prod_{l=1}^{K}\left(x_i^l\right)^{\gamma_{kl}} = \lambda_k \sum_{j=1}^{N} F_{ij}^k \prod_{l=1}^{K}\left(x_j^l\right)^{\beta_{kl}} \qquad \forall k \in \{1,\ldots,K\}$$

where

- $i, j \in \{1,\ldots,N\}$ are regions
- $k, l \in \{1,\ldots,K\}$ are the number of equilibrium equations/variables
- $F_{ij}^k \geq 0$, β_{kl}, γ_{kl} are exogenous
- x_i^l is equilibrium variable l in region i.

Denote $\mathbf{\Gamma}$ and $\mathbf{B}$ as the $K \times K$ matrices with elements $(\mathbf{\Gamma})_{kl} = \gamma_{kl}$ and $(\mathbf{B})_{kl} = \beta_{kl}$, respectively. Assuming that $\mathbf{\Gamma}$ is non-singular, define $\mathbf{A} = \mathbf{B}\mathbf{\Gamma}^{-1}$, and let $\boldsymbol{A^p}$ be the matrix constructed from the absolute values of the elements of $\mathbf{A}$, i.e. $(\boldsymbol{A^p})_{kl} = |\mathbf{A}_{kl}|$. Furthermore, let $\rho(\boldsymbol{A^p})$ be the spectral radius of $\boldsymbol{A^p}$. The theorem states that

1. If $F_{ij}^k > 0$ for all i, j, k, then there exists a strictly positive solution.
2. If $\rho(\boldsymbol{A^p}) \leq 1$ and $F_{ij}^k \geq 0$ for all i, j, k, then there is at most one strictly positive solution.
3. If $\rho(\boldsymbol{A^p}) < 1$ and $F_{ij}^k > 0$ for all i, j, k, then the unique solution can be computed by a simple iterative process.

In my model, $K = 3$ and the set of variables in the temporary equilibrium system is $\{L_{it}, W_{it}, \Pi_{it}\}_{i=1}^N$.

Existence Using the theorem above, I construct $F_{ij} \in \mathbb{R}^3$:

$$F_{ij} = \begin{pmatrix} M_{i,j}^{-1/\Omega} \\ \\ \frac{H_j}{H_i} M_{i,j}^{-1/\Omega} L_{j,t-1} \\ \\ \bar{u}_i^{\theta-\tilde{\theta}} H_i^{-\tilde{\theta}} \tau_{it}^{\tilde{\theta}} D_{ji}^{-\theta} \tau_{jt}^{1-\tilde{\theta}} \bar{u}_j^{\tilde{\theta}} H_j^{\tilde{\theta}} \end{pmatrix}$$

Recall that $\bar{u}_i, H_i, M_{ij}, D_{ij} > 0$ by assumption. Then $F_{ij}^k > 0$ for all $i, j = 1, 2, \ldots, N$ and $k = 1, 2, 3$. Given positive, finite vectors τ_t and L_{t-1}, there exists a strictly positive solution to this system, i.e., there exists a strictly positive temporary equilibrium.

Uniqueness Using the Theorem above, I construct matrices $\mathbf{B}, \mathbf{\Gamma} \in \mathbb{R}^{3\times 3}$:

$$\mathbf{B} = \begin{pmatrix} 0 & \frac{1}{\Omega} & 0 \\ 0 & 0 & -\frac{1}{\Omega} \\ \tilde{\theta}\left(1-\lambda\theta\right) + \left(1-\tilde{\theta}\right)\left(\alpha - \theta(1-\mu) + \frac{\theta\gamma_1}{\xi}\right) & \theta\tilde{\theta} & 0 \end{pmatrix}$$

$$\mathbf{\Gamma} = \begin{pmatrix} 0 & 0 & \frac{1}{\Omega} \\ 1 & -\frac{1}{\Omega} & 0 \\ \tilde{\theta}\left(1-\lambda\theta-\alpha+\theta(1-\mu)-\frac{\theta\gamma_1}{\xi}\right) + \lambda\theta & \theta(1+\tilde{\theta}) & 0 \end{pmatrix}$$

The parameters values I use in simulation are given in Table 1. For these values, $\rho\left(\mathbf{B}\mathbf{\Gamma}^{-1}\right) = 0.7557 < 1$. Thus, for given initial levels of the state variables, there is a unique strictly positive dynamic equilibrium.

4.2 Balanced growth path

In this subsection, I examine the balanced growth path properties of the model. This is valuable insofar as it enables me to draw conclusions about how changing the initial state of the economy (as in the case of my deportation experiments) affects the long-run behavior of the economy.

In what follows, I derive conditions for existence and uniqueness of a balanced growth path, using the argument presented in Appendix B.5 of DNRH as a guide. As innovation in my model is the same as in DNRH, the first part of my derivation follows DNRH quite closely. However, the fact that the population distribution in my model exhibits path dependence complicates the derivation.

From (8), the growth rate of τ_{it} can be written as

$$\frac{\tau_{i,t+1}}{\tau_{it}} = \phi_{it}^{\theta\gamma_1} \left[\sum_j \frac{1}{N} \frac{\tau_{jt}}{\tau_{it}} \right]^{1-\gamma_2}$$

From this equation, it is clear that the growth rate of τ_{it} will be constant across time and space only if $\frac{\tau_{it}}{\tau_{jt}}$ is constant across time. From this observation, it follows that the growth rate of τ_{it} must be constant across space along the BGP. Thus, for all i, j,

$$1 = \frac{\tau_{i,t+1}/\tau_{it}}{\tau_{j,t+1}/\tau_{jt}} = \left(\frac{\phi_{it}}{\phi_{jt}} \right)^{\theta\gamma_1} \left(\frac{\tau_{it}}{\tau_{jt}} \right)^{1-\gamma_2}$$

It follows that

$$\frac{\tau_{it}}{\tau_{jt}} = \left(\frac{\phi_{it}}{\phi_{jt}} \right)^{\frac{\theta\gamma_1}{1-\gamma_2}}$$

Then, from (26),

$$\frac{\tau_{it}}{\tau_{jt}} = \left(\frac{L_i}{L_j} \right)^{\frac{\theta\gamma_1}{\xi(1-\gamma_2)}}$$

Note that the time subscripts on L are dropped, acknowledging the fact that the population distribution is steady along the BGP. Then, from labor market clearing, I can derive an expression for τ_i in terms of L_i. The total labor supply is

$$\begin{aligned} \bar{L} &= \sum_{j=1}^{N} H_j L_j \\ &= L_i \sum_{j=1}^{N} \left(\frac{\tau_{it}}{\tau_{jt}} \right)^{\frac{\theta\gamma_1}{\xi(1-\gamma_2)}} H_j \end{aligned}$$

and rearranging gives

$$\tau_{it} = \kappa_t L_i^{\frac{\xi(1-\gamma_2)}{\theta\gamma_1}} \tag{29}$$

where $\kappa_t = \left(\bar{L}^{-1} \sum_j H_j \tau_{jt}^{\frac{\theta\gamma_1}{\xi(1-\gamma_2)}} \right)^{\frac{\xi(1-\gamma_2)}{\theta\gamma_1}}$.

This equation shows that, along the BGP, τ_i, the level of technology in region i, is a power

function of L_i, the population density of region i. As the exponent on L_i is positive, equation (29) implies that, in the very long run, more densely populated regions will have higher levels of technology and will exhibit higher productivity. This "market size" effect, the tendency toward investment in technologies that augment the productivity of a factor that is abundant, is well documented in the growth literature, and is explored by Acemoglu (2002) [18]. From this expression, one can also see that, for a given BGP population vector L, the extent to which the level of technology varies across space depends on the exponent $\frac{\xi(1-\gamma_2)}{\theta\gamma_1}$, with higher values of the exponent indicating greater technological heterogeneity. In particular, the exponent expression indicates that, all else equal, a more concave innovation production function or a lower elasticity of technology and productivity with respect to innovation, implies greater heterogeneity in technology.

Using equation (29), I derive a system of $3N$ equations in $3N$ unknowns that define the balanced growth path. As in the temporary equilibrium system, the set of unknowns is $\{L_i, W_{it}, \Pi_{it}\}_{i=1}^N$.

1. I substitute (29) into (28):

$$L_i^{\tilde{\theta}\left(1-\lambda\theta-\alpha+\theta(1-\mu)-\frac{\gamma_1\theta}{\xi}\right)+\lambda\theta-\tilde{\theta}\frac{\xi(1-\gamma_2)}{\theta\gamma_1}} W_{it}^{\theta\left(1+\tilde{\theta}\right)} \tag{30}$$
$$= \kappa_t\kappa_2\kappa_1^{-\theta}\sum_{j=1}^{N} D_{ji}^{-\theta}\left(\bar{u}_i^{\theta-\tilde{\theta}}H_i^{-\tilde{\theta}}\right)\left(\bar{u}_j^{\tilde{\theta}}H_j^{\tilde{\theta}}L_j^{\tilde{\theta}(1-\lambda\theta)+(1-\tilde{\theta})(\alpha-\theta(1-\mu)+\frac{\theta\gamma_1}{\xi})+(1-\tilde{\theta})\frac{\xi(1-\gamma_2)}{\theta\gamma_1}} W_{jt}^{\tilde{\theta}\theta}\right)$$

2. Dropping the time subscripts from variables that are constant along the BGP, (6) becomes:

$$L_i W_{it}^{-1/\Omega} = \sum_{j=1}^{N} \frac{H_j}{H_i} M_{j,i}^{-1/\Omega}\Pi_{jt}^{-1/\Omega} L_j \tag{31}$$

3. Equation (4) remains unchanged:

$$\Pi_{it}^{1/\Omega} = \sum_{j=1}^{N} \left(W_{jt}/M_{ij}\right)^{1/\Omega} \tag{32}$$

Notice that this system is in the general AAL form with $K = 3$. Thus, I can apply AAL Theorem 1:

Existence

$$F_{ij} = \begin{pmatrix} \bar{u}_i^{\theta-\tilde{\theta}} H_i^{-\tilde{\theta}} D_{ji}^{-\theta} \bar{u}_j^{\tilde{\theta}} H_j^{\tilde{\theta}} \\ \frac{H_j}{H_i} M_{j,i}^{-1/\Omega} \\ M_{j,i}^{-1/\Omega} \end{pmatrix} > 0$$

Thus, by AAL Theorem 1, there exists a strictly positive solution to this system. That is, a balanced growth path exists.

Uniqueness

$$\mathbf{B} = \begin{pmatrix} \tilde{\theta}(1-\lambda\theta) + (1-\tilde{\theta})(\alpha - \theta(1-\mu) + \frac{\theta\gamma_1}{\xi}) + (1-\tilde{\theta})\frac{\xi(1-\gamma_2)}{\theta\gamma_1} & \tilde{\theta}\theta & 0 \\ 1 & 0 & -1/\Omega \\ 0 & 1/\Omega & 0 \end{pmatrix}$$

$$\mathbf{\Gamma} = \begin{pmatrix} \tilde{\theta}\left(1 - \lambda\theta - \alpha + \theta(1-\mu) - \frac{\gamma_1\theta}{\xi}\right) + \lambda\theta - \tilde{\theta}\frac{\xi(1-\gamma_2)}{\theta\gamma_1} & \theta(1+\tilde{\theta}) & 0 \\ 1 & -1/\Omega & 0 \\ 0 & 0 & 1/\Omega \end{pmatrix}$$

For the parameter values used in my simulations (listed in Table 1), I find that $\rho\left(\mathbf{A^P}\right) = 1.2137 > 1$, with $\mathbf{A^P}$ defined as in Section 4.1. Thus, for the chosen parameter values, there may be multiple balanced growth paths. This, combined with the result that the dynamic equilibrium is unique for a given τ_0 and L_0, has interesting implications. For any given initial state of the economy, there is a unique path of temporary equilibria which converges to a BGP. However, since the model may have multiplicity of BGP, the long-run behavior of the model could depend on this initial state. This suggests that exogenously changing the population distribution, as occurs when a policy of deportation is implemented, may have effects that persist into the very long run.

