DYNAMIC SPATIAL GENERAL EQUILIBRIUM

Benny Kleinman
Ernest Liu
Stephen J. Redding

Dynamic Spatial General Equilibrium
Benny Kleinman, Ernest Liu, and Stephen J. Redding
NBER Working Paper No. 29101
July 2021
JEL No. F14,F15,R12

ABSTRACT

We develop a dynamic spatial general equilibrium model with forward-looking investment and migration decisions. We characterize analytically the transition path of the spatial distribution of economic activity in response to shocks. We apply our framework to the reallocation of US economic activity from the Rust Belt to the Sun Belt from 1965-2015. We find slow convergence to steady-state, with US states closer to steady-state at the end of our sample period than at its beginning. We find substantial heterogeneity in the effects of local shocks, which depend on capital and labor dynamics, and the spatial and sectoral incidence of these shocks.

Benny Kleinman
Department of Economics
Julis Romo Rabinowitz Building
Princeton, NJ 08544
binyamin@princeton.edu

Ernest Liu
Princeton University
Department of Economics
Julis Romo Rabinowitz Building
Princeton, NJ 08544
ernestliu@princeton.edu

Stephen J. Redding
Department of Economics &
School of Public and International Affairs
Princeton University
Princeton, NJ 08544
and CEPR
and also NBER
reddings@princeton.edu

1   Introduction

A key research question in economics is understanding the gradual response of the spatial distribution of economic activity to shocks. To address this question, we develop a dynamic spatial model with forward-looking investment and migration decisions. Despite accommodating a large state space and a rich geography, the model remains tractable, and we provide conditions for the existence and uniqueness of its general equilibrium. We linearize the model to derive closed-form solutions for the transition path of the entire spatial distribution of economic activity in response to shocks. Using these closed-form solutions, we characterize analytically the determinants of the economy's dynamic response to productivity, amenity, trade cost and migration cost shocks. We show that the speed of convergence to steady-state depends on the eigenvalues of a transition matrix that governs the evolution of the state variables. We undertake a spectral analysis of this transition matrix to distinguish shocks and exposure to these shocks. We apply our framework to the reallocation of US economic activity from the "Rust Belt" to the "Sun Belt" from 1965-2015. We find that US states are closer to steady-state by the end of our sample than at its beginning, contributing to a decline in geographical mobility over time. We find slow average speeds of convergence to steady-state, with an average half-life of around 20 years. We find substantial heterogeneity in the effects of local shocks, which depend on both capital and labor dynamics, and the spatial and sectoral incidence of shocks.

One of the key challenges in developing dynamic spatial models to address these questions is incorporating forward-looking dynamic investment decisions, because the investment decision in each location depends on economic activity in all locations in all future time periods. With high-dimensional state spaces, this can introduce a curse of dimensionality, which can make computing the equilibrium allocations challenging, or can make it difficult to distinguish between alternative possible future trajectories for the economy. We develop a framework that incorporates this forward-looking behavior for both investment decisions for the immobile factor (capital structures) and migration decisions for the mobile factor (labor). Even with these two sources of dynamics, we provide an analytical characterization of the existence and uniqueness of the steady-state equilibrium, and of the comparative statics of the spatial distribution of economic activity in each future time period in response to shocks.

To illustrate our approach, we begin with a baseline single-sector Armington model of trade, with dynamic discrete choice migration decisions, and investment determined as the solution to an intertemporal consumption-investment problem. The economy consists of many locations that differ in productivity, amenities, bilateral trade costs and bilateral migration costs. There are two types of agents: workers and landlords. Workers are geographically mobile but do not have access to an investment technology (and hence live "hand to mouth"). They make forward-

looking migration decisions, taking into account migration costs and the expected continuation value from optimal future location decisions, as in Caliendo et al. (2019). Landlords are geographically immobile but have access to an investment technology for accumulating local capital. They make forward-looking consumption-investment decisions to maximize intertemporal utility, as in Angeletos (2007) and Moll (2014).

We linearize the model to derive a closed-form solution for the economy's transition path in terms of an impact matrix (R), which captures the impact of shocks in the initial period in which they occur, and a transition matrix (P), which governs the evolution of the state variables from one period to the next in response to these shocks. The impact and transition matrices in turn depend on four observable trade and migration share matrices: (i) an expenditure share matrix (S) that reflects the expenditure share of each importer on each exporter; (ii) an income share matrix (T) that captures the share of each exporter's value-added derived from each importer; (iii) an outmigration share matrix (D) that reflects the share of people in a given origin that migrate to each destination; (iv) an inmigration share matrix (E) that corresponds to the share of people in a given destination that migrate from each origin. Given these observable matrices, initial values for the economy's state variables (population and the capital stock in each location) and the model's structural parameters, we can solve in closed-form for the entire spatial distribution of economic activity in each location in all future time periods.

This linearization has two key advantages. First, we show that our linear closed-form solutions can be used to undertake a spectral analysis of the economy's adjustment to shocks. We use an eigendecomposition of the transition matrix to distinguish shocks and exposure to shocks. We show that any empirical shocks can be expressed as a linear combination of eigen-shocks, defined as shocks for which the initial impact on the state variables corresponds to an eigenvector of the transition matrix. We show that the speed of convergence to steady-state, as measured by the half-life, is determined by the corresponding eigenvalues of this transition matrix. This speed of convergence is influenced by both capital and labor adjustment, and the dynamics in these two state variables interact with one another, because of the complementarity of capital and labor in production. For example, if both capital and labor are above steady-state in a location, this slows convergence, because the high capital stock raises the marginal product of labor, and hence retards its downward adjustment. Similarly, the high population raises the marginal product of capital and dampens its downward adjustment. In contrast, if capital is above steady-state whereas labor is below steady-state, this tends to speed convergence, because the high capital stock attracts labor, and the low population discourages capital accumulation. A lower intertemporal elasticity of substitution, greater dispersion in idiosyncratic mobility shocks and a lower labor share also all slow convergence towards steady-state.

Second, as our linearization uses methods from the dynamic stochastic general equilibrium

(DSGE) literature in macroeconomics, we are able to relax the assumption of perfect foresight commonly maintained in spatial models to allow for stochastic fundamentals and rational expectations. We report three main sets of results for the economy's transition path. First, we consider an economy that is somewhere along the transition path towards an unobserved steady-state at a given point in time, and show how to solve for the transition path to that steady-state in the absence of any further shocks to fundamentals. Second, starting with the same initial conditions, we show how to solve for the economy's transition path in response to any convergent sequence of expectations of future shocks. Third, we show how to solve for the economy's dynamic path when the fundamental shocks follow a first-order Markov process. Although we allow for high-dimensional state spaces with many locations or location-sectors, our use of conventional linear algebra techniques ensures that our approach is computationally efficient and quick and easy to implement. Finally, a caveat is that our closed-form solutions for the transition path are based on a linearization, and are thus only exact in theory for small shocks. However, we show that in practice they closely approximate the full non-linear model solution even for large shocks, such as the empirical distribution of decadal shocks observed during our sample period.

Although for simplicity we begin with our baseline single-sector Armington model, we show that our approach admits a large number of extensions and generalizations. We incorporate agglomeration forces in both production (productivity spillovers) and residential decisions (amenity spillovers). In the presence of these agglomeration forces, we show that the conditions for the existence of a unique steady-state require that these agglomeration forces are sufficiently weak relative to the model's dispersion forces, which include idiosyncratic worker preferences. More generally, we show that our results hold for an entire class of constant elasticity trade models, including models of perfect competition and constant returns to scale, and models of monopolistic competition and increasing returns to scale. We assume for simplicity that capital is only used in production, but we show that our results naturally extend to the case in which capital is also used residentially (housing). Finally, we demonstrate that our approach also extends to incorporate multiple sectors (as in Costinot et al. 2012) and both multiple sectors and input-output linkages (as in Caliendo and Parro 2015 and Caliendo et al. 2019).

In our main empirical application, we use data on U.S. states from 1965-2015 to examine the decline of the "Rust Belt" and the rise of the "Sun Belt." We show that this setting features convergence dynamics in both capital and net and gross migration, highlighting the relevance of a framework such as ours that features both of these adjustment margins. Our findings of slow average speeds of convergence to steady-state of around 20 years are consistent with reduced-form empirical evidence of persistent impacts of local shocks. Our results also emphasize the heterogeneity in the impact of these local shocks, depending on their size and incidence across locations, and the extent to which they affect both capital and labor state variables. We find that

average speeds of convergence are relatively constant over time, despite the modest observed decline in inter-state migration. These results highlight that this decline in inter-state migration need not imply a rise in geographical barriers to mobility, since it also influenced by the distance of the state variables from steady-state and the size and incidence of shocks.

In a final empirical exercise, we implement our multi-sector extension using region-sector data on U.S. states and foreign countries from 1999-2015. In this multi-sector extension, we continue to find relatively slow convergence towards steady-state, but the average half-life is notably lower than our baseline single-sector specification. This pattern of results reflects the property of the data that movements of people between sectors within U.S. states occur much more frequently than movements of people between U.S. states. An implication is that the speed with which the economy adjusts to shocks depends crucially on the extent to which they affect one industry relative to another industry within the same location, versus the extent to which that affect all industries in one location relative to all industries in another location.

Our research is related to several strands of existing work. First, our paper contributes to a long line of research on economic geography, including Krugman (1991b), Krugman and Venables (1995) and Helpman (1998), as synthesized in Fujita et al. (1999), and reviewed in Duranton and Puga (2004) and Redding (2020). Early theoretical research on economic geography considered static models or assumed myopic migration decisions, as in Krugman (1992).[1] In contrast, the more recent research on quantitative spatial models has often considered static specifications, including Redding and Sturm (2008), Allen and Arkolakis (2014), Ahlfeldt et al. (2015), Allen et al. (2017), Ramondo et al. (2016), Redding (2016), Caliendo et al. (2018) and Monte et al. (2018), as surveyed in Redding and Rossi-Hansberg (2017).

A key reason that quantitative spatial models have frequently focused on these static specifications is the challenge of modelling forward-looking investments in settings with high-dimensional state spaces. One approach to this challenge has been to consider specifications in which dynamic decisions reduce to static problems. In the innovation models of Desmet and Rossi-Hansberg (2014), Desmet et al. (2018) and Peters (2019), the incentive to invest in innovation each period depends on the comparison of static profits and innovation costs. In the overlapping generations model of Allen and Donaldson (2020), adults make migration decisions to maximize their own adult utility, and do not consider the utility of the next generation of youths. Another approach is to capture forward-looking migration decisions using dynamic discrete choice models, including Artuç et al. (2010); Artuc et al. (2021), Caliendo et al. (2019) and Caliendo and Parro (2020).[2] We in-

---

[1] Exceptions include a small number of theoretical trade and geography papers that have considered forward-looking decisions under perfect foresight in stylized settings with a small number of locations, including Krugman (1991a), Matsuyama (1991) and Baldwin (2001).

[2] See Glaeser and Gyourko (2005) and Greaney (2020) for models in which population dynamics are shaped by durable housing. See Walsh (2019) for a model in which innovation takes the form of the creation of new varieties.

corporate both forward-looking investment and migration decisions and use our linearization to derive closed-form solutions for the transition path of the entire spatial distribution of economic activity in response to shocks.

Second, our work is related to the literature on sufficient statistics in static international trade and economic geography models, including Arkolakis et al. (2012), Caliendo et al. (2017), Baqaee and Farhi (2019), Galle et al. (2018), Huo et al. (2019), Bartelme et al. (2019), Adão et al. (2019), Bonadio et al. (2020), and Kim and Vogel (2020).[3] Using a class of static constant elasticity trade models, Kleinman et al. (2020) show that the first-order comparative statics can be stacked as a matrix inversion problem, which yields closed-form solutions for the elasticity of the endogenous variables in each country with respect to shocks in any other country. Although these existing studies have developed sufficient statistics for *static* spatial models, our key contribution is to develop them for *dynamic* spatial models, incorporating both forward-looking investment decisions for the immobile factor and dynamic migration decisions for the mobile factor.

Third, our research is related to an empirical literature on local labor markets, including Autor et al. (2013), Kovak (2013), Kline and Moretti (2014), Dix-Carneiro and Kovak (2015), and Diamond (2016), Hornbeck and Moretti (2018) and Eriksson et al. (2019), as reviewed in Moretti (2011) and Autor et al. (2016). One strand of this literature has examined the reallocation of U.S. economic activity from the Rust Belt to the Sun Belt, such as Blanchard and Katz (1992), Feyrer et al. (2007), Rappaport (2007), Glaeser and Ponzetto (2010), Hartley (2013), Yoon (2017) and Alder et al. (2019). Another strand of this literature has emphasized the persistent impact of negative local labor market shocks, including in particular Dix-Carneiro and Kovak (2017), Amior and Manning (2018) and Autor et al. (2020). We contribute to this research by using our spectral analysis to analyze the determinants and heterogeneity of the economy's dynamic response to shocks.

The remainder of the paper is structured as follows. In Section 2, we introduce our baseline quantitative spatial model with dynamics from forward-looking investment and migration decisions. In Section 3, we linearize the model and derive our closed-form solutions for the economy's transition path. In Section 4, we show that our analysis admits a number of extensions and generalizations, including shocks to trade and migration costs, agglomeration forces, multiple sectors, and input-output linkages. In Section 5, we implement our baseline specification for U.S. states from 1965-2015 and our multi-sector extension for U.S. states and foreign countries from 1999-2015. In Section 6, we report a specification check, in which we examine the potential for non-linearities in the model. In Section 7, we summarize our conclusions.

---

[3]For sufficient statistics in heterogeneous agent models in macro, see Auclert et al. (2020) and Bilal (2021).

2 Dynamic Spatial Model

In this section, we introduce our baseline dynamic spatial general equilibrium model. We combine a specification of trade between locations with a constant trade elasticity, a formulation of migration decisions with a constant migration elasticity, optimal consumption-investment decisions with constant intertemporal elasticity of substitution preferences, and stochastic shocks to fundamentals under rational expectations.

We consider an economy with many locations indexed by $i \in \{1, \ldots, N\}$. Time is discrete and is indexed by t. There are two types of infinitely-lived agents: workers and landlords. Workers are endowed with one unit of labor that is supplied inelastically and are geographically mobile subject to migration costs. Workers do not have access to an investment technology, and hence live "hand to mouth," as in Kaplan and Violante (2014). Landlords are geographically immobile and own the capital stock in their location. They make forward-looking decisions over consumption and investment in this local stock of capital. We interpret capital as buildings and structures, which are geographically immobile once installed, and depreciate gradually at a constant rate δ.

The endogenous state variables are the population (ℓ_{it}) and capital stock (k_{it}) in each location. The key location characteristics that determine the spatial distribution of economic activity are the sequences of productivity (z_{it}), amenities (b_{it}), trade costs (τ_{nit}) and migration costs (κ_{nit}). Without loss of generality, we normalize the total population across all locations to one ($\sum_{i=1}^{N} \ell_{it} = 1$), so that ℓ_{it} can also be interpreted as the population share of location i at time t. Throughout the following, we use bold math font to denote a vector (lowercase letters) or matrix (uppercase letters). The derivations for all expressions and results in this section are reported in Section B of the online appendix.

To streamline the exposition, we focus in this section on shocks to productivities (z_{it}) and amenities (b_{it}), and assume that productivity and amenities are exogenous. In Section 4 below, we show that our approach extends to shocks to trade and migration costs, endogenous productivity and amenities through agglomeration forces, and a large number of other generalizations, including multiple sectors and input-output linkages.

2.1 Production

At the beginning of each period t, the economy inherits in each location i a mass of workers (ℓ_{it}) and capital stock (k_{it}). Firms in each location use labor and capital to produce output (y_{it}) of the variety supplied by that location. Production is assumed to occur under conditions of perfect competition and subject to the following constant returns to scale technology:

$$y_{it} = z_{it} \left(\frac{\ell_{it}}{\mu} \right)^{\mu} \left(\frac{k_{it}}{1-\mu} \right)^{1-\mu}, \qquad 0 < \mu < 1, \tag{1}$$

where z_{it} denotes productivity in location i at time t.

We assume that trade between locations is subject to iceberg variable trade costs, such that $\tau_{nit} \geq 1$ units of a good must be shipped from location i in order for one unit to arrive in location n, where $\tau_{nit} > 1$ for $n \neq i$ and $\tau_{iit} = 1$. From profit maximization, the cost to a consumer in location n of sourcing the good produced by location i depends solely on these iceberg trade costs and constant marginal costs:

$$p_{nit} = \tau_{nit} p_{iit} = \frac{\tau_{nit} w_{it}^{\mu} r_{it}^{1-\mu}}{z_{it}}, \tag{2}$$

where p_{iit} is the "free on board" price of the good supplied by location i before trade costs.

We choose the total labor income of all locations as our numeraire: $\sum_{i=1}^{N} w_{it} \ell_{it} = 1$.

2.2 Worker Consumption

Worker preferences within each period t are modeled as in the standard Armington model of trade with constant elasticity of substitution (CES) preferences. As workers do not have access to an investment technology, they spend their wage income and choose their consumption of varieties to maximize their utility each period. The indirect utility function of a worker in location n in period t depends on the worker's wage (w_{nt}), the cost of living (p_{nt}) and amenities (b_{nt}):

$$\ln u_{nt}^{w} = \ln b_{nt} + \ln w_{nt} - \ln p_{nt}, \tag{3}$$

where we use the superscript w to denote workers. The cost of living (p_{nt}) depends on the price of the variety sourced from each location i (p_{nit}):

$$p_{nt} = \left[\sum_{i=1}^{N} p_{nit}^{-\theta} \right]^{-1/\theta}, \tag{4}$$

where $\sigma > 1$ is the constant elasticity of substitution and $\theta = \sigma - 1 > 0$ is the trade elasticity.

2.3 Capital Accumulation

Landlords in each location choose their consumption and investment to maximize their intertemporal utility subject to their budget constraint. Landlords' intertemporal utility equals the expected present discounted value of their flow utility:

$$v_{it}^{k} = \mathbb{E}_t \sum_{s=0}^{\infty} \beta^{t+s} \frac{\left(c_{it+s}^{k} \right)^{1-1/\psi}}{1 - 1/\psi}, \tag{5}$$

where we use the superscript k to denote landlords; c_{it}^{k} is the consumption index dual to the price index (4); β is the discount rate; ψ is the elasticity of intertemporal substitution. Since landlords

are geographically immobile, we omit the term in amenities from their flow utility, because this does not affect the equilibrium in any way, and hence is without loss of generality.

We assume that the investment technology in each location uses the varieties from all locations with the same functional form as consumption. We assume that landlords can only invest in their own location and that one unit of capital can be produced using one unit of the consumption index in that location.[4] We interpret capital as buildings and structures, which are geographically immobile once installed. Capital is assumed to depreciate at the constant rate δ and we allow for the possibility of negative investment. The intertemporal budget constraint for landlords in each location requires that total income from the existing stock of capital ($r_{it}k_{it}$) equals the total value of their consumption ($p_{it}c_{it}^k$) plus the total value of net investment ($p_{it}(k_{it+1} - (1 - \delta)k_{it})$):

$$r_{it}k_{it} = p_{it}\left(c_{it}^k + k_{it+1} - (1 - \delta)k_{it}\right). \tag{6}$$

We begin by establishing a key property of landlords' optimal consumption-investment decisions. We use $R_{it} \equiv 1 - \delta + r_{it}/p_{it}$ to denote the gross return on capital.

Lemma 1. *The optimal consumption of location i's landlords satisfies $c_{it} = \varsigma_{it}R_{it}k_{it}$, where ς_{it} is defined recursively as*

$$\varsigma_{it}^{-1} = 1 + \beta^\psi \left(\mathbb{E}_t\left[R_{it+1}^{\frac{\psi-1}{\psi}}\varsigma_{t+1}^{-\frac{1}{\psi}}\right]\right)^\psi.$$

Landlords' optimal saving and investment decisions satisfy $k_{it+1} = (1 - \varsigma_{it})R_{it}k_{it}$.

Lemma 1 shows that the landlords in each location have a linear saving rate $(1 - \varsigma_{it})$ out of current period wealth $R_{it}k_{it}$, as in Angeletos (2007). The saving rate depends on the expected sequence of future returns on capital $\{\mathbb{E}_t R_{it+s}\}$, the discount rate β, and the elasticity of substitution ψ. In the special case of log-utility ($\psi = 1$), landlords have a constant saving rate β (i.e., $k_{it+1} = \beta R_{it}k_{it}$) as in the conventional Solow-Swan model and Moll (2014).

