Mechanics of Spatial Growth *

Sheng Cai
Yale University

Lorenzo Caliendo
Yale University, NBER

Fernando Parro
Pennsylvania State University, NBER

Wei Xiang
Yale University

October 31, 2022

Abstract

We study the role that trade and internal migration play in the process of spatial and aggregate growth. We consider an economy in which growth is shaped by the best global and local ideas that contribute to the local stock of knowledge. Global ideas diffuse to locations that are more exposed to international trade. Local ideas diffuse across space when workers move to another location. We embed the diffusion of ideas through trade and migration into a dynamic spatial framework with trade, forward-looking migration decisions, and capital accumulation. We characterize the equilibrium properties of the model, and apply the framework to study China's spatial and aggregate growth during the 1990s and 2000s. International trade and internal migration are important mechanisms for idea diffusion that contributed to China's spatial and aggregate growth, with heterogeneous effects across space. Using patent data we provide further evidence of idea diffusion through trade and migration.

*We thank Jonathan Eaton, Samuel Kortum, Ezra Oberfield, Esteban Rossi-Hansberg, Ana Maria Santacreu, and many seminar and conference participants for their useful conversations and comments. We thank Matthew Murillo and Xilin Yuan for excellent research assistance. Correspondence by e-mail: charles.cai@yale.edu, lorenzo.caliendo@yale.edu, fxp5102@psu.edu, wei.xiang@yale.edu.

1 Introduction

Understanding economic growth requires understanding how countries accumulate factors of production and increase the productivity of such factors. In recent decades, the world has witnessed the successful growth experiences of developing countries such as Vietnam, Laos, the Philippines, and China, among others, where high economic growth has occurred hand-in-hand with increased trade openness, more internal migration, and high productivity growth. Aggregate economic growth is shaped in part by the process of development across space within a country, namely, the dynamics of the distribution of economic activity in space, the extent to which locations have differential exposure to trade, the internal mobility of labor, the local evolution of productivity, and other local characteristics. In this paper we develop a tractable dynamic spatial growth model to study quantitatively the process of spatial development across locations in a country and how this process shapes aggregate growth.

We consider a world economy with multiple countries and multiple locations within a country. Growth in each location is shaped by the endogenous evolution of total factor productivity, which is the outcome of the diffusion of global and local ideas. In each location, a continuum of differentiated goods is produced, and there are many potential producers of each good. Producers have heterogeneous productivities (ideas) to produce goods and those who are actively producing contribute to the local pool of ideas that determines the local stock of knowledge, namely, the local fundamental productivity. Ideas diffuse across locations because of trade in goods and the migration of workers. To make the process of diffusion tractable, we model the diffusion of ideas across space as a stochastic process. The productivity of each idea is a combination of an random original component, a random insight drawn from the ideas of sellers to that location, and a random insight drawn from workers in that location. In particular, global ideas are embedded in imported intermediate goods and diffuse more to locations relatively more exposed to international trade. Workers learn about local ideas and diffuse them across space when they migrate and interact with local producers in the destination location. The local pool of insights from workers contains local ideas from non-migrants and ideas from migrants. Building on the results in Buera and Oberfield (2020), we show that the distribution of productivities at each location follows a Fréchet distribution and that the evolution of the stock of knowledge at each location can be characterized by a system of difference equations. The idea diffusion process in our framework is endogenous since it depends on how connected the particular location is to the other locations through trade and migration, and it also depends on the quality of insights from those locations.

Since the productivities at each location follow a Fréchet distribution, we apply the results in Eaton and Kortum (2002) to model the trade and production structure in our framework. Producers with heterogeneous productivities (ideas) in each location source intermediate goods from the lowest-cost suppliers across countries and combine them with labor and capital to produce goods. The labor supply across locations is shaped by the forward-looking migration decisions of workers as in Caliendo et al. (2019). At each moment in time, workers supply labor, purchase local goods, and sort into different locations. Workers carry ideas from their previous location with them as they migrate, and they provide insights to producers in the current location. The supply of capital in each location features forward-looking landlords making investment decisions in local capital as in Kleinman et al. (2021).

Our framework allows for a rich heterogeneity across space in terms of trade openness, mobility frictions, factor endowments, and initial stock of knowledge. With all the margins previously described, our paper provides a tractable dynamic spatial growth model to study the role of spatial development on aggregate growth in general equilibrium. We show the existence and uniqueness of the balanced growth path equilibrium, and we show how to compute the model without assuming that the economy is on the balanced growth path at the initial period.

We use our dynamic spatial growth framework to study quantitatively the role of our spatial mechanisms in shaping aggregate and spatial growth. To do so, we look at an economy that is in transition, with heterogeneous locations in terms of stock of knowledge, initial supplies of labor and capital, exposure to international trade, and mobility frictions. We apply our framework to the Chinese economy, which features such spatial heterogeneity, and we study the mechanics of spatial growth in China in the 1990s and 2000s.

During the 1990s and far into the 2000s, China experienced fast economic growth, sustained capital accumulation, relocation of factors and production across space, and increased trade openness. This growth experience is sometimes called the China shock in the literature and has been primarily used to study the effects of import competition on labor markets and other outcomes in the United States and other countries. Less work has been devoted to understanding the spatial dynamics that shaped aggregate growth in China. We study the mechanics of spatial growth in China that led to fast aggregate growth during this period. We answer questions such as what was the role of international trade and internal migration in China's spatial growth during the 1990s and 2000s? We use our spatial growth framework to provide a quantitative answer to this question.

To do so, we take the model to the year 1990. We divide China into 30 provinces, and we group the rest of the countries in a rest of the world. We construct gross migration flows across provinces in China using census data. We condition gross flows by Hukou type to take into account that the Hukou system affects incentives to migrate and to return migrate and therefore contributes to uneven spatial growth, an aspect of China's development experience that we capture in our framework. We estimate elasticities that govern the rate of idea flows from trade and migration as well as the rate of innovation. To estimate these elasticities, we first obtain cross-sectional measures of fundamental productivity using a model inversion and generate a set of time series moments. With the moment conditions, we then apply the generalized method of moments (GMM). With our data and these estimates in hand, we proceed to our quantitative assessment.

We first study how initial conditions and our mechanisms shaped spatial growth in China in the 1990s and 2000s. We take the model to the data without assuming that the economy is on a balanced growth path. With the initial allocation in 1990—namely, trade openness, spatial labor mobility, factor endowments, and stock of knowledge across locations—we apply the dynamic-hat algebra method (Caliendo et al. (2019)) to study the role of 1990 fundamentals in spatial development in China. We ask the following question: How would China have developed absent of the changes to trade and migration costs that occurred after 1990? We find an important role of initial conditions in the subsequent aggregate growth in China. We find that ideas from sellers contributed more to aggregate growth than ideas from migrants. The intuition comes from the fact that all provinces benefit from trade openness and access to better global ideas from the rest of the world. The contribution of ideas from people is more uneven. In the short run, the local stock of knowledge grows faster in locations that receive migrants from high-productivity places relative to locations that receive migrants from less-productive places. Over time, migrants learn about local ideas and contribute further to the local stock of knowledge. In the case of China, due in part to the Hukou system, return migration from high-productivity locations also shapes part of China's spatial development. We also find an important role of capital accumulation; aggregate growth would have declined by around half in the absence of capital accumulation.

Turning to the spatial growth effects, we find that aggregate growth is shaped by large heterogeneity in growth rates across space. During the 1990s, provinces located in coastal areas such as Shanghai, Guangdong, and Hainan benefited from access to better insights from the rest of the world and experienced higher growth rates. Over time, spatial growth moderated and tended to equalize as the economy moved closer to the balanced growth path. The engines of aggregate growth also changed over time; notably, Guangdong became the main contributor to aggregate growth in China while other provinces located in the central and eastern parts of China became less relevant engines for aggregate growth. We also discuss how the initial distribution of fundamentals across space shaped subsequent spatial growth in China. We find that provinces with a higher initial stock of knowledge and with more international trade openness experienced higher growth.

While initial conditions seem to be important for understanding the process of spatial development and aggregate growth in China in the 1990s and 2000s, China also experienced reforms and policies that resulted in further changes to international trade and internal migration costs during this period, most notably the country's accession to the World Trade Organization and the elimination of some Hukou restrictions. We estimate changes in bilateral trade and migration frictions after 1990 and ask how these changes in fundamentals contributed to spatial and aggregate growth in China. We find that the change in trade costs and migration costs contributed to extra aggregate growth by about one percentage point annually and that the growth effects were very heterogeneous across space.

We also provide reduced-form evidence of idea diffusion through trade and migration. Measuring the local stock of knowledge in the data is a difficult task. To construct a proxy for it, we obtain province-level patent data and patent data for the rest of the world and use it along with our trade and migration data to provide empirical evidence of the roles played by trade and migration to diffuse ideas and to contribute to the local stock of knowledge. We find evidence of the mechanism for spatial growth through idea diffusion, namely provinces more open to trade and with more migrants from locations with larger stock of knowledge experience a relative larger growth in their knowledge stock. In addition, guided by the structural relationship between the local knowledge stock and idea diffusion through trade and migration from our model, we run an instrumental variable regression and find evidence consistent with our reduced-form results.

Our research is related to different strands of existing work. While our paper contributes to a large body of quantitative spatial economics literature (see Redding and Rossi-Hansberg (2017) for a review), it mainly engages with recent work on dynamic spatial models. The general equilibrium trade structure and forward-looking migration decisions build on Caliendo et al. (2019), where locations trade goods as in Eaton and Kortum (2002). We model workers' mobility decisions subject to frictions as a dynamic discrete-choice problem as in Artuc et al. (2010). We introduce capital accumulation and spatial growth into a dynamic framework with labor market dynamics and trade. As described previously, capital accumulation in our framework features forward-looking atomistic landlords making investment decisions in local capital to maximize intertemporal utility following the structure in Kleinman et al. (2021).[1]

The distinctive feature of our dynamic spatial framework is the presence of spatial growth. The process of innovation and diffusion that gives rise to the theory of total factor productivity in our model is a discrete-time version of Buera and Oberfield (2020).[2] We extend Buera and Oberfield (2020) to a spatial setting where ideas diffuse through goods as well as through workers who move across space. Our paper also relates to Cai and Xiang (2022), who study global growth and technology diffusion through multinational production. In our context, ideas diffuse not only globally but also locally. Our paper also complements recent spatial frameworks with innovation, local diffusion of technology, and spatial growth, most notably in Desmet and Rossi-Hansberg (2014) and Desmet et al. (2018), and frameworks with frictional idea diffusion across space (e.g., Berkes et al. (2022)).[3] Our framework shares some aspects with these papers such as the spatial heterogeneity in fundamentals and the

[1]The distinction between landlords and workers also relates to the formulations in Angeletos (2007) and Moll (2014), and as discussed later on, adds tractability in the context of a dynamic spatial model with forward-looking mobile workers. Capital accumulation in our dynamic spatial framework also connects to dynamic models of capital accumulation and international trade (e.g., Eaton et al. (2016), Alvarez (2017), Ravikumar et al. (2019), Anderson et al. (2019)), with the important difference that labor is assumed to be immobile across countries in that strand of the literature.

[2]The model in Buera and Oberfield (2020) also relates to Kortum (1997) when there is no idea diffusion from insights, and to Jones (1995) and Atkeson and Burstein (2019) in a model with intertemporal knowledge spillovers that are not modeled explicitly as a function of insights.

[3]See also Cruz and Rossi-Hansberg (2022), which studies the spatial effects of climate change.

geographic aspect of local idea diffusion. However, in our framework, technology diffuses spatially through trade and migration, both of which are endogenous, instead of being dictated by geographical distance or technological frictions. Our framework also departs from these papers by introducing forward-looking migration and capital accumulation decisions.

Also related to our paper, Eaton and Kortum (1999) develops a model of idea diffusion across countries where the distribution of productivities in each country follows a Fréchet distribution and the evolution of the stock of knowledge is characterized by a system of differential equations. In their model, ideas diffuse across countries exogenously, and countries are assumed to be under autarky otherwise. Building on Eaton and Kortum (1999), Cai et al. (2022) develops a trade and growth model with dynamics through innovation and technology diffusion across countries and sectors. In their model, ideas diffuse with exogenous and heterogeneous speeds across all sectors and countries. In contrast, in our model the speed of diffusion across locations is endogenous and mediated by trade and migration. Our framework also departs from these papers as it incorporates spatial growth with forward-looking migration and capital accumulation decisions.[4]

The process of idea diffusion from migrants in our framework is motivated in part by a growing literature with empirical evidence on knowledge flows resulting from people interactions (e.g., Atkin et al. (2022), Buzard et al. (2020)), and with empirical evidence on the impact of immigrants on ideas, innovation, and growth in the United States and in other countries (e.g., Kerr (2008), Hunt and Gauthier-Loiselle (2010), Lewis (2011), Akcigit et al. (2017), Bernstein et al. (2018), Sequeira et al. (2019), Arkolakis et al. (2020), Burchardi et al. (2020), Prato (2021)).[5]

Our paper also contributes to a strand of the literature that studies the role of trade and migration in shaping spatial inequality in China in the 2000s through the lens of static spatial frameworks (e.g., Tombe and Zhu (2019) and Fan (2019)). We depart from this line of research by studying growth in China in the 1990s through the lens of a dynamic spatial growth model, which allows us to study not only the cross

[4]Idea diffusion through trade in our paper is also related to other recent frameworks modeling innovation and diffusion of technologies as stochastic processes to study the connection between trade and the diffusion of ideas (e.g., Lucas (2009), Perla et al. (2021), Sampson (2016)).

[5]There is also recent evidence on how internal migrants impact productivity and other related outcomes in their destination location in countries like China that have experienced large internal migration episodes (e.g., Facchini et al. (2019), Imbert et al. (2022)).

sectional but also the time series implication of spatial development in China. Finally, our paper also relates to other strands of the literature that have pointed to different determinants of the rise of China. Caliendo and Parro (2022) provides a review of the recent literature on the origins of the China shock. With our dynamic spatial framework we study how spatial development through our mechanisms contributed to China's growth.[6]

The rest of the paper is structured as follows. In Section 2 we describe the process of idea diffusion for a single economy; we then introduce locations and present the dynamic spatial growth framework, and we characterize the equilibrium properties of the model. Section 3 describes how we take the model to the Chinese economy at the province level, our estimation strategy of the relevant elasticities, and the method used to perform counterfactual analysis. Section 4 presents our quantitative results, and Section 5 provides reduced-form evidence on the contribution of idea diffusion through trade and migration to local knowledge. Section 6 concludes.

2 Dynamic Spatial Growth Model

In this section, we develop the dynamic spatial framework with idea diffusion. In the next subsection we start by describing the process of innovation and diffusion that gives rise to the evolution of an economy's stock of knowledge in a single economy given a general source distribution of insights that producers might access.

2.1 Innovation and Idea Diffusion

The building block in our framework is a discrete-time version of Buera and Oberfield (2020). To simplify the exposition, consider a single economy in which there is a continuum of intermediate varieties produced in the unit interval. For each variety, there is a large set of potential producers who have different technologies to produce the good. Each potential producer is characterized by the productivity of her idea, which we denote by q, to produce an intermediate variety. Between time t and time $t+1$, producers interact with other agents in the economy and are exposed to new ideas to produce a variety. The productivity of a new idea might or might not be higher than that of the ideas the producer already has so she only adopts a new idea

[6]More generally, the effects of China's trade expansion on U.S. labor markets as well as other outcomes in different countries has been the focus of an extensive body of literature (e.g., Autor et al. (2013), Acemoglu et al. (2016), Pierce and Schott (2016), Caliendo et al. (2019)).

if the new ideas' productivity is greater than q. Both the number of new ideas and the productivity of them are stochastic, which generates randomness in the usage of the new ideas. In particular, the number of new ideas to which a producer is exposed is stochastic and follows a Poisson distribution.

Each new idea corresponds to a new productivity to produce the variety and is given by zq'^{ρ}. This new idea has two random components: z is the original component of the producer, drawn from an exogenous distribution $H(z)$; and q' is an insight drawn from a source distribution $G_t(q')$ whose evolution we describe subsequently. The stochastic arrival of new ideas generates randomness in the exposure to new ideas that are originated by producers and by new insights. Producers generate new ideas originated from their internal source of ideas, drawn from their own distribution of original ideas. Diffusion is a component that is external to the producer and that allows her to be exposed to the ideas of other producers/sellers in the economy. These ideas diffuse at a rate that is captured by the parameter ρ. In this context, the original component of the producer's ideas can also be interpreted as randomness in the adaptation of insights from others to alternative uses.

To gain tractability, in Assumption 1 we specify the internal distribution of original ideas, the process for the arrival of ideas, and the parametric restrictions required to characterize the evolution of the knowledge frontier over time. We then impose these assumptions, and in Proposition 1 we characterize the frontier of knowledge in the economy and the evolution of the stock of knowledge over time.[7]

Assumption 1

a) The internal distribution of original ideas is Pareto; $H(z) = 1 - (z/\bar{z})^{-\theta}$, *where* $\bar{z}$ *is the lower bound of the support and* $\theta > 1$ *is the shape parameter of the distribution.*

b) The strength of idea diffusion, $\rho \in [0, 1)$, *is strictly less than 1.*

c) The number of new ideas that arrive between t *and* $t+1$ *follows a Poisson distribution with mean* $\Lambda_t = \alpha_t \bar{z}^{-\theta}$.

d) The source distribution has sufficiently thin tail; i.e. $\lim_{\bar{z} \to 0} \bar{z}^{-\theta} \left[1 - G_t\left(\left(\frac{q}{\bar{z}}\right)^{\frac{1}{\rho}}\right)\right] = 0$.