Note that the strategy employed here to study the BGP properties of the model is similar to that used by DNRH. However, due to the further tractability that results from DNRH Assumption 1, they are able to find a single integral equation which defines the BGP. Thus, applying AAL Theorem 1, they derive a single explicit condition on parameters for the uniqueness of BGP. Furthermore, for their choice of parameters, this condition is satisfied, so DNRH are able to prove that their model features a unique balanced growth path.

5 Calibration & Simulation

Regions and time periods To use this model to analyze US immigration policy, I consider a world comprising $N = 51$ regions, namely the 50 US states, plus Mexico. I set H_i to be the land area of each region, which I obtain from the US Census Bureau and World Bank.

To justify the generational migration structure featured in the model, I let each period have the interpretation of 20 years, an approximation of the length of one generation. I let $t = 0$ correspond to the year 2015. The calibration procedures outlined below require data from both $t = -1$ and $t = 0$, so I use population, wage, and subjective well-being data for the years 1995 and 2015.

Wage and population data I obtain initial wages and populations directly from the data. I set w_{i0} to 2015 GDP per capita in region i, obtained from the BEA and World Bank. Similarly, I set L_{i0} to the 2015 population density of region i, obtained from estimates by the US Census Bureau and World Bank. In calibrating migration costs (see below), I also employ the corresponding population density data for the year 1995.

Amenities, local productivity, and subjective well-being DNRH propose a method for backing out the exogenous local amenity using subjective well-being survey data. I outline a very similar method for backing out $\bar{u}_i$ in this model. I take state and country-level data from the results of a Gallup-Healthways survey and the Gallup World Poll in which participants were asked to rate their level of happiness using a "Cantril ladder" scale, on which responses of 0 and 10 represent, respectively, the lowest and highest levels of life satisfaction that the respondent can imagine. However, this subjective measure must be transformed into a cardinal measure of utility. Let W_i^{Cantril} denote the Catril well-being measure for region i. See Appendix C for details on the utility transformation.

This transformation results in an estimate of initial regional welfare: $W_{i0} = e^{1.8 \times W_{i0}^{\text{Cantril}}}$. Then, inverting equation (1) gives the exogenous local amenity as a function of local wage, population density, welfare, and prices:

$$\bar{u}_i = \mu w_{i0}^{-1} L_{i0}^{\lambda} W_{i0} P_{i0} \tag{33}$$

However, as the price index, P_{i0}, depends on the initial distribution of technology, τ_0, technology and amenities must be backed out together using the iterative procedure below.

The vectors w_0, L_0, and W_0 are now known. From equation (4), these imply Π_0 for a given matrix of migration costs M. Thus, all that remains is to find vectors τ_0 and P_0 such that, $\{w_0, L_0, W_0, \Pi_0\}$ form a temporary equilibrium. This can be accomplished by the following iterative process:

1. Set $\phi_{i0} = \left(\frac{\gamma_1}{\gamma_1+\mu\xi}\frac{L_{i0}}{\nu}\right)^{1/\xi}$.

2. Set $\text{mc}_{i0} = \left(\frac{1}{\mu}\right)^{\mu}\left(\frac{\nu\xi}{\gamma_1}\right)^{1-\mu}\phi_{i0}^{\xi(\mu-1)-\gamma_1}$

3. Guess τ_0.

4. Iterate over the following two steps until the process converges:

   (a) Update P_0:

$$P_{i0} := \left(\kappa_1\left[\frac{1}{\mu}\right]^{\mu}\left[\frac{\nu\xi}{\gamma_1}\right]^{1-\mu}\right)\left[\sum_{j=1}^{N}\tau_{j0}L_{r0}^{\alpha}\left(\text{mc}_{j0}D_{ij}\right)^{-\theta}\right]^{-\frac{1}{\theta}} \tag{34}$$

   (b) Use inversion of (23) to update τ_0:

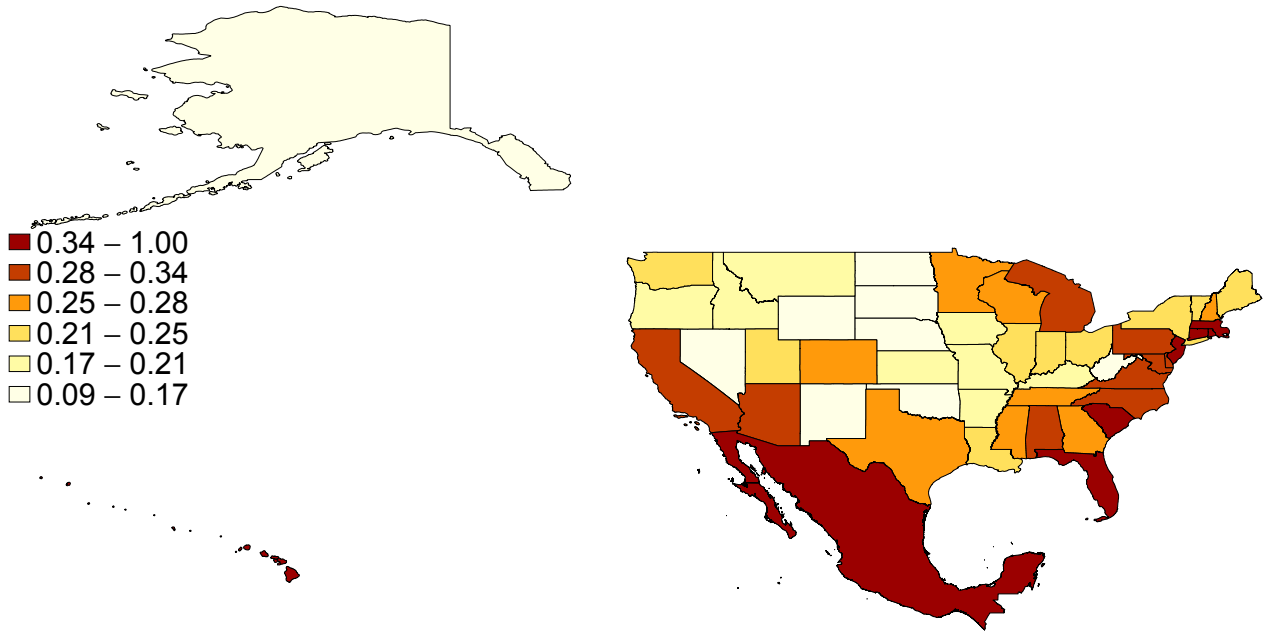
$$\tau_{i0} := \left(\kappa_1\left[\frac{1}{\mu}\right]^{\mu}\left[\frac{\nu\xi}{\gamma_1}\right]^{1-\mu}\right)^{\theta} H_i\left(\phi_{i0}\right)^{\theta(1-\mu)\xi-\theta\gamma_1} w_{i0}^{1+\theta}L_{i0}^{1-\alpha}\left[\sum_{j=1}^{N}\left(D_{j,i}\right)^{-\theta}\left(P_{j0}\right)^{\theta}\right]^{-1} \tag{35}$$

5. Then, use (33) to calculate local amenities: $\bar{u}_i := \mu w_{i0}^{-1}L_{i0}^{\lambda}W_{i0}P_{i0}$ (See results in Figure 2).

Note that the system (34) and (35) has exactly one positive solution though it is not guaranteed that a simple iterative process will converge to the solution. See proof in Appendix D. In practice, I find that I can reach a solution to this system iteratively.

After completing this process, I have initial vectors for all of the endogenous objects: $\{L_0, w_0, \tau_0, W_0, \Pi_0\}$. Thus, after calibrating migration costs, I will be able to simulate the model forward.

Figure 2: Calibration results: $\bar{u}_i$, exogenous amenity (Mexico = 1)



[5]

Migration costs Recall that, from the structure of $M_{i,j}$ which I specified in Section 3.1, calibrating migration costs requires calibrating three parameters: δ_1, δ_2, and δ_3.

From the data, L_{-1} is known, along with w_0, L_0, W_0. Recall that for all $i = 1, \ldots, N$,

$$L_{i0} W_{i0}^{-1/\Omega} = \sum_{j=1}^{N} \frac{M_{j,i}^{-1/\Omega}}{\sum_{k=1}^{N} \left(W_{k0}/M_{ik}\right)^{1/\Omega}} L_{j,-1}$$

From equation (3), the measure of people who move from i to j between periods -1 and 0 is

$$\ell_{ij,0} L_{i,-1} H_i = \left(\frac{W_{j0}/M_{ij}}{\Pi_{i0}}\right)^{1/\Omega} L_{i,-1} H_i$$

Let $\mathcal{U}$ be the set of regions in the US. I calibrate $\delta_1, \delta_2, \delta_3$ using an iterative procedure on the three equations below (note that iteration is required since Π_0 depends on δ_1, δ_2, and δ_3).

Suppose $G^{DATA}_{\text{interstate}}$ is the observed gross interstate migration flow between 1995 and 2015,

[5]For Figures 2-4, GIS shape files were obtained from the US Census-TIGER and DIVA-GIS.

and $G^{DATA}_{\text{US-Mex}}$ and $G^{DATA}_{\text{Mex-US}}$ are the corresponding flows from the US to Mexico and from Mexico to the US, respectively. Then, I calibrate δ_1, δ_2, and δ_3 by finding a solution to

$$\delta_1 = \left[\left(G^{DATA}_{\text{Interstate}}\right)^{-1} \sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{U}, j \neq i} \left(\frac{W_{j0}}{\Pi_{i0}}\right)^{1/\Omega} L_{i,-1} H_i \right]^{\Omega}$$

$$\delta_2 = W_{\text{Mex},0} \left[\left(G^{DATA}_{\text{US-Mex}}\right)^{-1} \sum_{i \in \mathcal{U}} \Pi_{i,0}^{-1/\Omega} L_{i,-1} H_i \right]^{\Omega}$$

$$\delta_3 = \left(\Pi_{\text{Mex},0}\right)^{-1} \left[\left(G^{DATA}_{\text{Mex-US}}\right)^{-1} L_{\text{Mex},-1} H_{\text{Mex}} \sum_{j \in \mathcal{U}} W_{j,0}^{1/\Omega} \right]^{\Omega}$$

These three conditions are derived in Appendix E.

I take the following gross migration moments from the data:

- From Molloy, Smith, & Wozniak (JEP 2011) [19], I have the lifetime interstate migration rates for each decade. I seek to match the gross interstate migration rate moment 0.318, which is an approximation of the 1995-2015 value, obtained by linear interpolation. Let $G^{DATA}_{\text{interstate}} = 0.318 \times \frac{\text{US population 1995}}{\text{US population 1995+Mexico population 1995}}$.
- From Pew Hispanic's November 2015 report [12], I can approximate the gross flows between the US and Mexico for the period 1995-2015. I seek to match the following two gross migration rates: US-Mexico migration rate of 0.0153 and Mexico-US migration rate of 0.0769. Let $G^{DATA}_{\text{US to Mex}} = 0.0153 \times \frac{\text{US population 1995}}{\text{US population 1995+Mexico population 1995}}$ and $G^{DATA}_{\text{Mex to US}} = 0.0769 \times \frac{\text{Mexico population 1995}}{\text{US population 1995+Mexico population 1995}}$.