We show below that there exists a steady-state equilibrium level of the capital-labor ratio in each location, towards which the economy gradually converges in the absence of further shocks. Along the transition path, the rental rate on capital can differ across locations, because of our assumption that capital corresponds to buildings and structures that are geographically immobile. However, the extent of these differences is limited by the use of tradeable consumption goods for investment, and quantitatively we find them to be relatively small. In steady-state, the real rental rate in terms of the consumption good is equalized across locations.[5]

---

[4]This specification can be extended to allow landlords to invest in other locations at the cost of additional complication. Although we make the standard assumption that consumption and investment use goods in the same proportions, similar results hold if investment uses only the good produced by each location.

[5]Introducing another type of capital (e.g. machines) that is perfectly mobile is straightforward. In steady-state: $r_i^*/p_i^* = (1 - \beta(1 - \delta))/\beta$. Steady-state differences in r_i^*/p_i^* can be accommodated by differences in the productivity of investment, where one unit capital is produced with θ_i units of the consumption good in location i.

2.4 Worker Migration Decisions

After supplying labor and spending wage income on consumption in each period t, workers observe idiosyncratic mobility shocks (ϵ_{gt}), and decide where to move. The value function for a worker in location i in period t ($\mathbb{V}_{it}^w$) is equal to the current flow of utility in that location plus the expected continuation value from the optimal choice of location:

$$\mathbb{V}_{it}^w = \ln u_{it}^w + \max_{\{g\}_1^N} \left\{ \beta \mathbb{E}_t \left[\mathbb{V}_{gt+1}^w \right] - \kappa_{git} + \rho \epsilon_{gt} \right\}, \tag{7}$$

where β is the discount rate; $\mathbb{E}_t [\cdot]$ denotes the expectation in period t over future location characteristics. Because workers are hand-to-mouth, it is without loss of generality to assume logarithmic utility ($\ln u_{it}^w$) over consumption. We make the conventional assumption that idiosyncratic mobility shocks are drawn from an extreme value distribution: $F(\epsilon) = e^{-e^{(-\epsilon - \overline{\gamma})}}$, where $\overline{\gamma}$ is the Euler-Mascheroni constant; the parameter ρ controls the dispersion of idiosyncratic mobility shocks; and we assume $\kappa_{iit} = 1$ and $\kappa_{nit} > 1$ for $n \neq i$.

2.5 Market Clearing

Goods market clearing implies that income in each location, which equals the sum of the income of workers and landlords, is equal to expenditure on the goods produced by that location:

$$(w_{it}\ell_{it} + r_{it}k_{it}) = \sum_{n=1}^{N} S_{nit} (w_{nt}\ell_{nt} + r_{nt}k_{nt}), \tag{8}$$

where we begin by assuming that trade is balanced, before later extending our analysis to incorporate trade imbalances in Section 4.

Capital market clearing implies that the rental rate for capital is determined by the requirement that landlords' income from the ownership of capital equals payments for its use. Using profit maximization and zero profits, this capital market clearing condition is given by:

$$r_{it}k_{it} = \frac{1-\mu}{\mu} w_{it}\ell_{it}. \tag{9}$$

2.6 General Equilibrium

Given the state variables $\{\ell_{i0}, k_{i0}\}$, the general equilibrium of the economy is the stochastic process of allocations and prices such that firms in each location choose inputs to maximize profits, workers make consumption and migration decisions to maximize utility, landlords make consumption and investment decisions to maximize utility, and prices clear all markets, with the appropriate measurability constraint with respect to the realizations of location fundamentals.

For expositional clarity, we collect the equilibrium conditions and express them in terms of a sequence of five endogenous variables $\{\ell_{it}, k_{it}, w_{it}, R_{it}, v_{it}\}_{t=0}^{\infty}$. All other endogenous variables of the model can be recovered as a function of these variables.

Capital Returns and Accumulation: Using capital market clearing (9), the gross return on capital in each location i must satisfy:

$$R_{it} = \left(1 - \delta + \frac{1-\mu}{\mu} \frac{w_{it}\ell_{it}}{p_{it}k_{it}}\right),$$

where the price index (4) of each location becomes

$$p_{nt} = \left[\sum_{i=1}^{N} \left(w_{it} \left(\frac{1-\mu}{\mu}\right)^{1-\mu} (\ell_{it}/k_{it})^{1-\mu} \tau_{ni}/z_i\right)^{-\theta}\right]^{-1/\theta}. \tag{10}$$

The law of motion for capital is

$$k_{it+1} = (1 - \varsigma_{it})\left(1 - \delta + \frac{1-\mu}{\mu}\frac{w_{it}\ell_{it}}{p_{it}k_{it}}\right)k_{it}, \tag{11}$$

where $(1 - \varsigma_{it})$ is the saving rate defined recursively as in Lemma 1:

$$\varsigma_{it}^{-1} = 1 + \beta^{\psi} \left(\mathbb{E}_t \left[R_{it+1}^{\frac{\psi-1}{\psi}} \varsigma_{t+1}^{-\frac{1}{\psi}}\right]\right)^{\psi}.$$

Goods Market Clearing: Using the CES expenditure share, the equilibrium pricing rule (2), and the capital market clearing condition (9) in the goods market clearing condition (8), the requirement that income equals expenditure on the goods produced by a location can be written solely in terms of labor income:

$$w_{it}\ell_{it} = \sum_{n=1}^{N} S_{nit} w_{nt}\ell_{nt}, \tag{12}$$

$$S_{nit} = \frac{\left(w_{it} (\ell_{it}/k_{it})^{1-\mu} \tau_{ni}/z_i\right)^{-\theta}}{\sum_{m=1}^{N} \left(w_{mt} (\ell_{mt}/k_{mt})^{1-\mu} \tau_{nm}/z_m\right)^{-\theta}}, \qquad T_{int} \equiv \frac{S_{nit} w_{nt}\ell_{nt}}{w_{it}\ell_{it}}, \tag{13}$$

where we have used the property that capital income is a constant multiple of labor income; S_{nit} is the expenditure share of importer n on exporter i at time t; we have defined T_{int} as the corresponding income share of exporter i from importer n at time t; and note that the order of subscripts switches between the expenditure share (S_{nit}) and the income share (T_{int}), because the first and second subscripts will correspond below to rows and columns of a matrix, respectively.

Worker Value Function: Using the value function (7), indirect utility function (3) and the properties of the extreme value distribution, the expected value from living in location n at time t after taking expectations with respect to the idiosyncratic mobility shocks $\{\epsilon_{gt}\}$ (that is, $v_{nt}^w \equiv \mathbb{E}_\epsilon [\mathbb{V}_{nt}^w]$), can be written as:

$$v_{nt}^w = \ln b_{nt} + \ln \left(\frac{w_{nt}}{p_{nt}} \right) + \rho \ln \sum_{g=1}^N \left(\exp \left(\beta \mathbb{E}_t v_{gt+1}^w \right) / \kappa_{gnt} \right)^{1/\rho}, \tag{14}$$

where the expectation $\mathbb{E}_t \left[v_{gt+1}^w \right] = \mathbb{E}_t \mathbb{E}_\epsilon \left[\mathbb{V}_{nt+1}^w \right]$ is taken over future fundamentals $\{z_{is}, b_{is}\}_{s=t+1}^\infty$.

Population Flow: Using the properties of the extreme value distribution, the population flow condition for the evolution of the population distribution over time is given by:

$$\ell_{gt+1} = \sum_{i=1}^N D_{igt} \ell_{it}, \tag{15}$$

$$D_{igt} = \frac{\left(\exp \left(\beta \mathbb{E}_t v_{gt+1}^w \right) / \kappa_{git} \right)^{1/\rho}}{\sum_{m=1}^N \left(\exp \left(\beta \mathbb{E}_t v_{mt+1}^w \right) / \kappa_{mit} \right)^{1/\rho}}, \qquad E_{git} \equiv \frac{\ell_{it} D_{igt}}{\ell_{gt+1}}, \tag{16}$$

where D_{igt} is the outmigration probability from location i to location g between time t and $t+1$; we have defined E_{git} as the corresponding inmigration probability to location g from location i between time t and $t+1$; again note that the order of subscripts switches between the outmigration probability (D_{igt}) and the inmigration probability (E_{git}), because the first and second subscripts will correspond below to rows and columns of a matrix, respectively.

Properties of General Equilibrium: Given the state variables $\{\ell_{it}, k_{it}\}$ and the realized location fundamentals $\{z_{it}, b_{it}\}$, the general equilibrium in each period is determined as in a standard static international trade model. Between periods, the evolution of the stock of capital $\{k_{it}\}$ is determined by the equilibrium saving rate, and the dynamics of the population distribution $\{\ell_{it}\}$ are determined by the gravity equation for migration. We now formally define equilibrium.

Definition 1. Equilibrium. Given the state variables $\{\ell_{i0}, k_{i0}\}$ in each location in an initial period $t = 0$, an *equilibrium* is a stochastic process of wages, capital returns, expected values, mass of workers and stock of capital in each location $\{w_{it}, R_{it}, v_{it}, \ell_{it+1}, k_{it+1}\}_{t=0}^\infty$ measurable with respect to the fundamental shocks up to time t ($\{z_{is}, b_{is}\}_{s=1}^t$), and solves the value function (14), the population flow condition (15), the goods market clearing condition (12), and the capital market clearing and accumulation condition (11), with the saving rate determined by Lemma 1.

We define a deterministic steady-state equilibrium as one in which the fundamentals $\{z_i^*, b_i^*\}$ and the endogenous variables $\{\ell_i^*, k_i^*, w_i^*, R_i^*, v_i^*\}$ are constant over time, where we use an asterisk to denote the steady-state value of variables.

Definition 2. Steady State. A *steady-state* of the economy is an equilibrium in which all location-specific fundamentals and endogenous variables (wages, expected values, mass of workers and stock of capital in each location) are time invariant: $\{z_i^*, b_i^*, \ell_i^*, k_i^*, w_i^*, R_i^*, v_i^*\}$.

Our model features rich spatial interactions between locations in both goods and factor markets and forward-looking investment and migration decisions. Nevertheless, the absence of agglomeration forces and diminishing marginal returns to capital accumulation ensure the existence of a unique steady-state spatial distribution of economic activity (up to a choice of numeraire), given time-invariant values of the locational fundamentals of productivity (z_i^*), amenities (b_i^*), goods trade costs (τ_{ni}^*) and migration frictions (κ_{ni}^*).

Proposition 1. Existence and Uniqueness. *There exists a unique steady-state spatial distribution of economic activity $\{\ell_i^*, k_i^*, w_i^*, R_i^*, v_i^*\}$ (up to a choice of numeraire) given time-invariant locational fundamentals $\{z_i^*, b_i^*, \tau_{ni}^*, \kappa_{ni}^*\}$ that is independent of initial conditions $\{\ell_{i0}, k_{i0}\}$.*

Proof. See Section B.7 of the online appendix. $\square$

Trade and Migration Share Matrices: We now introduce the trade and migration share matrices that we use in the next section to characterize the model's comparative statics. Let $\boldsymbol{S}$ be the $N \times N$ matrix with the ni-th element equal to importer n's expenditure on exporter i. Let $\boldsymbol{T}$ be the $N \times N$ matrix with the in-th element equal to the fraction of income that exporter i derives from selling to importer n. We refer to $\boldsymbol{S}$ as the *expenditure share* matrix and to $\boldsymbol{T}$ as the *income share* matrix. Intuitively, S_{ni} captures the importance of i as a supplier to location n, and T_{in} captures the importance of n as a buyer for country i. Note the order of subscripts: in matrix $\boldsymbol{S}$, rows are buyers and columns are suppliers, whereas in matrix $\boldsymbol{T}$, rows are suppliers and columns are buyers.

Similarly, let $\boldsymbol{D}$ be the $N \times N$ matrix with the ni-th element equal to the share of outmigrants from origin n to destination i. Let $\boldsymbol{E}$ be the $N \times N$ matrix with the in-th element equal to the share of inmigrants to destination i from origin n. We refer to $\boldsymbol{D}$ as the *outmigration* matrix and to $\boldsymbol{E}$ as the *inmigration* matrix. Intuitively, D_{ni} captures the importance of i as a destination for origin n, and E_{in} captures the importance of n as an origin for destination i. Again note the order of subscripts: in matrix $\boldsymbol{D}$, rows are origins and columns are destinations, whereas in matrix $\boldsymbol{E}$, rows are destinations and columns are origins.

For theoretical completeness, we maintain three assumptions on these matrices, which are satisfied empirically in all years of our data: **(i)** For any i, n, there exists k such that $\left[\boldsymbol{S}^k\right]_{in} > 0$

and $\left[\boldsymbol{D}^k\right]_{in} > 0$. **(ii)** For all i, $\boldsymbol{S}_{ii} > 0$ and $\boldsymbol{D}_{ii} > 0$. **(iii)** The matrices $\boldsymbol{S}$ and $\boldsymbol{D}$ are of rank $N - 1$. The first assumption states that all locations are connected with each other directly or indirectly, through flows of goods and migrants. The second assumption ensures that each location consumes a positive amount of domestic goods and has a positive amount of own migrants. The third assumption ensures that $N - 1$ columns of the trade and migration share matrices are linearly independent, with the final column determined by the requirement that the trade and migration shares sum to one.

3 Spatial Dynamics

We now linearize the model and derive closed-form solutions for the transition path of the spatial distribution of economic activity in response to productivity and amenity shocks. Our linearization has three main advantages for our empirical application. First, we are able to provide an analytical characterization of the properties of the economy's transition path, with for example the speed of convergence to steady-state depending on the eigenvalues of a transition matrix. Second, most existing research on spatial dynamics assumes perfect foresight. In contrast, our linearization allows us to solve for the economy's transition path under stochastic fundamentals and rational expectations, as well as under the conventional perfect foresight assumption. Third, we obtain our analytical characterization for constant intertemporal elasticity of substitution preferences, which are challenging to analyze in the non-linear model except in the special case of logarithmic preferences, because in general the equilibrium saving rate is determined as the solution to the fixed-point problem in Lemma 1.

In Section 3.1, we derive the linearized system of equations that characterizes the economy's transition path. In Section 3.2, we first illustrate our approach, by considering the special case of an economy that is on the transition path towards an initial steady-state and experiences a one-time, unexpected shock to fundamentals. We show the subsequent dynamic path of the economy can be characterized by an impact matrix ($\boldsymbol{R}$), which captures the impact of shocks in the initial period in which they occur, and a transition matrix ($\boldsymbol{P}$), which governs the evolution of the state variables from one period to the next in response to these shocks. We demonstrate that a spectral analysis of $\boldsymbol{P}$ and $\boldsymbol{R}$ provides a complete characterization of the model's transition dynamics: The eigenvalues of $\boldsymbol{P}$ govern the half-lives of convergence to steady-state; the eigenvectors reveal the locations exposed to particular shocks and the shocks that impact particular locations.

In Section 3.3, we show that our approach generalizes to an unexpected and convergent sequence of shocks to future fundamentals under perfect foresight. In Section 3.4, we show that our approach also generalizes to stochastic fundamentals, where agents form rational expectations about future fundamentals based on a known stochastic process for fundamentals. The derivation

for all results in this section is reported in Section B.8 of the online appendix.

One caveat is that our linearization is only exact in theory for small shocks. In principle, there could be important non-linearities for large shocks. In Section 6 below, we show that in practice our linearization provides a close approximation to the full non-linear model even for large shocks, such as the empirical distribution of shocks observed during our sample period.

3.1 Linearized Equilibrium Conditions

We now linearize the general equilibrium conditions of the model. We suppose that we observe the state variables $\{\boldsymbol{\ell}_t, \boldsymbol{k}_t\}$ and the trade and migration share matrices $\{\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}\}$ of the economy at time $t = 0$. The economy need not be in steady-state at $t = 0$, but we assume that it is on a convergence path towards a steady-state with constant fundamentals $\{\boldsymbol{z}, \boldsymbol{b}, \boldsymbol{\kappa}, \boldsymbol{\tau}\}$. We refer to the steady-state implied by these initial fundamentals as the *initial steady-state*. We use a tilde above a variable to denote a log deviation from the initial steady-state ($\widetilde{\chi}_{it+1} = \ln \chi_{it+1} - \ln \chi_i^*$) for all variables except for the worker value function, for which with a slight abuse of notation we use the tilde to denote a deviation in levels ($\widetilde{v}_{it} \equiv v_{it} - v_t^*$).

Totally differentiating the general equilibrium conditions around the implied initial steady-state, holding constant the aggregate labor endowment, trade costs and migration costs, we obtain the following system of equations that fully characterizes the economy's transition path in response to shocks to productivity and amenities:

$$\widetilde{\boldsymbol{k}}_{t+1} = \widetilde{\boldsymbol{k}}_t + (1 - \beta(1-\delta))\left(\widetilde{\boldsymbol{w}}_t - \widetilde{\boldsymbol{p}}_t - \widetilde{\boldsymbol{k}}_t + \widetilde{\boldsymbol{\ell}}_t\right) \tag{17}$$
$$+ (1 - \beta(1-\delta))\frac{1-\beta}{\beta}(\psi - 1)\mathbb{E}_t \sum_{s=1}^{\infty} \beta^s \left(\widetilde{\boldsymbol{w}}_{t+s} - \widetilde{\boldsymbol{p}}_{t+s} - \widetilde{\boldsymbol{k}}_{t+s} + \widetilde{\boldsymbol{\ell}}_{t+s}\right),$$

$$\widetilde{\boldsymbol{p}}_t = \boldsymbol{S}\left(\widetilde{\boldsymbol{w}}_t - \widetilde{\boldsymbol{z}}_t - (1-\mu)\left(\widetilde{\boldsymbol{k}}_t - \widetilde{\boldsymbol{\ell}}_t\right)\right), \tag{18}$$

$$[\boldsymbol{I} - \boldsymbol{T} + \theta(\boldsymbol{I} - \boldsymbol{T}\boldsymbol{S})]\widetilde{\boldsymbol{w}}_t = \left[-(\boldsymbol{I} - \boldsymbol{T})\widetilde{\boldsymbol{\ell}}_t + \theta(\boldsymbol{I} - \boldsymbol{T}\boldsymbol{S})\left(\widetilde{\boldsymbol{z}}_t + (1-\mu)\left(\widetilde{\boldsymbol{k}}_t - \widetilde{\boldsymbol{\ell}}_t\right)\right)\right], \tag{19}$$

$$\widetilde{\boldsymbol{\ell}}_{t+1} = \boldsymbol{E}\widetilde{\boldsymbol{\ell}}_t + \frac{\beta}{\rho}(\boldsymbol{I} - \boldsymbol{E}\boldsymbol{D})\mathbb{E}_t\widetilde{\boldsymbol{v}}_{t+1}, \tag{20}$$

$$\widetilde{\boldsymbol{v}}_t = \widetilde{\boldsymbol{w}}_t - \widetilde{\boldsymbol{p}}_t + \widetilde{\boldsymbol{b}}_t + \beta\boldsymbol{D}\mathbb{E}_t\widetilde{\boldsymbol{v}}_{t+1}. \tag{21}$$

In this system of first-order equations, there are no terms in the change in the trade and migration share matrices, because these terms are second-order in the underlying Taylor-series expansion, involving interactions between the changes in productivity and amenities and the changes in trade and migration shares. As we consider small changes in productivity and amenities, these second-order terms converge to zero. An important implication is that we can write the trade and migration share matrices with no time subscript ($\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}$), because they take

the same value for small changes in productivity and amenities, whether we use the initial values of the trade and migration share matrices at time $t = 0$, their values in the unobserved initial steady-state, or their values after adjustment to the shocks. In our empirical application, we use the initial values at time $t = 0$, because these are observed in the data.

From the capital accumulation condition (17), the capital stock in period $t + 1$ depends on current and future values of the real wage, the capital stock and population through the solution to landlords' consumption-investment problem. From the price index equation (18), the change in the cost of living component of the real wage depends on expenditure shares and the changes in wages, capital-labor ratios and productivities in each location. From the goods market clearing condition (19), the change in each location's per capita income depends on changes in population in all locations (first term) and changes in the price competitiveness of other locations's goods in all markets (second term). From the population flow equation (20), the change in each location's population depends on the changes in populations in all locations (first term) and the changes in expected values in all locations (second term). Finally, from the value function equation (20), the change in expected value in each location depends on the change in flow utility (first three terms) and the change in continuation value (last term).