In what follows we impose Assumption 1 to solve for the distribution of productivity in the economy. The next proposition presents the result.

Proposition 1. *Under Assumption 1, between* t *and* $t+1$, *the probability that the*

[7] We also refer the reader to Buera and Oberfield (2020) for a continuous-time-version derivation of the equilibrium evolution of technology in the economy.

best new idea has a productivity no greater than q, $F_t^{best\ new}(q)$, is given by

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty x^{\rho\theta} dG_t(x)\right).$$

Proof. See Online Appendix A.

Proposition 1 shows that the probability distribution of the best new idea is Fréchet with scale parameter θ and a location parameter determined by $\alpha_t \int_0^\infty x^{\rho\theta} dG_t(x)$. Note that in order to obtain this result there is no need to specify the external source distribution. This is an important result that we will use when we impose more structure over the source distribution. In addition, we can use the result of Proposition 1 to characterize the frontier of knowledge and its evolution over time. In particular, we denote by $F_t(q)$ the fraction of varieties whose best producer has productivity no greater than q. In a probabilistic sense, $F_t(q)$ is also the probability that the best productivity for a specific variety is no greater than q at time t. We call this object the *frontier of knowledge.* As the new ideas that arrive might have better productivity than the current best ideas, the evolution of $F_t(q)$ between t and $t+1$ follows

$$F_{t+1}(q) = F_0(q) \cdot \prod_{\tau=0}^{t} F_\tau^{best\ new}(q).$$

Proposition 2. *Assume that the initial frontier of knowledge at time 0 follows a Fréchet distribution given by $F_0(q) = \exp(-A_0 q^{-\theta})$. It follows that $F_t(\cdot)$ is Fréchet at any t given by*

$$F_t(q) = \exp\left[-\left(A_0 + \sum_{\tau=0}^{t-1} \alpha_\tau \int_0^\infty x^{\rho\theta} dG_\tau(x)\right) q^{-\theta}\right] = \exp\left(-A_t q^{-\theta}\right),$$

where the law of motion for the knowledge stock is given by

$$A_{t+1} = A_t + \alpha_t \int_0^\infty x^{\rho\theta} dG_t(x).$$

Proof. See Online Appendix A.

Proposition 2 establishes two results that we use in subsequent sections. First, the result indicating that at each moment in time the frontier of knowledge follows a Fréchet distribution allows us to specify the production and trade structure in our framework, as we describe in the next section. Second, we can see that both the arrival rate of new ideas α_t and the learning pool $G_t(\cdot)$ matter for the evolution of A_t. Later in the paper, after we describe the economic environment in our framework, we return to discuss how ideas diffuse over space and relate the learning pool $G_t(\cdot)$ to ideas from sellers and from migrants. Finally, note that it can also be shown that $F_t(\cdot)$ converges to a Fréchet asymptotically when $t \to \infty$, even without assuming that the initial frontier of knowledge follows a Fréchet distribution. Hence, the assumption about the initial frontier of knowledge is not strictly needed to obtain the result in Proposition 2.

2.2 Production, Factor Demand, and Trade

We now consider a world with N different locations indexed by i and n. At each location i there are heterogeneous and perfectly competitive producers of varieties of intermediate goods.[8] The technology to produce these intermediate goods requires labor and capital, which are the primary factors of production, and material inputs. The efficiency of an intermediate good producer is given by $q_{i,t}$, where we now index efficiencies by location. The output for a producer of an intermediate variety with efficiency $q_{i,t}$ in location i is given by

$$y_{i,t} = q_{i,t}\left(L_{i,t}^{\xi} K_{i,t}^{1-\xi}\right)^{\gamma} M_{i,t}^{1-\gamma},$$

where $L_{i,t}$, $K_{i,t}$, and $M_{i,t}$ are the demand for labor, capital, and material inputs. The parameters γ and $1-\gamma$ are the shares of value added and material inputs in output, and ξ and $1-\xi$ are the shares of labor and capital in value added. From the cost minimization problem of the producers, the unit price of an input bundle is given by $x_{i,t} = B\left(w_{i,t}^{\xi} r_{i,t}^{1-\xi}\right)^{\gamma} P_{i,t}^{1-\gamma}$, where $w_{i,t}$, $r_{i,t}$, and $P_{i,t}$ denote the price of labor, rental rate of capital, and the price of materials, and $B \equiv \left[\xi^{\xi}(1-\xi)^{1-\xi}\right]^{-\gamma}\gamma^{-\gamma}(1-\gamma)^{\gamma-1}$.

[8] As explained later on, at the beginning of the period producers get insights from sellers and migrants, with randomness in the productivity of those insights for alternative uses in the destination location. At the end of the period, technology to produce a variety can be imitated, and therefore, producers decide to charge a price equal to the marginal cost. Alternatively, we could have assumed producers engage in Bertrand competition so that the lowest-cost supplier of a variety either charge the optimal markup or set a limit price to just undercut the second-lowest cost supplier of the variety. As shown in Bernard et al. (2003) and Buera and Oberfield (2020), with Bertrand competition aggregate costs are a fraction $\theta/(1+\theta)$ of aggregate revenues in all locations, and under the assumption that profits from local producers are spent domestically, equilibrium conditions are isomorphic to those under perfect competition except for a constant in the price index.

We now use the results from the previous section in which we derived the law of motion of the stock of knowledge in an economy. Firms purchase intermediate goods from the lowest-cost supplier in the world. The frontier of knowledge in each location at each t is described by a Fréchet distribution with shape parameter θ and location-specific scale parameter $A_{i,t}$; namely, $F_{i,t}(q) = \exp(-A_{i,t}q^{-\theta})$.

Shipping goods across locations, from n to i, is subject to iceberg trade costs, $\kappa_{in,t}$, and therefore, the cost of purchasing an intermediate variety with efficiency q from n in location i is given by $\kappa_{in,t}x_{n,t}/q$. Hence, we can now follow the Eaton and Kortum (2002) formulation and derive the fraction of goods purchased by location i from location n as (see Online Supplement G.1 for the derivation), which is given by

$$\lambda_{in,t} = \frac{A_{n,t}\left(\kappa_{in,t}x_{n,t}\right)^{-\theta}}{\sum_{h=1}^{N} A_{h,t}\left(\kappa_{ih}x_{h,t}\right)^{-\theta}}. \tag{1}$$

Similarly, we can solve for the price index in location i, which is given by

$$P_{i,t} = T\left(\sum_{n=1}^{N} A_{n,t}\left(\kappa_{in,t}x_{n,t}\right)^{-\theta}\right)^{-1/\theta}, \tag{2}$$

where T is a constant.[9] Given this environment, total expenditure in location i, which we denote by $X_{i,t}$, is given by

$$X_{i,t} = (1-\gamma)\sum_{n=1}^{N} \lambda_{ni,t}X_{n,t} + I_{i,t},$$

which reflects that the total expenditure on goods is firms' expenditure on intermediate goods plus households' expenditure where a household's income is given by $I_{i,t} = w_{i,t}L_{i,t} + r_{i,t}K_{i,t}$. The term $\sum_n \lambda_{ni,t}X_{n,t}$ is the total demand for goods produced in i from all locations. The trade balance condition is given by

$$\sum_{n=1}^{N} \lambda_{in,t}X_{i,t} = \sum_{n=1}^{N} \lambda_{ni,t}X_{n,t},$$

where the left-hand side is the total imports by location i, and the right-hand side is the total exports from i (with domestic purchases entering both sides of the equation).

[9]Intermediate varieties are aggregated with a constant elasticity of substitution η, and T is a gamma function evaluated in the argument $T = \Gamma\left(1 + (1-\eta)/\theta\right)^{\theta/(1-\eta)}$.

Finally, using the expenditure equation, trade balance, and the relative demand for capital and labor, it follows that the labor market clearing condition can be expressed as

$$w_{i,t} L_{i,t} = \sum_{n=1}^{N} \lambda_{ni,t} w_{n,t} L_{n,t}. \tag{3}$$

2.3 Capital Accumulation Across Locations

We now turn to the supply side of the model. We start by describing capital accumulation decisions across space. At each location, we assume that there are atomistic landowners who consume local goods with logarithm preferences over consumption goods and whose source of income is from renting capital structures.[10] Landowners are forward-looking and seek to maximize the present discounted value of their utility by deciding how much to consume and invest at each moment in time. Landowners are geographically immobile, have access to an investment technology in local capital, and make their investment in units of consumption goods. We follow Kleinman et al. (2021) and interpret capital as buildings and structures that are geographically immobile once installed, and we specify the problem of a landowner in location i as

$$\begin{aligned} \max_{\{C_{i,t}, K_{i,t+1}\}_{t=0}^{\infty}} U &= \sum_{t=0}^{\infty} \beta^t \log(C_{i,t}), \\ \text{s.t. } r_{i,t} K_{i,t} &= P_{i,t} \left[C_{i,t} + K_{i,t+1} - (1-\delta) K_{i,t} \right] \text{ for all } t, \end{aligned}$$

where δ is the depreciation rate and $K_{i,0}$ is taken as given. The solution to this dynamic programming problem can be characterized by the policy functions on consumption and investment,

$$\begin{aligned} C_{i,t} &= (1-\beta) \left[r_{i,t}/P_{i,t} + (1-\delta) \right] K_{i,t}, \\ K_{i,t+1} &= \beta \left[r_{i,t}/P_{i,t} + (1-\delta) \right] K_{i,t}, \end{aligned} \tag{4}$$

which give rise to the law of motion of capital accumulation across locations. In Online Supplement G.3 we provide the detailed derivation of these policy functions. Note that since capital structures are accumulated locally and used for local production, the evolution of capital structures in part shapes the evolution of economic activity

[10] Our assumption regarding the logarithm preferences of landlords is consistent with the preferences we specify for workers in the next section.

across space. Similar to Kleinman et al. (2021), the immobility of landlords allows us to introduce forward-looking capital accumulation decisions in dynamic spatial economies with workers' mobility in a tractable way, and it prevents the number of state variables from increasing exponentially over time.[11]

We now turn to describe the dynamic labor supply decisions made by workers and migrants across locations in the model.

2.4 Dynamic Labor Supply Decisions

There is a continuum of heterogeneous forward-looking workers in the economy. Each worker observes the economic conditions and optimally decides where to locate in each period subject to mobility frictions and idiosyncratic taste shocks. We model this migration decision as a dynamic discrete-choice problem. In particular, workers maximize the present discounted value of their utility by deciding at each moment in time where to live. They supply one unit of labor inelastically at where they live, and they consume given their labor income ($w_{i,t}$) and the local price of goods ($P_{i,t}$). We denote by $U_{i,t}(c_{i,t})= log(c_{i,t})$ the current utility of a worker living in location i, where $c_{i,t} = w_{i,t}/P_{i,t}$. We assume that the decision of where to live the next period is affected by idiosyncratic amenity shocks that vary across locations denoted by $\epsilon_{n,t}$ and by mobility frictions of going from location i to location n, denoted by $m_{in,t}$. The presence of migration costs and idiosyncratic shocks generates a gradual adjustment of labor supply in response to changes in the economic environment.

As a result, the value of a worker in region i at time t is given by

$$v_{i,t} = log(w_{i,t}/P_{i,t}) + \max_{\{n\}_{n=1}^{N}} \{\beta E_t[v_{n,t+1}] - m_{in,t} + \nu\epsilon_{n,t}\}, \tag{5}$$

where β is the discount factor, which is assumed to be the same as the discount factor of landowners.

We assume that the idiosyncratic shocks $\epsilon_{n,t}$ are $i.i.d.$ realizations from a Gumbel (Type I Extreme Value) distribution with dispersion parameter ν. We denote by $E_t[v_{n,t+1}]$ the expectation at time t over the future realizations of the idiosyncratic shocks that shape the continuation value of each location. Using the properties of the Gumbel distribution, we can integrate both sides of equation (5) over $\epsilon_{n,t}$. We then

[11] As a result, this framework can accommodate alternative capital accumulation formulations such as assuming decreasing return to investment, as in Lucas and Prescott (1971) and Hercowitz and Sampson (1991).

obtain the value of location i for a representative worker in that location at time t, denoted by $V_{i,t} = E_t[v_{i,t}]$. The value of location i is given by

$$V_{i,t} = log(w_{i,t}/P_{i,t}) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta V_{n,t+1} - m_{in,t}\right)^{1/\nu}\right). \tag{6}$$

We denote by $\mu_{in,t}$ the fraction of workers that moves from location i to location n, which using the properties of the Gumbel distribution can be derived in closed form as

$$\mu_{in,t} = \frac{exp\left(\beta V_{n,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta V_{h,t+1} - m_{ih,t}\right)^{1/\nu}}. \tag{7}$$

This equilibrium condition determines the gross migration flows of workers across space (see Online Supplement G.2 for the derivation). It shows that individuals are forward-looking and decide where to supply labor tomorrow by evaluating the relative net future value of each location. The elasticity of the migration flow $(1/\nu)$ shapes how changes to migration costs affect migration flows. This expression for gross migration flows determines the evolution of the labor supply at each location i over time. In particular, the supply of workers at location i at time $t+1$ is given by the workers who decide to migrate to location i from all other locations n (including stayers in i) at time t. Therefore, the stock of workers at each location evolves according to

$$L_{i,t+1} = \sum_{n=1}^{N} \mu_{ni,t} L_{n,t}. \tag{8}$$

Having described the demand and supply sides of the model, in the next subsection we return to the idea diffusion process to specify the evolution of the local stock of knowledge across space as a result of trade and migration.

2.5 Idea Diffusion with Trade and Migration

We now specify the innovation and diffusion process described in Section 2.1 to allow for migrants and sellers to contribute to the local pool of ideas. To do this, we consider an economy in which producers in location n obtain new insights from two sources. First, producers obtain insights from sellers; namely, ideas from producers in other locations are embedded in imported intermediate varieties. Second, we assume that migrants carry insights with them when they arrive in a new location but the quality of those insights do not directly affect their wages or their migration decisions. The interpretation is that a migrant becomes exposed to the local ideas in their previous location, and then as they move across locations, they randomly meet a local producer. When they meet, the migrant shares ideas from her previous location and provides insights that can contribute to the local stock of knowledge. As a result, the productivity of a new idea that arrives can be generalized to

$$q = z q_\ell^{\rho_\ell} q_m^{\rho_m},$$

where q_ℓ is the insight drawn from a source distribution that is shaped by migration and q_m the insight drawn from a source distribution that is shaped by sellers. Note that under this functional form, having both the migrants and the foreign goods makes the new insight more productive than only having one of them, provided that migrants arrive from their origin locations with good insights.

The parameters $\rho_\ell, \rho_m \in [0, 1)$ capture the learning intensity from both types of insights (trade and migration) with $\rho_\ell + \rho_m < 1$. After imposing Assumption 1 and following the same steps as in Section 2.1, extending the notation by indexing the location by n, and given the results from Propositions 1 and 2, we obtain that the frontier of knowledge at each location is

$$F_{n,t}^{best\ new}(q) = exp\left(-A_{n,t} q^{-\theta}\right),$$

and the stock of knowledge evolves over time as

$$A_{n,t+1} - A_{n,t} = \alpha_t \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_{n,t}(q_\ell, q_m).$$

We assume that since q_ℓ is drawn from people and q_m is drawn from goods, they represent two different sources of (independent) ideas. Formally, when a worker from i at the end of period t decides to move to n, she carries with her an insight q_ℓ, which is a random draw from the frontier distribution in i, whose cumulative distribution function is $F_{i,t}(q_\ell)$. At the end time t, in location n, producers randomly meet a worker currently living in n, and the insight from this individual is the insight component of the new idea. Hence,

$$G_{n,t}(q_\ell) = \sum_{i=1}^{N} s_{in,t} F_{i,t}(q_\ell),$$

where $s_{in,t} = \frac{\mu_{in,t} L_{i,t}}{\sum_{h=1}^{N} \mu_{hn,t} L_{h,t}}$ is the share of workers in location n that arrived from i at the end of period t (see the derivation in Online Appendix A.2).

In the case of the source distribution of goods, we assume that there is learning from sellers as in Buera and Oberfield (2020); namely, that diffusion opportunities are randomly drawn from the set of best practices across all goods sold locally. In this way the source distribution $G_{n,t}(q_m)$ is given by the fraction of goods for which the lowest-cost provider of the good to location n is a producer in i with productivity less than or equal to q_m. Under these mechanisms for idea diffusion, we obtain that the difference equation that determines the evolution of the stock of knowledge at each location is given by

$$A_{n,t+1} - A_{n,t} = \alpha_t \Gamma_{\rho_\ell,\rho_m} \underbrace{\left[\sum_{i=1}^{N} s_{in,t} \left(A_{i,t} \right)^{\rho_\ell} \right]}_{people} \underbrace{\left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}} \right)^{\rho_m} \right]}_{goods}, \tag{9}$$

where $\Gamma_{\rho_\ell,\rho_m}$ is a constant given by $\Gamma(1-\rho_\ell) \times \Gamma(1-\rho_m)$ and where $\Gamma(x)$ is gamma function evaluated at x. In Online Appendix A.3 we present more details on the derivation of the law of motion of the stock of knowledge across locations with idea flows from people and goods.