These calibrations result in the following migration cost estimates:

$$\delta = \begin{bmatrix} 11.0027 \\ 3.9408 \\ 44.4809 \end{bmatrix}$$

These results imply that, as one would expect, the relative cost of moving from Mexico to the US is extremely high. They also suggest large interstate migration frictions, a result consistent with findings of previous work in international trade. For example, Autor, Dorn & Hanson (2013) [20] demonstrate that rising Chinese import competition has had strong differential effects across local labor markets in the US, a result that we would likely not observe if labor could reallocate frictionlessly across states in response to such shocks.

While it is tempting to say that $\delta_1 \approx 11$ has the interpretation that an agent pays $\frac{10}{11}$ of her welfare to move from one state to another, this is incorrect. The agents who actually move from their home state i to destination state j pay less than this, as the people who actually move have large ϵ_j. A more accurate way to interpret this value is to say that the *average* agent born in state i would have to pay $\frac{10}{11}$ of her welfare if she chose to move to state j. But those who actually choose to move from i to j have large ϵ_j draws and thus face lower effective frictions.

Simulation Given initial vectors $\{L_0, w_0, \tau_0, W_0, \Pi_0\}$, simulating the model into the future (i.e. computing the dynamic equilibrium) is achieved by the following procedure:

1. Using equation (26), population distribution L_t implies the level of innovation ϕ_{it} undertaken in each region. Then, from (8), one can obtain τ_{t+1}, the spatial distribution of technology for period $t+1$.

2. Given τ_{t+1} and L_t, iterating on the equations of the temporary equilibrium system gives $\{L_{t+1}, W_{t+1}, \Pi_{t+1}\}$. As shown above, there is a unique positive solution to this system and a simple iterative process is guaranteed to converge to this solution.

3. Repeat to obtain the dynamic equilibrium for periods 1 through T.

Trade costs DNRH calculate trade costs using the fast-marching procedure suggested by Allen & Arkolakis (2013)[6]. I take advantage of the parameters estimated by Allen & Arkolakis, but use a greatly simplified functional form, in which D_{ij} is an exponential function of the distance between the centroid of i and the centroid of j (I obtain centroids of US states from Rogerson (2015)[21] and the centroid of Mexico from INEGI.)

Parameters Table 1 lists the parameter values used in simulations and indicates how each was assigned.

Table 1: Parameter values

	Values as in DRNH:
$\sigma = 4$	Elasticity of substitution (Bernard et al., 2003) [22]
$\lambda = 0.32$	Governs strength of dispersion force (DNRH 2017)[3]
$\Omega = 0.5$	Fréchet parameter for idiosyncratic location shocks (Monte et al., 2015) [23]
$m = 0.55$	Relation between income and Cantril well-being measure (Deaton & Stone, 2003)[24]
$\alpha = 0.06$	Productivity spillover parameter (Carlino, Chatterjee & Hunt, 2007)[25]
$\theta = 6.5$	Fréchet parameter for local productivity shock (Eaton & Kortum, 2002 [5]; Simonovska & Waugh, 2014 [26])
$\mu = 0.8$	Labor share in production (Greenwood, Hercowitz, & Krusell, 1997 [27]; Desmet & Rappaport, 2015 [28])
$\gamma_1 = 0.319$	Relation between innovation and same-period productivity (DNRH 2017)
$\gamma_2 = 0.993$	Parameter governing diffusion of technology across space (DNRH 2017)
$\xi = 125$	Parameter governing convexity of innovation function (Desmet & Rossi-Hansberg, 2015) [29]
	Values calibrated using this model:
	Growth: Chosen to match initial aggregate growth rate of 2% for US and Mexico
$\nu = 3.08e - 71$	Scale parameter for innovation production function
	Migration costs: Chosen to match gross migration moments for US and Mexico
$\delta_1 = 11.0027$	Iceberg migration cost associated with interstate migration
$\delta_2 = 3.9408$	Iceberg migration cost associated with US-to-Mexico migration
$\delta_3 = 44.4809$	Iceberg migration cost associated with Mexico-to-US migration

6 Results

In this section, I first present and discuss simulation results for a baseline scenario in which there is no deportation and, second, compare these baseline results with those obtained from simulations in which a proportion η of undocumented immigrants for each US state are deported to Mexico in period 0. This comparison allows me to draw conclusions about the effects of a policy of deporting undocumented immigrants from the United States.

6.1 Baseline simulation

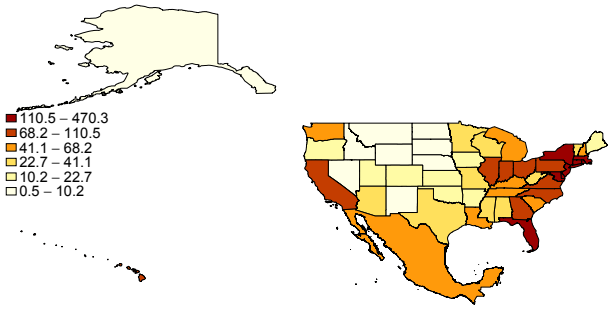
Figure 3 gives initial spatial distributions for population density, real wages, productivity, and welfare, which are calibrated or taken from the data as described in the previous section. Using the procedure above, I simulate the model forward 400 years (this corresponds to 20 periods). Outcomes for year 200 of this baseline simulation are presented in Figure 4. Scatterplots showing changes in the distributions of population and technology are presented in Figures 5 and 6. These figures illustrate that, between years 0 and 200, the baseline simulation predicts a significant spatial reallocation of labor and substantial changes in the distributions of real wages and productivity.

First, the convex shape the relation between initial population density and population density in year 200 (see Figure 5) indicates a tendency of labor to reallocate toward regions that already have a high population density. This feature of the simulation results is consistent with the trend of urbanization observed in historical data. Population data from the US Census Bureau shows that the percentage of Americans living in an urban area has increased steadily since the early 19th century. The mechanism underlying the tendency toward increasing concentrations of population in certain areas in this model differs somewhat from the explanation offered by other spatial models. In many models (including those of Fujita & Ogawa (1982)[30] and Ahlfeldt et al. (2015) [8]), this "urbanization" results from agglomeration economies, which incentivize workers and firms to concentrate their activities spatially in order to benefit from the productivity spillovers associated with spatial proximity. However, this cannot be the source of urbanization in my model because, as I discuss in Appendix A, the effective agglomeration force is negative in my model. This tendency of people to reallocate toward densely populated areas instead comes from the dynamic "market size" effect. All else equal, regions with a higher population density see larger investment in innovation and, thus, experience faster productivity growth. As I showed analytically in Section 4.2, this means that, in the long run, there is a strictly increasing relation between regional population density and productivity. This feature of the model is illustrated by Fig-

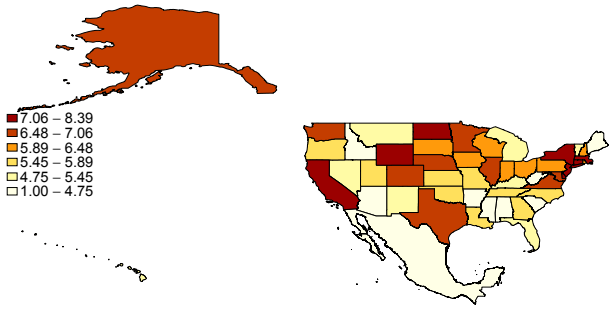
ure 7, which shows the correlation between log population density and log productivity over the course of the 400-year baseline simulation. As the economy moves toward the BGP over time, this correlation steadily approaches 1. This increasingly tight relationship between population density and productivity over the course of the simulation will have significant implications for the results of my deportation policy experiment. If higher population density means higher productivity, then a policy which exogenously reallocates labor across space has the potential to reshape the spatial distribution of productivity.

Second, these simulation results show a decrease in the heterogeneity of real wages and productivity. While Mexico still has lower real wages than each of the 50 states in year 200, it enjoys the fastest real wage growth of all 51 regions between the year 0 and year 200. The fact that Mexico enjoys more rapid progress than any US state in this simulation results from a "catch-up" effect. Mexico begins the simulation endowed with a far lower level of technology than other regions; in period $t = 0$, the average productivity of a worker in Mexico is just 23% of the average productivity of workers across the 50 US states. From equation (8), the technology evolution equation, one can see that the term capturing the diffusion of technology across space will, over time, result in τ_{Mex} converging toward the global mean level of technology. Regions that lag further behind get a larger boost from this technology diffusion effect, so Mexico experiences faster productivity growth than any other region. This catch-up effect is illustrated by Figure 6, which plots the initial productivity of each region (relative to New York, the region with highest initial productivity) versus its annualized productivity growth rate over the course of the 200-year simulation.

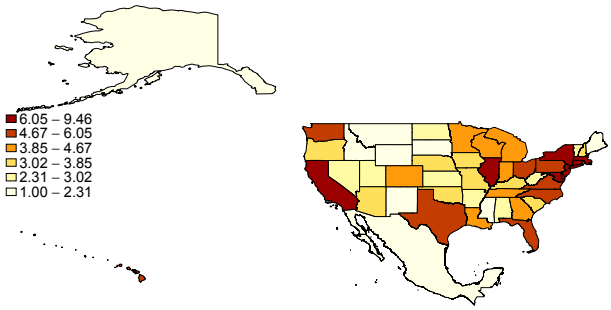
Figure 3: Calibration results: Year 0 ($t = 0$)



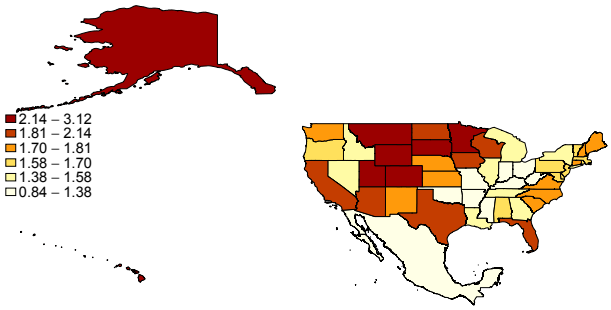
L_{i0}, population density (people / sq. km)



w_{i0}/P_{i0}, real wage (Mexico = 1)

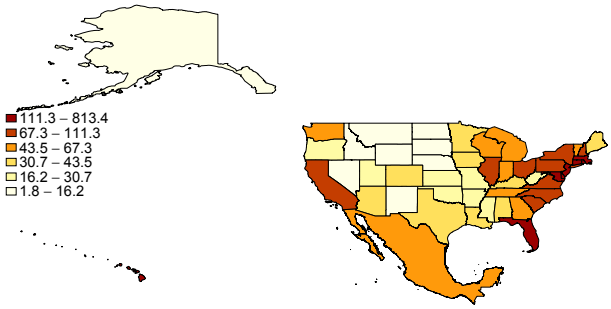


$(\tau_{i0} L_{i0}^{\alpha})^{1/\theta}$, productivity (Mexico = 1)

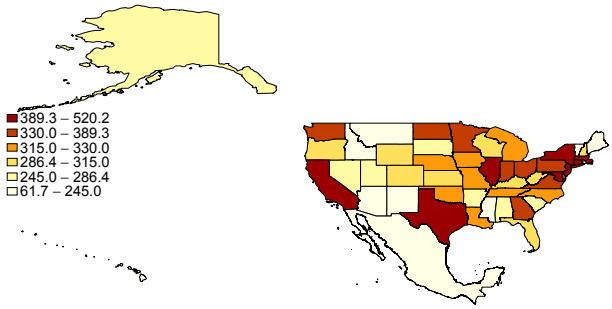


W_{i0}, welfare (Mexico = 1)