The state variables in this system of equations are the population ($\widetilde{\ell}_t$) and capital stock ($\widetilde{k}_t$) in each location. The laws of motion for both population (20) and capital (17) are forward-looking: the future population in each location depends on both the current population and the expected future value of living in each location; similarly, the optimal investment decision of landlords depends on future returns to capital (note $\widetilde{R}_t = (1 - \beta(1 - \delta))\left(\widetilde{w}_t - \widetilde{p}_t - \widetilde{k}_t + \widetilde{\ell}_t\right)$). In the special case of log-utility ($\psi = 1$), landlords have a constant saving rate β out of income net of depreciation, as discussed above. Therefore, the capital equation (17) becomes backward-looking in this special case, because the future capital stock is a function of only the current period variables. The value function (21) is also forward-looking, because the expected value of living in each location depends on current-period state variables, the productivity and amenity shocks, and the expected future value of living in each location. Finally, the wage equation (19) and price index equation (18) both depend only on current-period state variables and the productivity and amenity shocks, through the static trade model.

3.2 Transition Dynamics for a One-time Shock

As a first illustration of our approach, we begin by solving for the economy's transition path in response to a one-time shock. We suppose that agents learn at time $t = 0$ about a one-time, unexpected, and permanent change in productivity and amenities from time $t = 1$ onwards. Under this assumption, we can write the sequence of future fundamentals (productivities and amenities) relative to the initial level as $\left(\widetilde{z}_t, \widetilde{b}_t\right) = \left(\widetilde{z}, \widetilde{b}\right)$ for $t \geq 1$, and we can drop the

expectation operator in the system of equations (17) through (21).

Substituting the wage equation (19) and the value function (21) into the labor dynamics equation (20) and the capital accumulation condition (17), we show in Section B.8.4 of the online appendix that we can reduce can reduce the model's transition dynamics to the following linear system of second-order difference equations in the state variables:

$$\boldsymbol{\Psi}\widetilde{\boldsymbol{x}}_{t+2} = \boldsymbol{\Gamma}\widetilde{\boldsymbol{x}}_{t+1} + \boldsymbol{\Theta}\widetilde{\boldsymbol{x}}_t + \boldsymbol{\Pi}\widetilde{\boldsymbol{f}}, \tag{22}$$

where $\widetilde{\boldsymbol{x}}_t = \begin{bmatrix} \widetilde{\boldsymbol{\ell}}_t \\ \widetilde{\boldsymbol{k}}_t \end{bmatrix}$ is a $2N \times 1$ vector of the state variables; $\widetilde{\boldsymbol{f}} = \begin{bmatrix} \widetilde{\boldsymbol{z}} \\ \widetilde{\boldsymbol{b}} \end{bmatrix}$ is a $2N \times 1$ vector of the shocks to fundamentals; and $\boldsymbol{\Psi}, \boldsymbol{\Gamma}, \boldsymbol{\Theta}$, and $\boldsymbol{\Pi}$ are $2N \times 2N$ matrices that depend only on the structural parameters of the model $\{\psi, \theta, \beta, \rho, \mu, \delta\}$ and the observed trade and migration share matrices $\{\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}\}$.

This matrix system of second-order difference equations (22) can be solved using standard techniques from the time-series macroeconomics literature. Using the method of undetermined coefficients following Uhlig (1999), we obtain a closed-form solution for the transition path of the entire spatial distribution of economic activity in terms of a transition matrix ($\boldsymbol{P}$), which governs the evolution of the state variables from one period to the next, and an impact matrix ($\boldsymbol{R}$), which captures the initial impact of the shocks to fundamentals.[6]

Proposition 2. Transition Path. *Suppose that the economy at time $t = 0$ is on a convergence path towards an initial steady-state with constant fundamentals ($\boldsymbol{z}, \boldsymbol{b}, \boldsymbol{\kappa}, \boldsymbol{\tau}$). At time $t = 0$, agents learn about one-time, permanent shocks to productivity and amenities ($\widetilde{\boldsymbol{f}} = \begin{bmatrix} \widetilde{\boldsymbol{z}} \\ \widetilde{\boldsymbol{b}} \end{bmatrix}$) from time $t = 1$ onwards. There exists a $2N \times 2N$ transition matrix ($\boldsymbol{P}$) and a $2N \times 2N$ impact matrix ($\boldsymbol{R}$) such that the second-order difference equation system in (22) has a closed-form solution of the form:*

$$\widetilde{\boldsymbol{x}}_{t+1} = \boldsymbol{P}\widetilde{\boldsymbol{x}}_t + \boldsymbol{R}\widetilde{\boldsymbol{f}} \quad \text{for } t \geq 0. \tag{23}$$

Proof. See Section B.8.4 of the online appendix. □

From Proposition 2, we have a closed-form solution for the evolution of the state variables in each location in all future time periods. Therefore, we can compute the time path of the entire spatial distribution of economic activity in response to productivity and amenity shocks, using only the initial observed values of the state variables $\{\boldsymbol{\ell}_t, \boldsymbol{k}_t\}$, the transition matrix ($\boldsymbol{P}$) and the impact matrix ($\boldsymbol{R}$). Both the transition matrix ($\boldsymbol{P}$) and impact matrix ($\boldsymbol{R}$) can be computed from our observed trade and migration share matrices ($\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}$) and the structural parameters of the model.

---

[6]Relative to the time-series macro literature, our dynamic spatial model features a larger state space of many locations or location-sectors over time. Nevertheless, the use of standard linear algebra techniques allows our approach to accommodate large state spaces, while remaining computationally efficient and easy to implement.

3.2.1 Convergence Dynamics Versus Fundamental Shocks

Before providing a further analytical characterization of the economy's transition path, we show that our linearization permits an additive decomposition of the observed change in the spatial distribution of economic activity across locations into the contributions of convergence to steady-state and shocks to fundamentals.

Although Proposition 2 states the law of motion of the state variables in terms of log-deviations relative to the initial steady-state, note that we can take first-differences between two time periods to eliminate the initial steady-state values of variables and obtain a law of motion in log changes between time periods (noting that $\Delta \ln x_{t+1} \equiv \ln x_{t+1} - \ln x_t = \widetilde{x}_{t+1} - \widetilde{x}_t$ and $\Delta \ln \boldsymbol{f}_1 \equiv \ln \boldsymbol{f}_1 - \ln \boldsymbol{f}_0 = \widetilde{\boldsymbol{f}}$):

$$\Delta \ln x_{t+1} = \begin{cases} \boldsymbol{P} \Delta \ln x_t & \text{if } t > 0, \\ \boldsymbol{P} \Delta \ln x_t + \boldsymbol{R} \Delta \ln \boldsymbol{f}_{t+1} & \text{if } t = 0. \end{cases} \tag{24}$$

Summing (24) across time periods, we thus obtain the following additive decomposition of the change in the state variables between years -1 and t into the contributions of convergence dynamics and shocks to fundamentals:

$$\ln \boldsymbol{x}_t - \ln \boldsymbol{x}_{-1} = \underbrace{\sum_{s=0}^{t} \boldsymbol{P}^s \left(\ln \boldsymbol{x}_0 - \ln \boldsymbol{x}_{-1} \right)}_{\substack{\text{convergence given} \\ \text{initial fundamentals}}} + \underbrace{\sum_{s=0}^{t-1} \boldsymbol{P}^s \boldsymbol{R} \widetilde{\boldsymbol{f}}}_{\substack{\text{dynamics from} \\ \text{fundamental shocks}}} \quad \text{for all } t \geq 1. \tag{25}$$

In the absence of any shocks to fundamentals ($\widetilde{\boldsymbol{f}} = \boldsymbol{0}$), the second term on the right-hand side of equation (25) is zero, and the state variables converge to:

$$\ln \boldsymbol{x}^*_{\text{initial}} = \lim_{t \to \infty} \ln \boldsymbol{x}_t = \ln \boldsymbol{x}_{-1} + \left(\boldsymbol{I} - \boldsymbol{P} \right)^{-1} \left(\ln \boldsymbol{x}_0 - \ln \boldsymbol{x}_{-1} \right), \tag{26}$$

where $\left(\boldsymbol{I} - \boldsymbol{P} \right)^{-1} = \sum_{s=0}^{\infty} \boldsymbol{P}^s$ is well-defined under the condition that the spectral radius of $\boldsymbol{P}$ is smaller than one, a property that we verify empirically.

A first implication is that we can solve for the unobserved steady-state in each year, given the observed state variables for two initial periods ($t = -1, 0$), the trade and migration share matrices for one initial period $\{\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}\}$, and the structural parameters $\{\psi, \theta, \beta, \rho, \mu, \delta\}$. While one can also solve for the unobserved steady-state by extending existing dynamic exact-hat algebra approaches to incorporate forward-looking capital investments, no additive decomposition analogous to (25) exists, because convergence dynamics and shocks to fundamentals interact through the non-linear structure of the model.

A second implication of Proposition 2 is that we can solve for the impact of shocks to fundamentals on the steady-state values of the state variables. To see this, note that if the economy is

initially in a steady-state at time 0, then the first term on the right-hand side of equation (25) is zero, and the path of the state variables follows:

$$\widetilde{x}_t = \ln x_t - \ln x_0 = \sum_{s=0}^{t-1} P^s R \widetilde{f} = \left(I - P^t\right)\left(I - P\right)^{-1} R \widetilde{f} \quad \text{for all } t \geq 1. \qquad (27)$$

In this case, the initial response is $\widetilde{x}_1 = R\widetilde{f}$, and, taking the limit as $t \to \infty$, we obtain the comparative steady-state response to fundamental shocks $\widetilde{f}$:

$$\lim_{t \to \infty} \widetilde{x}_t = \ln x_{\text{new}}^* - \ln x_{\text{initial}}^* = \left(I - P\right)^{-1} R \widetilde{f}, \qquad (28)$$

where we show that this comparative steady-state response can be written as the elasticity of the endogenous variables in each location with respect to a shock in any location in Proposition A.1 in Section B.8.3 in the online appendix.

A third implication is is that we can solve for the impact of counterfactual shocks to fundamentals on the state variables at various time horizons, even without data on the initial values of the state variables. Given any set of fundamental shocks ($\widetilde{f}$), the R matrix captures the initial impact of these shocks $\widetilde{f}$ on the state variables, and the P matrix governs the propagation of this impact over time, such that $\sum_{s=0}^{t-1} P^s R \widetilde{f}$ is the cumulative impact up till time t. These P and R matrices depend on only the observed trade and migration share matrices (S, T, D, E) and the structural parameters of the model. This property can be useful in settings where the initial values of the state variables (such as the capital stock in each location) are hard to observe. In contrast, solving for the impact of shocks to fundamentals in the non-linear model requires observed values for the initial state variables.

3.2.2 Spectral Analysis of the Transition Matrix P

We now provide a further analytical characterization of the economy's transition path. Our analysis so far in Proposition 2 is based on the entire transition matrix (P). We now show that we can further characterize the economy's transition path in terms of lower-dimensional components, namely the eigenvectors and eigenvalues of this transition matrix (P). Such a lower-dimensional representation is not possible within the full non-linear model, for which only numerical solutions for the transition path exist. We use this lower-dimensional representation to characterize the economy's speed of convergence to steady-state and the heterogeneous impact of local shocks.

In equation (25) above, we have already shown that we can decompose the dynamic path of the economy into one component capturing shocks to fundamentals and another component capturing convergence to the initial steady-state. Therefore, for the remainder of this subsection, we focus for expositional simplicity on an economy that is initially in steady-state.

Eigendecomposition of the Transition Matrix   We begin by undertaking an eigendecomposition of the transition matrix, $\boldsymbol{P} \equiv \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues arranged in decreasing order by absolute values, and $\boldsymbol{V} = \boldsymbol{U}^{-1}$. For each eigenvalue λ_k, the k-th column of $\boldsymbol{U}$ ($\boldsymbol{u_k}$) and the k-th row of $\boldsymbol{V}$ ($\boldsymbol{v'_k}$) are the corresponding right- and left-eigenvectors of $\boldsymbol{P}$, respectively, such that

$$\lambda_k \boldsymbol{u}_k = \boldsymbol{P}\boldsymbol{u}_k, \qquad \lambda_k \boldsymbol{v}'_k = \boldsymbol{v}'_k \boldsymbol{P}.$$

That is, $\boldsymbol{u}_k$ ($\boldsymbol{v}'_k$) is the vector that, when left-multiplied (right-multiplied) by $\boldsymbol{P}$, is proportional to itself but scaled by the corresponding eigenvalue λ_k.[7] We refer to $\boldsymbol{u}_k$ simply as eigenvectors. Both $\{\boldsymbol{u}_k\}$ and $\{\boldsymbol{v}'_k\}$ are bases that span the $2N$-dimensional vector space.

We next introduce a particular type of shock to productivity and amenities that proves useful for characterizing the model's transition dynamics. We define an *eigen-shock* as a shock to productivity and amenities ($\widetilde{\boldsymbol{f}}_{(k)}$) for which the initial impact of these shocks on the state variables ($\boldsymbol{R}\widetilde{\boldsymbol{f}}_{(k)}$) coincides with a real eigenvector of the transition matrix ($\boldsymbol{u}_k$). Assuming that the impact matrix is invertible, which we verify empirically, we can recover these eigen-shocks from the impact matrix ($\boldsymbol{R}$) and the eigenvectors of the transition matrix ($\boldsymbol{u}_k$), using $\widetilde{\boldsymbol{f}}_{(k)} = \boldsymbol{R}^{-1}\boldsymbol{u}_k$. Recall that both the impact matrix ($\boldsymbol{R}$) and the eigenvectors of the transition matrix ($\boldsymbol{u}_k$) can be computed using only our observed trade and migration share matrices ($\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}$) and the structural parameters of the model $\{\psi, \theta, \beta, \rho, \mu, \delta\}$. Therefore, we can solve for the eigen-shocks from these observed data and the structural parameters of the model.

Using our eigendecomposition and definition of an eigen-shock, we can undertake a spectral analysis of the economy's dynamic response to shocks. In particular, we can express the transition path of the state variables in response to any empirical productivity and amenity shocks ($\boldsymbol{f}$) in terms of a linear combination of the eigenvectors and eigenvalues of the transition matrix. Additionally, the weights or loadings in this linear combination can be recovered from a linear projection (regression) of the observed shocks ($\widetilde{\boldsymbol{f}}$) on the eigen-shocks ($\widetilde{\boldsymbol{f}}_{(k)}$).

Proposition 3. *Spectral Analysis. Consider an economy that is initially in steady-state at time $t = 0$ when agents learn about one-time, permanent shocks to productivity and amenities ($\widetilde{\boldsymbol{f}} = \begin{bmatrix} \widetilde{\boldsymbol{z}} \\ \widetilde{\boldsymbol{b}} \end{bmatrix}$) from time $t = 1$ onwards. The transition path of the state variables can be written as a linear combination the eigenvalues (λ_k) and eigenvectors ($\boldsymbol{u}_k$) of the transition matrix:*

$$\widetilde{\boldsymbol{x}}_t = \sum_{s=0}^{t-1} \boldsymbol{P}^s \boldsymbol{R}\widetilde{\boldsymbol{f}} = \sum_{k=1}^{2N} \frac{1 - \lambda_k^t}{1 - \lambda_k} \boldsymbol{u}_k \boldsymbol{v}'_k \boldsymbol{R}\widetilde{\boldsymbol{f}} = \sum_{k=1}^{2N} \frac{1 - \lambda_k^t}{1 - \lambda_k} \boldsymbol{u}_k a_k, \tag{29}$$

---

[7] Note that $\boldsymbol{P}$ need not be symmetric. This eigendecomposition can be undertaken as long the transition matrix has distinct eigenvalues, a condition that we verify empirically. We construct the right-eigenvectors such that the 2-norm of $\boldsymbol{u}_k$ is equal to 1 for all k, where note that $\boldsymbol{v}'_i \boldsymbol{u}_k = 1$ for $i = k$ and $\boldsymbol{v}'_i \boldsymbol{u}_k = 0$ otherwise.

where the weights in this linear combination (a_k) can be recovered as the coefficients in a linear projection (regression) of the observed shocks ($\widetilde{\boldsymbol{f}}$) on the eigen-shocks ($\widetilde{\boldsymbol{f}}_{(k)}$).

Proof. The proposition follows from the eigendecomposition of the transition matrix: $\boldsymbol{P} \equiv \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}$, as shown in Section B.8.7 of the online appendix. $\square$

Proposition 3 is one of our main analytical results and plays a key role in our empirical analysis. We now show how it can be used to characterize both the speed of convergence to steady-state and the heterogeneous impact of shocks across locations.

Speed of Convergence We measure the speed of convergence to steady-state using the conventional measure of the half-life. In particular, we define the half-life of a shock $\widetilde{\boldsymbol{f}}$ for the i-th state variable as the time it takes for that state variable to converge half of the way to steady-state:

$$t_i^{(1/2)}\left(\widetilde{\boldsymbol{f}}\right) \equiv \arg\max_{t\in\mathbb{Z}_{>0}} \frac{|\widetilde{x}_{it} - \widetilde{x}_{i\infty}|}{\max_s |\widetilde{x}_{is} - \widetilde{x}_{i\infty}|} \geq \frac{1}{2},$$

where $\widetilde{x}_{i\infty} = x^*_{i,\text{new}} - x^*_{i,\text{initial}}$.

We begin by considering the speed of convergence for eigen-shocks, for which the initial impact on the state variables corresponds to a real eigenvector of the transition matrix ($\widetilde{\boldsymbol{f}}_{(k)} = \boldsymbol{R}^{-1}\boldsymbol{u}_k$). For these eigen-shocks, the state variables converge exponentially towards steady-state, and the speed of convergence depends solely on the corresponding eigenvalue (λ_k).

Proposition 4. Speed of Convergence. *Consider an economy that is initially in steady-state at time $t = 0$ when agents learn about one-time, permanent shocks to productivity and amenities ($\widetilde{\boldsymbol{f}} = \begin{bmatrix} \widetilde{\boldsymbol{z}} \\ \widetilde{\boldsymbol{b}} \end{bmatrix}$) from time $t = 1$ onwards. Suppose that these shocks are an eigen-shock ($\widetilde{\boldsymbol{f}}_{(k)}$), for which the initial impact on the state variables at time $t = 1$ coincides with a real eigenvector ($\boldsymbol{u}_k$) of the transition matrix ($\boldsymbol{P}$): $\boldsymbol{R}\widetilde{\boldsymbol{f}}_{(k)} = \boldsymbol{u}_k$. The transition path of the state variables ($\boldsymbol{x}_t$) in response to such an eigen-shock ($\widetilde{\boldsymbol{f}}_{(k)}$) is:*

$$\widetilde{\boldsymbol{x}}_t = \sum_{j=1}^{2N} \frac{1-\lambda_j^t}{1-\lambda_j} \boldsymbol{u}_j \boldsymbol{v}_j' \boldsymbol{u}_k = \frac{1-\lambda_k^t}{1-\lambda_k}\boldsymbol{u}_k \implies \ln\boldsymbol{x}_{t+1} - \ln\boldsymbol{x}_t = \lambda_k^t\boldsymbol{u}_k,$$

and the half-life is given by:

$$t_i^{(1/2)}\left(\widetilde{\boldsymbol{f}}\right) = -\left\lceil \frac{\ln 2}{\ln \lambda_k} \right\rceil$$

for all state variables $i = 1, \cdots, 2N$, where $\lceil \cdot \rceil$ is the ceiling function.

Proof. The proposition follows from the eigendecomposition of the transition matrix ($\boldsymbol{P} \equiv \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}$), for the case in which the initial impact of the shocks to productivity and amenities on the state variables at time $t = 1$ coincides with a real eigenvector ($\boldsymbol{R}\widetilde{\boldsymbol{f}}_{(k)} = \boldsymbol{u}_k$) of the transition matrix ($\boldsymbol{P}$), as shown in Section B.8.7 of the online appendix. $\square$

From Proposition 4, the impact of eigen-shocks ($\widetilde{\boldsymbol{f}}_{(k)}$) on the state variables in each time period is always proportional to the corresponding eigenvector ($\boldsymbol{u}_k$), and decays exponentially at a rate determined by the associated eigenvalue (λ_k), as the economy converges to the new steady-state.[8] These eigenvalues fully summarize the economy's dynamic response to eigen-shocks, even in our setting with a high-dimensional state space, a rich geography of trade and migration costs, and multiple sources of dynamics from investment and migration.

Each eigen-shock ($\widetilde{\boldsymbol{f}}_{(k)}$) has a different speed of convergence (as captured by the associated eigenvalue λ_k), because this speed of convergence to steady-state depends not only on the structural parameters $\{\psi, \theta, \beta, \rho, \mu, \delta\}$, but also on the incidence of the shock on the population and capital stock in each location (as captured by the elements of the associated eigenvector $\boldsymbol{u}_k = \boldsymbol{R}\widetilde{\boldsymbol{f}}_{(k)}$). From Proposition 3, any empirical shock ($\widetilde{\boldsymbol{f}}$) can be expressed as a linear combination of the eigen-shocks, which hence implies that the speed of convergence will also vary across empirical shocks, depending on their incidence on the population and capital stock in each location.