Equilibrium condition (9) is quite intuitive. It shows that the local stock of knowledge evolves over time according to the arrival rate of new ideas α_t, according to how the location is connected and exposed to ideas from migrants, $s_{in,t}$, and according to how open the location is to trade, $\lambda_{ni,t}$. The diffusion of ideas from migrants and sellers is endogenous since both migration and trade patterns are equilibrium objects in our framework. The relative strength of idea diffusion, governed by the diffusion parameters ρ_ℓ and ρ_m, shapes the importance of learning from people or goods. The fact that there are diminishing returns to technological improvement from insights, given that the strength of idea diffusion is less than one, makes it harder to obtain insights that are good enough over time. Hence, if α_t is time-invariant, then as the knowledge frontier evolves over time, the growth rate of the stock of knowledge falls with a limiting value of zero. As a result, as the knowledge frontier evolves, ideas need to arrive faster over time in order to sustain a constant growth rate. This feature is shared by semi-endogenous growth models in Buera and Oberfield (2020), Jones (1995), Kortum (1997), and Atkeson and Burstein (2019). Given this, we make the following assumption about the arrival rate.

Assumption 2 *α_t has constant growth rate g_α given by*

$$\alpha_t = \alpha_0(1+g_\alpha)^t.$$

We now define formally the equilibrium of the dynamic spatial growth model.

Definition 1. Equilibrium. *Given an initial distribution of the local stock of knowledge $\{A_{i,0}\}_{i=1}^{N}$, factor endowments $\{L_{i,0}, K_{i,0}\}_{i=1}^{N}$, evolution of fundamentals $\{\alpha_0, \kappa_{in,t}, m_{in,t}\}_{i=1,n=1,t=0}^{N,N,\infty}$, and parameters and elasticities $(\rho_\ell,\ \rho_m,\ \theta,\ \nu,\ \gamma,\ \xi,\ \beta)$, the sequential competitive equilibrium of the dynamic spatial growth model is characterized by a sequence of values, factor prices, goods prices, labor allocations, capital stocks, and stock of knowledge, $\left\{V_{i,t}, w_{i,t}, r_{i,t}, P_{i,t}, L_{i,t}, K_{i,t}, A_{i,t}\right\}_{i=1,t=0}^{N,\infty}$, that satisfies the equilibrium conditions determined by the bilateral trade shares (1), the equilibrium location prices (2), the labor market clearing condition (3), the capital accumulation condition (4), the location value function (6), the worker gross flow condition (7), the law of motion of labor (8), and the evolution of the stock of knowledge (9).*

In the long run, as the economy evolves over time, it approaches a balanced growth path equilibrium in which all equilibrium variables grow at a constant long-run rate. We now characterize the balanced growth path of the model. We first formally define the balanced growth path. We then express all equilibrium variables in the model relative to their balanced growth rate (what we refer to as the detrended variables) and then show that the equilibrium conditions of the detrended model give rise to a unique solution. Namely, we show that there exists a unique balanced growth path of the dynamic spatial growth model.

Definition 2. Balanced Growth Path. *Along the balanced growth path all equilibrium variables grow at a constant rate. In particular, denote by g_y the growth rate of a generic variable y at the balanced growth path. At the balanced growth path the stock of knowledge grows at a rate $1+g_A = (1+g_\alpha)^{\frac{1}{(1-\rho_l-\rho_m)}}$, capital grows at a rate $1+g_k = (1+g_A)^{\frac{1}{\theta\xi\gamma}}$, and values grow at a rate $1+g_v = (1+g_A)^{\frac{1}{\theta\xi\gamma(1-\beta)}}$.*

Online Appendix B solves for the equilibrium long-run growth rates of all variables along the balanced growth path. The appendix also shows how to detrend all the equilibrium variables and equilibrium conditions, namely, how to express them relative to their balanced long-run growth. In particular,

Definition 3. Detrended Economy. *Denote with a "˜" the variable relative to its long-run growth. In the detrended economy* $\tilde{y}_t \equiv y_t/\left(1+g_y\right)^t$ *for all variables* y_t, *where* g_y *is the growth rate of variable* y_t *at the balanced growth path.*

The next proposition establishes the existence and uniqueness of the balanced growth path equilibrium. At the balanced growth path all the detrended variables are not growing, and as a result, the equilibrium variables of the detrended model reach a steady state. Hence, at the balanced growth path, $\tilde{y}_{t+1} = \tilde{y}_t = \bar{y}$, and it remains constant for all t. We use an upper bar to express the detrended equilibrium variables at the balanced growth path.

Proposition 3. Existence and Uniqueness. *Given parameters and elasticities* $(\rho_\ell, \rho_m, \theta, \nu, \gamma, \xi, \beta)$, *and fundamentals* $\{\alpha_0, \bar{\kappa}_{in}, \bar{m}_{in}\}_{i=1,n=1}^{N,N}$, *there exists a unique (up to scale) solution given by* $\{\bar{w}_i, \bar{r}_i, \bar{L}_i, \bar{K}_i, \bar{V}_i, \bar{A}_i\}_{i=1}^{N}$ *that satisfies the equilibrium conditions of the detrended model at the balanced growth path.*

Proof. See Online Appendix C.

Proposition 3 establishes that the model has a unique balanced growth path. The proof extends the results of Kleinman et al. (2021), and shows that the spectral radius of the matrix of elasticities of the non-linear system at the balanced growth path is equal to one, which establishes the uniqueness of the balanced growth path equilibrium in our spatial growth model up to a normalization.[12]

We now turn to quantitatively study the importance of our mechanisms for aggregate and spatial growth. To do so, we apply our framework to China, an economy that features heterogenous locations in terms of stock of knowledge, initial supply of labor and supply of capital, exposure to international trade, and mobility flows.

[12] In the proof we solve for the six eigenvalues that characterize the system of equilibrium conditions. The eigenvalues are $(1, 1, \frac{-b\pm\sqrt{b^2-4ac}}{2a}, 0, \rho_\ell + \rho_m)$ where $a = \beta + \nu + \theta\gamma\nu\xi$, $b = -\nu(1+\beta-\gamma\xi(1+\theta(1-\beta)))$, $c = \beta(\nu - 1 - \gamma\xi\nu(1+\theta))$. To gain further intuition of this result, Online Supplement H presents a series of partial results that helps us solve for the general model. For example, we present results for a version of the model with no idea flows. In this case we show that there are four eigenvalues, which are given by $(1, 1, \frac{-b\pm\sqrt{b^2-4ac}}{2a})$. We then consider the case of an economy with idea flows from sellers. We show that the equilibrium eigenvalues are $(1, 1, \frac{-b\pm\sqrt{b^2-4ac}}{2a}, \rho_m)$, where the additional eigenvalue compared to the model with no idea flows is exactly given by ρ_m, the strength of idea flows from trade. Similarly, in a model with only idea flows from migration, one obtains that the new eigenvalue is given by ρ_ℓ, the strength of idea flows from migration. Finally, the online supplement also presents the results of the general model.

3 Quantitative Analysis

During the 1990s and far into the 2000s, China experienced fast economic growth, considerable capital accumulation, shifts in the distribution of economic activity and factors of production across space, increased productivity, and trade openness. Caliendo and Parro (2022) reviews recent literature that describes the macroeconomic performance of China during the 1990s and 2000s and the different factors that contributed to China's growth.

We now turn to study spatial growth in China in the 1990s and 2000s through the lens of the dynamic spatial growth model developed in the previous section. We take the model to year 1990 in a world composed of 30 Chinese provinces and a constructed rest of the world. In doing so, we use migration, production, and value added data. We also use trade data between provinces and the rest of the world. Importantly in the case of China, where there are well-defined mobility frictions across provinces, we condition gross migration flows across provinces by Hukou status. To understand how the Hukou system works, think about a province-level "passport" that identifies an individual based on their province of origin and restricts non-locals' access to certain amenities.

Accordingly, in the quantitative analysis we extend our framework to take into account these considerations. In particular, we allow for workers with different Hukou statuses to value locations differently, as Hukou restrictions give them access to different amounts of amenities, and we also allow workers to face different mobility restrictions. In equilibrium, this implies different mobility rates across provinces for individuals with different Hukou statuses that we discipline in the data. Hence, the equilibrium conditions of the dynamic labor supply decisions of workers are now given by

$$V_{i,t}^{H} = log(\psi_i^H w_{i,t}/P_{i,t}) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta V_{n,t+1}^{H} - m_{in,t}^{H}\right)^{1/\nu}\right), \tag{10}$$

$$\mu_{in,t}^{H} = \frac{exp\left(\beta V_{n,t+1}^{H} - m_{in,t}^{H}\right)^{1/\nu}}{\sum_{g=1}^{N} exp\left(\beta V_{g,t+1}^{H} - m_{ig,t}^{H}\right)^{1/\nu}}, \tag{11}$$

$$L_{i,t+1} = \sum_{H}\sum_{n=1}^{N} \mu_{ni,t}^{H} L_{n,t}^{H}, \tag{12}$$

where the H index denotes the Hukou status and ψ_i^H is the amenity parameter of location i for an individual with Hukou status H. Once in the same location, workers with different Hukou statuses consume the same basket of goods and earn the same real wages although their levels of utility are different because they have access to different amenities. In this way, we aim to capture a characteristic of this economy in transition: that is, that migrants to a given province registered in a different province have access to different amounts of amenities, face different mobility costs, and as a result, make different migration decisions compared with migrants registered in the destination province. We later provide some descriptive evidence of the importance of two-way migration across provinces in China in part due to the Hukou restrictions.

We now proceed to describe the data sources we use in our quantitative analysis. In Online Appendix E we further describe data sources and data construction.

3.1 Data

To bring the model to the data, we need data across provinces in China and for the rest of the world on bilateral trade shares $\lambda_{in,t}$, total expenditure $X_{i,t}$, value added $w_{i,t}L_{i,t} + r_{i,t}K_{i,t}$, the distribution of employment $L_{i,t}$, and migration flows across provinces $\mu_{in,t}$, conditional on Hukou type. We also need the share of value added in gross output γ, the share of labor in value added ξ, and the initial capital stocks $K_{i,0}$. In addition, we need estimates of the trade elasticity θ, the migration elasticity $1/\nu$, the discount factor β, and the depreciation rate δ. We later describe how we discipline the elasticities that govern innovation and idea diffusion $(\alpha_0, \rho_l, \rho_m)$.

We consider a model in which each period represents five years. Hence, we use a discount factor β of 0.86, equivalent to an annual discount factor of 0.97, which implies a yearly interest rate of roughly 4 percent. The trade elasticity $\theta = 4.55$ is obtained from Caliendo and Parro (2015). We set a migration elasticity of $1/\nu = 0.15$, which is the value estimated by Cruz (2021) for a five-year period in a sample of developing countries. We set a depreciation rate $(1-\delta) = 0.95^5$, which corresponds to an annual depreciation rate of 5 percent. We compute the values of $\gamma = 0.38$ and $\xi = 0.54$, which correspond to the parameter values for the year 1990 from the world's aggregates in the Eora multi-region input-output table. Finally, we set a value of $g_\alpha = 0.013$ that matches the long-term productivity growth in the U.S. economy that during the great moderation period in the 1990s was arguably on a balanced growth path.

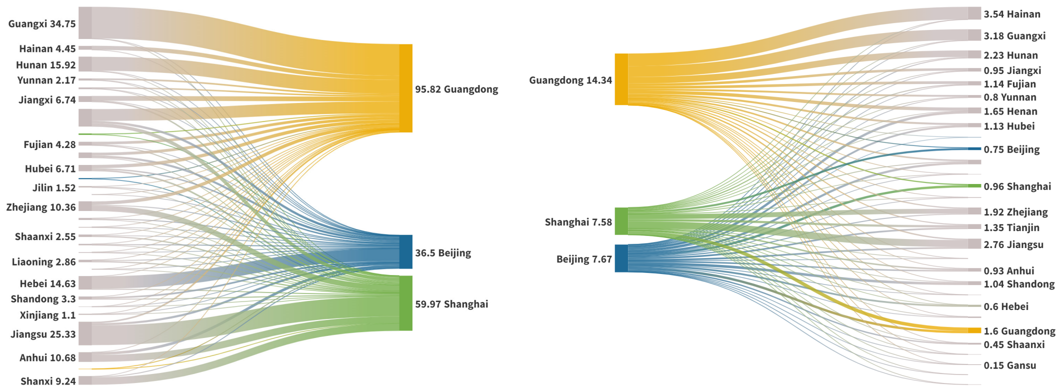
Gross Migration Flows. We obtain five-year mobility rates across provinces in China from the 1 percent sample of the 1990 census from IPUMS. The census data contains both the location (province) in 1990 and the location (province) five years ago. We take the working-age (15-64 years old) population as our sample. Furthermore, we keep respondents who are actively employed in 1990. To check the representativeness of our sample, we compute the employment share of each province out of the nationwide employment, and we compare it to the data counterpart provided in the 1991 China Statistics Yearbook.

To condition the gross flows on Hukou type, we proceed as follows. We use information from the 1990 census on the status and nature of registration. In particular, if the individual has the status "residing and registered here", we use the current location in 1990 as the registration location. If the individual has the category "residing here over 1 year, but registered elsewhere", "living here less than 1 year and absent from registration place over 1 year", or "living here with registration unsettled", we use the person's location in 1985 as the registration location; otherwise, the individual was living abroad, and we drop such observations (less than 0.02 percent of the observations). For those who registered non-locally yet resided in the same province in 1985 and 1990, we assign their registration location with a probability given by the immigration share from that location.

As an illustration, Figure 1 displays the mobility flows from all provinces to Beijing, Shanghai, and Guangdong (left-hand panel) and from these provinces to the rest of the provinces in China (right-hand panel). Origin provinces are on the left axis and destination provinces are on the right axis, and a thicker line in the figure means a larger flow. As we describe in the next section, these three provinces have higher initial measured productivity, and as expected, in the left-hand panel we see how they receive migrants from all provinces in China. In the right-hand panel, we also observe how migrants move from these high-productivity places to the rest of China, which is an indication of the importance of return migration in China due in part to the Hukou restrictions as well as how return migrants diffuse knowledge from high-productivity places.

Figure 1: Mobility across provinces in China (1985-1990)

Mobility flows across provinces



Note: The figure illustrates mobility flows across provinces in China. The panels display the mobility flows (in 10,000 people) for the selected provinces, where the left axis presents the origin provinces and the right axis shows the destination provinces.

Trade and Production Data. We obtain export and import data between Chinese provinces and the rest of the world from the China Compendium of Statistics, 1949-2008. We also obtain GDP and employment data across provinces from the same source. The GDP for the rest of the world is obtained from the Penn World Table 10.0 (PWT). The PWT reports real GDP at constant 2017 national prices; hence, we convert real GDP for the rest of the world to 1990 prices using the world GDP deflator from the World Bank's World Development Indicators. To estimate the series of capital stock across provinces, we follow Shan (2008) and apply the perpetual inventory method, using fixed capital formation from the China Compendium of Statistics as the measure of investment and estimates of capital stocks at a base year from Young (2003). For the rest of the world, we obtain the capital stock at constant 2017 national prices from the PWT, which we convert to 1990 prices using GDP deflators from the same source. Using our constructed series of capital stock and equation (4), we obtain the initial real rental rates across locations.

Finally, we point out that in the quantitative analysis we abstract from trade across provinces and sectoral heterogeneity given the lack of data along these dimensions in the Chinese statistics for the 1990s and even for more recent years. As a result, our quantitative analysis will center on the role of local idea diffusion through internal migration and global idea diffusion through international trade.[13]

[13]Still, mobility across provinces in part captures the mobility of workers between sectors (e.g,

3.2 Initial Stock of Knowledge

To estimate the initial stock of knowledge across locations, we start with the definition of real GDP. In our model, real GDP in location n at $t = 0$ is given by

$$Real\ GDP_{n,0} = \frac{w_{n,0}L_{n,0} + r_{n,0}K_{n,0}}{P_{n,0}} = \left(A_{n,0}/(\lambda_{nn,0}\Upsilon)\right)^{\frac{1}{\gamma\theta}} \left(K_{n,0}\right)^{(1-\xi)} \left(L_{n,0}\right)^{\xi}, \quad (13)$$

where $\Upsilon = (BT)^{\theta} (1-\xi)^{(1-\xi)\gamma\theta} (\xi)^{\xi\gamma\theta}$.[14] Real GDP in our model is determined by factor accumulation (capital, labor) and by measured productivity. In particular, measured productivity is captured by the term $\left(A_{n,0}/(\lambda_{nn,0}\Upsilon)\right)^{\frac{1}{\gamma\theta}}$. It has two main components: fundamental productivity $A_{n,o}$, and trade openness captured by the inverse of the domestic expenditure share $\lambda_{nn,0}$. The intuition is that in a closed economy— namely, when $\lambda_{nn,0} = 1$—measured productivity is the same as fundamental productivity $A_{n,o}$, which is the average efficiency of the set of goods produced and consumed in n. In an open economy, firms purchase a fraction of goods from abroad and produce only that set of goods of which they are the lowest-cost supplier in the world. Hence, a smaller domestic expenditure share $\lambda_{nn,0}$ results in firms in n producing a smaller set of goods with higher marginal efficiency.