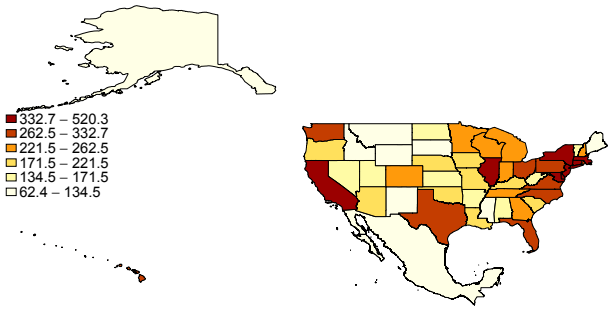
Figure 4: Baseline simulation results: Year 200 ($t = 10$)



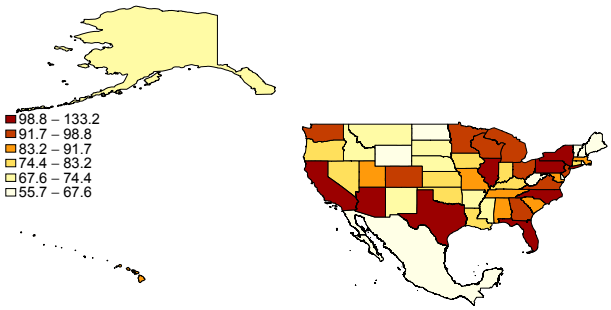
$L_{i,10}$, population density (people / sq. km)



$w_{i,10}/P_{i,10}$, real wage ($w_{\text{Mex},0}/P_{\text{Mex},0} = 1$)



$\left(\tau_{i,10} L_{i,10}^{\alpha}\right)^{1/\theta}$, productivity ($\text{prod}_{\text{Mex},0} = 1$)



$W_{i,10}$, welfare ($W_{\text{Mex},0} = 1$)

Figure 5: Labor reallocation and urbanization: Years 0 to 200

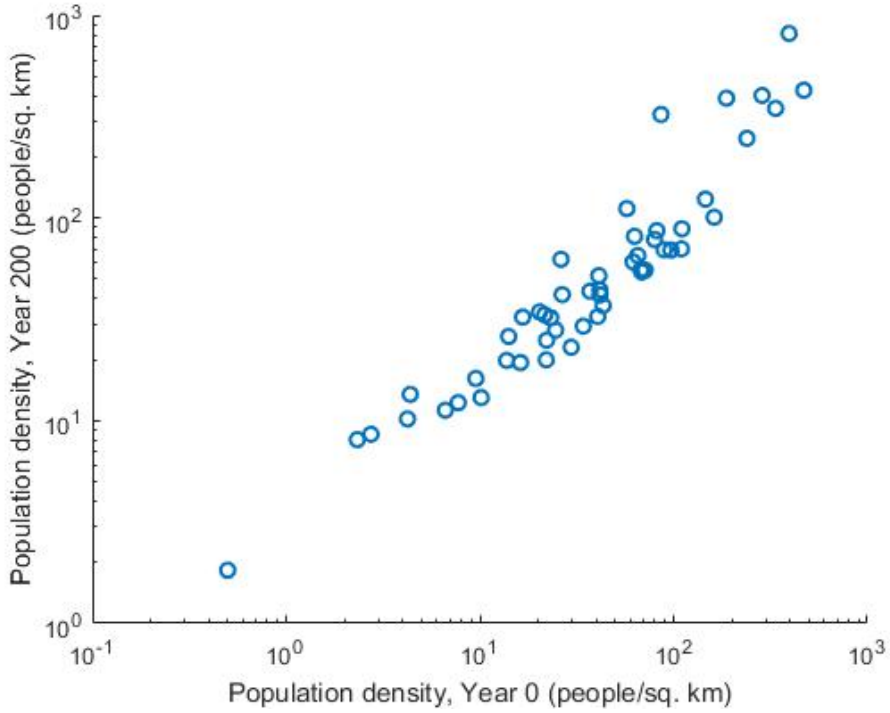


Figure 6: Productivity catch-up: Years 0 to 200

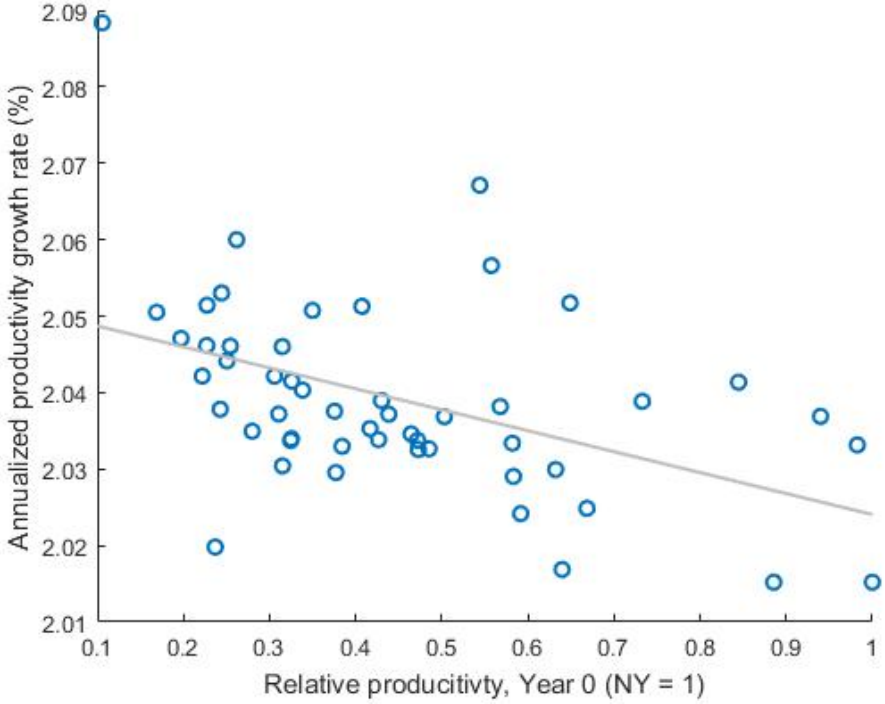
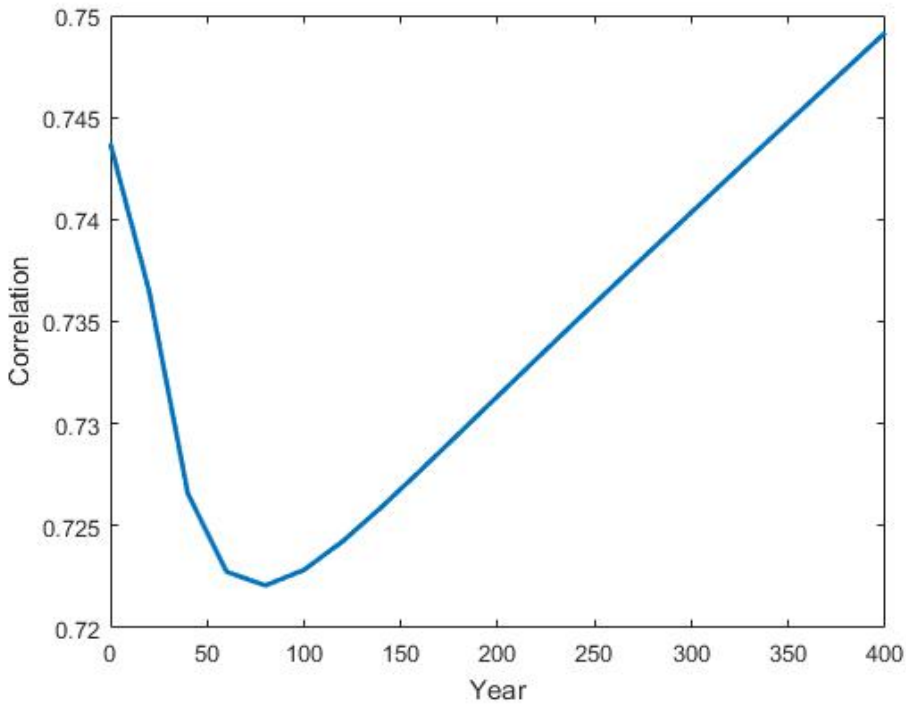


Figure 7: Correlation between log of population density and log of productivity: Years 0 to 400



6.2 Evaluating deportation policy

I study the effects of a policy of deporting undocumented immigrants from the US to Mexico by comparing the baseline simulation results above to those of counterfactual simulations in which the initial population distribution, L_0, is altered to reflect the effects of a policy which deports a proportion η of all undocumented immigrants from their current US states to Mexico. I conduct these experiments at various levels of "deportation intensity," with each scenario representing the deportation of, in the aggregate, 1 million, 2 million, 3 million, and 5 million undocumented immigrants, respectively. Recall that L_{i0}^U denotes the initial undocumented immigrants per unit land in region i, obtained from Pew Hispanic estimates.[15] In each deportation scenario, the post-deportation population density of region i is

$$L'_{i0} = \begin{cases} L_{i0} - \eta L_{i0}^U & \text{if } i \in \mathcal{U} \\ L_{i0} + \eta \sum\limits_{j \in \mathcal{U}} L_{j0}^U \frac{H_j}{H_i} & \text{if } i = \text{Mexico} \end{cases} \tag{36}$$

where $\eta = \frac{\text{Total number of deportees}}{\sum_{i \in \mathcal{U}} L_{i0}^U H_i}$.

Aggregate effects Figures 8 and 9 present aggregate outcomes from these experiments, showing the percent difference in country-level real GDP per capita in each scenario as compared with the baseline simulation. In the scenarios with most intensive deportation, Mexicans see significantly lower real incomes (as large as 0.9%) in the period following the deportation, while the US experiences a smaller boost in GDP p.c. (no larger than 0.23%). The effect of deportation on Mexican GDP p.c. shrinks over time, eventually turning positive, so that, after 360 years, average income in Mexico surpasses the level in the baseline simulation. Similarly, the gains in per capita GDP that the US experiences in the periods following deportation not only deteriorate over time, but also reverse sign in the long-run; after 280 years, Americans are, on average, worse off in real income terms as a result of a policy of deportation. This result is interesting from both a policy standpoint and a theoretical one. First, the result of opposite short- and long-run effects on US GDP p.c. adds a potential complication to the decision of a policymaker seeking to maximize the average incomes of US citizens, as, at some point, the positive wage effects of reduced competition in US labor markets are overwhelmed by the negative innovation effects of deportation. Second, the fact that the effects of the deportation policy do not simply revert to zero in the long-run suggests that, in altering the distribution of labor in period $t = 0$, the deportation changes the balanced growth path to which the economy converges. While I knew from my analysis in Section 4.2 that there was the potential for multiplicity of BGP, it was not obvious, *ex ante*, that changing the initial population distribution would change the BGP and thus affect outcomes in the very long run.

Figure 8: Effects of various intensities of deportation of US real GDP per capita.