Heterogeneous Impact of Shocks   Using our spectral analysis in Proposition 3, we can decompose the observed evolution of the state variables ($\widetilde{\boldsymbol{x}}_t$) in response to any empirical productivity and amenity shocks ($\widetilde{\boldsymbol{f}}$) into the contributions of each eigencomponent ($\boldsymbol{u}_k$).

Using this decomposition, we can distinguish shocks from exposure to these shocks. Each eigencomponent can be interpreted as the response of the state variables to a different type of shock ($\boldsymbol{u}_k = \boldsymbol{R}^{-1}\widetilde{\boldsymbol{f}}_{(k)}$). Different empirical productivity and amenity shocks ($\widetilde{\boldsymbol{f}}$) place different weights on these eigencomponents (as captured by the loadings, a_k). The state variables of the economy (the population and the capital stock in each location) are unevenly exposed to these different types of shocks. The exposure of the state variable in location i to the k-th eigen-shock is given by the relevant element of the eigencomponent (u_{ik}), while its exposure to any empirical shock $\widetilde{\boldsymbol{f}}$ depends on the products of the relevant elements of each eigencomponent (u_{ik}) and the weights or loadings on the different eigencomponents (a_k).

The impact of these different shocks on the state variables evolves over time, because different eigencomponents ($\boldsymbol{u}_k$) have different associated eigenvalues (λ_k), and hence different speeds of convergence to steady-state. Therefore, we can use our spectral analysis to isolate different types of shocks (as captured by different eigencomponents), the exposure of the state variables to these different types of shocks (as captured by the elements of these eigencomponents and the loadings), and the time horizons over which these effects occur (as captured by the eigenvalues).

---

[8]In general, these eigenvectors and eigenvalues can be complex-valued. If the initial impact is the real part of a complex eigenvector $\boldsymbol{u}_k$ ($\boldsymbol{R}\widetilde{\boldsymbol{f}} = \mathrm{Re}\,(\boldsymbol{u}_k)$), then $\ln \boldsymbol{x}_{t+1} - \ln \boldsymbol{x}_t = \mathrm{Re}\,(\lambda_k^t \boldsymbol{u}_k) \neq \mathrm{Re}\,(\lambda_k) \cdot \mathrm{Re}\,(\lambda_k^{t-1}\boldsymbol{u}_k)$. That is, the impact no longer decays at a constant rate λ_k. Instead, the complex eigenvalues introduce oscillatory motion as the dynamical system converges to the new steady-state. For expositional purposes, we focus on real-valued eigenvalues and eigenvectors. In our empirical application, the imaginary components of $\boldsymbol{P}$'s eigenvalues are small, implying that oscillatory effects are small relative to the effects that decay exponentially.

As such, this decomposition has both similarities and differences with empirical shift-share Bartik decompositions. A key similarity is that both distinguish shocks and exposure to these shocks. A first key difference is that our decomposition is derived from our closed-form solution for transition path of the spatial distribution of economic activity in the general equilibrium of our model. A second key difference is that our decomposition is dynamic, and hence evolves over time, as the economy gradually adjusts to the shock.

3.3 Convergent Sequence of Shocks Under Perfect Foresight

We now generalize our analysis of the model's transition dynamics to any convergent sequence of future shocks to productivities and amenities under perfect foresight. In particular, we consider an economy that is somewhere on a convergence path towards an initial steady-state with constant fundamentals at time $t = 0$, when agents learn about a convergent sequence of future shocks to productivity and amenities $\left\{ \widetilde{f}_s \right\}_{s \geq 1}$ from time $t = 1$ onwards, where $\widetilde{f}_s$ is a vector of log differences in fundamentals between times s and 0 for each location.

Proposition 5. *Sequence of Shocks Under Perfect Foresight. Consider an economy that is somewhere on a convergence path towards steady-state at time $t = 0$, when agents learn about a convergent sequence of future shocks to productivity and amenities $\left\{ \widetilde{f}_s \right\}_{s \geq 1} = \left\{ \begin{bmatrix} \widetilde{z}_s \\ \widetilde{b}_s \end{bmatrix} \right\}_{s \geq 1}$ from time $t = 1$ onwards. There exists a $2N \times 2N$ transition matrix $(\boldsymbol{P})$ and a $2N \times 2N$ impact matrix $(\boldsymbol{R})$ such that the second-order difference equation system in (22) has a closed-form solution of the form:*

$$\widetilde{x}_t = \sum_{s=t+1}^{\infty} \left(\boldsymbol{\Psi}^{-1} \boldsymbol{\Gamma} - \boldsymbol{P} \right)^{-(s-t)} \boldsymbol{R} \left(\widetilde{f}_s - \widetilde{f}_{s-1} \right) + \boldsymbol{R} \widetilde{f}_t + \boldsymbol{P} \widetilde{x}_{t-1} \qquad \text{for all } t \geq 1, \qquad (30)$$

with initial condition $\widetilde{x}_0 = \boldsymbol{0}$ and where $\boldsymbol{\Psi}, \boldsymbol{\Gamma}$ are matrices from the second-order difference equation (22) and are derived in Section B.8.7 of the online appendix.

Proof. See Section B.8.7 of the online appendix. □

Therefore, even though we consider a general convergent sequence of shocks to productivity and amenities in a setting with many locations connected by a rich geography, and with multiple sources of dynamics from investment and migration, we are again able to obtain a closed-form solution for the transition path of the spatial distribution of economic activity. Our linearization allows us to take log deviations between the state variables at each future point in time $s \geq 1$ and the initial steady-state at time $t = 0$, taking into account the intervening changes in fundamentals between those times. The closed-form solution in equation (30) holds regardless of the economy's initial distance from steady-state at time $t = 0$, which is captured in the lagged log difference in the state variables from the initial steady-state ($\widetilde{x}_{t-1}$).

Proposition 5 encompasses a number of special cases. First, the permanent shock to productivity and amenities in period 1 considered in the previous subsection is naturally a special case of the convergent sequence of shocks to fundamentals considered here: if $\widetilde{\boldsymbol{f}}_s = \widetilde{\boldsymbol{f}}$ for all $s \geq 1$, equations (30) and (23) coincide. Second, we encompass a permanent shock to productivity and amenities in any period $\hat{t} \geq 1$, with $\widetilde{\boldsymbol{f}}_s = \widetilde{\boldsymbol{f}}$ for $s \geq \hat{t}$ and zero before $\hat{t}$. In this case, the dynamic evolution of the state variables in equation (30) simplifies to

$$
\widetilde{\boldsymbol{x}}_t =
\begin{cases}
\boldsymbol{R}\widetilde{\boldsymbol{f}}_t + \boldsymbol{P}\widetilde{\boldsymbol{x}}_{t-1} & t \geq \hat{t}, \\
(\boldsymbol{\Psi}^{-1}\boldsymbol{\Gamma} - \boldsymbol{P})^{-(\hat{t}-t)} \, \boldsymbol{R}\widetilde{\boldsymbol{f}} + \boldsymbol{P}\widetilde{\boldsymbol{x}}_{t-1} & t < \hat{t}.
\end{cases}
$$

The transition matrix $\boldsymbol{P}$ and impact matrix $\boldsymbol{R}$ are exactly the same as those in the previous subsection, and can be recovered from our observed trade and migration share matrices $\{\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{D}, \boldsymbol{E}\}$ and the structural parameters of the model $\{\psi, \theta, \beta, \rho, \mu, \delta\}$.

3.4   Stochastic Location Characteristics and Rational Expectations

While we have illustrated our approach in the previous two subsections under the assumption of perfect foresight, we now return to the more general specification in Section 2, in which agents form rational expectations about stochastic fundamentals. In this more general case, we show that we can use a similar approach to solve in closed-form for the model's transition dynamics, given an assumed stochastic process for fundamentals. In particular, using $\Delta \ln$ to denote log changes between two periods (e.g. $\Delta \ln z_{it} \equiv \ln z_{it} - \ln z_{it-1}$), we assume the following AR(1) process for fundamentals, which allows shocks to productivity and amenities to have permanent effects on the level of these variables, and hence to affect the steady-state equilibrium:

$$
\begin{aligned}
\Delta \ln z_{it+1} &= \rho^z \Delta \ln z_{it} + \varpi_{it}^z, & |\rho^z| &< 1, \\
\Delta \ln b_{it-1} &= \rho^b \Delta \ln b_{it} + \varpi_{it}^b, & |\rho^b| &< 1,
\end{aligned}
\tag{31}
$$

where $\rho^z = \rho^b = 0$ corresponds to the special case in which fundamentals follow a random walk; and ϖ_{it}^z and ϖ_{it}^b are mean zero and independently and identically distributed innovations.

The linearization methods that we used to derive closed-form solutions in the previous subsections are taken from the Dynamic Stochastic General Equilibrium (DSGE) literature in macroeconomics. Therefore, they continue to apply with stochastic fundamentals, and we can adapt our earlier results to compute the entire transition path of the *expected* value of the models' endogenous variables over time.

Upon observing fundamental shocks at time t, agents form rational expectations about the future evolution of fundamentals. Given our assumed AR(1) process for fundamentals in equation (31), agents expect future shocks to fundamentals to decay geometrically to zero over time.

Therefore, we can write the expected values of these future shocks to fundamentals as:

$$\mathbb{E}_t \left[\Delta \ln \boldsymbol{f}_{t+s} \right] = \boldsymbol{N}^s \Delta \ln \boldsymbol{f}_t, \tag{32}$$

where $\boldsymbol{N} \equiv \begin{bmatrix} \rho^z \cdot \boldsymbol{I}_{N \times N} & \boldsymbol{0}_{N \times N} \\ \boldsymbol{0}_{N \times N} & \rho^b \cdot \boldsymbol{I}_{N \times N} \end{bmatrix}$ and $\mathbb{E}_t \left[\cdot \right]$ is the expectation conditional on the realizations of shocks up to time t.

Using this property, we can extend our earlier results to obtain a closed-form solution for the economy's expected transition path under stochastic fundamentals and rational expectations. In particular, consider an economy at time t with state variables $\boldsymbol{x}_t$ and fundamentals $\boldsymbol{f}_t$. To first-order, the changes in the state variables between periods t and $t+1$ reflect both convergence to the steady-state implied by time-t fundamentals and the convergent sequence of expected changes in fundamentals from equation (32).

Proposition 6. Stochastic Fundamentals and Rational Expectations. *Suppose that productivity and amenities evolve stochastically according to the AR(1) process in equation (31) and agents have rational expectations. There exists a $2N \times 2N$ transition matrix ($\boldsymbol{P}$) and a $2N \times 2N$ impact matrix ($\boldsymbol{R}$) such that the evolution of the economy's state variables (x_t) has the following closed-form solution:*

$$\Delta \ln \boldsymbol{x}_{t+1} = \boldsymbol{P} \Delta \ln \boldsymbol{x}_t + \boldsymbol{R} \Delta \ln \boldsymbol{f}_t + \sum_{s=0}^{\infty} \left(\boldsymbol{\Psi}^{-1} \boldsymbol{\Gamma} - \boldsymbol{P} \right)^{-s} \boldsymbol{R} \boldsymbol{N}^{s+1} \left(\Delta \ln \boldsymbol{f}_t - \Delta \ln \boldsymbol{f}_{t-1} \right). \tag{33}$$

Proof. See Section B.8.7 of the online appendix. □

Proposition 6 characterizes the equilibrium stochastic process of the endogenous state variables. For each period t, the population and capital stock in each location in period $t + 1$ (x_{t+1}) are determined by past state variables (x_t, x_{t-1}), realizations of fundamental shocks in period t ($\Delta \ln \boldsymbol{f}_t$), and expectations of future fundamental shocks. Combining Proposition 6 with equation (32) for expected future fundamentals, we can also solve for the expected future path of the state variables $\mathbb{E}_t \left[\ln \boldsymbol{x}_{t+s} \right]$ for all t and s.

Proposition 6 relates closely to our earlier results for perfect foresight in the previous two subsections. The first two terms on the right-hand-side of equation (33) capture how the state variables would evolve if location fundamentals realized at time t were expected to be permanent, as in Proposition 2.[9] This coincides with the solution under rational expectations if location fundamentals follow a random walk. In this case, both autoregressive coefficients ρ^z and ρ^b are equal to zero, and $\boldsymbol{N}$ is a zero matrix. The final term, which is a summation over all future periods, captures the impact of the sequence of future fundamental shocks expected at time t,

---

[9]Note that we can take the first-difference of log-deviations from steady-state to obtain log-changes, as in Section 3.2.1 above, where $\Delta \ln x_t = \widetilde{x}_t - \widetilde{x}_{t-1}$.

relative to the sequence expected in the previous period $t-1$. This last term can be compared with the corresponding terms in Proposition 5, where the sequence of future fundamental shocks is known under perfect foresight. Intuitively, the innovations in fundamental shocks at time t (ϖ_{it}^z and ϖ_{it}^b in equation (31)) affect the current-period state variables ($\boldsymbol{\ell}_t$, $\boldsymbol{k}_t$) in ways similar to unexpected "MIT" shocks. These innovations in fundamental shocks at time t also affect the entire sequence of future fundamental shocks expected at time t, because of serial correlation in fundamental shocks (ρ^z and ρ^b not equal to zero).

Therefore, our linearization allows us to relax the assumption of perfect foresight commonly maintained in quantitative spatial research to allow for stochastic fundamentals and rational expectations, and yet still solve analytically for the transition path of the expected spatial distribution of economic activity. In contrast, solving the full non-linear model for this case of stochastic fundamentals and rational expectations is considerably more challenging, because there is no analytical solution for the mapping from expected future fundamental shocks to expected future state variables.

4   Extensions

We now show that our approach accommodates a large number of extensions and generalizations. In Subsection 4.1, we incorporate shocks to trade and migration costs. In Subsection 4.2, we allow productivity and amenities to have an endogenous component that reflects agglomeration forces, as well as an exogenous component of fundamentals. In Subsection 4.3, we discuss extensions with multiple final goods sectors (Costinot et al. 2012) and input-output linkages (Caliendo and Parro 2015 and Caliendo et al. 2019). In Subsection 4.4, we discuss a number of additional extensions, including trade deficits and residential capital (housing).

4.1   Shocks to Trade and Migration Costs

In this subsection, we show that our analysis naturally incorporates shocks to trade and migration costs, where the derivation for all results in this subsection is reported in Section D.1 of the online appendix. Whereas productivity and amenity shocks are common across all partner locations, trade and migration cost shocks are bilateral, which implies that our comparative static results now have a representation as a three tensor. To reduce these three tensors down to a matrix (two tensor) representation, we aggregate bilateral trade and migration shocks across partner locations, using the appropriate weights implied by the model.

In particular, we define two measures of outgoing and incoming trade costs, which are trade-share weighted averages of the bilateral trade costs across all export destinations and import sources, respectively. We define *outgoing* trade costs for location i as $\ln \tau_{it}^{\text{out}} \equiv \sum_{n=1}^{N} T_{int} \ln \tau_{nit}$,

where the weights are the income share (T_{int}) that location i derives from selling to each export destination n. We specify *incoming* trade costs for location n as $\ln \tau_{nt}^{\text{in}} \equiv \sum_{i=1}^{N} S_{nit} \ln \tau_{nit}$, where the weights are the expenditure share (S_{nit}) that location n devotes to each import source i. Similarly, we define *outgoing* migration costs for location i as $\ln \kappa_{it}^{\text{out}} \equiv \sum_{n=1}^{N} D_{int} \ln \kappa_{nit}$, where the weights are the outmigration shares from location i to each destination n. We specify *incoming* migration costs for location n as $\ln \kappa_{nt}^{\text{in}} \equiv \sum_{i=1}^{N} E_{nit} \ln \kappa_{nit}$, where the weights are the inmigration shares (E_{nit}) to location n from each origin i.

Using these definitions of outgoing and incoming trade and migration costs, we obtain an analogous system of equations for the model's transition dynamics:

$$\widetilde{\boldsymbol{k}}_{t+1} = \begin{bmatrix} [\beta(1-\delta)\,\boldsymbol{I} + (1-\beta(1-\delta))(1-\mu)\,\boldsymbol{S}]\left(\widetilde{\boldsymbol{k}}_t - \widetilde{\boldsymbol{\ell}}_t\right) - (1-\beta(1-\delta))\,\widetilde{\boldsymbol{\tau}}^{\text{in}} \\ + (1-\beta(1-\delta))(\boldsymbol{I} - \boldsymbol{S})\,\widetilde{\boldsymbol{w}}_t + (1-\beta(1-\delta))\,\boldsymbol{S}\widetilde{\boldsymbol{z}} + \widetilde{\boldsymbol{\ell}}_t \end{bmatrix}, \quad (34)$$

$$[\boldsymbol{I} - \boldsymbol{T} + \theta(\boldsymbol{I} - \boldsymbol{TS})]\,\widetilde{\boldsymbol{w}}_t = \begin{bmatrix} \theta(\boldsymbol{I} - \boldsymbol{TS})\left(\widetilde{\boldsymbol{z}} + (1-\mu)\left(\widetilde{\boldsymbol{k}}_t - \widetilde{\boldsymbol{\ell}}_t\right)\right) \\ - (\boldsymbol{I} - \boldsymbol{T})\,\widetilde{\boldsymbol{\ell}}_t + \theta\,[\boldsymbol{T}\widetilde{\boldsymbol{\tau}}^{\text{in}} - \widetilde{\boldsymbol{\tau}}^{\text{out}}] \end{bmatrix}, \quad (35)$$

$$\widetilde{\boldsymbol{\ell}}_{t+1} = \boldsymbol{E}\widetilde{\boldsymbol{\ell}}_t + \frac{\beta}{\rho}(\boldsymbol{I} - \boldsymbol{ED})\,\widetilde{\boldsymbol{v}}_{t+1} - \frac{1}{\rho}\left(\widetilde{\boldsymbol{\kappa}}^{\text{in}} - \boldsymbol{E}\widetilde{\boldsymbol{\kappa}}^{\text{out}}\right), \quad (36)$$

$$\widetilde{\boldsymbol{v}}_t = \begin{bmatrix} (\boldsymbol{I} - \boldsymbol{S})\,\widetilde{\boldsymbol{w}}_t + \boldsymbol{S}\widetilde{\boldsymbol{z}} + (1-\mu)\,\boldsymbol{S}\left(\widetilde{\boldsymbol{k}}_t - \widetilde{\boldsymbol{\ell}}_t\right) - \widetilde{\boldsymbol{\tau}}^{\text{in}} + \widetilde{\boldsymbol{b}} - \widetilde{\boldsymbol{\kappa}}^{\text{out}} + \beta\boldsymbol{D}\widetilde{\boldsymbol{v}}_{t+1} \end{bmatrix}, \quad (37)$$

where recall that we use a tilde above a variable to denote a log deviation from the initial steady-state. Note that equations (17), (19), (20) and (21) in our baseline specification above correspond to the special case of equations (36)-(37) in which $\widetilde{\boldsymbol{\tau}}^{in} = \widetilde{\boldsymbol{\tau}}^{out} = \widetilde{\boldsymbol{\kappa}}^{in} = \widetilde{\boldsymbol{\kappa}}^{out} = \boldsymbol{0}$.

4.2 Agglomeration Forces

In this subsection, we generalize our baseline specification to introduce agglomeration forces, where the derivation for all results in this subsection is reported in Section D.2 of the online appendix. We allow productivity and amenities to have both exogenous and endogenous components. The exogenous component captures locational fundamentals, such as climate and access to natural water, while the endogenous component captures agglomeration forces. Following the standard approach in the economic geography literature, we model these agglomeration forces as constant elasticity functions of a location's own population: $z_{it} = \bar{z}_{it}\ell_{it}^{\eta^z}$ and $b_{it} = \bar{b}_{it}\ell_{it}^{\eta^b}$, where $\eta^z > 0$ and $\eta^b > 0$ parameterize the strength of agglomeration forces for productivity and amenities, respectively.[10]

---

[10]Although for simplicity we assume that agglomeration and dispersion forces only depend on a location's own population, our framework can be further generalized to incorporate spillovers across locations, as in Ahlfeldt et al. (2015) and Allen et al. (2020). While we focus on agglomeration forces ($\eta^z > 0$ and $\eta^b > 0$), it is straightforward to also allow for additional dispersion forces ($\eta^z < 0$ and $\eta^b < 0$).