Inverting equation (13), and solving for fundamental productivity $A_{n,0}$, we obtain

$$A_{n,0} = \Upsilon \left(\frac{Real\ GDP_{n,0}}{\left(K_{n,0}\right)^{1-\xi} \left(L_{n,0}\right)^{\xi}} \right)^{\gamma\theta} \lambda_{nn,0}. \quad (14)$$

Using the data described in the previous subsection, we compute the initial stock of knowledge across provinces in China as well as for the rest of the world.[15] Figure 2 presents the initial stock of knowledge (year 1990) across locations. In the upper panel, we see that the 1990 stock of knowledge for provinces in China is smaller than that for the rest of the world. Across provinces in China, the initial stock of knowledge is very heterogeneous, with Shanghai, Liaoning, and Guangdong being the top three provinces in terms of the initial stocks of knowledge, and Gansu, Guizhou, and

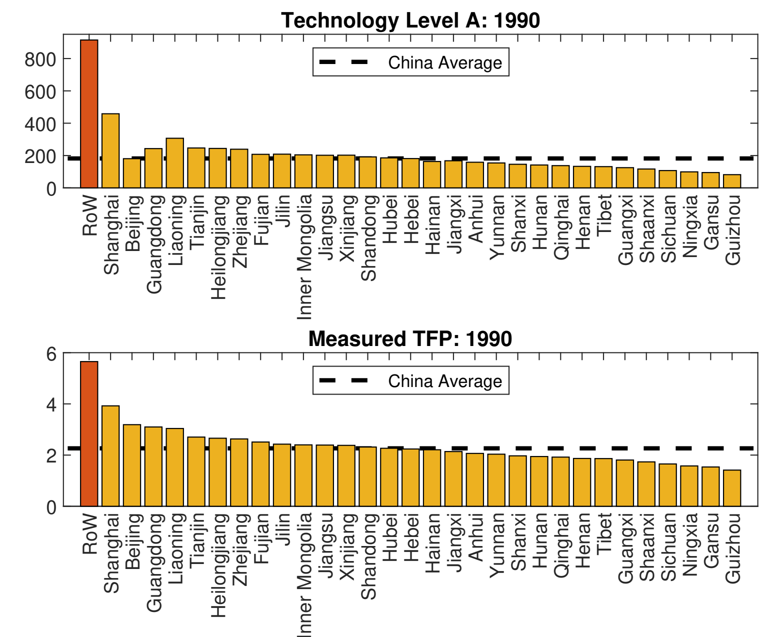
agriculture, non-agriculture) given the very uneven distribution of economic activity in China. Previous work that has incorporated internal trade across provinces has imputed these data in different ways (e.g., Tombe and Zhu (2019), Poncet (2003)).

[14]See Online Supplement G.4 for the details of this derivation.

[15]We set a value of $\eta = 2$ in the gamma function in equation (14).

Ningxia the bottom three provinces. The bottom panel presents the 1990 measured productivity across locations, which corrects for the impact of trade as previously explained. Again we observe that the rest of the world has higher measured productivity in 1990 than the provinces in China. We can see that Shanghai, Beijing, and Guangdong are the top three provinces with the highest measured productivity, whereas Gansu, Guizhou, and Ningxia are the bottom three provinces.

Figure 2: Initial stock of knowledge and measured productivity across locations (1990)



Note: The figures present the initial stock of knowledge (upper panel), computed as described in this section, and measured TFP (bottom panel), computed as $(A_{n,0}/(\lambda_{nn,0}/\Upsilon))^{1/\gamma\theta}$.

3.3 Estimation of Idea Diffusion from Trade and Migration

In our dynamic spatial growth model, three parameters discipline productivity growth and idea diffusion across locations: the strength of idea diffusion through sellers ρ_m, the strength of idea diffusion through migrants ρ_l, and the arrival rate of insights α_0. There are no benchmark values for these parameters in the literature that we can use as points of comparison.[16] To discipline these parameters in the our dynamic spatial model, we proceed as follows.

[16] Buera and Oberfield (2020) present an estimate of ρ_m by using aggregate cross-country data and obtain a value of $\rho_m = 0.6$. The authors calculate the arrival rate of ideas as the residual needed to explain the evolution of TFP backed out by their calculation.

We first measure fundamental productivity, $A_{n,t}$, by geography for different periods of time. We do this using equation (14), as described in Section 3.2. It is important to emphasize that our estimated fundamental productivities for the various periods of time are cross-sectional measures; we do not impose any structure on how each of these measures might be related over time. After obtaining cross-sectional estimates of fundamental productivity, we then create moment conditions related to the evolution of fundamental productivity over time and compute the same moments using the model-implied fundamental productivity from equation (9). We then use the GMM to estimate the parameters of interest following Hansen and Singleton (1982) and Newey (1985). One key advantage of this method is that we do not need to make assumptions about the statistical distribution of the data. Instead, we calibrate the parameters to match moment conditions.

We denote by $\Theta \equiv (\alpha_0, \rho_m, \rho_l)'$ the set of parameters that we want to estimate. We define the moment conditions or orthogonality conditions as $g_N\left(\hat{\Theta}\right) \equiv (1/N)\ h\left(\varpi_t, \hat{\Theta}\right)$, where $h\left(\varpi_t, \hat{\Theta}\right)$ is a vector of moment conditions and $\varpi_t \equiv (\lambda_{in,t}, s_{ni,t}, A_{i,t}, A_{i,t+1})'$ is the vector containing all the available information used for the estimation. We use fundamental productivity estimates for the period 1990-2000 and five moment conditions: the first moment is the average change in fundamental productivity levels across locations; the second moment is the average growth rate in fundamental productivities; the third moment is the variance in the time changes in fundamental productivity levels; the fourth moment is the covariance between the initial fundamental productivities and the change in fundamental productivity levels; and the fifth moment is the covariance between the initial fundamental productivities and the growth rate in fundamental productivities.

To provide further intuition on how these five moments help identify the innovation and diffusion parameters, we note that the first two moments help us identify α_0 since the initial arrival rate of ideas scales up productivity everywhere. The third moment helps us separate α_0 from the diffusion parameters ρ_m and ρ_l since they provide information about how heterogeneity in trade openness and mobility flows result in cross-province variations in the stock of knowledge over time. The last two moments provide information to disentangle ρ_m from ρ_l. The intuition is that provinces with a higher initial stock of knowledge tend to be more open to trade and therefore benefit more from the global diffusion of ideas from the rest of the world. Hence, ideas from sellers tend to generate a positive covariance between the initial stock of knowledge and subsequent changes in productivity. On the other hand, ideas from people are not necessarily associated with a positive covariance; this depends on whether locations receive migrants from places with relatively good insights.

GMM searches for the parameters that minimize all moment conditions by solving the optimization problem $\hat{\Theta} = \arg\min : g_N\left(\hat{\Theta}\right)' W g_N\left(\hat{\Theta}\right)$, where W is a weighting matrix that weights how different linear combinations of moments account for the data. The parameters are then estimated following an iterative process. We start by solving first the minimization problem setting $W = I$ (identity matrix), and after that, we construct the weighting matrix with the estimated parameters and solve the problem again until the estimation is approximately equal to the one from the previous iteration.[17] Finally, in our GMM estimation we allow for an error term that together with α_0 captures the effects of determinants of the evolution of the knowledge stock other than ideas from goods and ideas from people. In Online Appendix E.1 we display the empirical moment conditions and the model-implied moments predicted by the evolution of fundamental productivity using equation (9). Using this procedure, we obtain our preferred estimates of $\rho_l = 0.2$, $\rho_m = 0.61$ and $\alpha_0 = 0.18$.

3.4 Computing Counterfactuals

To compute the dynamic spatial growth model, we apply dynamic-hat algebra techniques developed in Caliendo et al. (2019) and show that by expressing the equilibrium conditions in relative time differences, we are able to compute the model without needing to estimate the levels of exogenous fundamentals or assuming that the economy is in the balanced growth path in the initial period. The intuition is that solving the model in relative time differences requires conditioning the model on observable allocations, which contain all the information about the fundamentals, and matching the cross-section of the actual economy in the initial year that does not need to be in a balanced growth path. The next proposition establishes the result.

Proposition 4. Dynamic-Hat Algebra. Define $\hat{y}_{t+1}$ as the time difference of the detrended endogenous variable $\tilde{y}$; namely, $\hat{y}_{t+1} = \left(\tilde{y}_{t+1}/\tilde{y}_t\right)$. *Given an initial observed allocation* $\left\{\{\lambda_{in,0}\}_{i=1,n=1}^{N,N}, \{\mu_{in,0}\}_{i=1,n=1}^{N,N}, \{w_{i,0}L_{i,0}\}_{i=1}^{N}, \{K_{i,0}\}_{i=1}^{N}, \{L_{i,0}\}_{i=1}^{N}\right\}$,

[17] Hansen (1982) shows that the optimal weighting matrix is given by the inverse of spectral density at frequency zero of the error terms, the inverse of the long-run variance-covariance of $h(\varpi, \hat{\Theta})$. We use the Newey-West correction to estimate the long-run variance-covariance.

parameters and elasticities $(\rho_\ell, \rho_m, \theta, \nu, \gamma, \xi, \beta)$*, initial rate and growth rate in the arrival of ideas* (α_0, g_α) *and a convergent sequence of future changes in fundamentals under perfect foresight* $\{\hat{\kappa}_{in,t}, \hat{m}_{in,t}\}_{i=1,n=1,t=1}^{N,N,\infty}$*, the solution for the sequence of changes in the model's endogenous variables in the detrended model* $\{\hat{y}_{t+1}\}_{t=1}^{\infty}$ *does not require information on the level of fundamentals (trade and migration costs).*

Proof. See Online Appendix D

In the detrended balanced growth path, $\hat{A}_n = 1$, and therefore $\hat{y} = 1$ for all variables $\tilde{y}$. We use this property of the detrended model to develop an algorithm to compute counterfactuals in the dynamic spatial growth model, which is described in Online Supplement J. In addition, as the proposition establishes, solving the model in relative time differences requires conditioning the model on the initial observable allocations $\lambda_{in,0}$, $w_{i,0}L_{i,0}$, $L_{i,0}$, $\mu_{in,0}$, and $K_{i,0}$, and elasticities θ, ν, β, δ, ρ_ℓ, ρ_m, and α_0. The previous sections have described our process for collecting these initial allocations and disciplining the parameters and elasticities in our framework.

4 Mechanics of Spatial Growth in China

In this section we present our quantitative analysis of the mechanics of spatial growth in China. We first study the role of initial conditions in shaping spatial and aggregate development in China in the 1990s and 2000s. We then explore the effects of changes in fundamentals (international trade costs and internal migration restrictions) after 1990 on spatial and aggregate growth.

4.1 Role of Initial Conditions

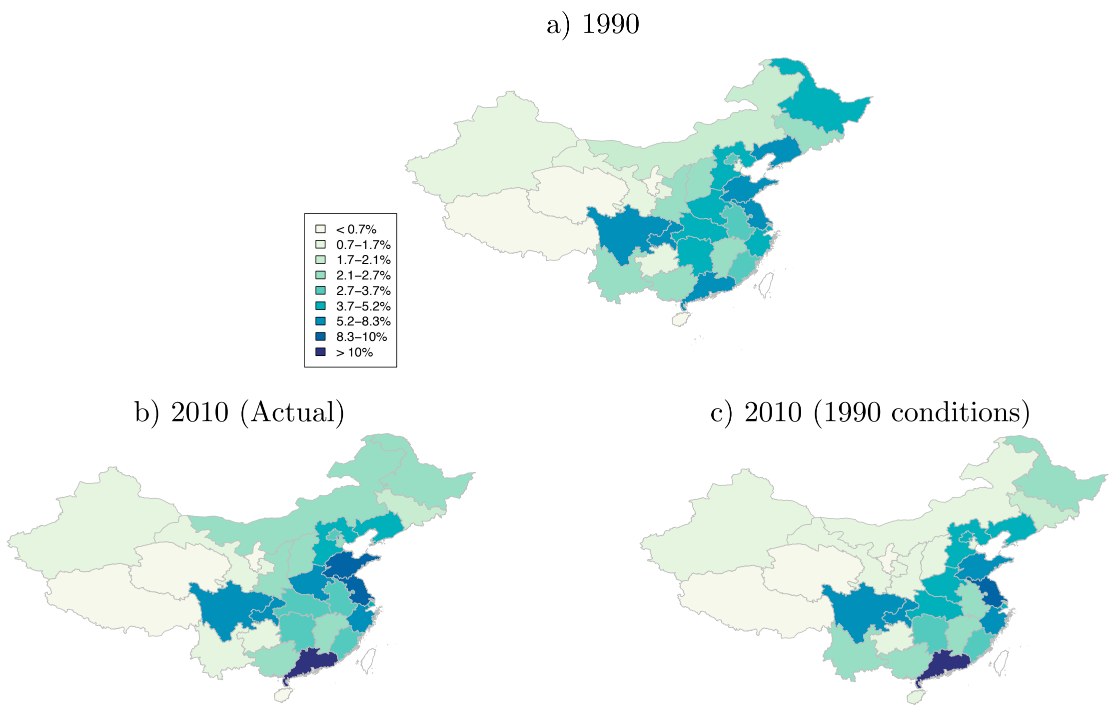
We first use our framework to study how initial conditions shaped spatial growth in China in the 1990s and 2000s. To do so, as we described before, we take the model to the data in the year 1990. The model does not assume that the economy is in a balanced growth path in the initial year since the model is taken to the data in the initial period conditioning on observed allocations. We then compute the economy with 1990 fundamentals and the endogenous evolution of productivity; namely, we answer the counterfactual question: How would the provinces in China and the rest of the world have looked if fundamentals (trade and migration costs) had stayed at their 1990 levels and the changes in the stock of knowledge operate due to the mechanisms in the spatial growth model?

4.1.1 Regional Distribution of Economic Activity

We start by describing the initial distribution of economic activity in China and its evolution in subsequent decades, and we explore the role of initial conditions in shaping the distribution of economic activity in China across the 1990s and 2000s.

Figure 3 presents the actual GDP shares across provinces in the year 1990 and the GDP shares twenty years later in 2010. The left column of the bottom panels shows the actual GDP shares across provinces in China in 2010, and the right panel presents the predicted GDP shares in 2010 under the initial conditions.

Figure 3: Regional distribution of economic activity (GDP shares)



Note: The figures show the distribution of economic activity across provinces in China, measured as GDP shares, in the data and with 1990 fundamentals over the period 1990-2010.

Starting with the initial GDP shares across provinces in China, we can see that economic activity tends to concentrate in the center and coastal areas of China, with Guangdong, Shandong, and Jiangsu being the largest provinces in terms of GDP in 1990. In 2010, in the bottom left panel, we can see a persistent concentration of economic activity in the same areas of China, but unlike in 1990, in 2010, economic activity tends to move from the central part of China to the coastal areas. For instance, Guangdong, Jiangsu, Zhejiang, Fujian, and Shandong are all provinces that increased their GDP shares, and these provinces are located in the coastal areas of China. In the bottom right panel, we see a similar pattern predicted by the model, pointing to the role of initial conditions in shaping the redistribution of economic activity across space during the subsequent two decades. In Online Supplement K.1 we present the evolution of the GDP shares across provinces in China every five years, which displays the same pattern described in this subsection.

4.1.2 Aggregate Growth in China

We next study the implication of initial conditions and spatial development described in the previous subsection for aggregate growth in China during the 1990s and 2000s. Table 1 presents the annual real GDP growth in China for different time frames over the period 1990-2020 in the data as well as the real GDP growth over the same periods if fundamentals (trade and migration costs) had stayed constant at the 1990 levels. We also quantify the contribution of idea diffusion through goods and people, and the contribution of capital accumulation, to aggregate growth in China.

Comparing the actual real GDP growth and that under the initial 1990 conditions in the first two rows of the table, we find that initial conditions play an important role in explaining subsequent growth in China during the 1990s and 2000s. As discussed in the review by Caliendo and Parro (2022), many reforms in China that involve changes to trade and industrial policy took place before the 1990s. In our model, these reforms are captured by the initial conditions. As previously described, our methodology does not assume that the economy is on the balanced growth path in the initial period. The framework is taken to the actual data in 1990, and therefore, the actual initial allocations contain information about fundamentals and policies in the Chinese economy and the rest of the world up to that year. We find an important role of these initial conditions in aggregate growth in China in the decades after 1990.