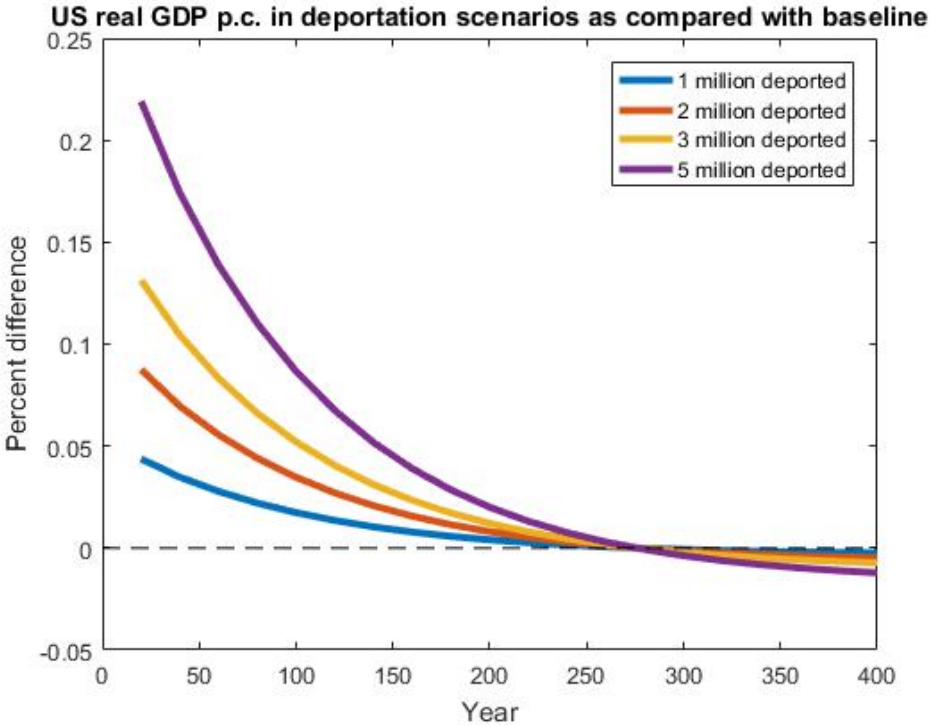
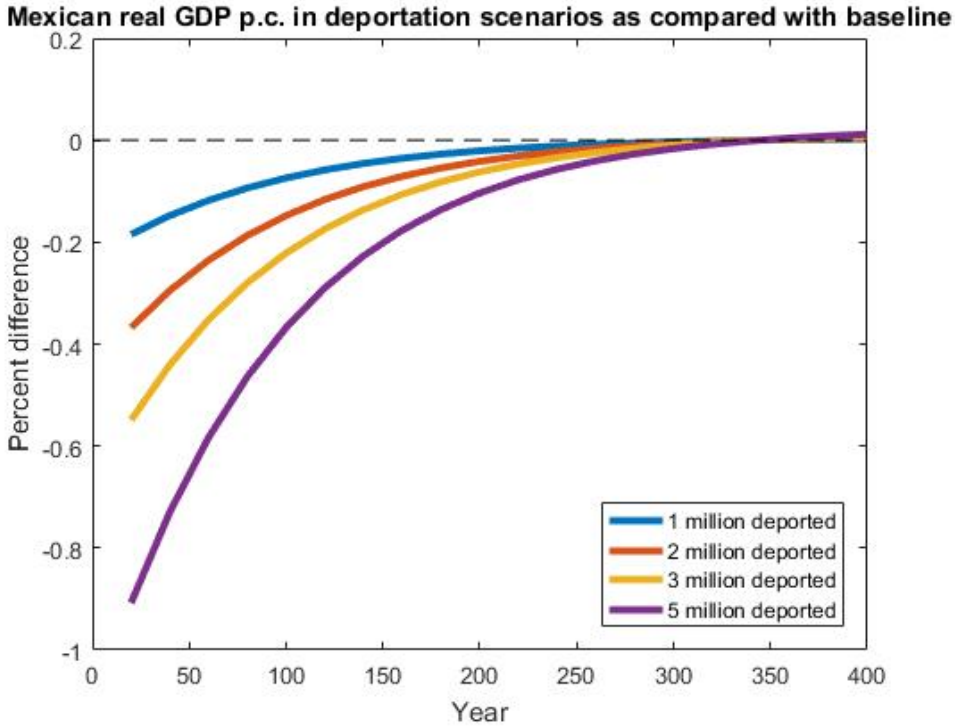


Figure 9: Effects of various intensities of deportation of Mexican real GDP per capita.



Local effects The fact that US states differ both in the proportion of undocumented immigrants in their populations and in other economic dimensions results in differential effects of deportation across the 50 states. Figure 10 shows the state-level effects of deportation (for the 3 million deportee scenario) on real wages, welfare, and population density over time. Similarly, Figure 11 presents a map displaying short-run and long-run effects on real wages across the 50 states. Figure 12 shows the corresponding maps for welfare. Unsurprisingly, the largest short-run increases in real wages as a result of deportation occur in states which have relatively large proportions of undocumented immigrants. Texas and California, the states with the largest proportion of undocumented immigrants, experience the most significant real wage increases (.3035% and .2867%, respectively, in the 3 million deportee scenario) as a result of deportation. After 400 years, the effects of deportation of real wages are, in percentage terms, much smaller and, interestingly, negative. All 50 states see lower real wages in the deportation scenario after 320 years. The corresponding analysis of welfare effects shows largely similar results, although the congestion term in the amenity means that Americans enjoy an additional benefit from decreased population density as a result of the deportation policy. In the period following deportation, each of the 50 states sees higher welfare, with increases ranging from 0.187% in Ohio to .455% in California, as compared with the scenario without deportation. In the long run, these welfare improvements deteriorate. In year 400, all states see welfare within 0.01% of the levels in the baseline simulation. Despite the fact that workers in most states see lower real wages than in the baseline simulations, only 21 states see lower welfare in year 400 of the deportation scenario, as agents benefit from persistently lower population densities across all states.

Figure 10: Percent difference in outcomes for each of the 50 US states, deportation simulation (3 million deportee scenario) as compared with the baseline scenario.

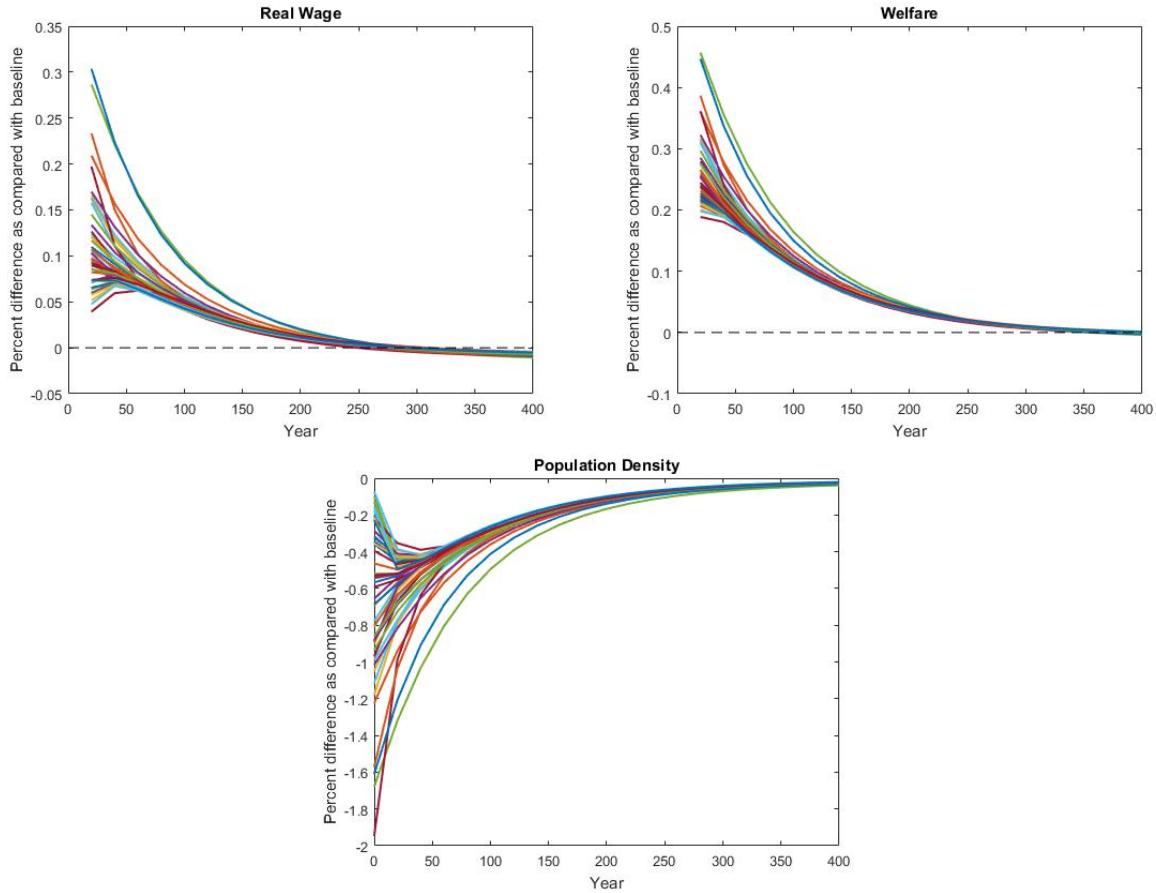


Figure 11: Percent difference in real wages for the 50 US states, deportation simulation (3 million deportee scenario) as compared with the baseline scenario. These maps present the distributions of effects across states for years 20 and 400.

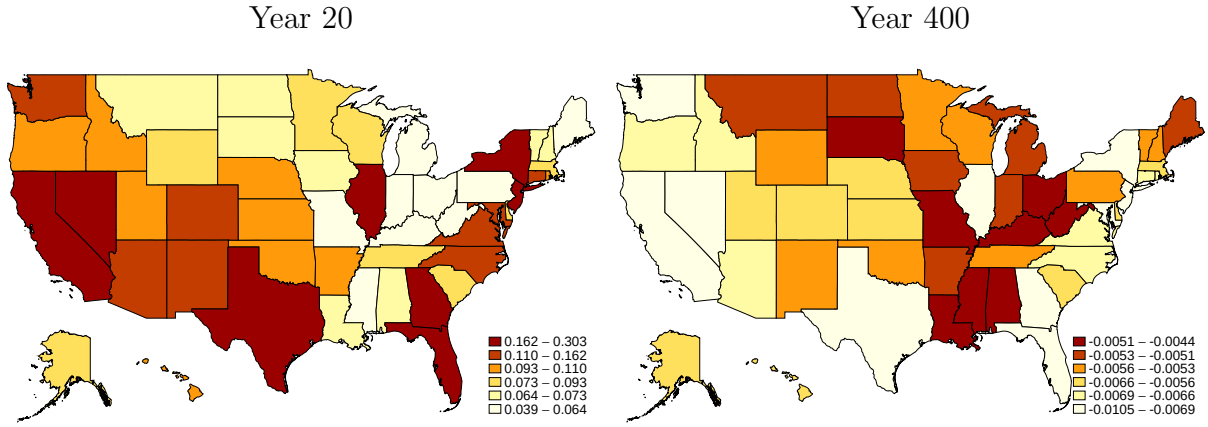
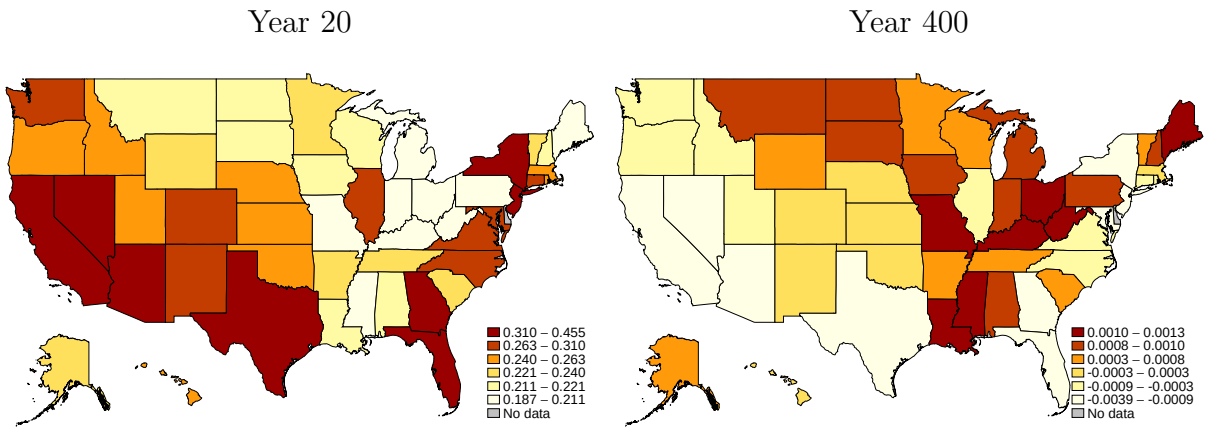


Figure 12: Percent difference in welfare for the 50 US states for deportation simulation (3 million deportee scenario) as compared with the baseline scenario. These maps give the distributions of effects across states for years 20 and 400.