In this extension, the general equilibrium conditions of the model remain the same as in Section 2.6 above, substituting for the terms in productivity and amenities (z_{it} and b_{it}) using the terms in location fundamentals and agglomeration forces ($z_{it} = \bar{z}_{it}\ell_{it}^{\eta^z}$ and $b_{it} = \bar{b}_{it}\ell_{it}^{\eta^b}$). The introduction of these agglomeration forces magnifies the impact of exogenous differences in fundamentals on the spatial distribution of economic activity. Furthermore, depending on the strength of these agglomeration forces, there can either be a unique steady-state equilibrium or multiple steady-state equilibria for the spatial distribution of economic activity. Again we obtain an analytical characterization of the conditions for the existence and uniqueness of the general equilibrium of the model, as summarized in the following proposition.

Proposition 7. **(A)** *There exists a unique steady-state spatial distribution of economic activity* $\{w_i^*, v_i^*, \ell_i^*, k_i^*\}$ *(up to a choice of numeraire) given the exogenous fundamentals* $\{\bar{z}_i, \bar{b}_i, \tau_{ni}, \kappa_{ni}\}$ *if the largest absolute value of the following eigenvalues is less than one.*

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \frac{1 - \frac{\beta\eta^b}{\rho}}{\frac{\beta}{\rho}\left(1 - \eta^b - \eta^z\theta - \eta^b\mu\theta\right) + (1 + \mu\theta)} \\ \frac{(1 + \mu\theta)}{\frac{\beta}{\rho}\left(1 - \eta^b - \eta^z\theta - \eta^b\mu\theta\right) + (1 + \mu\theta)} \end{pmatrix}.$$

(B) *For sufficiently small agglomeration forces* ($\eta^b + \eta^z\theta + \eta^b\mu\theta < 1$), *the largest absolute value of these eigenvalues is necessarily less than one.*

Proof. See Section D.2 of the web appendix. $\square$

In general, from part (A) of Proposition 7, whether there is an unique steady-state in the model depends on the strength of agglomeration forces (η^z, η^b), trade elasticity (θ), capital intensity (μ), the discount rate (β) and the dispersion of idiosyncratic preferences (ρ). From part (B) of Proposition 7, a sufficient condition for the existence of a unique steady-state is that the agglomeration forces (η^z, η^b) are sufficiently small. As these parameters converge towards zero, we obtain our baseline specification without agglomeration forces, in which there is necessarily a unique steady-state equilibrium, as shown in Section 2.6 above. Additionally, an increase in the dispersion of idiosyncratic preferences (a larger value for ρ) necessarily increases the range of values for the other parameters for which a unique steady-state equilibrium exists, as shown in the proof of the proposition.

After incorporating the additional agglomeration parameters in the model's general equilibrium conditions in Section 2.6 above, we obtain analogous closed-form solutions for the transition path of the spatial distribution of economic activity in terms of a transition matrix (P) and an impact matrix (R), as in Proposition 2 for our baseline specification above.

4.3 Multi-sectors and Input-Output Linkages

Our sufficient statistics also extend to environments with multiple sectors and input-output linkages. We briefly discuss these extensions here and derive them in the online appendix.

As in our baseline specification in Section 2, we assume that capital is geographically immobile across locations once installed. For the multi-sector model, we consider two different assumptions about capital mobility across sectors. In Section D.3 of the online appendix, we assume that installed capital is specific to a location, but mobile across sectors within locations. In Section D.4 of the online appendix, we assume installed capital is specific to both a location and sector. This second specification corresponds to a dynamic spatial version of the traditional specific-factors model, in which there are migration frictions for the mobile factor across locations and sectors, and there is endogenous accumulation of the factor specific to each location and sector over time. In both cases, we assume Cobb-Douglas preferences across sectors and CES preferences across the goods supplied by locations within sectors, as in Costinot et al. (2012).

In Section D.5 of the online appendix, we further generalize these multi-sector specifications to allow for input-output linkages, where the production technology in each sector now uses the two primary factors of labor and capital together with intermediate inputs according to a Cobb-Douglas functional form, as in Caliendo and Parro (2015).

In all of these specifications, consumption, production, trade and migration are modelled in a similar way as in our baseline, single-sector specification. The set of equilibrium conditions for the value function (14), population flow (15), and goods market clearing (12) extends naturally from describing each location to describing each location-sector, and the capital market clearing and accumulation condition (11) generalizes in a form that depends on the capital mobility assumption. We again obtain a closed-form solution for the transition path of the entire spatial distribution of economic activity in terms of a transition matrix (P) and an impact matrix (R), as in Proposition 2 for our baseline specification above.

In these multi-sector extensions, the trade and migration share matrices are defined from location-sector to location-sector. Additionally, in the input-output specification, the expenditure share (S) and income share (T) matrices must be adjusted to take into account the network structure of production. First, the gross value of trade from exporter i to importer n in industry k includes not only the direct value-added created in this exporter and sector but also the indirect value added created in previous stages of production. Second, the effect of a productivity shock in one location on any other location now differs depending on the extent to which it reduces input prices (and hence production costs) or reduces output prices.

4.4 Other Extensions

Our approach also accommodates a number of other extensions and generalizations. In our baseline specification, we model trade between locations as in Armington (1969), in which goods are differentiated by location of origin. In Section C of the online appendix, we establish a number of isomorphisms, in which we show that our results hold throughout the class of trade models with a constant trade elasticity. In Section D.6 of the online appendix, we incorporate trade deficits following the conventional approach of the quantitative international trade literature of treating these deficits as exogenous. In Section D.7 of the online appendix, we allow capital to be used residentially (for housing) as well as commercially (in production). In each case, we derive an analogous closed-form solution for the economy's transition path in terms of a transition matrix (P) and an impact matrix (R), as in Proposition 2 for our baseline specification above.

5 Quantitative Analysis

In this section, we report our main empirical results for the dynamics of the spatial distribution of economic activity across U.S. states from 1965-2015. We choose U.S. states as our spatial units, because of the availability of data on bilateral shipments of goods, bilateral migration flows and capital stocks over this long historical time period, and because of the substantial observed changes in the spatial distribution of economic activity over time. For the same reasons, we focus for most of our empirical analysis on a version of our baseline single-sector model, augmented to take account of the empirically-relevant distinction between traded and non-traded goods.[11] To examine the extent to which sectoral specialization influences exposure to shocks, we also implement our multi-sector extension with location-sector-specific capital from Section 4.3 for the shorter time period from 1999-2015 for which the sector-level data are available.

In Subsection 5.1, we introduce our data sources and definitions. In Subsection 5.2, we discuss the parameterization of our model. In Subsection 5.3, we provide reduced-form evidence on the substantial reorientation of economic activity across U.S. states over our sample period, including the decline of the Rust Belt and rise of the Sun Belt. In Subsection 5.4, we use our additive decomposition of the economy's transition path to examine the respective contributions of convergence to steady-state and fundamental shocks. In Subsection 5.5, we use our spectral analysis to provide evidence on the speed of convergence to steady-state, the respective contributions of capital and labor dynamics, and the heterogeneous impact of shocks. In Subsection 5.6, we examine the distributional consequences of shocks to fundamentals. In Subsection 5.7, we report the results of our multi-sector extension, and provide evidence on the role of sectoral specialization

---

[11]Therefore, our single-sector model in our empirical implementation features a single traded sector and a single non-traded sector, as developed in detail in Section E of the online appendix.

in influencing the dynamic response of the economy to shocks.

5.1 Data

Our main source of data for our baseline quantitative analysis from 1965-2015 is the national economic accounts of the Bureau of Economic Analysis (BEA), which report population, gross domestic product (GDP) and the capital stock for each U.S. state.[12] We focus on the 48 contiguous U.S. states plus the District of Columbia, excluding Alaska and Hawaii, because they only became U.S. states in 1959 close to the beginning of our sample period, and could be affected by idiosyncratic factors as a result of their geographical separation. We deflate GDP and the capital stock to express them in constant (2012) prices.

We construct bilateral five-year migration flows between U.S. states from the U.S. population census from 1960-2000 and from the American Community Survey (ACS) after 2000. We define a period in the model as equal to five years to match these observed data. We interpolate between census decades to obtain five-year migration flows for each year of our sample period. To take account of international migration to each state and fertility/mortality differences across states, we adjust these migration flows by a scalar for each origin and destination state, such that origin population in year t pre-multiplied by the migration matrix equals destination population in year $t + 1$, as required for internal consistency.

We construct the value of bilateral shipments between U.S. states from the Commodity Flow Survey (CFS) from 1993-2017 and its predecessor the Commodity Transportation Survey (CTS) for 1977. We again interpolate between reporting years and extrapolate the data backwards in time before 1977 using relative changes in the income of origin and destination states, as discussed in further detail in Section H of the online appendix. For our baseline quantitative analysis with a single traded and non-traded sector, we abstract from direct shipments to and from foreign countries, because of the relatively low level of U.S. trade openness, particularly towards the beginning of our sample period.

For our multi-sector extension from 1999-2015, we construct data for the 48 contiguous U.S. states, 43 foreign countries and 23 economic sectors, yielding a total of 2,093 region-sector combinations, where a region is either a U.S. state or a foreign country. We allow for trade across all region-sectors, and for migration across all U.S. states and sectors. We obtain sector-level data on value added, employment and the capital stock for each U.S. state from the national economic accounts of the Bureau of Economic Analysis (BEA).

We construct migration flows between U.S. states in each sector by combining data from the U.S. population census, American Community Survey (ACS), and Current Population Survey (CPS), as discussed in Section H of the online appendix. We use the value of bilateral shipments

---

[12]For further details on the data sources and definitions, see the data appendix in Section H of the online appendix.

between U.S. states in each sector from the Commodity Flow Survey (CFS), interpolating between reporting years. We measure foreign trade for each U.S. state and sector using the data on foreign exports by origin of movement (OM) and foreign imports by state of destination (SD) from the U.S. Census Bureau.[13] For each foreign country and sector, we obtain data on value added, employment and the capital stock from the World Input-Output Tables (WIOT).

5.2 Parameterization

To implement our approach empirically, we assume central values of the model's structural parameters from the existing empirical literature. We assume a standard value of the trade elasticity of $\theta = 5$, as in Costinot and Rodríguez-Clare (2014). We set the 5-year discount rate equal to the conventional value of $\beta = (0.95)^5$. We assume an elasticity of intertemporal substitution of $\psi = 1$, which corresponds to logarithmic intertemporal utility. We assume a value for the migration elasticity of $\rho = 3\beta$, which is in line with the values in Bryan and Morten (2019), Caliendo et al. (2019) and Fajgelbaum et al. (2019). We set the share of labor in value added to $\mu = 0.65$, as a central value in the macro literature. We assume a five percent annual depreciation rate, such that the 5-year depreciation rate is $\delta = 1 - (0.95)^5$, which is again a conventional value in the macro and productivity literatures. We report comparative statics for how these model parameters affect the economy's transition path in our quantitative analysis below.

5.3 Reduced-Form Evidence on the Rust and Sun-Belt

One the most striking features of economic activity in the United States over our sample period is its reorientation away from the "Rust Belt" in the mid-west and north-east towards the "Sun Belt" in the south and west. Although we implement our quantitative analysis at the level of individual states, we begin by reporting some aggregate results for four groupings of states to illustrate this large-scale reorientation. Following Alder et al. (2019), we define the Rust Belt as the states of Illinois, Indiana, Michigan, New York, Ohio, Pennsylvania, West Virginia and Wisconsin, and the Sun Belt as the states of Arizona, California, Florida, New Mexico and Nevada. We group the remaining states into two categories to capture longstanding differences between the North and South: Other Southern States, which includes all former members of the Confederacy, except those in the Sun Belt; and Other Northern States, which comprises all the Union states from the

---

[13]The Census Bureau constructs these data from U.S. customs transactions, aiming to measure the origin of the movement of each export shipment and the destination of each import shipment. Therefore, these data differ from measures of exports and imports constructed from port of exit/entry, and from the data on the exports of manufacturing enterprises (EME) from the Annual Survey of Manufactures (ASM). See https://www.census.gov/foreign-trade/aip/elom.html and Cassey (2009).

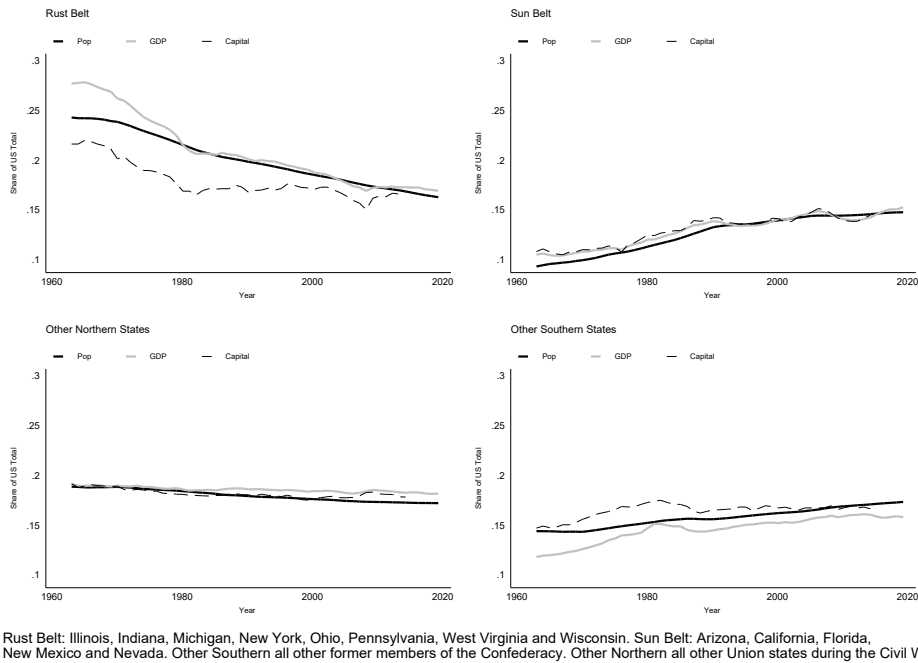U.S. Civil War, except those in the Rust Belt or Sun Belt.[14]

In Figure 1, we display the shares of these four groups of states in the U.S. population over time. As shown in the top-left panel, the Rust Belt exhibits by far the largest decline in population share, which falls by around 8 percentage points from 24.2-16.5 percent from 1965-2015. As shown in the top-right panel, the Sun Belt displays the largest increase in population share, which rises by around 5 percentage points from 9.3-14.7 percent over the same period. In contrast, the trends in population shares for the other two groups of states are much flatter. The population share of Other Northern States falls by 1.6 percentage points from 18.8-17.2 percent, while that for Other Southern States rises by 2.8 percentage points from 14.4-17.2 percent.

Figure 1 also shows the shares of these four groups of states in real GDP and the real capital stock in the United States over time.[15] We find that GDP and capital stock shares show some differences from population shares, highlighting the potential relevance of capital accumulation and productivity growth for the observed reorientation of economic activity. In the Rust Belt, we find that capital and GDP shares fall more rapidly than population in the 1960s and 1970s. In contrast, in the Sun Belt, GDP and capital shares lie above population shares from the early 1960s to the 1990s, before population shares ultimately converge towards them. In Other Northern States, population shares fall marginally below GDP and capital shares from the mid-1980s onwards. Finally, in Other Southern States, GDP and capital shares rise substantially more sharply than population shares in the 1960s and 1970s, consistent with a role for income convergence.

---

[14]Therefore, "Other Southern" includes Alabama, Arkansas, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. "Other Northern" includes Colorado, Connecticut, Delaware, District of Columbia, Idaho, Iowa, Kansas, Kentucky, Maine, Maryland, Massachusetts, Minnesota, Missouri, Montana, Nebraska, New Hampshire, New Jersey, North Dakota, Oklahoma, Oregon, Rhode Island, South Dakota, Utah, Vermont, Washington and Wyoming.

[15]We find similar patterns whether we use real or nominal shares of GDP and the capital stock.

Figure 1: Shares of Population, Gross Domestic Product (GDP) and the Capital Stock in the United States over Time
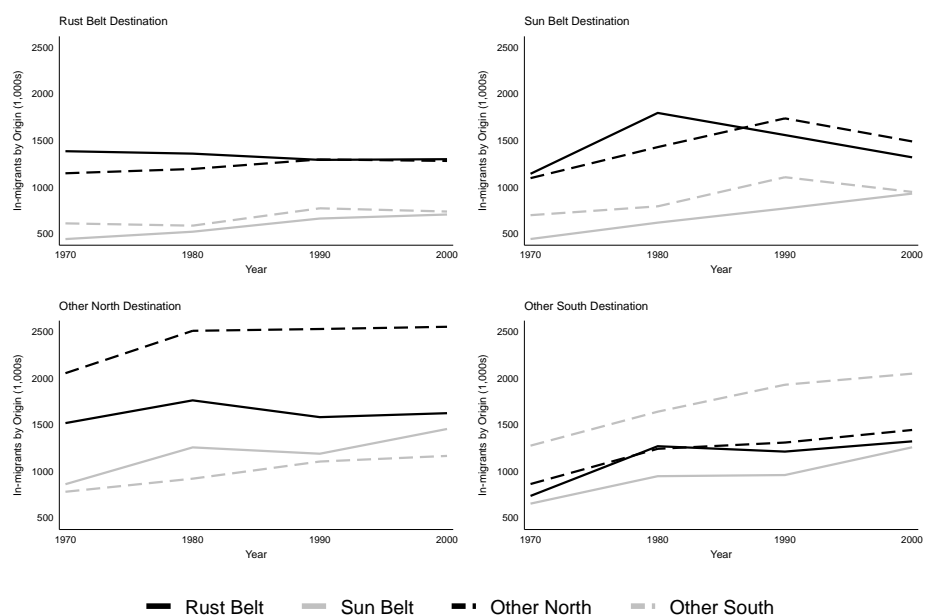

Rust Belt: Illinois, Indiana, Michigan, New York, Ohio, Pennsylvania, West Virginia and Wisconsin. Sun Belt: Arizona, California, Florida, New Mexico and Nevada. Other Southern all other former members of the Confederacy. Other Northern all other Union states during the Civil War

Notes: Shares of total population, real gross domestic product (GDP) and the real capital stock in the United States over time; real GDP and capital stock in 2012 prices.

In Figure 2, we provide evidence on the role of internal migration as a source of population changes for the four groups of states. Internal migration is measured as movements of people between states within the United States and excludes international migration. We focus for brevity on in-migrants, measured as inflows of internal migrants (in thousands) into each destination region, separated out by origin region. Three features are noteworthy. First, geographical proximity matters for migration flows, such that other Rust Belt states are one of the leading sources of in-migrants for the Rust Belt (top-left panel), consistent with our model's gravity equation predictions. Second, all groups of states receive non-negligible in-migration flows, such that gross migration flows are larger than net migration flows, in line with the idiosyncratic mobility shocks in our model. Third, despite the role for geography, the Rust Belt and Other Northern states are the two largest sources of in-migrants for the Sun Belt, consistent with internal migration contributing to the observed reorientation of population shares.

Finally, although not shown in these figures, we find a modest decline in rates of internal migration between states in the later years of our sample, which is in line the findings of a number of studies, including Kaplan and Schulhofer-Wohl (2017) and Molloy et al. (2011). Consistent with the comparison of several different sources of administrative data in Hyatt et al. (2018), we find that this decline in rates of internal migration between states is smaller in the population census

data than in Current Population Survey (CPS) data.

Figure 2: Internal In-migrants for each Destination Region by Origin Region over Time

Rust Belt: Illinois, Indiana, Michigan, New York, Ohio, Pennsylvania, West Virginia and Wisconsin. Sun Belt: Arizona, California, Florida, New Mexico and Nevada. Other Southern all other former members of the Confederacy. Other Northern all other Union states during the Civil War

Notes: Internal in-migration to each destination region by source region from 1960-2000; internal migration includes all movements of people between states within the United States and excludes international migration.

5.4   Convergence to Steady-state and Fundamental Shocks

We begin by using our additive decomposition from Section 3.2.1 to provide evidence on the contributions of convergence to steady-state and shocks to fundamentals to the observed reorientation of economic activity across U.S. states over our sample period.

First, we use equation (26) to compute the steady-state values of the state variables for each U.S. state and year $\{\ell_t^*, k_t^*\}$ in the absence of further changes in unobserved fundamentals $\{z_t, b_t, \tau_t, \kappa_t\}$, given the observed state variables for two previous years and the observed trade and migration share matrices for each year (S, T, D, E).

In Figure 3, we summarize the state-level results by displaying actual and steady-state population shares for each of our four groupings of states. The black line shows the actual data for each year; the dashed red line shows the implied steady-state assuming no further changes in fundamentals. From the top-left panel, we find that actual population in the Rust Belt states was already substantially above its steady-state value in 1965, and only begins to approach its steady-state value towards the end of our sample period. In the top-right panel, actual population in the Sun Belt states was substantially below its steady-state value in 1965, but converges towards its steady-state value by the late 2000s. In the bottom-right panel, actual population in Other

34

Southern states rises substantially above its steady-state value during our sample period, before the two sets of population shares converge towards one another by the end of our sample period. Finally, in the bottom-left panel, actual and steady-state population shares in Other Northern states lie relatively close to one another throughout our sample period.