In the next two rows of the table, we evaluate the contribution of idea flows from people and from goods to aggregate growth in China. In the third row, we compute the model assuming $\rho_l = 0$; namely, that productivity evolves endogenously only due to idea flows from goods. We find that without idea diffusion through people, aggregate growth in China would have been smaller but still significant. In the fourth row of the table, we quantify aggregate growth in China with idea diffusion from people only; namely, in a model with $\rho_m = 0$. We find that without idea flows from goods,

Table 1: Annual GDP growth rate

	90-95	90-00	90-05	90-10	90-15	90-20
Actual GDP growth	12.3%	10.4%	10.2%	10.5%	9.9%	9.3%
With fundamentals in 1990	10.6%	10.1%	9.6%	9.2%	8.9%	8.6%
W/o ideas from people ($\rho_l = 0$)	8.7%	8.0%	7.4%	6.9%	6.5%	6.2%
W/o ideas from goods ($\rho_m = 0$)	6.4%	5.4%	4.7%	4.1%	3.7%	3.4%
W/o capital accumulation	5.0%	4.9%	4.7%	4.6%	4.5%	4.5%
W/o idea diffusion ($\rho_l = 0$, $\rho_m = 0$)	6.1%	5.1%	4.4%	3.8%	3.4%	3.0%

Note: Aggregate real GDP growth is obtained from the World Development Indicators. GDP growth with 1990 fundamentals is computed by solving the dynamic spatial growth model with constant fundamentals. The growth rate without idea flows from people is obtained by computing the model with $\rho_l = 0$, and the growth rate without idea flows from goods is obtained by computing the model with $\rho_m = 0$. The second to last row presents the aggregate GDP growth in the absence of capital accumulation. The last row presents the aggregate growth with capital accumulation and no idea diffusion, obtained by computing the model with $\rho_l = 0$ and $\rho_m = 0$.

aggregate growth in China would have been even smaller. Intuitively, two factors explain the larger importance of idea flows from goods for aggregate growth in China. First, as described in Section 3.2, the initial stock of knowledge across provinces in China is lower than it is in the rest of the world; hence, international trade makes an important contribution to growth through the diffusion of good ideas from the rest of the world to all provinces in China. At the same time, the contribution of idea flows from people can have offsetting effects on growth since return migration from high-productivity places fosters growth in the stock of knowledge in the destination province but receiving migrants from low-productivity locations slows down the process of knowledge accumulation. As we described in Section 3.1, both migration patterns are salient in China. Second, the estimated elasticity that governs the diffusion of ideas through people is smaller than the one that disciplines the diffusion of ideas through goods.

In the second to last row of the table, we quantify the importance of capital accumulation for aggregate growth in China. We find that in a model without capital accumulation, initial conditions in 1990 would have resulted in around a half of the aggregate growth in subsequent decades. In the last row of the table, we compute the aggregate growth with capital accumulation and no idea diffusion. Comparing the last two rows of the table, we can see that capital accumulation played a more important role than idea diffusion as an engine for aggregate growth in the early 1990s, but idea diffusion became more important over time. In the absence of changes in fundamentals, capital accumulation had a relatively stable contribution to aggregate growth. Different from capital accumulation, the contribution of idea diffusion to aggregate growth increased over time. As knowledge diffused and locations increased their stock of knowledge, people contributed with better insights to their locations and to other locations when moving, and provinces more opened to international trade also benefited more from better global insights as the stock of knowledge in the rest of the world also increased.

We next dig into the process of spatial development that shapes the aggregate growth described in this section.

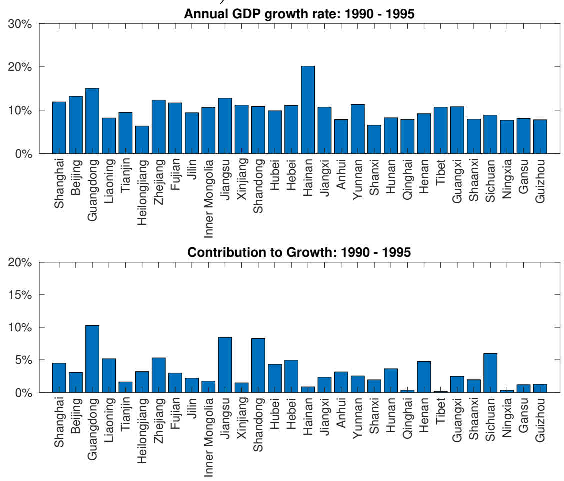
4.1.3 Spatial Growth

We turn to describe spatial growth across provinces in China. In particular, in this subsection we study how the aggregate growth in China described in Section 4.1.2 was shaped by spatial growth. To do so, Figure 4 displays the real GDP growth across provinces in China during different time frames over the period 1990-2020 as predicted by our model under the initial conditions in 1990. In each panel, the upper figure presents the annual real GDP growth and the lower panel displays the contribution of each province to the aggregate real GDP growth in China during that period.

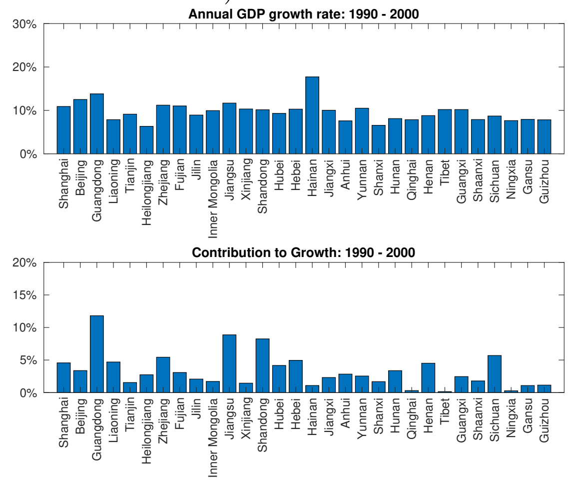
Several results emerge from the figure. Looking first at the upper figures in each panel, in the 1990s we find large heterogeneity in spatial growth. We can see that Hainan, Shanghai, Beijing, and Guangdong are the provinces with the highest growth rates in the 1990s. Among them, the last three are the ones with the highest initial measured productivity, as described in Section 3.2. At the same time, these are provinces located in the coastal areas with better access to foreign goods. Hence, we observe some divergence after 1990 in the form of higher growth in provinces with better initial technology, which is in part shaped by the idea diffusion through goods purchased from the rest of the world. Over the 2000s, we can see that growth rates tend to moderate and converge across provinces, as shown in the remaining upper figures, as the Chinese economy starts approaching the balanced growth path. Also, ideas from people diffuse across space, and insights from migrants coming from provinces with a relatively low stock of knowledge slow down growth in destinations with a relatively higher stock of knowledge, fostering some convergence in the stock of knowledge across space.

Figure 4: Spatial growth (annual, percent)

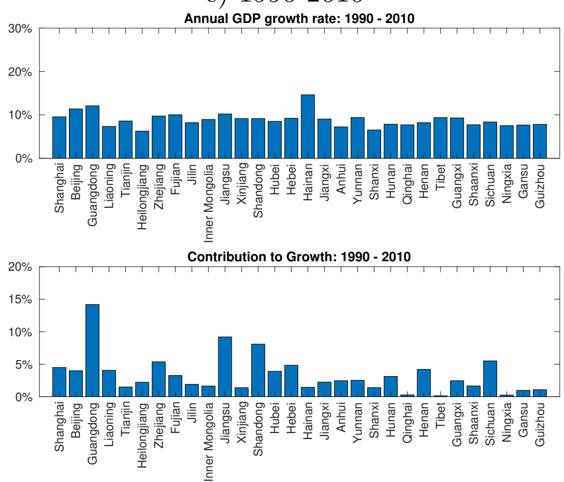
a) 1990-1995



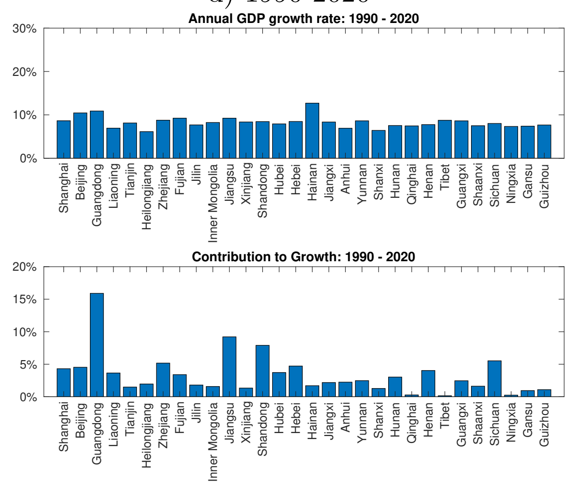
b) 1990-2000



c) 1990-2010



d) 1990-2020



Note: The figures show the annual real GDP growth across provinces and the contribution of each province to the aggregate growth in China in different time frames over the period 1990-2020. Spatial growth in each figure is computed in the model under the initial 1990 conditions.

In the lower figures of each panel, we present the contribution of each province to aggregate growth in China in each period of time. From these figures, we can see how Guangdong became a much more important engine of aggregate growth in China over time, as it benefits relatively more from ideas from the rest of the world, especially given its advantaged location on the coast near Hong Kong. The figures also show that other provinces like Beijing became more important contributors to aggregate growth in China. At the same time, other large provinces like Shandong, Henan, and Hubei decreased their importance for aggregate growth over time.[18]

[18] Online Supplement K.2 presents spatial growth effects for additional time frames, which deliver

Overall, this subsection illustrates how aggregate growth was shaped by heterogeneous spatial growth across provinces in China over time. The mechanics of spatial growth across provinces is in turn shaped by initial conditions as well as the dynamics of productivity as a result of the idea diffusion through goods and people, changes in trade openness and migration, and the dynamics of labor markets and capital accumulation across space. In the next subsection, we study further the role of initial conditions in shaping spatial growth in China.

4.1.4 Initial Distribution of Fundamentals and Spatial Development

In this subsection we study how the spatial growth across provinces during the 1990s and 2000s that we described in the previous section correlates with the initial distribution of fundamentals; namely, the initial stock of knowledge, the initial level of trade openness, and initial mobility frictions.

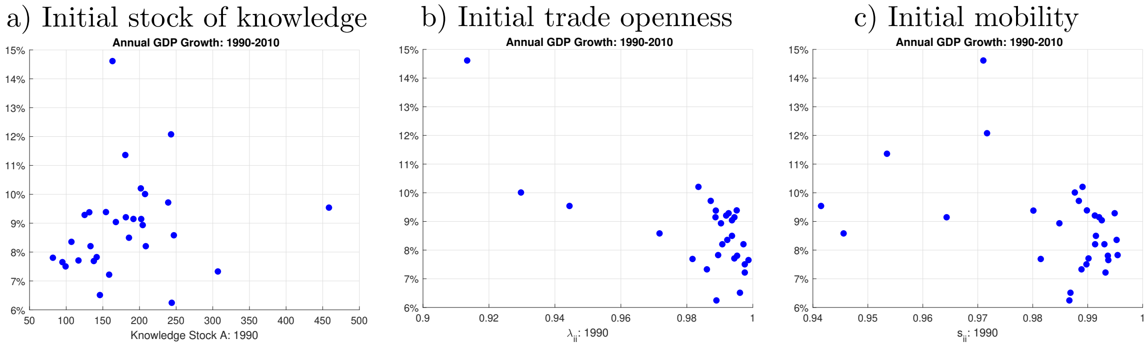
The upper panel of Figure 5 shows a scatter plot between the initial stock of knowledge and the real GDP growth across provinces over the period 1990-2010. We can see a somewhat positive correlation, meaning that provinces with a relatively higher initial fundamental productivity tend to grow more in the subsequent decades. It is important to emphasize that our dynamic spatial growth model allows for this correlation to have any sign. On the one hand, provinces with a higher initial stock of knowledge might also be more open to trade and as a result, they can benefit more from ideas diffused from the rest of the world relative to provinces with a lower initial stock of knowledge. On the other hand, provinces with relatively higher fundamental productivity might be attractive to migrants from other provinces who bring relatively few good-quality insights, which might slow down growth in those provinces.[19]

Panel (b) presents a scatter plot between real GDP growth and the initial level of trade openness measured by domestic expenditure share λ_{ii}, where a smaller value of λ_{ii} means higher trade openness. The figure shows a clear negative correlation, meaning that a higher initial level of trade openness leads to higher growth in subsequent decades. This correlation is intuitive since provinces more open to trade benefit more from idea diffusion through goods purchased from the rest of the world.

the same conclusions discussed in this section. In addition, to complement the analysis of Section 4.1.2, Online Supplement K.3 shows the spatial heterogeneity in the relative contributions of ideas from goods and ideas from people to growth across provinces in China.

[19]This positive relationship between the initial stock of knowledge and spatial growth is also in line with the positive correlation between GDP growth and initial per capita GDP across provinces observed in the data during the same period.

Figure 5: Real GDP growth versus initial conditions



Note: The figures show scatter plots of annual GDP growth across provinces in China over the period 1990-2010 against initial conditions in 1990: initial knowledge stock in Panel (a), initial level of trade openness, λ_{ii}, in Panel (b) (two outlier provinces were trimmed), and initial mobility, s_{ii}, in Panel (c).

Finally, Panel (c) presents a scatter plot between growth and initial mobility measured by the fraction of stayers in a province denoted by s_{ii}, where a smaller value of s_{ii} means lower mobility frictions and therefore higher mobility. In this case, the negative correlation is less clear than it is in the case of trade openness. As explained before, provinces that receive migrants experience faster growth in knowledge only if the insights from the provinces from which they migrate are of sufficiently good quality. All these scatter plots deliver the same message for the period 1990-2020.

4.2 Changes in Fundamentals

Our quantitative analysis in the previous sections shows that initial conditions seem to be important for understanding and quantifying the process of spatial development and aggregate economic growth in China. It also shows the importance of the general equilibrium interactions of the mechanisms in our framework such as idea diffusion through trade and migration, labor market dynamics, and capital accumulation.

Over the 1990s and 2000s, China undertook other reforms related to changes in trade costs and migration frictions that might have also impacted spatial growth. In particular, when China joined the World Trade Organization, provinces more exposed to trade might have developed more relative to the less exposed provinces. Likewise, Hukou reforms might have fostered idea flows by increasing mobility across provinces. In this section, we explore quantitatively through the lens of our framework the impact of these reforms on spatial growth in China.

In terms of trade costs, we capture the changes in trade costs between China and the rest of the world using the time variation in bilateral trade shares relative to domestic expenditure shares across provinces. In other words, from our model we can back up changes in trade costs as $\frac{\hat{\lambda}_{in,t}\hat{\lambda}_{ni,t}}{\hat{\lambda}_{ii,t}\hat{\lambda}_{nn,t}} = \left(\hat{\kappa}_{in,t}\hat{\kappa}_{ni,t}\right)^{-\theta}$.[20]

We estimate changes in bilateral trade frictions over the period 1990-2010 between provinces in China and the rest of the world and then ask how spatial growth in China would have looked if the only change in fundamentals over the period 1990-2010 had been bilateral international trade costs. To do so, we compare this counterfactual economy with the evolution of the economy with the 1990 initial conditions described in the previous sections.

We also explore Hukou reforms in a simple way. We capture the changes in migration frictions across provinces in China using the cross-variation in five-year mobility rates from 1985-1990 to 1995-2000 as $\frac{\hat{\mu}_{in,t}\hat{\mu}_{ni,t}}{\hat{\mu}_{ii,t}\hat{\mu}_{nn,t}} = \left(\hat{m}_{in,t}\hat{m}_{ni,t}\right)^{-\frac{1}{v}}$, and apply this change in mobility frictions to all Hukou types.[21] We then ask the counterfactual question of how spatial growth in China would have looked if the only change in fundamentals after 1990 were the change in mobility frictions to Hukou types. To do so, we compare this counterfactual economy with the evolution of the economy with 1990 initial conditions described in the previous sections.

The spatial growth effects of changes in trade costs and mobility frictions are presented in Figure 6. The left-hand panels display the effects of changes in trade costs, and the right-hand panels present the effects of Hukou restrictions. We present the results for the periods 1990-2000 and 1990-2020, and in Online Supplement K.4 we present results for additional time frames.

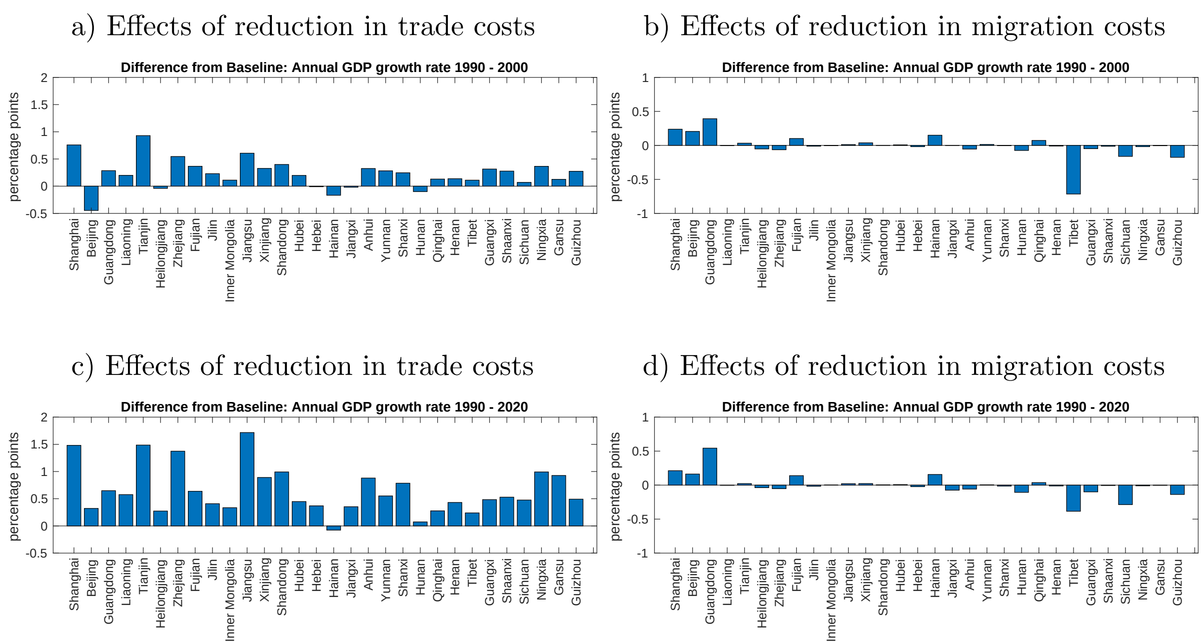
We find that the change in trade costs contributed to growth effects that were very heterogeneous across space. Panel (a) shows the spatial growth effects of changes in trade costs over the 1990s. We can see that changes in trade costs increase growth in almost all provinces in China relative to the baseline economy with initial trade and migration costs. The coastal provinces of Shanghai, Tianjin, and Jiangsu experience the largest growth effects from changes in trade costs as they benefit relatively more from trade openness compared with provinces located farther away. Some provinces that initially are relatively more open to trade, such as Beijing, Hainan, and Hunan,

[20]This statistic is known as the Head-Ries index (Head and Ries (2001)) and is widely used in the trade and spatial literature to measure bilateral trade frictions.