7 Conclusion

In this paper, I have developed a dynamic spatial growth model which combines the technology and innovation structure of Desmet, Nagy & Rossi-Hansberg (2017) [3] with the generational migration structure of Allen & Donaldson (2017) [4]. I have explored the equilibrium and balanced growth path properties of this model and used it to conduct policy experiments studying the long-run economic effects of deporting undocumented immigrants from the US to Mexico, with the goal of evaluating the claim of the Trump campaign that US citizens would be better off if not for the presence of undocumented immigrants in the US labor force. Analyzing the results of these experiments, I find that US workers do experience small short-run gains in real wages and welfare as a result of the deportation policy, but that these gains deteriorate the long run. In the very long run, workers in all US states receive lower real wages as a result of the deportation.

It is worth considering the extent to which these conclusions depend on my modeling choices. Altering certain features of the model will not change the conclusions qualitatively. For instance, if one eliminates the dispersion force (i.e. sets $\lambda = 0$), then US workers enjoy slightly more persistent gains from the deportation policy, but still, after 400 years, receive lower wages on average than in the baseline scenario. However, the long-run results of the deportation are quite sensitive to changes in the strength of the productivity spillover (α); the parameters of the functions relating labor supply, innovation, and productivity (particularly γ_1 and ξ); and the relative factor input shares of land and labor (governed by μ). As I

explain in Appendix A, altering any one of these parameters changes the effective agglomeration force. Such a change not only affects the impact of deportation on productivity and real wages in the short-run, but also determines the development of the spatial distribution of labor and, therefore, technology in the long-run. To illustrate the importance of one of these features, I present in Appendix F results of the deportation policy experiment in the case where land is not a scarce factor of production (i.e. $\mu = 1$). In this experiment, American workers on average receive lower real wages not only in the long run, but also in the short run, as a result of deportation. This sensitivity of my results to certain parameter values and modeling choices indicates that a high level of precision is required in estimates of production functions and productivity spillovers in order to obtain highly credible predictions from such a model.

8 Limitations and further work

This paper highlights a number of opportunities for further work in this area are research. First, there are several ways in which the deportation experiment conducted using my model could be improved to offer more accurate predictions. In my experiments, treating Mexico as a single region represents a strong assumption, as the model has no frictions to the movement of goods, people, and technology within a region. One approach to resolving this issue would be to treat each of Mexico's 31 states as a separate region. However, in order to carry out the deportation experiment in a setting in which Mexico is broken into its 31 states, one would need more refined data. While population and GDP data for each Mexican state could be obtained from INEGI, additional data on both subjective well-being and undocumented immigrants would also be required. Gallup provides state-level subjective well-being data for the US but provides only a country-level measure for Mexico. A state-level Cantril measure would be necessary to calibrate amenities and initial technology levels for the Mexican states. In addition, if one were to treat each Mexican state as a separate region, carrying out deportation simulations would require taking a stance on which Mexican states undocumented immigrants would be deported to. The scarcity of detailed data on undocumented immigrants would make it challenging to find a reasonably accurate breakdown by state of origin of undocumented immigrants from Mexico living in the US.

A second shortcoming of the way in which the exercises presented in this paper are carried out is that, in considering only the US and Mexico, my experiments offer an incomplete picture of the effects of a policy of deporting undocumented immigrants from the United States. Though just over half of all undocumented immigrants living in the US are of Mexi-

can origin, millions more come from elsewhere in Latin America or from Asia, Europe, Africa, and the Middle East [31]. A more complete analysis of this policy problem would, therefore, expand the exercise to the entire world, allowing for results that describe the deportation policy's effects both on the complete set of countries to which undocumented immigrants would be deported and the dynamic effects on the United States through trade and migration.

In addition to improving the way in which the deportation policy experiment is carried out in the context of this model, further theoretical work could also lead to the development of a model that captures important dimensions of the immigration policy problem which are omitted from the model I have presented here. In particular, my model makes the strong assumptions that undocumented immigrants and legal residents/citizens are identical in preferences and productivity, and that immigrant and native labor are perfect substitutes. These assumptions could be relaxed in a number of ways. First, one could introduce a production function which treats the labor supplied by undocumented immigrants and the labor supplied by US citizens/legal residents as factors which are imperfect substitutes. This would enrich the model not only by capturing the fact that these two group differ, on average, in skills and human capital, but also by allowing for citizens and undocumented immigrants earning different wages. A second, related, improvement to the model could be achieved by introducing multiple sectors. In the data, the industry makeup of undocumented workers employed in the United States differs significantly from that of US-born workers [32]. Using this data, one could calibrate sector-specific production function parameters, allowing sectors to differ in their input elasticities for each of the two labor factors, as well as in the extent to which the two factors are substitutable. While formulating and calibrating such a model may be challenging, the result would be a framework that would offer much richer predictions for the effects of a deportation policy. Such a model could be used to not only more accurately predict the effects of deportation on American citizens, but also to analyze effects on undocumented immigrants remaining in the US, changes in sector-specific prices, and the impact on sectoral employment shares in the short- and long-run.

A Isomorphism with Allen-Donaldson model

The equilibrium system of the AD model consists of the following $4N$ equations in $4N$ unknowns: w_{it}, the wage; L_{it}, the population (not density); W_{it}, the welfare measure; and Π_{it}, the expected utility of an agent born in region i.

$$\Pi_{it}^{\theta} = \sum_{j} \mu_{ij}^{-\theta} W_{jt}^{\theta} \tag{37}$$

$$L_{it} W_{it}^{-\theta} = \sum_{j} -\mu_{ji}^{-\theta} \Pi_{jt}^{-\theta} L_{j,t-1} \tag{38}$$

$$w_{it}^{1-\sigma} L_{it}^{\beta(1-\sigma)} W_{it}^{\sigma-1} = \sum_{j} \left(\frac{\tau_{ij}}{\bar{A}_{jt} \bar{u}_{it}} \right)^{1-\sigma} L_{jt}^{\alpha(\sigma-1)} w_{jt}^{1-\sigma} \tag{39}$$

$$w_{it}^{\sigma} L_{it}^{1-\alpha(\sigma-1)} = \sum_{j} \left(\frac{\tau_{ij}}{\bar{A}_{it} \bar{u}_{jt}} \right)^{1-\sigma} L_{jt}^{\beta(\sigma-1)} W_{jt}^{1-\sigma} w_{jt}^{\sigma} L_{jt} \tag{40}$$

Table 2: "Cross-walk" for isomorphism

Allen-Donaldson	**My model**	
Simple notational differences:		
μ_{ij}	M_{ij}	Iceberg migration costs
τ_{ij}	D_{ij}	Iceberg trade costs
θ	$1/\Omega$	Migration elasticity
$\sigma - 1$	θ	Trade elasticity
β	$-\lambda$	Dispersion force
α	$\frac{\alpha}{\theta} - (1-\mu) + \frac{\gamma_1}{\xi}$	Productivity spillover
$\bar{A}_{it}$	$\kappa_2^{1/\theta} \kappa_1^{-1} \tau_{it}^{1/\theta}$	Exogenous productivity

Using Table 2, one can easily see that the AD equilibrium system comprising equations (37)-(40) can be rewritten as the temporary equilibrium system of my model presented on pages 16-17 (assuming $H_i = H_j$ for all i, j). While the first five rows of Table 2 simply

indicate changes in notation, the final two rows allow one to gain insight about both local productivity and the effective level of agglomeration forces in my model.

In the AD model, α is the parameter governing spillovers in productivity, with higher values of α indicating a stronger agglomeration force. In mapping my model into the AD structure, I find that the effective productivity spillover parameter in my model is $\frac{\alpha}{\theta} - (1-\mu) + \frac{\gamma_1}{\xi}$. This expression allows me to draw conclusions about the implicit agglomeration or dispersion forces generated by various features of my model. First, the subtraction of the $(1-\mu)$ term reflects the fact that firms located in more densely populated regions face higher rental rates, an effect which increases the marginal cost (effectively lowering the productivity) of firms located in congested areas. Thus, in my model, competitive local land markets act as a dispersion force. Second, the addition of the $\frac{\gamma_1}{\xi}$ term reflects the fact that firms located in densely populated regions invest in higher levels of innovation, boosting current-period productivity. Thus, innovation incentives in my model act as an agglomeration force. For the parameter values used in my simulations (see Table 1), $\frac{\alpha}{\theta} - (1-\mu) + \frac{\gamma_1}{\xi} = -0.1882$. Thus, the rent effect dominates, resulting in an effective dispersion, rather than agglomeration force in productivity spillover.

It is worth noting that while this effective productivity spillover is negative, my model still features a positive "market size" effect. The negative productivity spillover means that for a given level of technology, τ_{it}, an increase in the population of region i will lead to a decrease in region i's period-t productivity. However, the positive market size effect means that this increase in the population density of region i increases the incentive for firms in region i to invest in innovation that augments the productivity of labor. Thus, while an exogenous increase in population would lower region i's current-period productivity, it may lead to higher productivity in the long run. This tension between short- and long-run effects is consequential for the deportation policy experiments I perform. The market size effect is discussed further in Section 4.2 and is demonstrated in the baseline simulation results in Section 6.1.

This expression for the "effective" productivity spillover is also interesting in that it suggests, for a particular choice of parameter values, an empirical strategy for quantifying the strength of such spillovers. Consider the case with $\alpha = 0$ and $\mu = 1$ (i.e., labor is the only factor of production). Then, the effective productivity spillover is $\frac{\gamma_1}{\xi}$. Thus, in this case, estimating the innovation function (as is done in Desmet & Rossi-Hansberg (2015) [29]) suffices to determine the magnitude of the agglomeration force.

B Proof of AAT Theorem 2 (special case)

What follows is a proof of AAT Theorem 2 in the special case with symmetric (rather than just "quasi-symmetric") trade costs. This proof follows closely the one presented by AAT.