Across these four panels, we find that actual population shares can remain persistently either above or below their steady-state values for decades, implying slow convergence to steady-state. Actual and steady-state population shares are also closer together at the end of our sample period than at its beginning, suggesting that one potential reason for the modest observed decline in population mobility over time could be that the economy is now closer to steady-state.

Figure 3: Actual and Implied Steady-State Population Shares by Region
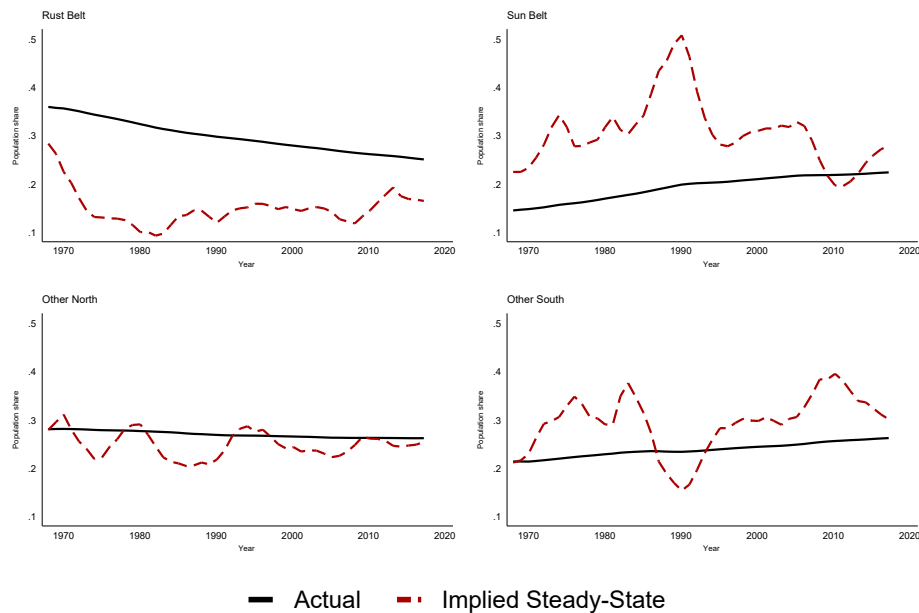
Rust Belt: Illinois, Indiana, Michigan, New York, Ohio, Pennsylvania, West Virginia and Wisconsin. Sun Belt: Arizona, California, Florida, New Mexico and Nevada. North and South definitions based on Federal and Confederacy states

Note: Black line shows actual population shares; dashed line shows implied steady-state population share from equation (25), assuming no further changes in fundamentals, given the observed state variables for two previous years and the observed trade and migration share matrices for each year.

Second, we use equation (26) to compute the counterfactual transition path of economic activity implied by convergence towards an initial steady-state with unchanged fundamentals. In Figure 4, we display these counterfactual transition paths for population shares for 1965, 1975, 1985 and 1995 (dashed lines), as well as the actual evolution of population shares (solid black line). We extend the counterfactual transition paths beyond the end of our sample period to again highlight the model's implied slow convergence towards steady-state. From our additive decomposition in equation (25), the difference between the actual and counterfactual transition paths corresponds to contributions from shocks to fundamentals.

We find that the transition path implied by convergence towards an initial steady-state has substantial predictive power for the trajectory of actual population shares. We find the largest residual contributions from shocks to fundamentals in the late 1960s and early 1970s, which is consistent with this period including substantial structural change in manufacturing and the first oil price shock. From the mid-1970s onwards, we find smaller values for these residuals, with the counterfactual transition paths from convergence towards an initial steady-state closely tracking the actual evolution of population shares (solid black line).

Figure 4: Actual Population Shares and Counterfactual Population Shares Implied by Convergence Towards an Initial Steady-State with Unchanged Fundamentals
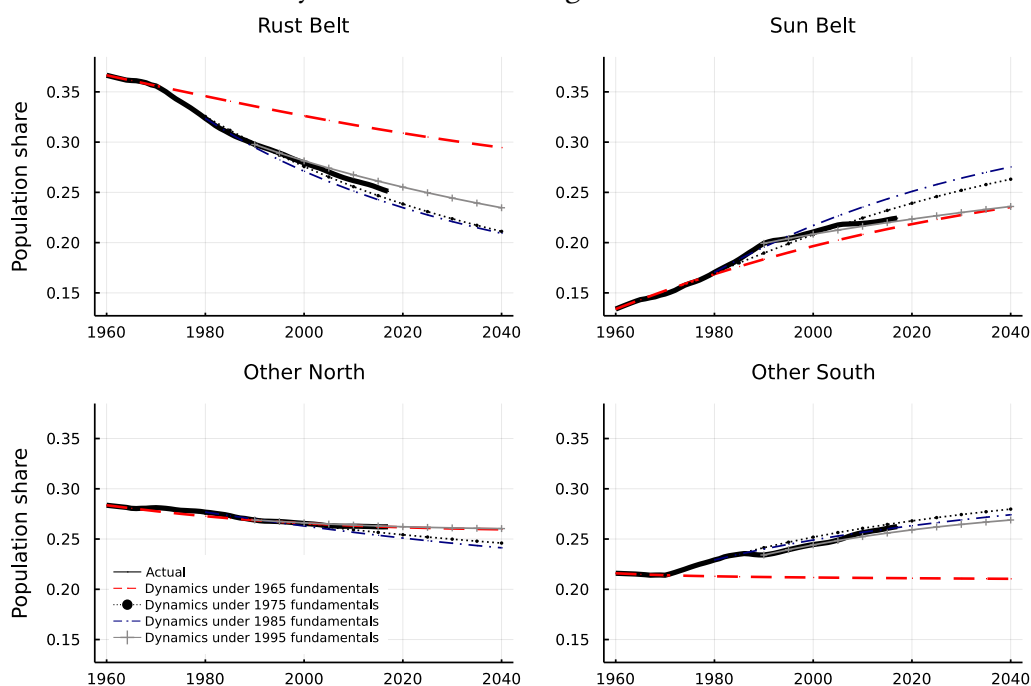
Note: Solid black line shows actual population shares; dashed lines show counterfactual transition paths for population shares based on convergence to an initial steady-state with unchanged fundamentals from equation (26), starting in either 1965, 1975, 1985 or 1995.

To provide further evidence on this predictive power of convergence dynamics, we regress actual population growth on predicted population growth based on convergence towards an initial steady-state with unchanged fundamentals. We estimate these regressions for each of the periods 1965-2015, 1975-2015, 1985-2015 and 1995-2015, as reported in Section F.1 of the online appendix. Even for the period from 1965-2015, for which there is the greatest residual contribution from shocks to fundamentals, we find a strong positive and statistically significant relationship between actual and predicted population growth, with a regression slope (standard error) of 0.64 (0.18) and R-squared of 0.19. For all periods starting in the mid-1970s onwards, we find an even stronger relationship between actual and predicted population growth. For example, for the pe-

riod from 1975-2015, we estimate a regression slope (standard error) of 0.99 (0.095) and R-squared of 0.82.

In Section F.1 of the online appendix, we show that controlling for initial log population, log capital stock, and log population growth has little impact on either the estimated coefficient or R-squared in these regressions. Therefore, the strength of the relationship between actual and predicted population growth does not simply reflect mean reversion, because we continue to find substantial independent information in predicted population growth, even after including these controls for initial levels and rates of growth of economic activity.

5.5   Spectral Analysis

We next provide a further characterization of the economy's transition path using our eigende-composition of the transition matrix (P) from Propositions 3 and 4 in Section 3.2.2 above.

Speed of Convergence to Steady-State   We begin by using Proposition 4 to compute half-lives of convergence to steady-state. In particular, we compute the number of years for the state variables to converge half of the way towards steady-state for eigen-shocks ($\widetilde{\boldsymbol{f}}_{(k)}$), for which the initial impact of these shocks on the state variables ($\boldsymbol{R}\widetilde{\boldsymbol{f}}$) corresponds to a real eigenvector of the transition matrix ($\boldsymbol{u}_k$). We have as many half-lives as eigen-shocks, where each of these eigen-shocks corresponds to a different pattern of shocks to productivity and amenities across the state variables in each location.
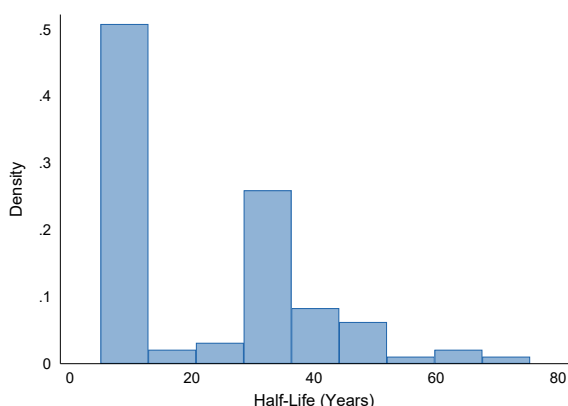
In the left panel of Figure 5, we display a histogram of these half-lifes across the eigencomponents. We find that the speed of convergence to steady-state is typically slow, with an average half life of around 20 years. This pattern of results is consistent with reduced-form findings of persistent impacts of local shocks (e.g. Dix-Carneiro and Kovak 2017; Autor et al. 2020) and with our empirical findings in the previous subsection that states can remain persistently above or below steady-state for decades. We also find substantial heterogeneity in the speed of convergence to steady-state, with half lives ranging from less than 10 years to more than 70 years, depending on the size and incidence of the shocks to productivity and amenities across the state variables in each location. This heterogeneity in speeds of convergence is consistent with reduced-form evidence of uneven impacts of local shocks, as in Eriksson et al. (2019).

In the right panel of Figure 5, we show that capital and labor dynamics interact with one another to shape this heterogeneous speed of convergence. On the vertical axis, we display the half-life of convergence to steady-state (in years) for each eigenvalue of the transition matrix. On the horizontal axis, we display the regression slope coefficient between the gaps from steady-state for labor and capital for the corresponding eigen-shock. As apparent from the figure, we find a strong, positive and non-linear relationship between the half-life of convergence to steady-state

and the correlation between the gaps from steady-state for the two state variables. We observe low half-lifes (fast convergence) for negative correlations and high half-lifes (slow convergence) for positive correlations. We find that a tight relationship between these two variables, with little dispersion in half-lives for a given correlation between the steady-state gaps.

Figure 5: Half-lifes for Convergence Towards Steady-State

(a) Histogram of Half-lifes for Shocks to Productivity and Amenities that Correspond to Eigenvectors of the Transition Matrix (P) in 2000

(b) Relationship Between Half-Lifes of Convergence Towards Steady-State and the Correlation between the Initial Gaps of the Labor and Capital State Variables from Steady-State in 2000
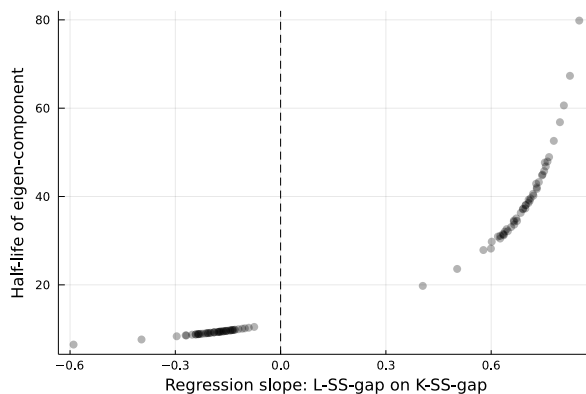
Note: Half-life corresponds to the time in years for the state variables to converge half of the way towards steady-state for an eigen-shock, for which the initial impact of the shock to productivity and amenities on the state variables ($R\widetilde{f}$) corresponds to an eigenvector (u_k) of the transition matrix (P); left panel shows the distribution of half-lifes across eigencomponents of the transition matrix in 2000; right panel plots these half-lifes of convergence to steady-state for the year 2000 against the regression slope coefficient between the gaps from steady-state for the labor and capital state variables for the corresponding eigen-shock.

This role of an interaction between capital and labor dynamics in shaping the speed of convergence arises from the complementarity between capital and labor in the production technology. A high stock of capital in a location raises the marginal productivity of labor, while a high population in a location raises the marginal productivity of capital. Therefore, if both capital and labor are above steady-state in a location, this slows convergence, because the high capital stock retards the downward adjustment of labor, and the high population dampens the downward adjustment of capital. Similarly, if both capital and labor are below steady-state, this also slows convergence, by an analogous line of reasoning. In contrast, if capital is above steady-state whereas labor is below steady-state, this speeds up convergence, because the high capital stock attracts labor, and the low population discourages capital accumulation.

The transition matrix (P) changes over time with underlying changes in the trade and migration share matrices (S, T, D, E). Therefore, we can compute eigenvalues of this transition matrix, and hence half-lives of convergence to steady-state, for each year of our sample period.

Although we observe a modest decline in geographical mobility in the data on state-to-state migration, we find that the speed of convergence towards steady-state is relatively constant over time, with the mean and maximum half-life declining somewhat towards the end of our sample period. This juxtaposition of a decline in geographical mobility and faster convergence towards steady-state again highlights the idea that a decline in geographical mobility does not necessarily imply a rise in migration frictions and hence slower convergence towards steady-state. This decline in geographic mobility instead can be explained by the economy being closer to steady-state at the end of our sample period than at its beginning and/or by a change to the pattern of shocks to productivity and amenities across locations.

Figure 6: Half-lifes of Convergence to Steady-State for Alternative Parameter Values
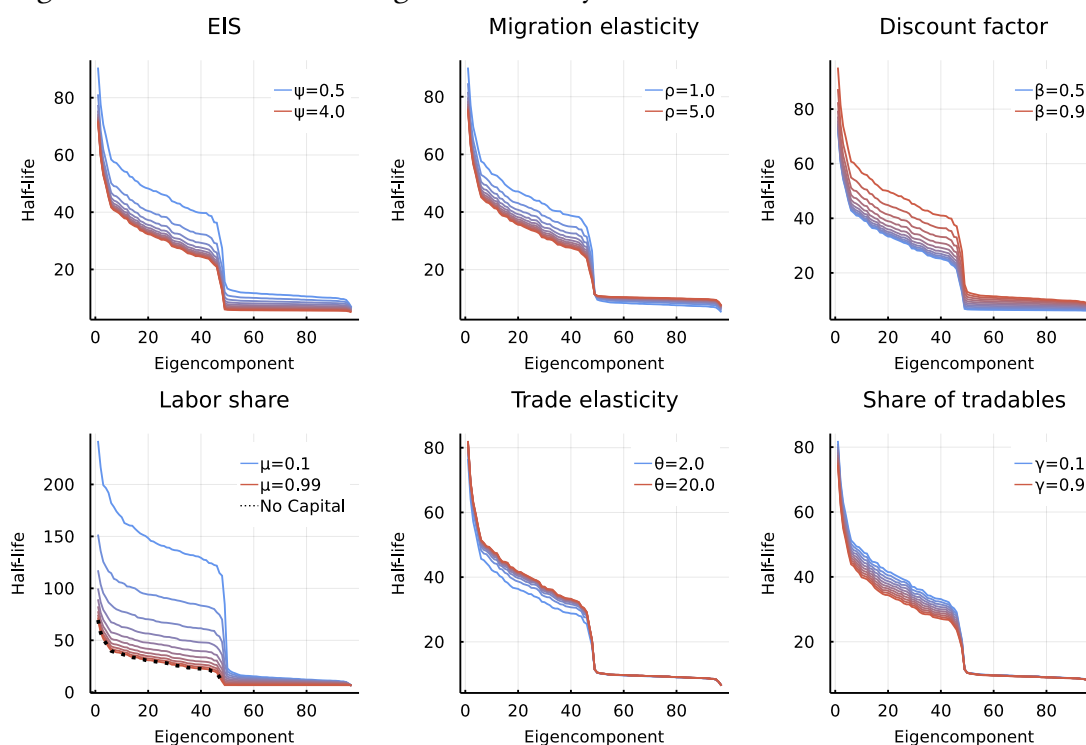
Note: Half-lifes of convergence to steady-state for each eigen-shock for alternative parameter values in 2000; vertical axis shows half-life in years; horizontal axis shows the rank of the eigen-shocks in terms of their half-lifes for our baseline parameter values (with one the highest half life); each panel varies the noted parameter, holding constant the other parameters at their baseline values; the blue and red solid lines denote the lower and upper range of the parameter values considered, respectively; each of the other eight lines in between varies the parameters uniformly within the stated range; the thick black dotted line in the bottom-left panel displays half-lifes for the special case of our model without capital, which corresponds to the limiting case in which μ converges to one.

Comparative Statics of the Speed of Convergence Using our closed-form solution for the transition matrix (P) as a function of the observed data $\{S, T, D, E\}$ and the structural parameters of the model $\{\psi, \theta, \beta, \rho, \mu, \delta\}$, we can also examine the comparative statics of the speed of convergence with respect to parameter values. We vary each parameter individually,

holding constant the other parameters, and recompute the transition matrix ($\boldsymbol{P}$), eigen-shocks ($\widetilde{\boldsymbol{f}}_{(k)} = \boldsymbol{R}^{-1}\boldsymbol{u}_k$), and corresponding eigenvalues (λ_k). An advantage of our linearization is that we can examine the speed of convergence for the entire spectrum of eigen-shocks, where any empirical shocks can be written as linear combinations of these eigen-shocks. Undertaking an analogous exercise using the non-linear model would be challenging, since one would need to undertake numerical counterfactuals for the set of all possible fundamental shocks, which is not well defined.

In Figure 6, we display half-lifes across eigen-shocks for different parameter values, where the eigen-shocks are sorted in terms of decreasing half-life for our baseline parameter values. Each panel varies the noted parameter, holding constant the other parameters at their baseline values. In the top-left panel, a lower intertemporal elasticity of substitution (ψ) implies a longer half-life (slower convergence), because consumption becomes less substitutable across time for landlords, which reduces their willingness to respond to investment opportunities. In the top-middle panel, a lower migration elasticity (ρ) also implies a longer half-life (slower convergence), because it increases the dispersion of idiosyncratic tastes, which leads workers to become less responsive to differences in the expected value of living in each location. In the top-right panel, a higher discount rate (β) also implies a longer half-live (slower convergence), because it increases the steady-state rate of saving out of income net of depreciation.

In the bottom-left panel, a lower labor share (μ) implies a longer half-life (slower convergence), which reflects the interaction between capital and labor dynamics. Note that the special case of our model without capital corresponds to the limiting case in which the labor share converges to one. We show half-lifes of convergence to steady-state in this special case using the thick black dotted line, again sorted in terms of decreasing half-life for our baseline parameter values. In this special case, we only only have N state variables and eigencomponents, compared to $2N$ state variables and eigencomponents in the general model.

Comparing these two specifications, we see that incorporating capital investments introduces a new adjustment margin, which can increase the economy's speed of convergence in response to some shocks, as shown by the additional eigencomponents with low half-lifes to the right of the end of the thick dotted black line. However, this new capital adjustment margin also interacts with the existing labor adjustment margin, because of the complementarity between capital and labor in the production technology. As labor reallocates between regions, this induces endogenous responses in the capital stock, which in turn induces further labor reallocation, and hence slows the speed of convergence. As a result, the first N eigencomponents in the model with both labor and capital lie above those in the special case with only labor, as shown by the eigencomponents above the thick dotted black line.

In the bottom-middle panel, a higher trade elasticity (θ) implies a longer half-life (slower

40

convergence), because it increases the responsiveness of production and consumption in the static trade model, and hence requires greater reallocation of capital and labor across locations. In the bottom-right panel, a lower share of tradeables (γ) implies a longer half life (slower convergence), because it makes the impact of shocks more concentrated locally, which requires greater labor and capital reallocation between locations.

Heterogeneous Impact of Shocks   We next use our eigendecomposition of the transition matrix to provide evidence on the heterogeneous impact of local shocks. We compute the loading ($\boldsymbol{\alpha}_i'$) of an individual location i on the eigenvectors (u_k) using a vector of weights (w_i') equal to one for its state variables and zero otherwise: $\boldsymbol{\alpha}_i' = \boldsymbol{w}_i'\boldsymbol{U}$. Using these loadings for two different locations i and n, we compute the similarity of their exposure to productivity and amenity shocks as the correlation between these loadings:

$$\text{Similarity}_{i,n} \equiv \frac{\boldsymbol{\alpha}_i'\boldsymbol{\alpha}_n}{||\boldsymbol{\alpha_i}|| \times ||\boldsymbol{\alpha_n}||}, \tag{38}$$

where the stronger this correlation, the greater the extent to which these two locations are exposed to more similar productivity and amenity shocks.