[21]Note that since Hukou type is assigned to either the origin or the destination province in the data, changes in mobility frictions are isomorphic to changes in amenities by Hukou type.

see growth slightly decline in the 1990s, as they face increased competition from other provinces, especially in the coastal areas that benefit from trade with the rest of the world. Moving down to the bottom panel, we can see that changes in trade costs foster growth across all provinces over time as ideas from the rest of the world continue to diffuse to provinces in China.

Figure 6: Effects of trade and migration costs on spatial growth (percentage points)



Note: The figures show the percentage point change in real GDP growth across provinces due to the trade and migration restrictions in different time frames. The left-hand panels present the effects of changes in trade costs and the right-hand panels show the effects of migration restrictions. All effects are computed relative to the baseline economy with 1990 trade and migration costs.

In the right-hand panels, we present the spatial growth effects of changes to migration restrictions. We can see that changes in mobility restrictions lead to smaller growth effects than changes in trade costs. However, the growth effects of changes in mobility restrictions are more heterogeneous than the growth effects of trade costs, with increases in growth in some provinces and decreases in growth in others relative to the baseline economy with 1990 conditions. Consistent with the intuition provided in previous sections, changes in mobility frictions benefit more open provinces that have scaled-up production and provinces that benefit from the ideas from migrants coming from places with a higher stock of knowledge, while they slow down growth in provinces left behind by international trade and internal migration. In Online Supplement K.4 we present the same figures with the spatial growth effects from changes in trade costs and Hukou restrictions together across provinces in China.

Table 2 presents the aggregate growth in China due to both 1990 fundamentals and the changes in international trade costs and migration restrictions over the 1990s and 2000s. To facilitate the analysis, the first two rows of the table reproduce the actual aggregate annual growth rate, and the growth rate with 1990 fundamentals displayed in Table 1. The rest of the rows present the growth effects with the 1990 fundamentals and the changes in trade costs and migration restrictions. We can see from the table that both changes in trade costs and changes in mobility restrictions added about one percentage point of extra annual aggregate growth in China by the 2000s, mostly coming from the changes in international trade costs. The changes in migration restrictions did not have a significant impact on aggregate growth, although they had more significant effects in particular locations, as described previously.

Table 2: Annual GDP growth rate: Changes in fundamentals

	90-95	90-00	90-05	90-10	90-15	90-20
Actual GDP growth	12.3%	10.4%	10.2%	10.5%	9.9%	9.3%
With fundamentals in 1990	10.6%	10.1%	9.6%	9.2%	8.9%	8.6%
Fund. in 1990 & change in trade cost	10.7%	10.3%	10.1%	9.9%	9.6%	9.4%
Fund. 1990 & change in mig. restrictions	10.6%	10.1%	9.7%	9.3%	9.0%	8.7%
Fund. in 1990 & change in fundamentals	10.7%	10.4%	10.2%	10.0%	9.7%	9.4%

Note: The first two rows of the table reproduce the actual annual growth rate of China and the growth rate with 1990 fundamentals displayed in Table 1. The third row presents the annual growth rate with 1990 fundamentals and changes in international trade costs. The fourth row presents the annual growth rate with 1990 fundamentals and changes in migration restrictions. The last row presents the annual growth rate with 1990 fundamentals and changes in international trade costs and migration restrictions.

5 Empirical Evidence of Idea Diffusion

In the previous sections we highlighted the importance of the spatial mechanisms in our framework for shaping spatial development and aggregate growth, and in particular, the role of idea diffusion through trade and migration. We did so through the lens of the spatial dynamic growth model developed in Section 2. In this section, we complement the structural analysis by providing reduced-form evidence of idea diffusion from trade and migration. To do so, we obtain province-level patent data as a proxy for the local stock of knowledge and use it along with our trade and migration data to provide empirical evidence of the role played by trade and migration in the diffusion of ideas.

We obtain province-level patent data from the China Statistics Yearbooks. To proxy the measure of knowledge stock, $A_{n,t}$, we obtain the cumulative approved patents at the province level for each year over the period 1985-2010. We then compute the change in the stock of knowledge every five years from 1985 to 2010. For the approved patents of the rest of the world, we obtain data from Google Patent from 1985-2010 following Liu and Ma (2021).[22] In Online Appendix F we provide more details on these patent data.

Using our patent data, as well as the migration and trade data described in Section 3.1 and in Online Appendix E, we run the following empirical specification:

$$\log\left(A_{n,t+1} - A_{n,t}\right) = \tau + \beta_m \log \lambda_{nn,t} + \beta_l \log\left(migration_{n,t}\right) + \tau_n + \tau_t + \epsilon_{n,t}.$$

The term $\lambda_{nn,t}$ is the domestic expenditure share and captures the (inverse) level of trade openness of province n. We expect the coefficient β_m to be negative, indicating that provinces more exposed to international trade benefit more from the global diffusion of ideas and experience higher growth in the stock of knowledge as a result. We define $\log\left(migration_{n,t}\right) = \log\left[\sum_{i=1}^{N} s_{in,t} A_{i,t}\right]$, which equals the weighted average of the stock of knowledge diffusing to location n at time t through both migrants and locals. This term in our model captures the idea diffusion through people, and as a result, we expect the coefficient β_l to be positive. Finally, the term τ_n controls for province fixed effects, τ_t is a year fixed effect, and $\epsilon_{n,t}$ is an error term following an i.i.d. standard normal distribution. Since individual-level population census data for the year 1995 do not exist, migration flows are unavailable for the period 1990-1995; hence, we run the regression using data for five-year intervals from 1995 to 2010.

Table 3 reports the results. In Column (1) we first show that faster growth in the stock of knowledge is associated with a higher degree of trade openness, and Column (2) shows that the growth in knowledge stock positively correlates with idea diffusion through people, controlling for year and province fixed effects. In Column (3), we show that trade openness and idea diffusion through people together contribute to the growth in the stock of knowledge in the way that our theory suggests. In each

[22] We are grateful to Song Ma for sharing the Google Patent data.

of these specifications, the coefficients of interest, β_m and β_l, have the expected signs and are statistically significant.[23]

Table 3: Estimates of knowledge diffusion through trade and people

	$\log(A_{n,t+1} - A_{n,t})$		
	(1)	(2)	(3)
$\log \lambda_{nn,t}$	-6.138***		-5.782**
	(2.222)		(2.296)
$\log(migration_{n,t})$		0.353*	0.274*
		(0.201)	(0.153)
Constant	8.692***	5.448**	5.960***
	(0.113)	(2.003)	(1.535)
Observations	90	90	90
R-squared	0.544	0.473	0.564
Year FE	✓	✓	✓
Province FE	✓	✓	✓

Note: Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

In the previous regression we documented the contribution of idea diffusion through people to the local stock of knowledge. One might wonder whether this contribution is entirely driven by local ideas, as a significant share of locals stay in the same location during each five-year window. To address this question, we distinguish between idea diffusion through the locals (stayers) and that through non-locals (immigrants) with the following specification,

$$\log(A_{n,t+1} - A_{n,t}) = \tau + \beta_m \log \lambda_{nn,t} + \beta_i \log(immigration_{n,t}) + \beta_s \log A_{n,t} + \tau_n + \tau_t + \epsilon_{n,t},$$

where $\log(immigration_{n,t}) = \log\left[\sum_{i \neq n} s_{in,t} A_{i,t}\right]$ captures the weighted sum of the knowledge brought by migrants to n from locations other than n. In this specification we control for $\log A_{n,t}$, which is the measure of local knowledge. Therefore, the coefficient β_i captures the idea diffusion through non-locals (immigrants), and the coefficient β_s captures the knowledge diffusion through locals (stayers).

[23]For completeness, in Online Supplement K.5 we present scatter plots with correlations between the change in the local stock of knowledge constructed with our patent data, and our measures of trade openness and idea diffusion through people. These correlations also show a positive relationship between trade openness and idea diffusion through people, and growth in the local stock of knowledge.

Table 4: Estimates of knowledge diffusion through trade and migration

	$\log(A_{n,t+1} - A_{n,t})$				$\log\left(\frac{A_{n,t+1}-A_{n,t}}{A_{n,t}}\right)$
	(1)	(2)	(3)	(4)	(5)
$\log\lambda_{nn,t}$		-5.234**		-5.200**	-4.549*
		(2.412)		(2.399)	(2.591)
$\log(immigration_{n,t})$			0.133*	0.128*	0.170**
			(0.076)	(0.070)	(0.073)
$\log A_{n,t}$	0.655***	0.549***	0.703***	0.596***	
	(0.211)	(0.185)	(0.218)	(0.189)	
Constant	3.121*	3.825**	2.012	2.753	-0.817*
	(1.775)	(1.548)	(2.049)	(1.724)	(0.406)
Observations	90	90	90	90	90
R-squared	0.537	0.610	0.550	0.622	0.827
Year FE	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓

Note: Robust standard errors in parentheses.*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4 reports the results. In Column (1), we start by showing the positive and significant coefficient of $\log A_{n,t}$, β_s, which reveals that idea diffusion through local knowledge contributes to growth in the local knowledge. Column (2) suggests that in addition to the local knowledge, higher knowledge growth tends to be seen in locations with more exposure to international trade. In Column (3), the positive and significant coefficient of the term $\log(immigration_{n,t})$, β_i, shows that after controlling for the local knowledge stock, the knowledge brought by non-local immigrants also contributes to the growth in the stock of knowledge. Column (4) shows that international trade openness, ideas brought by migrants, and local knowledge stock all contribute to the growth in local knowledge stock, in line with the spatial mechanisms in our model. Alternatively, we could replace the dependent variable by the growth rate in the knowledge stock. In Column (5), the dependent variable is $\log\left(\frac{A_{n,t+1}-A_{n,t}}{A_{n,t}}\right)$, and the purpose of this alternative specification is to normalize the change in knowledge stock in each location by the local knowledge stock. Again, the results suggest that international trade and ideas diffused by non-locals contribute significantly to the growth in the stock of knowledge.

We have provided reduced-form evidence of the contribution of idea diffusion through both international trade and migration to the stock of local knowledge. In doing so, we do not aim at establishing a causal effect since our theory is consistent with two-way causality between knowledge and idea diffusion through trade and migration due to general equilibrium effects. Still, in Online Appendix F we provide further evidence of idea diffusion by implementing an instrumental variable strategy to estimate the effect of idea diffusion on the local stock of knowledge. We also derive empirical specifications using the model's structure. Consistent with the reduced-form evidence presented in this section, we find a statistically significant contribution of idea diffusion through trade and migration to the local stock of knowledge.

6 Concluding Remarks

In this study we have developed a dynamic spatial growth model to study, understand, and quantify the role of spatial growth on aggregate economic activity. In our spatial growth model, internal migration and trade provide the mechanics for spatial growth. Producers and migrants share ideas with other producers, and the flow of ideas across space and time serves as the main mechanism that generates spatial growth. We characterized the equilibrium properties of our model and showed that it has a unique balanced growth path. We also showed how to take the model to the data in order to perform quantitative exercises without assuming that the economy is initially in a balanced growth path.

As an application, we took our model to the data and studied the importance of trade and migration as engines of growth for the Chinese economy after 1990. Initial conditions and our spatial mechanisms that operate through international trade and internal migration played a considerable role in the spatial development and aggregate growth during the 1990s and 2000s in China. The changes in fundamentals due to trade openness and Hukou restrictions also contributed to aggregate growth and heterogeneous spatial development in China.

References

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2016): "Import Competition and the Great US Employment Sag of the 2000s," *Journal of Labor Economics*, 34, S141–S198.

Akcigit, U., J. Grigsby, and T. Nicholas (2017): "Immigration and the Rise of American Ingenuity," *American Economic Review*, 107, 327–31.

Allen, T., C. Arkolakis, and X. Li (2020): "On the Equilibrium Properties of Network Models with Heterogeneous Agents," Working Paper 27837, NBER.

Alvarez, F. (2017): "Capital Accumulation and International Trade," *Journal of*

Monetary Economics, 91, 1–18.

Anderson, J. E., M. Larch, and Y. V. Yotov (2019): “Trade and Investment in the Global Economy: A Multi-Country Dynamic Analysis,” *European Economic Review*, 120, 1–26.

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

Arkolakis, C., S. K. Lee, and M. Peters (2020): “European Immigrants and the United States’ Rise to the Technological Frontier,” Yale University, mimeo.

Artuc, E., S. Chaudhuri, and J. McLaren (2010): “Trade Shocks and Labor Adjustment: A Structural Empirical Approach,” *The American Economic Review*, 100, 1008–1045.

Atkeson, A. and A. Burstein (2019): “Aggregate Implications of Innovation Policy,” *Journal of Political Economy*, 127, 2625–2683.

Atkin, D., K. Chen, and A. Popov (2022): “The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley,” MIT, mimeo.

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.

Bai, C.-E., C.-T. Hsieh, and Y. Qian (2006): “The Return to Capital in China,” NBER Working Papers 12755, NBER, Inc.

Berkes, E., R. Gaetani, and M. Mestieri (2022): “Technological Waves and Local Growth,” Federal Reserve Bank of Chicago, mimeo.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. S. Kortum (2003): “Plants and Productivity in International Trade,” *American Economic Review*, 93, 1268–90.

Bernstein, S., R. Diamond, T. J. McQuade, and B. Pousada (2018): “The Contribution of High-Skilled Immigrants to Innovation in the United States,” Stanford University, mimeo.

Buera, F. J. and E. Oberfield (2020): “The Global Diffusion of Ideas,” *Econometrica*, 88, 83–114.

Burchardi, K. B., T. Chaney, T. A. Hassan, L. Tarquinio, and S. J. Terry (2020): “Immigration, Innovation, and Growth,” Working Paper 27075, NBER.

Buzard, K., G. A. Carlino, R. M. Hunt, J. K. Carr, and T. E. Smith (2020): “Localized Knowledge Spillovers: Evidence from the Spatial Clustering of R&D Labs and Patent Citations,” *Regional Science and Urban Economics*, 81, 103490.

Cai, J., N. Li, and A. M. Santacreu (2022): “Knowledge Diffusion, Trade, and Innovation across Countries and Sectors,” *American Economic Journal: Macroeconomics*, 14, 104–45.

Cai, S. and W. Xiang (2022): “Multinational Production, Technology Diffusion, and Economic Growth,” Yale University, mimeo.

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," *The Review of Economic Studies*, 82, 1–44.

——— (2022): "Lessons from U.S.-China Trade Relations," *Annual Review of Economics*, forth.

Chen, W., X. Chen, C.-T. Hsieh, and Z. Song (2019): "A Forensic Examination of China's National Accounts," NBER Working Papers 25754, NBER.

Cruz, J. L. (2021): "Global Warming and Labor Market Reallocation," Princeton Univ., mimeo.

Cruz, J. L. and E. Rossi-Hansberg (2022): "The Economic Geography of Global Warming," University of Chicago, mimeo.

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," *The American Economic Review*, 104, 1211–1243.

Eaton, J. and S. Kortum (1999): "International technology diffusion: Theory and measurement," *International Economic Review*, 40, 537–570.

——— (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741–79.

Eaton, J., S. Kortum, B. Neiman, and J. Romalis (2016): "Trade and the Global Recession," *American Economic Review*, 106, 3401–38.

Facchini, G., M. Liu, A. Mayda, and M. Zhou (2019): "China's Great Migration: The Impact of the Reduction in Trade Policy Uncertainty," *Journal of International Economics*, 120, 126–144.

Fan, J. (2019): "Internal Geography, Labor Mobility, and the Distributional Impacts of Trade," *American Economic Journal: Macroeconomics*, 11, 252–88.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hansen, L. P. and K. J. Singleton (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

Head, K. and J. Ries (2001): "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of US-Canada trade," *American Economic Review*, 91, 858–876.

Hercowitz, Z. and M. Sampson (1991): "Output Growth, the Real Wage, and Employment Fluctuations," *The American Economic Review*, 81, 1215–1237.

Hunt, J. and M. Gauthier-Loiselle (2010): "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, 2, 31–56.

Imbert, C., M. Seror, Y. Zhang, and Y. Zylberberg (2022): "Migrants and Firms: Evidence from China," *American Economic Review*, 112, 1885–1914.

Jones, C. (1995): "R&D-Based Models of Economic Growth," *Journal of Political*

Economy, 103, 759–84.