Assuming trade balance gives

$$Y_i = \sum_j X_{ij} = \sum_j X_{ji}$$

And, given the gravity structure of trade flows, we have

$$X_{ij} = M_{ij}\gamma_i\delta_j$$

Assume M is a *symmetric* matrix of trade costs. Then,

$$Y_i = \sum_j M_{ij}\gamma_i\delta_j = \gamma_i \sum_j M_{ij}\delta_j$$
$$Y_i = \sum_j M_{ji}\gamma_j\delta_i = \delta_i \sum_j M_{ij}\gamma_j$$

So,

$$1 = \frac{\delta_i \sum_j M_{ij}\gamma_j}{\gamma_i \sum_j M_{ij}\delta_j}$$

$$\frac{\gamma_i}{\delta_i} = \frac{\sum_j M_{ij}\gamma_j}{\sum_j M_{ij}\delta_j} = \sum_j \frac{M_{ij}\delta_j}{\sum_j M_{ij}\delta_j} \times \left(\frac{\gamma_j}{\delta_j}\right)$$

For any $\kappa > 0$, this equation is solved by $\gamma_i = \kappa\delta_i$. Using the Perron-Frobenius Theorem, the solution is unique up to a scale. □

C Conversion of Cantril well-being measure to initial welfare

Deaton and Stone (2013) [24] find the following relation between subjective well-being and income:

$$W_i^{k,\text{Cantril}} = m \ln y_i^k + v_i + \epsilon_{i,DS}^k \tag{41}$$

where v_i is a location fixed-effect.

Ignoring migration costs, the welfare of an individual k living in region i in the model is

$$W_{it}^k = u_{it} y_i \epsilon_i^k$$

Taking logs and multiplying through by m gives

$$m \ln W_{it}^k = m \ln u_{it} + m \ln y_i + m \ln \epsilon_i^k \tag{42}$$

Equations (41) and (42) imply that

$$W_{i0}^k = e^{\frac{1}{m} W_{i0}^{k,\text{Cantril}}}$$

Deaton & Stone estimate m to be 0.55.

D Solution to (34) and (35): Existence & uniqueness

I use AAL Theorem 1 to prove that the system (34) and (35) has exactly one positive solution.

I rewrite equation (34) as

$$P_{i0}^{-\theta} = \left(\kappa_1 \left[\frac{1}{\mu} \right]^{\mu} \left[\frac{\nu \xi}{\gamma_1} \right]^{1-\mu} \right)^{-\theta} \sum_{j=1}^{N} \left(D_{ij}^{-\theta} L_{i0}^{\alpha} \text{mc}_{j0}^{-\theta} \right) \tau_{j0}$$

and I rewrite (35) as

$$\tau_{i0}^{-1} = \left(\kappa_1 \left[\frac{1}{\mu}\right]^{\mu} \left[\frac{\nu\xi}{\gamma_1}\right]^{1-\mu}\right)^{-\theta} \sum_{j=1}^{N} \left(H_i^{-1} \left(\phi_{i0}\right)^{-\theta(1-\mu)\xi+\theta\gamma_1} w_{i0}^{-1-\theta} L_{i0}^{-1+\alpha} \left(D_{j,i}\right)^{-\theta}\right) \left(P_{j0}\right)^{\theta}$$

Note that this is a system of $2N$ equations in $2N$ unknowns (namely P_0 and τ_0), which is in the form specified by AAL Theorem 1.

$$F_{ij} = \begin{pmatrix} D_{ij}^{-\theta} L_{r0}^{\alpha} \mathrm{mc}_{j0}^{-\theta} \\ H_i^{-1} \left(\phi_{i0}\right)^{-\theta(1-\mu)\xi+\theta\gamma_1} w_{i0}^{-1-\theta} L_{i0}^{-1+\alpha} \left(D_{j,i}\right)^{-\theta} \end{pmatrix} > 0$$

Therefore, there exists a strictly positive solution to the system.

$$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ \theta & 0 \end{pmatrix}$$

$$\mathbf{\Gamma} = \begin{pmatrix} -\theta & 0 \\ 0 & -1 \end{pmatrix}$$

Since $\det \mathbf{\Gamma} = \theta > 0$, $\mathbf{\Gamma}$ is invertible. In particular,

$$\mathbf{\Gamma}^{-1} = \frac{1}{\theta} \begin{pmatrix} -1 & 0 \\ 0 & -\theta \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1/\theta \end{pmatrix}$$

$$\mathbf{A} = \mathbf{B}\mathbf{\Gamma}^{-1} = \begin{pmatrix} 0 & 1 \\ \theta & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1/\theta \end{pmatrix} = \begin{pmatrix} 0 & -1/\theta \\ -\theta & 0 \end{pmatrix}$$

$$\mathbf{A}^P = \begin{pmatrix} 0 & 1/\theta \\ \theta & 0 \end{pmatrix}$$

which has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$. So, $\rho\left(\mathbf{A}^P\right) = 1$. Thus, there exists at most one strictly positive solution (to scale). However, since $\rho\left(\mathbf{A}^P\right) \not< 1$, the Theorem does not guarantee the existence of an algorithm that converges uniformly to the solution. Note that this result is independent of the choice of θ.

E Derivation of gross migration moment expressions

From equation (3), the total measure of people who migrate from region i to region j between periods -1 and 0 is

$$\ell_{ij0} \times H_i L_{i,-1} = \left(\frac{W_{j0}/M_{ij}}{\Pi_{i0}} \right)^{1/\Omega} H_i L_{i,-1} \tag{43}$$

First, I find an expression for gross interstate migration in the model by summing this expression across all pairs of states $(i, j) \in \mathcal{U} \times \mathcal{U}$:

$$\begin{aligned} G^{\text{MODEL}}_{\text{Interstate}} &= \sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{U}, j \neq i} \left(\frac{W_{j0}/M_{ij}}{\Pi_{i0}} \right)^{1/\Omega} H_i L_{i,-1} \\ &= \delta_1^{-1/\Omega} \sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{U}, j \neq i} \left(\frac{W_{j0}}{\Pi_{i0}} \right)^{1/\Omega} H_i L_{i,-1} \end{aligned}$$

Setting this expression equal to gross interstate migration from the data and solving for δ_1 gives

$$\delta_1 = \left[\left(G^{DATA}_{\text{Interstate}} \right)^{-1} \sum_{i \in \mathcal{U}} \sum_{j \in \mathcal{U}, j \neq i} \left(\frac{W_{j0}}{\Pi_{i0}} \right)^{1/\Omega} L_{i,-1} H_i \right]^{\Omega} \tag{44}$$

I similarly derive an expression for gross migration from the US to Mexico by setting $j =$ Mexico and summing (43) across all states $i \in \mathcal{U}$:

$$\begin{aligned} G^{\text{MODEL}}_{\text{US-Mex}} &= \sum_{i \in \mathcal{U}} \left(\frac{W_{\text{Mex},0}/M_{i,\text{Mex}}}{\Pi_{i0}} \right)^{1/\Omega} H_i L_{i,-1} \\ &= \left(\frac{W_{\text{Mex},0}}{\delta_2} \right)^{1/\Omega} \sum_{i \in \mathcal{U}} \Pi_{i0}^{-1/\Omega} H_i L_{i,-1} \end{aligned}$$

Setting this expression equal to the gross US-to-Mexico migration moment from the data implies

$$\delta_2 = W_{Mex,0} \left[\left(G^{DATA}_{\text{US-Mex}} \right)^{-1} \sum_{i \in \mathcal{U}} \Pi_{i,0}^{-1/\Omega} L_{i,-1} H_i \right]^{\Omega} \tag{45}$$

Finally, I derive an expression for gross migration from Mexico to the US by setting $i =$

Mexico and summing across all states $j \in \mathcal{U}$:

$$\begin{aligned} G^{\text{MODEL}}_{\text{Mex-US}} &= \sum_{j \in \mathcal{U}} \left(\frac{W_{j,0}/M_{\text{Mex},j}}{\Pi_{\text{Mex},0}} \right)^{1/\Omega} L_{\text{Mex},-1} H_{\text{Mex}} \\ &= \left(\delta_3 \Pi_{\text{Mex},0}\right)^{-1/\Omega} H_{\text{Mex}} L_{\text{Mex},-1} \sum_{j \in \mathcal{U}} W_{j,0}^{1/\Omega} \end{aligned}$$

Setting this expression equal to the gross Mexico-to-US migration moment from the data and solving for δ_3 gives

$$\delta_3 = \left(\Pi_{\text{Mex},0}\right)^{-1} \left[\left(G^{DATA}_{\text{Mex-US}}\right)^{-1} L_{\text{Mex},-1} H_{\text{Mex}} \sum_{j \in \mathcal{U}} W_{j,0}^{1/\Omega} \right]^{\Omega} \tag{46}$$

Iterating over equations (44)-(46) leads to estimates for parameters $\delta_1, \delta_2, \delta_3$ (see Section 5).

F Deportation experiment results for the $\mu = 1$ case

In this section, I consider the case in which land is not a scarce factor (i.e. $\mu = 1$). I repeat the migration policy experiments from Section 6.2 using this specification and present results illustrating both the aggregate and local effects of deportation.

Figure 13 illustrates that this specification predicts aggregate effects of deportation for the US that are vastly different from those computed in the case in which land is a scarce factor (see Figure 8). Whereas deportation resulted in higher US GDP per capita for the first 200+ years in the original model specification, effects in the $\mu = 1$ case are negative and very persistent. The stark difference in short-run effects in the two specifications of the model results from the fact that when land is no longer a scarce factor, the negative effective productivity spillover that resulted from high rental rates in densely populated areas is eliminated. Thus, in the $\mu = 1$ case, the effective productivity spillover is positive, so when undocumented immigrants are exogenously moved from the US, this results in an immediate decrease in US productivity. The corresponding aggregate effects for Mexico, illustrated in Figure 14, are not very different qualitatively from those in the original specification (compare with Figure 9), but do differ in the magnitude of effects. While Mexicans experienced a decrease in real income as large as 0.9% in the period following deportation in the original specification, these losses are not nearly so large are in the $\mu = 1$ case. Even in the most extreme deportation scenario I consider, Mexico now sees just a 0.37% decrease in real GDP p.c. as a result of the deportation policy. In addition, these losses are recovered more quickly than in the original

specification.

These substantial differences between the simulated effects of deportation in the cases with $\mu = 0.8$ and $\mu = 1$ underline the fact that the results of these experiments are somewhat sensitive to model specifications and parameter values. This demonstrates that great care should be taken when specifying features and choosing parameters for a quantitative spatial model.

Figure 13: Effects of various intensities of deportation on US real GDP per capita. Specification in which land is not a scarce factor (i.e. $\mu = 1$).