In Figure 7, we display the similarity of exposure to productivity and amenity shocks for U.S. states in 1965 (top panel) and 2015 (bottom panel) using a network graph. The nodes correspond to U.S. states, with the size of these nodes reflecting the population shares of the states. The thickness of each edge captures the degree of similarity in exposure to productivity and amenity shocks, where we focus on the 200 edges with the highest degrees of similarity for reasons of legibility. States are grouped to maximize modularity (the fraction of edges within the groups minus the expected fraction if the edges were distributed at random). The different colored shading indicates the distinct groups of states.

Figure 7: Network of State Exposure to Productivity and Amenity Shocks

(a) 1965
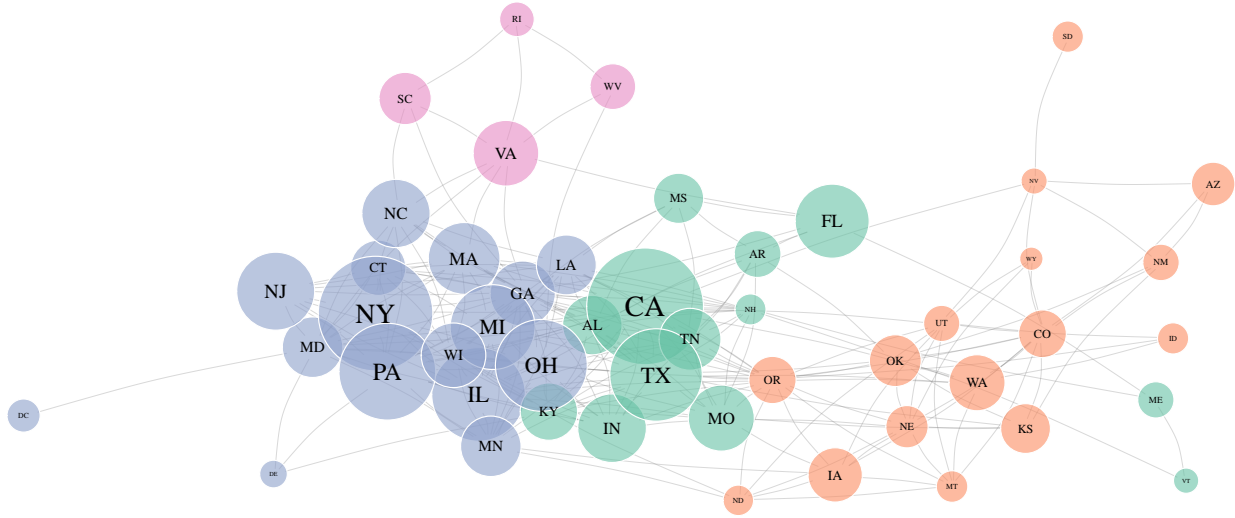
(b) 2015
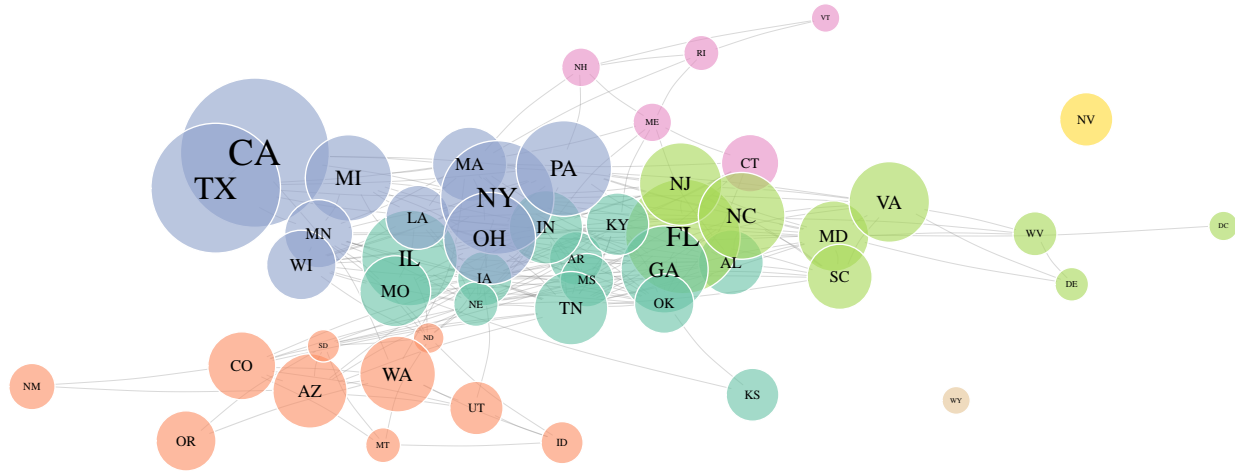
Note: Network of bilateral similarity of state exposure to productivity and amenity shocks in 1965 (Panel A) and 2015 (Panel B), as defined in the main text; states grouped to maximize modularity (the fraction of edges within the groups minus the expected fraction if the edges were distributed at random); the two-letter codes correspond to the postal codes for each U.S. state (e.g. CA represents California); colors indicate distinct groups of states.

We find a powerful role for state size and geography in shaping the similarity of exposure to productivity and amenity shocks. In 1965, we find four main groups of states, as indicated by the blue, green, orange and purple colors. The blue group includes most of the traditional industrial states (e.g. New York, Pennsylvania, Michigan and Ohio), while the green group incorporates Southern and Western states (e.g. California and Texas). Between 1965 and 2015, we find that a fifth group of states emerges, which consists largely of Mid or Southern-Atlantic states (e.g. New Jersey, North Carolina, Maryland and Virginia), and is shown in light green. We also find that California and Texas move to the blue group of states between these two years, as they increase in

size and population, and become more connected with these traditional industrial states. Overall, we find that the network is more tightly clustered in 2015 than in 1965, consistent with an increase in the integration of the U.S. economy over time.

5.6 Distributional Consequences

We now turn to the distributional consequences of shocks to fundamentals. Our framework implies a rich pattern of heterogeneous welfare effects of these shocks, both across landlords because they are geographically immobile, and across workers because of migration frictions, which imply that a worker's initial location matters for welfare effects. For both landlords and workers, the welfare effects of the shock depend not only on the change in steady-state, but also on the economy's dynamic response along the entire transition path.

In Section B.9 of the online appendix, we provide an analytical characterization of these heterogenous welfare effects. First, we show that the relative welfare effects of shocks across workers initially located in each region can be inferred from the initial population movements. Second, we use our closed-form solutions for the economy's transition path to solve for the average welfare effect for all workers weighted by initial population shares. We show that this average welfare effect can be decomposed into direct effects from fundamental shocks and indirect effects from the changes in state variables in response to these fundamental shocks.

To illustrate these heterogeneous welfare effects, we start at the observed values of the state variables (population and the capital stock) at the beginning of our sample period in 1965, and undertake a counterfactual for a one-time permanent productivity shock equal to the accumulated empirical distribution of productivity shocks from 1965-2015. We recover this empirical distribution of productivity shocks from the inversion of the non-linear model, using an extension of existing dynamic exact-hat algebra approaches to incorporate forward-looking capital investments, as discussed in Section G.1 of the online appendix. We solve for the empirical distribution of relative changes in productivity across states, using the normalization that the productivity shocks are mean zero in logs. In this counterfactual, we solve for the transition path of the spatial distribution of economic activity towards the new steady-state in response to this one-time shock. We evaluate the uneven impact of these productivity shocks on the welfare of workers depending on their location.

In Figure 8, we contrast the effects of these shocks on worker flow utility ($w_i b_i / p_i$) versus worker expected value (v_i^w). In both panels, the vertical axis shows the change in the flow utility ($w_i b_i / p_i$) in the initial period for workers located in each state in that period. In the left-panel, the horizontal axis shows the change in the expected value (v_i^w) in the initial period for workers located in each state in that period. In contrast, in the right-panel, the horizontal axis shows the change in the expected value (v_i^w) of living in each state in the new steady-state equilibrium. Both

the change in expected value and the change in flow utility are normalized to have a mean of zero across states.

In both panels, we find substantial heterogeneity in the welfare effects of the productivity shocks, depending on the state in which workers are initially located. This heterogeneity is driven by the migration frictions, which imply that it is costly and takes time for workers in states that experience relative reductions in expected values to reallocate towards those that experience relative increases in expected values. As population reallocates across states, this leads to endogenous changes in the stock of capital in each state, which in turn induces further labor reallocation. In general, the changes in relative expected values are much larger than the changes in relative flow utilities in the initial period, because these expected values correspond to the net present value of the stream of expected future flow utilities.

Comparing the two panels, the change in the initial period flow utility is much more strongly correlated with the change in the initial period expected value (left panel) than the change in the expected value in the new steady-state equilibrium (right panel). Again this intuitive pattern reflects the reallocation of some workers from states that initially experience relative reductions in expected values towards other states that initially experience relative increases in expected values, and the resulting endogenous changes in the stock of capital in each state.

Figure 8: Counterfactual Impact of Productivity and Amenity Shocks on Flow Utility and Expected Value in the Initial Period and the New Steady-State
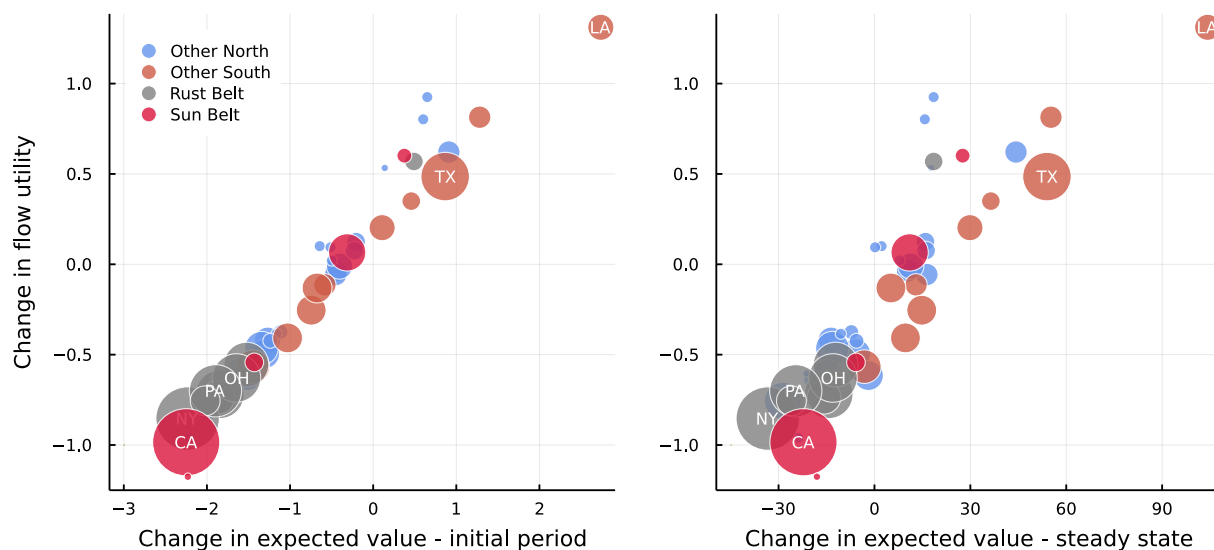
Note: Starting at the observed values of the state variables at the beginning of our sample in 1965, we undertake a counterfactual for a one-time permanent productivity shock in each state, equal to the accumulated empirical productivity shocks from 1965-2015; the left panel shows that change in flow utility and expected value in the initial period; the right panel shows the change in flow utility in the initial period and the expected value in the new steady-state; both the change in flow utility and expected value are normalized to have a mean of zero across U.S. states; the two-letter codes correspond to the postal codes for each U.S. state (e.g. CA represents California); the colors indicate the states in each of our four groups from Figure 1 above (e.g. gray corresponds to Rust Belt).

5.7 Multi-sector Quantitative Analysis

In a final empirical exercise, we implement our multi-sector extension with region-sector specific capital from Section 4.3 above, using our region-sector data from 1999-2015. We compute our closed-form solutions for the comparative statics of economic activity in each region-sector with respect to productivity and amenity shocks in any region-sector, along the entire transition path. In the interests of brevity, we focus on our analytical characterization of the economy's transition path using our spectral analysis of the transition matrix (P).
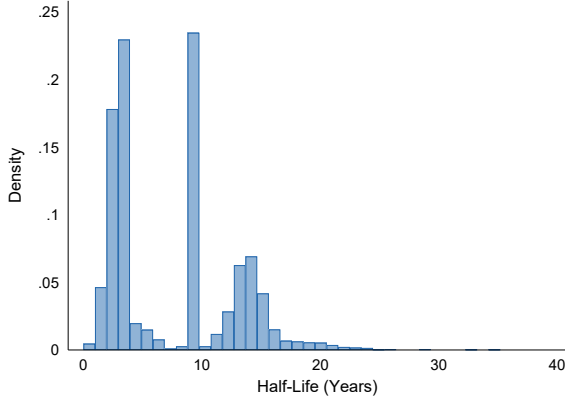
We begin by using our generalization of Proposition 4 in the multi-sector model to compute half-lifes of convergence towards steady-state for shocks to productivity or amenities for which the initial impact on the state variables ($R\widetilde{f}$) corresponds to an eigenvector (u_k) of the transition matrix (P). In Figure 9, we display the distribution of these half-lifes across eigenvectors of the transition matrix in the year 2000. As apparent from the figure, we find substantially more rapid convergence to steady-state in our multi-sector extension, with an average half-life of 7 years and a maximum half-life of 35 years (compared to around 20 and 85 years in our baseline single-sector model). This finding is driven by the property of the region-sector migration matrices that flows of people between sectors within the same U.S. state are much larger than those between different U.S. states. A key implication is that there is likely to be heterogeneity in the persistence of the impact of local shocks, depending on whether they induce reallocation across industries within the same location or reallocation across different locations.

In the right panel of Figure 9, we show that capital and labor dynamics again interact with one another to shape the speed of convergence towards steady-state in the multi-sector model. On the vertical axis, we display the half-life of convergence to steady-state (in years) for each eigenvalue of the transition matrix. On the horizontal axis, we display the regression slope coefficient between the gaps from steady-state for labor and capital for the corresponding eigen-shock. As for the single-sector model above, we find a strong, positive and non-linear relationship between the half-life of convergence to steady-state and the correlation between the gaps from steady-state for the two state variables. We observe low half-lifes (fast convergence) for negative correlations and high half-lifes (slow convergence) for positive correlations.

Under our assumption of no international migration, the deviation of labor from steady-state is always equal to zero for foreign countries in our multi-sector model. Therefore, they adjust to fundamental shocks through capital accumulation alone. As a result, we observe a mass of eigen-shocks with intermediate half-lifes in both panels of the figure. These half-lifes lie in-between eigen-shocks with fast convergence (negative correlation between the gaps from steady-state) and those with slow convergence (positive correlation between the gaps from steady-state).

Figure 9: Half-lifes for Convergence Towards Steady-State in the Multi-Sector Model

(a) Histogram of Half-lifes for Shocks to Productivity and Amenities that Correspond to Eigenvectors of the Transition Matrix ($\boldsymbol{P}$) in 2000

(b) Relationship Between Half-Lifes of Convergence Towards Steady-State and the Correlation between the Initial Gaps of the Labor and Capital State Variables from Steady-State in 2000
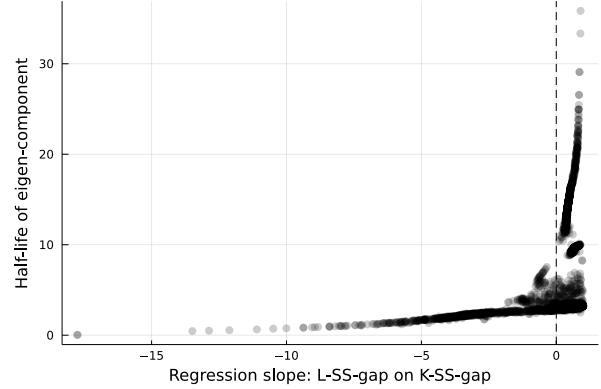
Note: Half-life corresponds to the time in years for the state variables to converge half of the way towards steady-state for an eigen-shock, for which the initial impact of the shock to productivity and amenities on the state variables ($\boldsymbol{R}\widetilde{\boldsymbol{f}}$) corresponds to an eigenvector ($\boldsymbol{u_k}$) of the transition matrix ($\boldsymbol{P}$) in the multi-sector model; left panel shows the distribution of half-lifes across eigencomponents of the transition matrix in 2000 in the multi-sector model; right panel plots these half-lifes of convergence to steady-state for the year 2000 in the multi-sector model against the regression slope coefficient between the gaps from steady-state for the labor and capital state variables for the corresponding eigen-shock.

6 Non-linearities

Our quantitative analysis so far has demonstrated three main advantages of our linearization. First, we derive closed-form solutions for the economy's transition path, which permits our analytical analysis of the determinants of the economy's dynamic response to shocks. Second, we obtain these closed-form solutions under stochastic fundamentals and rational expectations, which are challenging to analyze in the non-linear model, and hence most prior research on spatial dynamics has assumed perfect foresight. Third, we obtain our analytical characterization for constant intertemporal elasticity of substitution preferences, which are also challenging to analyze in the non-linear model, except in the special case of logarithmic preferences, because in general the equilibrium saving rate is determined as the solution to the fixed-point problem in Lemma 1 in Section 2 above.

A caveat is that our linearization in only exact in theory for small shocks and in principle there could be important non-linearities. To examine the potential scope for these non-linearities, we compare the predictions of our linearization to those of the full non-linear model solution. In particular, we extend the numerical shooting algorithm for dynamic migration decisions in Caliendo et al. (2019) to incorporate forward-looking capital investments for the special case of

our model with logarithmic preferences and perfect foresight. This shooting algorithm involves first guessing the entire transition path for population and the capital stock towards the new steady-state, and then iterating over subsequent updates of this entire transition path, until the solution and guess converge to one another. We show that in practice our linearization provides a close approximation to the full non-linear model solution even for large shocks, such as the empirical distribution of shocks observed during our sample period. For brevity, we report results for our baseline single-sector specification from Section 2 above.

In Subsection 6.1, we invert the full non-linear model to recover the unobserved changes in productivity, amenities, trade costs and migration frictions (z_{it}, b_{it}, τ_{nit}, κ_{git}) implied by the observed changes in the state variables (ℓ_{it}, k_{it}) and the trade and migration shares (S_{nit}, T_{int}, D_{igt}, E_{git}). In Subsection 6.2, we examine the potential role for non-linearities, by undertaking counterfactuals using the empirical distribution of productivity shocks, and comparing the predictions of our linearization to those of the full non-linear model solution.

6.1 Empirical Distribution of Fundamental Shocks

We begin by inverting the non-linear model to recover the unobserved changes in fundamentals (z_{it}, b_{it}, τ_{nit}, κ_{git}) implied by the observed data. We make the conventional assumption of perfect foresight and use our extension of existing dynamic exact-hat algebra approaches to incorporate forward-looking capital investments. We solve for the unobserved changes in fundamentals from the general equilibrium conditions of the model, using the observed data on bilateral trade and migration flows, population, capital stock and labor income per capita. We recover these unobserved fundamentals, without making assumptions about where the economy lies on the transition path or the specific trajectory of fundamentals, as shown in Section G.1 of the online appendix.

In the left and right panels of Figure 10, we show the recovered empirical distributions of relative changes in productivity ($\widehat{z}_i = z_{i2000}/z_{i1990}$) and amenities ($\widehat{b}_i = b_{i2000}/b_{i1990}$) across U.S. states from 1990-2000, where both variables are normalized to have a geometric mean of one. We find that relative changes in productivity and amenities are clustered around their geometric mean of one, but individual states can experience substantial changes in relative productivity and amenities over a period of a decade, of up to 50 percent. In Section G.2 of the online appendix, we provide further evidence that we find an intuitive pattern of changes in productivity, amenities, bilateral trade costs, and bilateral migration frictions.