KERR, W. R. (2008): "Ethnic Scientific Communities and International Technology Diffusion," *The Review of Economics and Statistics*, 90, 518–537.

KLEINMAN, B., E. LIU, AND S. J. REDDING (2021): "Dynamic Spatial General Equilibrium," Working Paper 29101, NBER.

KORTUM, S. S. (1997): "Research, Patenting, and Technological Change," *Econometrica: Journal of the Econometric Society*, 1389–1419.

LEWIS, E. (2011): "Immigration, Skill Mix, and Capital Skill Complementarity*," *The Quarterly Journal of Economics*, 126, 1029–1069.

LIU, E. AND S. MA (2021): "Innovation Networks and Innovation Policy," Working Papers 29607, NBER.

LUCAS, R. E. (2009): "Ideas and Growth," *Economica*, 76, 1–19.

LUCAS, R. E. AND E. C. PRESCOTT (1971): "Investment Under Uncertainty," *Econometrica*, 39, 659–681.

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

NEWEY, W. (1985): "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, 29, 229–256.

PERLA, J., C. TONETTI, AND M. E. WAUGH (2021): "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 111, 73–128.

PIERCE, J. R. AND P. K. SCHOTT (2016): "The Surprisingly Swift Decline of US Manufacturing Employment," *American Economic Review*, 106, 1632–62.

PONCET, S. (2003): "Measuring Chinese Domestic and International Integration," *China Economic Review*, 14, 1–21.

PRATO, M. (2021): "The Global Race for Talent: Brain Drain, Knowledge Transfer and Growth," University of Chicago, mimeo.

RAVIKUMAR, B., A. M. SANTACREU, AND M. SPOSI (2019): "Capital Accumulation and Dynamic Gains from Trade," *Journal of International Economics*, 119, 93–110.

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

SAMPSON, T. (2016): "Dynamic Selection: An Idea Flows Theory of Entry, Trade, and Growth," *The Quarterly Journal of Economics*, 131, 315–380.

SEQUEIRA, S., N. NUNN, AND N. QIAN (2019): "Immigrants and the Making of America," *The Review of Economic Studies*, 87, 382–419.

SHAN, H. (2008): "Reestimating the Capital Stock of China: 1952-2006," *The Journal of Quantitative and Technical Economics*, 11–30.

TOMBE, T. AND X. ZHU (2019): "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 109, 1843–72.

YOUNG, A. (2003): "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period," *Journal of Political Economy*, 111, 1220–1261.

Online Appendix: Mechanics of Spatial Growth

A Proofs and Derivations of Idea Diffusion

In this appendix, we derive the idea diffusion process with a generic source distribution of insights. We then endogenize the source distribution as a result of idea diffusion from migrants and sellers.

A.1 Law of Motion of the Stock of Knowledge

We start by providing a proof of Proposition 1 and then derive the law of motion with idea flows after specifying the external source of ideas. In this section we use uppercase letters for random variables, and lowercase letters for their realized values.

Proposition 1. *Under Assumption 1, between t and $t+1$, the probability that the best new idea has productivity no greater than q, $F_t^{best\ new}(q)$, is given by*

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty x^{\rho\theta} dG_t(x)\right)$$

in the limiting case when $\bar{z} \to 0$.

Proof. For any new idea that arrives between time t and $t+1$, the probability at time t that its productivity is no greater than q is given by

$$\begin{aligned}
&F_t^{new}(q) \\
&= \Pr[ZQ'^{\rho} \leq q] \\
&= \int_0^\infty \Pr\left[Z \leq \frac{q}{Q'^{\rho}}\middle| q'\right] dG_t(q') \\
&= \int_0^{(q/\bar{z})^{1/\rho}} \Pr\left[Z \leq \frac{q}{Q'^{\rho}}\middle| q'\right] dG_t(q') + \int_{(q/\bar{z})^{1/\rho}}^\infty \Pr\left[Z \leq \frac{q}{Q'^{\rho}}\middle| q'\right] dG_t(q') \\
&= \int_0^{(q/\bar{z})^{1/\rho}} \Pr\left[Z \leq \frac{q}{Q'^{\rho}}\middle| q'\right] dG_t(q'). \\
&= \int_0^{(q/\bar{z})^{1/\beta}} H\left(\frac{q}{q'^{\beta}}\right) dG_t(q'),
\end{aligned}$$

where the fourth equality follows from the fact that $\Pr\left[Z \leq \frac{q}{Q'^{\beta}}\middle| Q' > (q/\bar{z})^{1/\beta}\right] = \Pr\left[Z \leq \bar{z}\right] = 0$. Using Assumption 1 a), on the functional form of $H(\cdot)$, we obtain

$$F_t^{new}(q) = \int_0^{(q/\bar{z})^{1/\rho}} \left[1 - \left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q').$$

Note that in order to derive this expression, we do not need to specify the source distribution of the insights. Assumption 1 c) implies that between t and $t+1$, the probability that the best new idea has productivity no greater than q is given by

$$
\begin{aligned}
&F_t^{best\ new}(q) \\
&= \Pr\left[\text{all new ideas are no greater than } q\right] \\
&= \sum_{s=0}^{\infty} \Pr[\#\text{ new ideas} = \text{s}] \cdot \Pr\left[\text{all new ideas are no greater than } q \mid \#\text{ new ideas} = \text{s}\right] \\
&= \sum_{s=0}^{\infty} \frac{(\alpha_t \bar{z}^{-\theta})^s e^{-(\alpha_t \bar{z}^{-\theta})}}{s!} \cdot F_t^{new}(q)^s \\
&= \underbrace{\sum_{s=0}^{\infty} \frac{[\alpha_t \bar{z}^{-\theta} F_t^{new}(q)]^s \cdot e^{-(\alpha_t \bar{z}^{-\theta}) F_t^{new}(q)}}{s!}}_{=1} \cdot e^{-(\alpha_t \bar{z}^{-\theta})(1-F_t^{new}(q))},
\end{aligned}
$$

and therefore we obtain that

$$F_t^{best\ new}(q) = e^{-(\alpha_t \bar{z}^{-\theta})(1-F_t^{new}(q))}.$$

In order to characterize the probability distribution of the best new ideas, we hold α_t constant and investigate the limiting case where $\bar{z} \to 0$. We then have that

$$
\begin{aligned}
\lim_{\bar{z}\to 0} \alpha_t \bar{z}^{-\theta}(1-F_t^{new}(q)) &= \lim_{\bar{z}\to 0} \alpha_t \bar{z}^{-\theta} \left(1 - \int_0^{(q/\bar{z})^{1/\rho}} \left[1 - \left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q')\right) \\
&= \lim_{\bar{z}\to 0} \alpha_t \bar{z}^{-\theta} \left(1 - G_t\left((\frac{q}{\bar{z}})^{\frac{1}{\rho}}\right) + \int_0^{(q/\bar{z})^{1/\rho}} \left[\left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q')\right) \\
&= \alpha_t \lim_{\bar{z}\to 0} \bar{z}^{-\theta} \left[1 - G_t\left((\frac{q}{\bar{z}})^{\frac{1}{\rho}}\right)\right] \\
&+ \alpha_t \lim_{\bar{z}\to 0} \bar{z}^{-\theta} \int_0^{(q/\bar{z})^{1/\rho}} \left[\left(\frac{q/\bar{z}}{q'^{\rho}}\right)^{-\theta}\right] dG_t(q') \\
&= \alpha_t \lim_{\bar{z}\to 0} \bar{z}^{-\theta} \left[1 - G_t\left((\frac{q}{\bar{z}})^{\frac{1}{\rho}}\right)\right] + \alpha_t \int_0^{\infty} \left(\frac{q}{q'^{\rho}}\right)^{-\theta} dG_t(q'),
\end{aligned}
$$

where the first term on the right-hand side is zero by Assumption 1 d). In the limiting case when $\bar{z} \to 0$, the expression is equal to the second term only, which is $-\alpha_t q^{-\theta} \int_0^{\infty} x^{\rho\theta} dG_t(x)$.

Henceforth, we assume $\bar{z} \to 0$ and focus on the limiting case. The best new idea then follows

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^{\infty} x^{\rho\theta} dG_t(x)\right).$$

The productivity of the economy depends on the frontier of knowledge, $F_t(q)$. The frontier of knowledge denotes the fraction of varieties whose best producer has productivity no greater than q. In a probabilistic sense, $F_t(q)$ is also the probability that the best productivity for a specific variety is no greater than q at time t.

Proposition 2. *Assume that the initial frontier of knowledge at time* 0 *follows a Fréchet distribution given by* $F_0(q) = \exp(-A_0 q^{-\theta})$.

Imposing this assumption, then it follows that $F_t(\cdot)$ *is Fréchet at any* t *given by*

$$\begin{aligned} F_t(q) = \exp &\left[- \left(A_0 + \sum_{\tau=0}^{t-1} \alpha_\tau \int_0^\infty x^{\rho\theta} dG_\tau(x) \right) q^{-\theta} \right] \\ &= \exp\left(-A_t q^{-\theta}\right), \end{aligned}$$

where the law of motion for the knowledge stock is given by

$$A_{t+1} = A_t + \alpha_t \int_0^\infty x^{\rho\theta} dG_t(x).$$

Proof. The frontier $F_t(q)$ changes from t to $t+1$ because some new ideas might have better productivity than the current best. At $t+1$, we then have

$$\begin{aligned} F_{t+1}(q) &= \Pr\left[\text{the best productivity is no greater than } q \text{ at } t+1\right] \\ &= \Pr\left[\text{the best productivity is no greater than } q \text{ at } t\right] \cdot \\ &\quad \Pr[\text{no new ideas greater than } q \text{ between } t \text{ and } t+1] \\ &= F_t(q) \cdot F_t^{best\ new}(q) \\ &= F_0(q) \cdot \prod_{\tau=0}^{t} F_\tau^{best\ new}(q), \end{aligned}$$

where the last line follows from iteration back to $t = 0$.

Assume that the initial distribution at time 0 follows a Fréchet distribution; namely,

$$F_0(q) = \exp(-A_0 q^{-\theta}).$$

Then it follows that $F_t(\cdot)$ is Fréchet at any t:

$$\begin{aligned} F_t(q) = \exp &\left[- \left(A_0 + \sum_{\tau=0}^{t-1} \alpha_\tau \int_0^\infty x^{\rho\theta} dG_\tau(x) \right) q^{-\theta} \right] \\ &= \exp\left(-A_t q^{-\theta}\right). \end{aligned}$$

It also follows that the law of motion of the knowledge stock is

$$A_{t+1} = A_t + \alpha_t \int_0^\infty x^{\rho\theta} dG_t(x).$$

As we can see from this equation, both the arrival rate of new ideas α_t and the learning pool $G_t(\cdot)$ matter for the evolution of A_t.

A.2 Migration and the Source Distribution of Insights

Assume that at time t in location n, when a new idea arrives, the insight from a randomly drawn person currently living in n is the insight component of the new idea. Then

$$\begin{aligned} G_{n,t}(q') &= \Pr[\text{the insight component is no greater than } q'] \\ &= \sum_{i=1}^{N} \Pr[\text{the person with the insight lives in } i \text{ at } t] \cdot \\ &\qquad \Pr[\text{the insight is no greater than } q' | \text{the person with the insight lives in } i \text{ at } t] \\ &= \sum_{i=1}^{N} s_{in,t} F_{i,t}(q'), \end{aligned}$$

where $s_{in,t}$ is the share of households from location i living in location n. In particular, we denote by $\mu_{in,t}$ the fraction of households that relocate from from i to n. We then have $s_{in,t} = \frac{\mu_{in,t} L_{i,t}}{\sum_{h=1}^{N} \mu_{hn,t} L_{h,t}}$ and

$$\int_0^\infty x^{\rho_\ell \theta} dG_t(x) = \Gamma(1-\rho_\ell) \sum_{i=1}^{N} s_{in,t} (A_{i,t})^{\rho_{\ell l}}.$$

Finally, the law of motion of the stock of knowledge with ideas from people is given by

$$A_{n,t+1} - A_{n,t} = \alpha_{n,t} \Gamma(1-\rho_\ell) \sum_{i=1}^{N} s_{in,t} (A_{i,t})^{\rho_\ell}.$$

A.3 Derivation of the Law of Motion of Knowledge with Ideas from Migrants and Sellers

Now we derive the law of motion of the knowledge stock with idea flows from both trade and migration.

We impose the following version of Assumption 1 to incorporate both sources of idea flows:

Assumption 1'

a) The same as Assumption 1 a)

b) The strength of idea diffusion, $\rho_m + \rho_l \in [0,1)$, is strictly less than 1.
c) The same as Assumption 1 c)
d) The source distribution has a sufficiently thin tail such that for any monotonically decreasing sequence of $\bar{z}_n \to 0$, $\alpha_t \lim_{n\to\infty} \bar{z}_n^{-\theta} \left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] = 0$, where $B(\bar{z}) := \{(x_1, x_2) : \bar{z} x_1^{\rho_l} x_2^{\rho_m} < q\} \subset \mathbb{R}^2$. In addition, the integral $\int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m)$ exists.

Proposition 1'. *Under Assumption 1', between t and $t+1$, the probability that the best new idea has productivity no greater than q, $F_t^{best\ new}(q)$, is given by*

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_t(q_\ell, q_m)\right)$$

in the limiting case where $\bar{z} \to 0$.

Proof: For any new idea that arrives between time t and $t+1$, the probability at t that its productivity is no greater than q is given by

$$\begin{aligned}
&F_t^{new}(q) \\
&= \Pr[Z Q_\ell^{\rho_\ell} Q_m^{\rho_m} \le q] \\
&= \int_{\mathbb{R}_+^2} \Pr\left[Z \le \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m) \\
&= \int_{B(\bar{z})} \Pr\left[Z \le \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m) \\
&+ \int_{\mathbb{R}_+^2 \setminus B(\bar{z})} \Pr\left[Z \le \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m) \\
&= \int_{B(\bar{z})} \Pr\left[Z \le \frac{q}{Q_\ell^{\rho_\ell} Q_m^{\rho_m}} \middle| q_\ell, q_m\right] dG_t(q_\ell, q_m),
\end{aligned}$$

where $B(\bar{z})$ is defined in Assumption 1' d). Using Assumption 1' a), we obtain

$$F_t^{new}(q) = \int_{B(\bar{z})} \left[1 - \left(\frac{q/\bar{z}}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta}\right] dG_t(q_\ell, q_m).$$

The probability that the best new idea has productivity no greater than q is the same as before: $F_t^{best\ new}(q) = e^{-(\alpha_t \bar{z}^{-\theta})(1-F_t^{new}(q))}$. Consider a monotonically decreasing sequence of $\bar{z}_n \to 0$. We prove by the dominated convergence theorem that $\lim_{n\to\infty} \alpha_t \bar{z}_n^{-\theta}(1 - F_t^{new}(q)) = \alpha_t \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m)$. The integral exists under Assumption 1' d).