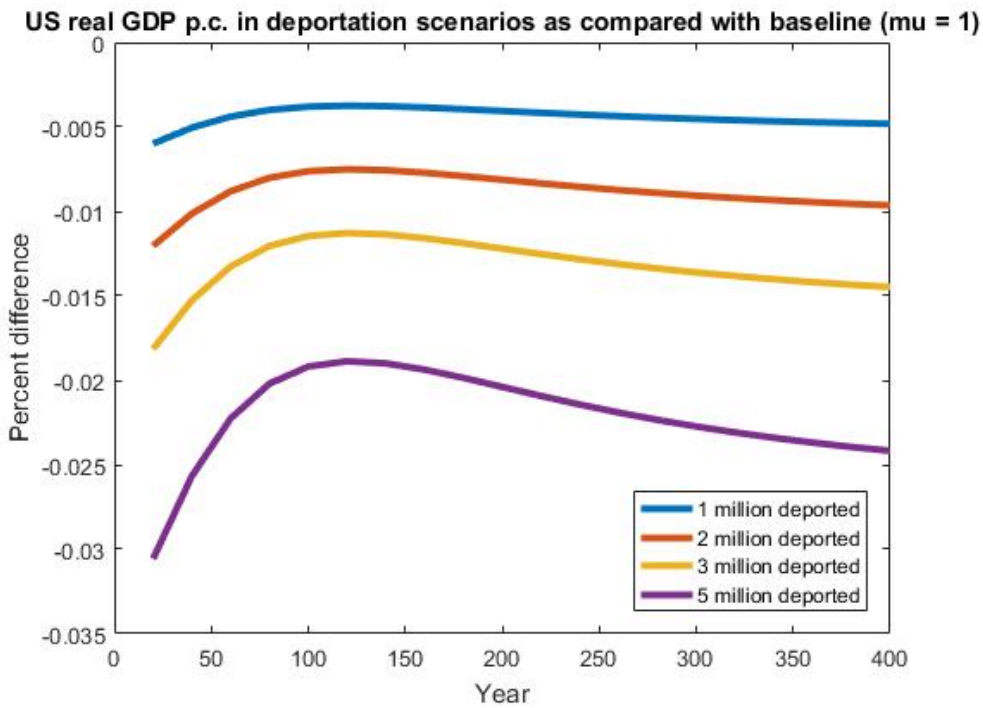


Figure 14: Effects of various intensities of deportation on Mexican real GDP per capita. Specification in which land is not a scarce factor (i.e. $\mu = 1$).

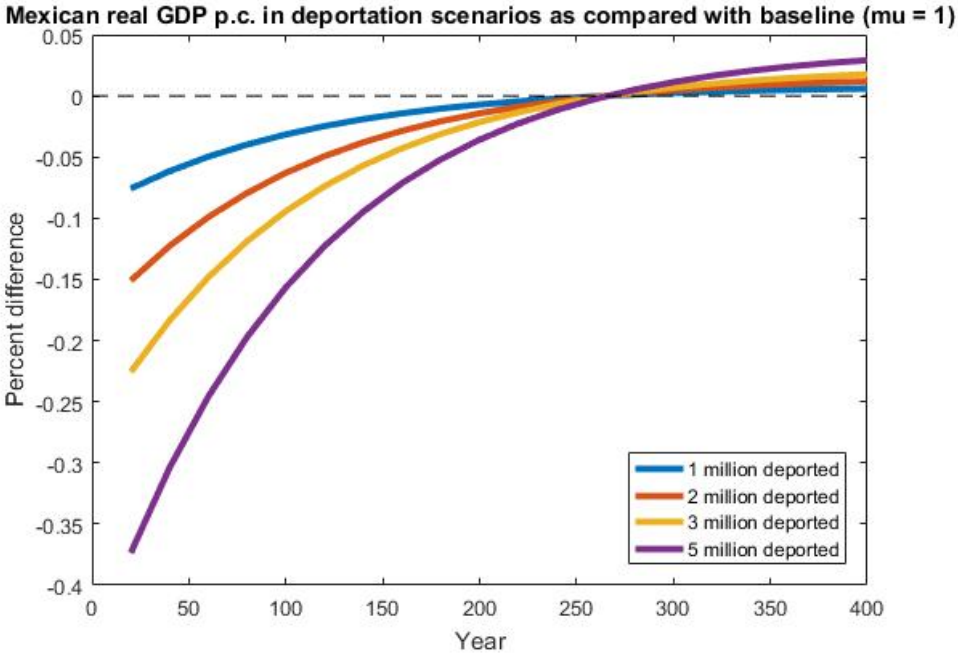
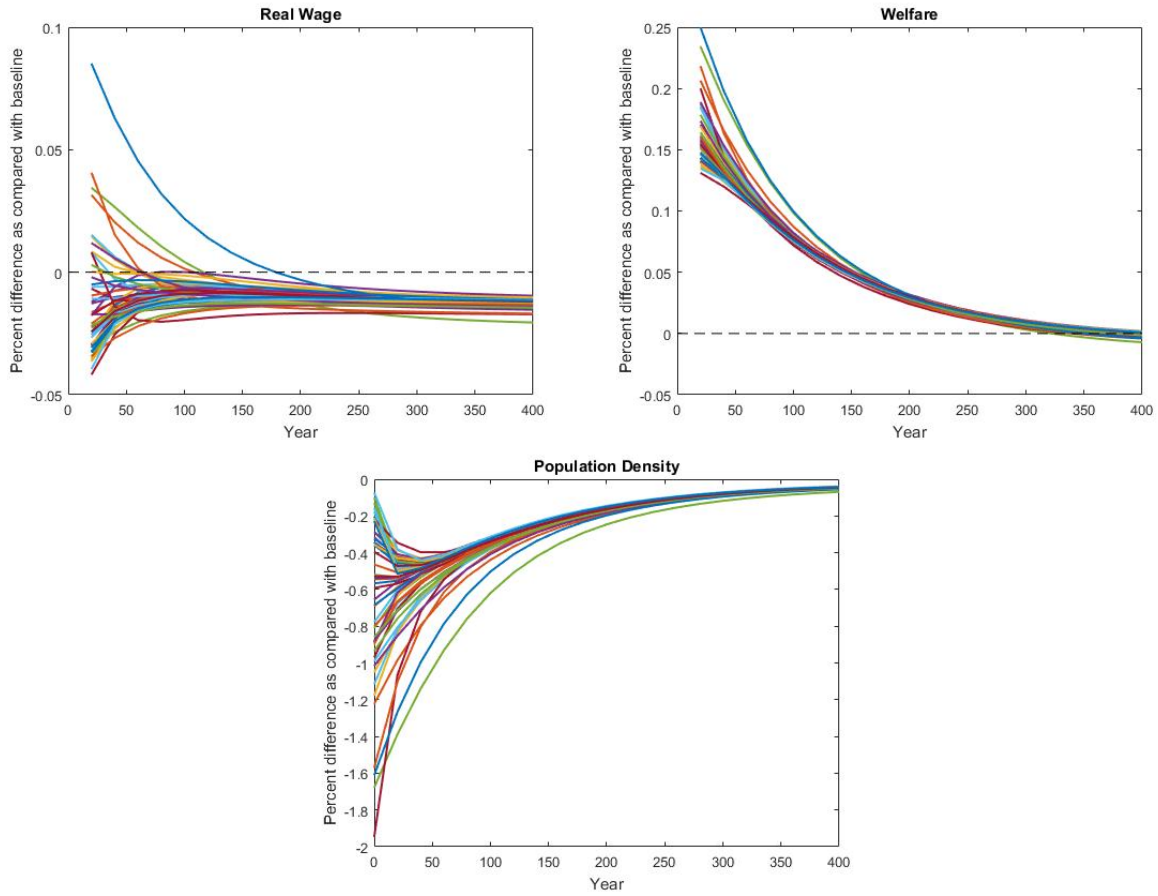


Figure 15: Percent difference in outcomes for the US 50 states of deportation simulation (3 million deportee scenario) as compared with the baseline scenario. Specification in which land is not a scarce factor (i.e. $\mu = 1$).



References

[1] Christian Dustmann, Uta Schönberg, and Jan Stuhler. The impact of immigration: Why do studies reach such different results? *The Journal of Economic Perspectives*, 30(4):31–56, 2016.

[2] Paul M Romer. Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102, 1990.

[3] Klaus Desmet, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg. The geography of development. *Journal of Political Economy (forthcoming)*, 2017.

[4] Treb Allen and David Donaldson. Notes on geography and path dependence. Unpublished.

[5] Jonathan Eaton and Samuel Kortum. Technology, geography, and trade. *Econometrica*, 70(5):1741–1779, 2002.

[6] Treb Allen and Costas Arkolakis. Trade and the topography of the spatial economy. Technical report, National Bureau of Economic Research, 2013.

[7] Stephen J Redding and Esteban Rossi-Hansberg. Quantitative spatial economics. Technical report, National Bureau of Economic Research, 2016.

[8] Gabriel M Ahlfeldt, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015.

[9] Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598, 2010.

[10] Glenn Ellison, Edward L Glaeser, and William R Kerr. What causes industry agglomeration? evidence from coagglomeration patterns. *The American Economic Review*, 100(3):1195–1213, 2010.

[11] Jeffrey S. Krogstad, Jens Manuel; Passel and D'Vera Cohn. 5 facts about illegal immigration in the u.s. Technical report, Pew Research Center, 2016.

[12] Ana Gonzalez-Barrera. More mexicans leaving than coming to the us. Technical report, Pew Research Center, 2015.

[13] Brian Bennett. High deportation figures are misleading. *Los Angeles Times*, April 1, 2014.

[14] Department of Homeland Security. Yearbook of immigration statistics 2015. Technical report, Department of Homeland Security, 2015.

[15] Jeffrey S. Passel and D'Vera Cohn. Overall number of u.s. unauthorized immigrants holds steady since 2009. Technical report, Pew Research Center, 2016.

[16] Treb Allen, Costas Arkolakis, and Yuta Takahashi. Universal gravity. *Mimeo Northwestern and Yale University*, 2014.

[17] Treb Allen, Costas Arkolakis, and Xiangliang Li. On the existence and uniqueness of trade equilibria. Technical report, Citeseer, 2015.

[18] Daron Acemoglu. Directed technical change. *The Review of Economic Studies*, 69(4):781–809, 2002.

[19] Raven Molloy, Christopher L Smith, and Abigail Wozniak. Internal migration in the united states. *The Journal of Economic Perspectives*, 25(3):173–196, 2011.

[20] H David, David Dorn, and Gordon H Hanson. The china syndrome: Local labor market

effects of import competition in the united states. *The American Economic Review*, 103(6):2121–2168, 2013.

[21] Peter A Rogerson. A new method for finding geographic centers, with application to us states. *The Professional Geographer*, 67(4):686–694, 2015.

[22] Andrew B Bernard, Jonathan Eaton, J Bradford Jensen, and Samuel Kortum. Plants and productivity in international trade. *The American Economic Review*, 93(4):1268–1290, 2003.

[23] Ferdinando Monte, Stephen J Redding, and Esteban Rossi-Hansberg. Commuting, migration and local employment elasticities. Technical report, National Bureau of Economic Research, 2015.

[24] Angus Deaton and Arthur A Stone. Two happiness puzzles. *The American economic review*, 103(3):591–597, 2013.

[25] Gerald A Carlino, Satyajit Chatterjee, and Robert M Hunt. Urban density and the rate of invention. *Journal of Urban Economics*, 61(3):389–419, 2007.

[26] Ina Simonovska and Michael E Waugh. The elasticity of trade: Estimates and evidence. *Journal of international Economics*, 92(1):34–50, 2014.

[27] Jeremy Greenwood, Zvi Hercowitz, and Per Krusell. Long-run implications of investment-specific technological change. *The American Economic Review*, pages 342–362, 1997.

[28] Klaus Desmet and Jordan Rappaport. The settlement of the united states, 1800–2000: The long transition towards gibrats law. *Journal of Urban Economics*, 2015.

[29] Klaus Desmet and Esteban Rossi-Hansberg. On the spatial economic impact of global warming. *Journal of Urban Economics*, 88:16–37, 2015.

[30] Masahisa Fujita and Hideaki Ogawa. Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional science and urban economics*, 12(2):161–196, 1982.

[31] Jeffrey S. Passel and D'Vera Cohn. Unauthorized immigration totals rise in 7 states, fall in 14. Technical report, Pew Research Center, 2014.

[32] Jeffrey S. Passel and D'Vera Cohn. Share of unauthorized immigrant workers in production, construction jobs falls since 2007. Technical report, Pew Research Center, 2015.