Figure 10: Relative Productivity and Amenity Shocks from 1990-2000 from our Model Inversion

(a) Productivity Shocks ($\widehat{z}_i = z_{i2000}/z_{i1990}$)
(b) Amenity Shocks ($\widehat{b}_i = b_{i2000}/b_{i1990}$)
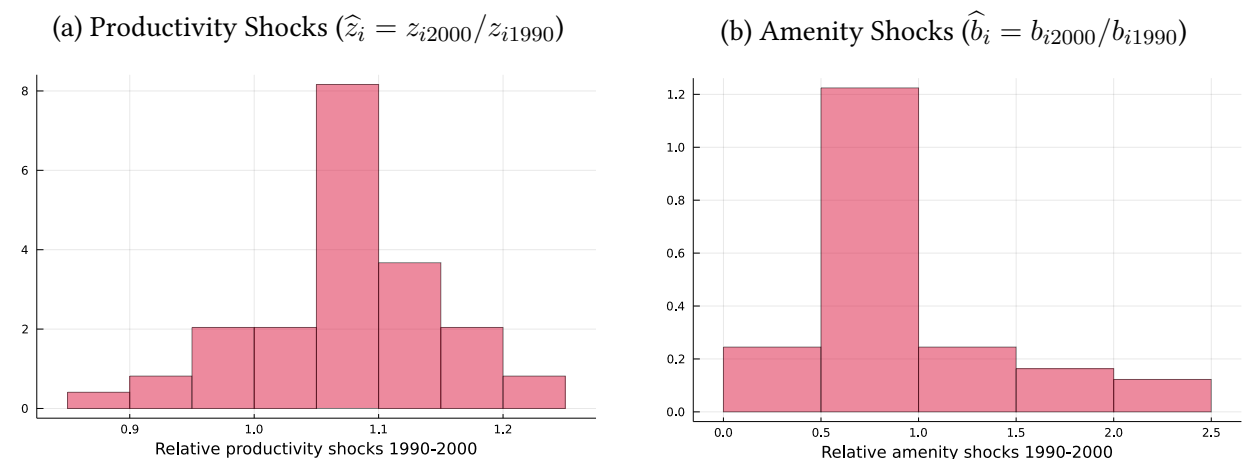
Note: Histograms of the distributions of relative changes in productivity ($\widehat{z}_i = z_{i2000}/z_{i1990}$) and amenities ($\widehat{b}_i = b_{i2000}/b_{i1990}$) from 1990-2000 from our model inversion, as discussed in Section G.1 of the online appendix. Relative changes in productivity ($\widehat{z}_i = z_{i2000}/z_{i1990}$) and amenities ($\widehat{b}_i = b_{i2000}/b_{i1990}$) are both normalized to have a geometric mean of one.

6.2 Transition Paths

To examine the potential scope for non-linearities, we compare the transition paths for the spatial distribution of activity implied by our linearization and the full non-linear model solution. First, we solve for the steady-state in the full non-linear model implied by the 1990 values of fundamentals $\{z_i, b_i, \tau_{ni}, \kappa_{ni}\}$ from the model inversion. Second, we undertake counterfactuals for the empirical distribution of productivity shocks from 1990-2000. We solve for the economy's transition path in response to these productivity shocks using both our linearization and our extended dynamic exact hat algebra algorithm for the full non-linear model.

In Figure 11, we show the transition path for population shares (left panel) and population relative to the initial steady-state (right panel) for each US state. In both panels, the solid blue line denotes full non-linear model, and the red dashed-line corresponds to our linearization. We find that the two sets of predictions track one another relatively closely along the entire transition path of more than one hundred years, implying a limited role for non-linearities even for the empirical distribution of productivity shocks over a decade. This approximation is somewhat better for population shares (left panel) than for population relative to the initial steady-state (right panel), but remains close in both cases. We find a similar pattern of results for the capital stock and for the response of both state variables to amenity shocks.[16]

---

[16]While omitted in the interests of brevity, we also find a close relationship between the predictions of our linearization and the non-linear model solution for changes in steady-states, with a regression slope coefficient of 1.003 and a coefficient of correlation of 0.999. Following the approach of Kleinman et al. (2020) for a static trade model, we can also derive analytical bounds for the quality of the approximation for changes in steady-state.

Figure 11: Transition Path Predictions of Our Linearization and the Full Non-Linear Model Solution for Counterfactual Changes in Productivity

(a) Transition Path for Population Shares

(b) Transition Path for Population Relative to Initial Steady-State Levels
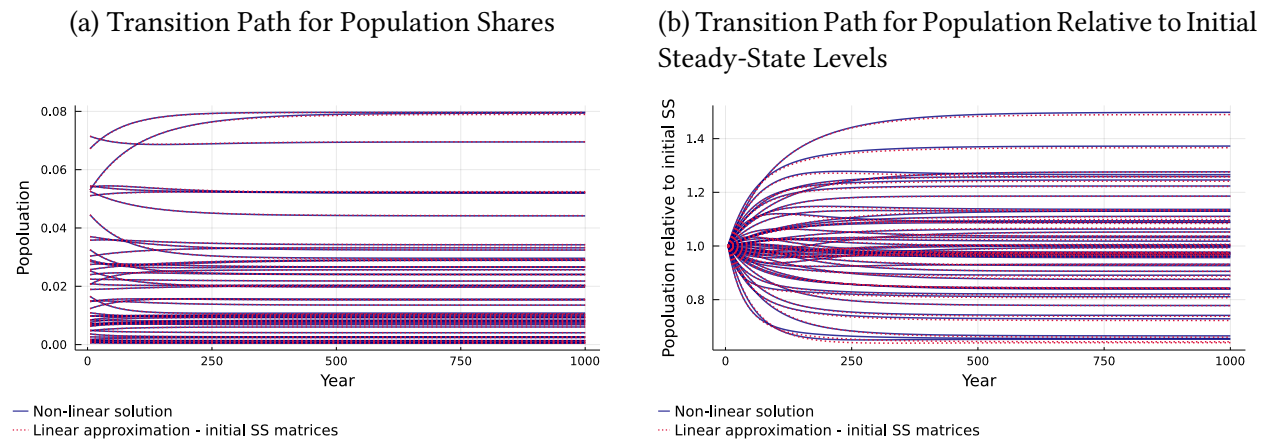
Note: We first solve for the steady-state values of the state variables $\{\ell_i^*, k_i^*\}$ implied by the 1990 values of the fundamentals from inverting the full non-linear model under perfect foresight $\{z_i, b_i, \tau_{ni}, \kappa_{ni}\}$. We next undertake counterfactuals using the empirical distribution of relative changes in productivity over the period 1990-2000, as discussed in Section G.1 of the online appendix. We compare the predicted transition path in the state variables from our linearization in Proposition 2 to those from the full non-linear model solution using an extension of the conventional dynamic exact-hat algebra approach to incorporate capital investments.

Therefore, our linearization is not only exact in theory for small changes, but also provides a close approximation in practice to the economy's transition path for the empirical distribution of decadal productivity and amenity shocks. As discussed above, the main advantages of our linearization are that it yields closed-form solutions for the transition path, encompasses stochastic fundamentals and rational expectation, and allows for constant intertemporal elasticity of substitution preferences. A secondary advantage is that we can use our closed-form solutions to evaluate (to first-order) any number of counterfactuals for different combinations of shocks in different locations. In contrast, using the full non-linear model, one must instead solve each counterfactual separately using the extended numerical shooting algorithm discussed above. In Section G.3 of the online appendix, we show that this shooting algorithm is substantially more costly computationally, particularly for high-dimensional state spaces.

7   Conclusions

A classic question in economics is the response of economic activity to local shocks. In general, this response can be gradual, because of investments in capital structures and migration frictions. A key challenge in modeling these dynamics, is that agents' forward-looking decisions depend on the entire spatial distribution of economic activity across all locations in all future time periods. Our first main contribution is to develop a tractable dynamic spatial general equilibrium

framework that incorporates both forward-looking capital investments and dynamic migration decisions. Despite the many locations connected by a rich geography of trade and migration frictions, and the multiple sources of dynamics, we provide an analytical characterization of the conditions for the existence and uniqueness of the steady-state equilibrium.

Our second main contribution is to derive closed-form solutions for the first-order general equilibrium effect of shocks to fundamentals (productivities, amenities, trade costs and migration costs) along the entire transition path of the spatial distribution of economic activity. These sufficient statistics depend on four observable matrices for expenditure shares, income shares, outmigration shares and inmigration shares, the initial values of the state variables (population and the capital stock) in each location, and the model's parameters. We show that the economy's transition path is fully characterized by a second-order difference equation in these state variables. This second-order difference equation has a closed-form solution in terms of an impact matrix, which captures the initial impact of the shocks, and a transition matrix, which governs the subsequent evolution of the state variables in response to these shocks.

We use our closed-form solutions for the economy's transition path to provide an analytical characterization of the economy's dynamic response to shocks. We use an eigendecomposition of the transition matrix to distinguish between shocks and exposure to shocks. We introduce the concept of an eigen-shock for which the initial impact of the shock on the state variables is equal to a eigenvector of the transition matrix. We show that these eigen-shocks can be recovered from the observed data; any empirical shocks can be written as a linear combination of the eigen-shocks; and the weights or loadings in this linear combination can be recovered by regressing the empirical shocks on the eigen-shocks. We use these results to decompose the evolution of the state variables in response to empirical shocks into the contributions of each eigen-shock. We show that the speed of convergence to steady-state, as measured by the half-life, is determined by the corresponding eigenvalues of the transition matrix.

In our main empirical application, we use data on U.S. states from 1965-2015 to examine the decline of the "Rust Belt" in the North-East and Mid-West and the rise of the "Sun Belt" in the South and West. We show that this setting features convergence dynamics in capital and both net and gross migration, highlighting the relevance of a framework such as ours that incorporates both forward-looking investments and dynamic migration decisions. Already at the beginning of our sample period in 1965, we find that Rust Belt and Sun Belt states were substantially below and above their steady-state populations, respectively. By the end of sample period, all states are much closer to their steady-state population than they were at its beginning. We show that the initial distance of a state's population from its steady-state has substantial predictive power for subsequent population growth from 1965-2015, even after controlling for the initial levels of population and the capital stock and initial population growth.

We find slow average convergence to steady-state, with an average half-life in our baseline specification of around 20 years, which is consistent with recent empirical findings of persistent impacts of local shocks. We also find substantial heterogeneity in this speed of convergence, which depends on the size and incidence of shocks to productivity and amenities, and is in line with reduced-form evidence of uneven impacts of local shocks. Since capital and labor are complementary in the production technology, we show that these two sources of dynamics can either offset or reinforce one another, depending on the correlation between the impact of the shocks on the capital and labor state variables. Although our closed-form solutions for the economy's transition path are only exact in theory for small shocks, we show that in practice they provide a close approximation to the full non-linear model solution even for large shocks, such as the empirical distribution of decadal shocks observed during our sample period.

In a final empirical exercise, we implement our multi-sector extension using region-sector data on U.S. states and foreign countries from 1999-2015. We find lower average half-lifes in our multi-sector extension, which reflects the property of the data that there is greater mobility of labor across sectors within states than across states. Nevertheless, we again find substantial variation in these half-lifes, reinforcing the heterogeneity in the impacts of local shocks, which now depend on both their spatial and sectoral incidence. We find that these average half-lifes are relatively constant over time, which contrasts with the modest observed decline in geographical mobility. This pattern of results highlights that this decline in geographical mobility does not necessarily imply a rise in spatial frictions, since it is also influenced by convergence towards steady-state, and changes in the pattern and magnitude of shocks.

References

Adão, R., C. Arkolakis, and F. Esposito (2019): "Spatial Linkages, Global Shocks, and Local Labor Markets: Theory and Evidence," *NBER Working Paper*, 25544.

Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189.

Alder, S., D. Lagakos, and L. Ohanian (2019): "Labor Market Conflict and the Decline of the Rust Belt," Boston University, mimeograph.

Allen, T. and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 129, 1085–1140.

Allen, T., C. Arkolakis, and X. Li (2017): "Optimal City Structure," Yale University, mimeograph.

——— (2020): "On the Equilibrium Properties of Network Models with Heterogeneous Agents," *NBER Working Paper*, 27837.

Allen, T. and D. Donaldson (2020): "The Geography of Path Dependence," Dartmouth College, mimeograph.

AMIOR, M. AND A. MANNING (2018): "The Persistence of Local Joblessness," *American Economic Review*, 108, 1942–1970.

ANGELETOS, G.-M. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10, 1–30.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2012): "New Trade Models, Same Old Gains," *American Economic Review*, 102, 94–130.

ARMINGTON, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 16, 159–178.

ARTUC, E., I. BRAMBILLA, AND G. PORTO (2021): "Patterns of Labor Market Adjustment to Trade Shocks with Imperfect Capital Mobility," Universidad San Andres, Buenos Aires.

ARTUÇ, E., S. CHAUDHURI, AND J. MCLAREN (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–45.

AUCLERT, A., M. ROGNLIE, AND L. STRAUB (2020): "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model," Harvard University, mimeo.

AUTOR, D., D. DORN, AND G. HANSON (2020): "On the Persistence of the China Shock," MIT, mimeograph.

AUTOR, D., D. DORN, AND G. H. HANSON (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68.

AUTOR, D. H., D. DORN, AND G. H. HANSON (2016): "The China Shock: Learning from Labor Market Adjustment to Large Changes in Trade," *Annual Review of Economics*, 8, 205–240.

BALDWIN, R. (2001): "Core-Periphery Model with Forward-Looking Expectations," *Regional Science and Urban Economics*, 31, 21–49.

BAQAEE, D. R. AND E. FARHI (2019): "Networks, Barriers, and Trade," Harvard University, mimeograph.

BARTELME, D., T. LAN, AND A. LEVCHENKO (2019): "Specialization, Market Access and Real Income," University of Michigan, mimeograph.

BILAL, A. (2021): "Solving Heterogeneous Agent Models with the Master Equation," Harvard University, mimeograph.

BLANCHARD, O. J. AND L. F. KATZ (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 1, 1–61.

BONADIO, B., Z. HUO, A. A. LEVCHENKO, AND N. PANDALAI-NAYAR (2020): "Global Supply Chains in the Pandemic," *NBER Working Paper*, 27224.

BRYAN, G. AND M. MORTEN (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 127, 2229–2268.

CALIENDO, L., M. DVORKIN, AND F. PARRO (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835.

CALIENDO, L. AND F. PARRO (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 82, 1–44.

——— (2020): "The Quantitative Effects of Trade Policy on Industrial and Labor Location," Yale University, mimeograph.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *Review of Economic Studies*, 85, 2042–2096.

Caliendo, L., F. Parro, and A. Tsyvinski (2017): "Distortions and the Structure of the World Economy," *NBER Working Paper*, 23332.

Cassey, A. J. (2009): "State Export Data: Origin of Movement Versus Origin of Production," *Journal of Economic and Social Measurement*, 34, 241–268.

Costinot, A., D. Donaldson, and I. Komunjer (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo?s Ideas," *Review of Economic Studies*, 79, 581–608.

Costinot, A. and A. Rodríguez-Clare (2014): "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in *Handbook of International Economics*, ed. by E. Helpman, K. Rogoff, and G. Gopinath, Amsterdam: Elsevier North Holland, vol. 4, 197–262.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018): "The Geography of Development," *Journal of Political Economy*, 126, 903–983.

Desmet, K. and E. Rossi-Hansberg (2014): "Spatial Development," *American Economic Review*, 104, 1211–1243.

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

Dix-Carneiro, R. and B. K. Kovak (2015): "Trade Liberalization and the Skill Premium: A Local Labor Markets Approach," *American Economic Review*, 105, 551–557, papers and Proceedings.

——— (2017): "Trade Liberalization and Regional Dynamics," *American Economic Review*, 107, 1908–2946.

Duranton, G. and D. Puga (2004): "Micro-Foundations of Urban Agglomeration Economies," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson and J.-F. Thisse, Amsterdam: Elsevier, vol. 4, 2063–2117.

Eriksson, K., K. N. Russ, J. C. Shambaugh, and M. Xu (2019): "Trade Shocks and the Shifting Landscape of U.S. Manufacturing," *NBER Working Paper*, 25646.

Fajgelbaum, P. D., E. Morales, J. C. Suárez-Serrato, and O. Zidar (2019): "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 86, 333–376.

Feyrer, J., B. Sacerdote, and A. D. Stern (2007): "Did the Rust Belt Become Shiny? A Study of Cities and Counties that Lost Steel and Auto Jobs in the 1980s," *Brookings-Wharton Papers on Urban Affairs*, 41–89.

Fujita, M., P. Krugman, and A. J. Venables (1999): *The Spatial Economy: Cities, Regions, and International Trade*, Cambridge MA: MIT Press.

Galle, S., M. Yi, and A. Rodriguez-Clare (2018): "Slicing the Pie: Quantifying the Aggregate and Distributional Consequences of Trade," University of Calirfornia, Berkeley, mimeograph.

Glaeser, E. L. and J. Gyourko (2005): "Urban Decline and Durable Housing," *Journal of Political Economy*, 113, 345–375.

GLAESER, E. L. AND G. A. PONZETTO (2010): "Did the Death of Distance Hurt Detroit and Help New York?" in *Agglomeration Economies*, ed. by E. L. Glaeser, Chicago: Chicago University Press, chap. 10.

GREANEY, B. (2020): "The Distributional Effects of Uneven Regional Growth," University of Washington, mimeograph.

HARTLEY, D. (2013): "Urban Decline in Rust-Belt Cities," *Economic Commentary*, 2013-06.

HELPMAN, E. (1998): "The Size of Regions," in *Topics in Public Economics: Theoretical and Applied Analysis*, ed. by D. Pines, E. Sadka, and I. Zilcha, Cambridge: Cambridge University Press, 33–54.

HORNBECK, R. AND E. MORETTI (2018): "Who Benefits From Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality," *NBER Working Paper*, 24661.

HUO, Z., A. A. LEVCHENKO, AND N. PANDALAI-NAYAR (2019): "International Comovement in the Global Production Network," *NBER Working Paper*, 25978.

HYATT, H., E. MCENTARFER, K. UEDA, AND A. ZHANG (2018): "Interstate Migration and Employer-to-Employer Transitions in the United States: New Evidence from Administrative Records Data," *Demography*, 55, 2161–2180.

KAPLAN, G. AND S. SCHULHOFER-WOHL (2017): "Understanding the Long-run Decline in Interstate Migration," *International Economic Review*, 58, 57–94.

KAPLAN, G. AND G. L. VIOLANTE (2014): "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82, 1199–1239.

KIM, R. AND J. VOGEL (2020): "Trade and Welfare (Across Local Labor Markets)," *NBER Working Paper*, 27133.

KLEINMAN, B., E. LIU, AND S. J. REDDING (2020): "International Friends and Enemies," *NBER Working Paper*, 27587.

KLINE, P. AND E. MORETTI (2014): "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 129, 275–331.

KOVAK, B. K. (2013): "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?" *American Economic Review*, 103, 1960–1976.

KRUGMAN, P. (1991a): "History Versus Expectations," *Quarterly Journal of Economics*, 106, 651–667.

——— (1991b): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, 483–499.

——— (1992): "A Dynamic Spatial Model," *NBER Working Paper Series*, 4219.

KRUGMAN, P. AND A. VENABLES (1995): "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110, 857–880.

MATSUYAMA, K. (1991): "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics*, CVI, 617–50.

MOLL, B. (2014): "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation," *American Economic Review*, 104, 3186–3221.

MOLLOY, R., C. L. SMITH, AND A. WOZNIAK (2011): "Internal Migration in the United States," *Journal of Economic Perspectives*, 25, 173–196.

MONTE, F., S. REDDING, AND E. ROSSI-HANSBERG (2018): "Commuting, Migration and Local Employment Elasticities," *American Economic Review*, 108, 3855–3890.

MORETTI, E. (2011): "Local Labor Markets," in *Handbook of Labor Economics*, ed. by D. Card and O. Ashenfelter, Amsterdam: Elsevier North Holland, vol. 4b, 1238–1303.

PETERS, M. (2019): "Market Size and Spatial Growth: Evidence from Germany's Post-War Population Expulsions," Yale University, mimeograph.

RAMONDO, N., A. RODRÍGUEZ-CLARE, AND M. SABORÍP-RODRÍGUEZ (2016): "Trade, Domestic Frictions and Scale Effects," *American Economic Review*, 106, 3159–3184.

RAPPAPORT, J. (2007): "Moving to Nice Weather," *Regional Science and Urban Economics*, 37, 375–398.

REDDING, S. J. (2016): "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148–167.

——— (2020): "Trade and Geography," in *Handbook of Regional and Urban Economics*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, Amsterdam: Elsevier, forthcoming.

REDDING, S. J. AND E. ROSSI-HANSBERG (2017): "Quantitative Spatial Models," *Annual Review of Economics*, 9, 21–58.

REDDING, S. J. AND D. M. STURM (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98, 1766–1797.

UHLIG, H. (1999): "A Toolkit for Analyzing Non-linear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon and A. Scott, New York: Oxford University Press, 30–61.

WALSH, C. (2019): "Firm Creation and Local Growth," Yale University, mimeograph.

YOON, C. (2017): "Estimating a Dynamic Spatial Equilibrium Model to Evaluate the Welfare Implications of Regional Adjustment Processes," *International Economic Review*, 58, 473–497.