Define $g_n : \mathbb{R}_+ \to \mathbb{R}$,

$$\begin{aligned}
g_n(q) &= \bar{z}_n^{-\theta}(1 - F_t^{new}(q)) \\
&= \bar{z}_n^{-\theta}\left(1 - \int_{B(\bar{z}_n)} \left[1 - \left(\frac{q/\bar{z}}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta}\right] dG_t(q_\ell, q_m)\right) \\
&= \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] + \int_{B(\bar{z}_n)} \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m) \\
&= \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] + \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} dG_t(q_\ell, q_m).
\end{aligned}$$

By Assumption 1' d), we have $\lim_{n\to\infty} \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] = 0$. Since $\forall q \geq 0$, $\forall n$,

$$\left|\left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)}\right| \leq \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta},$$

and

$$\lim_{n\to\infty} \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} = \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta},$$

by the dominated convergence theorem, we have

$$\lim_{n\to\infty} \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} dG_t(q_\ell, q_m) = \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m),$$

so

$$\begin{aligned}
\lim_{n\to\infty} g_n(q) &= \lim_{n\to\infty} \alpha_t \bar{z}_n^{-\theta}(1 - F_t^{new}(q)) \\
&= \lim_{n\to\infty} \bar{z}_n^{-\theta}\left[1 - \int_{B(\bar{z}_n)} dG_t(q_\ell, q_m)\right] + \lim_{n\to\infty} \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} \mathbf{1}_{B(\bar{z}_n)} dG_t(q_\ell, q_m) \\
&= \int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t(q_\ell, q_m).
\end{aligned}$$

Henceforth, we assume $\bar{z} \to 0$ and focus on the limiting case. The best new idea then follows

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^{\theta} dG_t(q_\ell, q_m)\right)$$

or, using the Riemann integral,

$$F_t^{best\ new}(q) = \exp\left(-\alpha_t q^{-\theta} \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_t(q_\ell, q_m)\right).$$

As in the previous section, in this section it follows that the frontier distribution $F_{n,t}(\cdot)$ follows a Fréchet distribution with location parameter $A_{n,t}$ and shape parameter θ, and the law of motion of $A_{n,t}$ is

$$A_{n,t+1} = A_{n,t} + \alpha_t \int_0^\infty \int_0^\infty \left(q_\ell^{\rho_\ell} q_m^{\rho_m}\right)^\theta dG_{n,t}(q_\ell, q_m).$$

Since we assume that q_l, q_m come from independent distributions, the law of motion becomes

$$A_{n,t+1} = A_{n,t} + \alpha_t \int_0^\infty q_\ell^{\theta\rho_\ell} dG_{n,t}^l(q_\ell) \int_0^\infty q_m^{\theta\rho_m} dG_{n,t}^m(q_m).$$

The first integral,

$$\int_0^\infty q_\ell^{\theta\rho_\ell} dG_{n,t}^l(q_\ell) = \Gamma(1-\rho_\ell) \sum_{i=1}^N s_{in,t}(A_{i,t})^{\rho_\ell},$$

is the same as in the previous section. The derivation of this term follows the previous section of this appendix. For the second integral, we assume learning from sellers as in Buera and Oberfield (2020). Namely, the insights from goods are randomly drawn from the set of goods sold locally. To simplify the notation, we omit intermediate goods in the derivation that follows. In this case,

$$\begin{aligned}
G_{n,t}^m(x) &= \sum_i \mathbb{P}\left[q_i \leq x,\ \text{i is the lowest-cost supplier to n at t}\right] \\
&= \sum_i \mathbb{P}\left[q_i \leq x, q_j \leq \frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} q_i\ \forall j\right] \\
&= \sum_i \int_0^x f_{i,t}(q) \left(\prod_{j\neq i} F_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} q\right)\right) dq,
\end{aligned}$$

where $F_{i,t}(\cdot)$ and $f_{i,t}(\cdot)$ are the cumulative distribution function (CDF) and probability density function (PDF) of a Fréchet distribution with location parameter $A_{i,t}$ and shape parameter θ, respectively:

$$\begin{aligned}
F_{i,t}(q) &= \exp\left(-A_{i,t} q^{-\theta}\right), \\
f_{i,t}(q) &= A_{i,t}\theta q^{-\theta-1} \exp\left(-A_{i,t} q^{-\theta}\right).
\end{aligned}$$

Therefore,

$$
\begin{aligned}
G_{n,t}^{m}(x) &= \sum_{i}\int_{0}^{x} f_{i,t}(q)\left(\prod_{j\neq i} F_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}q\right)\right)dq \\
&= \sum_{i}\int_{0}^{x} A_{i,t}\theta q^{-\theta-1}\exp\left(-A_{i,t}q^{-\theta}\right)\exp\left(-\sum_{j\neq i}A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}q^{-\theta}\right)dq \\
&= \sum_{i}\int_{0}^{x} A_{i,t}\theta q^{-\theta-1}\exp\left(-\sum_{j}A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}q^{-\theta}\right)dq \\
&= \sum_{i}\frac{A_{i,t}\left(w_{i,t}\kappa_{ni}\right)^{-\theta}}{\sum_j A_{j,t}\left(w_{j,t}\kappa_{nj}\right)^{-\theta}}\exp\left(-\sum_{j}A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}x^{-\theta}\right) \\
&= \sum_{i}\pi_{ni,t}\exp\left(-\sum_{j}A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}x^{-\theta}\right).
\end{aligned}
$$

It follows that the second integral, which represents the learning from goods, is given by

$$
\begin{aligned}
\int_{0}^{\infty} q_m^{\theta\rho_m}dG_{n,t}^{m}(q_m) &= \int_{0}^{\infty} q_m^{\theta\rho_m}d\sum_{i}\pi_{ni,t}\exp\left(-\sum_{j}A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}q_m^{-\theta}\right) \\
&= \sum_{i}\pi_{ni,t}\int_{0}^{\infty} q_m^{\theta\rho_m}d\exp\left(-\sum_{j}A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}q_m^{-\theta}\right)
\end{aligned}
$$

Using change of variables, define $x=\sum_j A_{j,t}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta}q_m^{-\theta}$, and we have

$$
\begin{aligned}
\int_{0}^{\infty} q_m^{\theta\rho_m}dG_{n,t}^{m}(q_m) &= \sum_{i}\pi_{ni,t}\int_{0}^{\infty}\sum_{j}A_{j,t}^{\rho_m}\left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}}\right)^{-\theta\rho_m}x^{-\rho_m}d\exp(-x) \\
&= \Gamma(1-\rho_m)\sum_{i}\pi_{ni,t}\left(\frac{A_{i,t}}{\pi_{ni,t}}\right)^{\rho_m}.
\end{aligned}
$$

Therefore, the law of motion of $A_{n,t}$ is given by

$$
A_{n,t+1}-A_{n,t}=\alpha_t\Gamma(1-\rho_\ell)\Gamma(1-\rho_m)\left[\sum_{i=1}^{N}s_{in,t}(A_{i,t})^{\rho_\ell}\right]\left[\sum_{i=1}^{N}\pi_{ni,t}\left(\frac{A_{i,t}}{\pi_{ni,t}}\right)^{\rho_m}\right].
$$

B Detrended Model and Balanced Growth Path

In this appendix we characterize the long-run growth rates of the equilibrium variables of the model at the balanced growth path. In what follows, we denote the long-run growth rate of any variable y_t by $(1+g_y)$, and we also refer to a variable with a "~" as a detrended variable. In particular, $\tilde{y}_t = y_t / (1+g_y)^t$.

The equilibrium conditions of the detrended model are given by

$$\tilde{V}_{i,t} = \beta log\,(1+g_v) + log\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\right) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}\right), \tag{B.1}$$

$$\tilde{P}_{i,t} = T\left(\sum_{n=1}^{N} \tilde{A}_{n,t}\left(\kappa_{in,t}\tilde{x}_{n,t}\right)^{-\theta}\right)^{-1/\theta}, \tag{B.2}$$

$$\tilde{w}_{i,t}L_{i,t} = \sum_{n=1}^{N} \tilde{A}_{i,t}\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}}{\tilde{P}_{n,t}/T}\right)^{-\theta} \tilde{w}_{n,t}L_{n,t}, \tag{B.3}$$

$$\tilde{r}_{i,t}\tilde{K}_{i,t} = \sum_{n=1}^{N} \tilde{A}_{i,t}\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}}{\tilde{P}_{n,t}/T}\right)^{-\theta} \tilde{r}_{n,t}\tilde{K}_{n,t}, \tag{B.4}$$

$$L_{i,t+1} = \sum_{n=1}^{N} \mu_{ni,t}L_{n,t}, \tag{B.5}$$

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)}\left(\tilde{r}_{i,t}/\tilde{P}_{i,t} + (1-\delta)\right)\tilde{K}_{i,t}, \tag{B.6}$$

$$\tilde{A}_{n,t+1} - \frac{\tilde{A}_{n,t}}{(1+g_A)} = \frac{\alpha_0 \Gamma_{\rho_\ell,\rho_m}}{(1+g_A)} \sum_{i=1}^{N} s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}, \tag{B.7}$$

where we note that since there is no population growth, employment does not have a long-run growth rate; namely, $\tilde{L}_{n,t} = L_{n,t}$. Since values grow at the same rate in the long run, it follows that $\tilde{\mu}_{ni,t} = \mu_{ni,t}$, as we show below.

We start with the evolution of the stock of knowledge. At the balanced growth path, $A_{n,t}$ for all n grow at a rate $1+g_A$. From the law of motion of the stock of knowledge, we have

$$A_{n,t+1} - A_{n,t} = \alpha_t \Gamma_\rho \sum_{i=1}^{N} s_{in,t}\left(A_{i,t}\right)^{\rho_l} \sum_{i=1}^{N} \lambda_{ni,t}\left(\frac{A_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m},$$

using Assumption 2 and after detrending the variables, we obtain

$$
\begin{aligned}
&\tilde{A}_{n,t+1}\left(1+g_A\right)^{t+1}-\tilde{A}_{n,t}\left(1+g_A\right)^{t}\\
&=\alpha_0\left(1+g_\alpha\right)^t\Gamma_\rho\sum_{i=1}^{N}s_{in,t}\left(\tilde{A}_{i,t}\left(1+g_A\right)^t\right)^{\rho_l}\sum_{i=1}^{N}\lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}\left(1+g_A\right)^t}{\lambda_{ni,t}}\right)^{\rho_m}
\end{aligned}
$$

or

$$
\tilde{A}_{n,t+1}\left(1+g_A\right)-\tilde{A}_{n,t}=\left(1+g_\alpha\right)^t\left(1+g_A\right)^{t(\rho_l+\rho_m-1)}\alpha_0\Gamma_\rho\sum_{i=1}^{N}s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l}\sum_{i=1}^{N}\lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m},
$$

which then implies that the long-run growth rate of the stock of knowledge is related to the growth rate of the arrival of ideas in the following way:

$$
1+g_A=\left(1+g_\alpha\right)^{\frac{1}{(1-\rho_l-\rho_m)}}.
$$

As a result, the detrended equilibrium evolution of the local stock of knowledge evolves according to

$$
\tilde{A}_{n,t+1}-\frac{\tilde{A}_{n,t}}{\left(1+g_A\right)}=\frac{\tilde{\alpha}_0\Gamma_\rho}{\left(1+g_A\right)}\sum_{i=1}^{N}s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l}\sum_{i=1}^{N}\lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}
$$

or

$$
\frac{\tilde{A}_{n,t+1}}{\tilde{A}_{n,t}}=\frac{1}{1+g_A}+\frac{\alpha_0\Gamma_\rho}{\left(1+g_A\right)\bar{A}_{n,t}}\sum_{i=1}^{N}s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l}\sum_{i=1}^{N}\lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}.
$$

We now consider the detrended value functions of the workers. Let $e^{V_{i,t}}=e^{\tilde{V}_{i,t}}\left(1+g_v\right)^t$. We then have

$$
\tilde{V}_{i,t}+log\left(1+g_v\right)^t=log\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\left(1+g_{w/p}\right)^t\right)+\nu log\left(\sum_{n=1}^{N}exp\left(\beta\tilde{V}_{n,t+1}+\beta log\left(1+g_v\right)^{t+1}-m_{in,t}\right)^{1/\nu}\right), \tag{B.8}
$$

where $g_{w/p}$ is the growth rate of $\tilde{w}_{i,t}/\tilde{P}_{i,t}$ at the balanced growth path. It follows that

$$
\tilde{V}_{i,t}+log\left(1+g_v\right)^t=log\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\right)+log\left(1+g_{w/p}\right)^t+log\left(1+g_v\right)^{\beta(t+1)}+\nu log\left(\sum_{n=1}^{N}exp\left(\beta\tilde{V}_{n,t+1}-m_{in,t}\right)^{1/\nu}\right),
$$

which immediately implies that

$$
\left(1+g_v\right)^{(1-\beta)t}=\left(1+g_{w/p}\right)^t,
$$

$$
1+g_v=\left(1+g_{w/p}\right)^{\frac{1}{(1-\beta)}}.
$$

Hence, the detrended equilibrium values become

$$\tilde{V}_{i,t} = log\left(\tilde{w}_{i,t}/\tilde{P}_{i,t}\right) + log\left(1+g_v\right)^{\beta} + \nu log\left(\sum_{n=1}^{N} exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}\right).$$

This result immediately implies that $\mu_{in,t}$ is not growing in the long run since

$$\mu_{in,t} = \frac{exp\left(\beta V_{n,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta V_{l,t+1} - m_{il,t}\right)^{1/\nu}} = \frac{exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta \tilde{V}_{l,t+1} - m_{il,t}\right)^{1/\nu}}.$$

It also implies that $L_{i,t}$ does not have long-run growth since

$$\begin{aligned} L_{i,t+1} &= \sum_{n=1}^{N} \mu_{ni,t} L_{n,t}, \\ &= \sum_{n=1}^{N} \frac{exp\left(\beta V_{i,t+1} - m_{ni,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta V_{l,t+1} - m_{nl,t}\right)^{1/\nu}} L_{n,t} \\ &= \sum_{n=1}^{N} \frac{exp\left(\beta \tilde{V}_{i,t+1} - m_{ni,t}\right)^{1/\nu}}{\sum_{l=1}^{N} exp\left(\beta \tilde{V}_{l,t+1} - m_{nl,t}\right)^{1/\nu}} L_{n,t}. \end{aligned}$$

Let us now consider the labor market clearing condition,

$$w_{i,t} L_{i,t} = \sum_{n=1}^{N} A_{i,t} \left(\frac{\kappa_{ni,t} x_{i,t}}{P_{n,t}/T}\right)^{-\theta} w_{n,t} L_{n,t}.$$

First note that

$$\begin{aligned} x_{i,t} = \tilde{x}_{i,t}\left(1+g_x\right)^t &= B\left(\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\left(1+g_{w/p}\right)^t\right)^{\xi}\left(\frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}}\left(1+g_{r/p}\right)^t\right)^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t}\left(1+g_p\right)^t \\ &= \tilde{x}_{i,t}\left(1+g_{w/p}\right)^{t\xi\gamma}\left(1+g_{r/p}\right)^{t(1-\xi)\gamma}\left(1+g_p\right)^t. \end{aligned}$$

Using this expression, we can express the labor market clearing condition in a detrended form as

$$\begin{aligned} &\tilde{w}_{i,t}\left(1+g_w\right)^t L_{i,t} \\ &= \sum_{n=1}^{N} \tilde{A}_{i,t}\left(1+g_A\right)^t \left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}\left(1+g_{w/p}\right)^{t\xi\gamma}\left(1+g_{r/p}\right)^{t(1-\xi)\gamma}\left(1+g_p\right)^t}{\tilde{P}_{n,t}\left(1+g_p\right)^t/T}\right)^{-\theta} \tilde{w}_{n,t}\left(1+g_w\right)^t L_{n,t}, \end{aligned}$$

where we use the fact that $L_{i,t}$ does not grow in the long run. It follows that

$$1 = (1+g_A)^t \left(\left(1+g_{w/p}\right)^{t\xi\gamma} \left(1+g_{r/p}\right)^{t(1-\xi)\gamma} \right)^{-\theta},$$

$$\left(1+g_{w/p}\right)^{\theta\xi\gamma} \left(1+g_{r/p}\right)^{\theta(1-\xi)\gamma} = (1+g_A)\,. \tag{B.9}$$

We follow the same steps for the capital accumulation equation. Then the detrended labor and capital market equilibrium conditions become

$$\tilde{w}_{i,t} L_{i,t} = \sum_{n=1}^{N} \tilde{A}_{i,t} \left(\frac{\kappa_{ni,t} \tilde{x}_{i,t}}{\tilde{P}_{n,t}/T} \right)^{-\theta} \tilde{w}_{n,t} L_{n,t},$$

$$\tilde{r}_{i,t} \tilde{K}_{i,t} = \sum_{n=1}^{N} \tilde{A}_{i,t} \left(\frac{\kappa_{ni,t} \tilde{x}_{i,t}}{\tilde{P}_{n,t}/T} \right)^{-\theta} \tilde{r}_{n,t} \tilde{K}_{n,t},$$

where $\tilde{K}_{n,t}$ is the detrended value of capital that we subsequently characterize.

We now detrend the price index equilibrium condition,

$$P_{i,t} = T \left(\sum_{n=1}^{N} A_{n,t} \left(\kappa_{in,t} x_{n,t}\right)^{-\theta} \right)^{-1/\theta},$$

which once detrended can be expressed as

$$\tilde{P}_{i,t} (1+g_p)^t = T \left(\sum_{n=1}^{N} \tilde{A}_{n,t} (1+g_A)^t \left(\kappa_{in,t} \tilde{x}_{n,t} \left(1+g_{w/p}\right)^{t\xi\gamma} \left(1+g_{r/p}\right)^{t(1-\xi)\gamma} (1+g_p)^t \right)^{-\theta} \right)^{-1/\theta}.$$

Using equation (B.9), we obtain the detrended equilibrium condition for the price index:

$$\tilde{P}_{i,t} = T \left(\sum_{n=1}^{N} \tilde{A}_{n,t} \left(\kappa_{in,t} \tilde{x}_{n,t}\right)^{-\theta} \right)^{-1/\theta}.$$

Now note that since in equilibrium we have that

$$w_{i,t} L_{i,t} = \frac{\xi}{1-\xi} r_{i,t} K_{i,t},$$

then

$$\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}} \left(1+g_{w/p}\right)^t L_{i,t} = \frac{\xi}{1-\xi} \frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} \left(1+g_{r/p}\right)^t \tilde{K}_{i,t} (1+g_k)^t,$$

which immediately implies that

$$1+g_{w/p} = \left(1+g_{r/p}\right)(1+g_k)\,. \tag{B.10}$$

We now detrend the law of motion of capital accumulation,

$$K_{i,t+1} = \beta\left(r_{i,t}/P_{i,t} + (1-\delta)\right)K_{i,t},$$

which can be written as

$$\tilde{K}_{i,t+1}\left(1+g_k\right)^{t+1} = \beta\left(\left(1+g_{r/p}\right)^t \frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} + (1-\delta)\right)\bar{K}_{i,t}\left(1+g_k\right)^t$$

or

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)}\left(\left(1+g_{r/p}\right)^t \frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} + (1-\delta)\right)\tilde{K}_{i,t}.$$

We then require that

$$g_{r/p} = 0,\ \Rightarrow,\ g_r = g_p,$$

and in this way, the detrended capital accumulation equation becomes

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)}\left(\frac{\tilde{r}_{i,t}}{\tilde{P}_{i,t}} + (1-\delta)\right)\tilde{K}_{i,t}.$$

From equation (B.9) we obtain that

$$1+g_{w/p} = (1+g_A)^{\frac{1}{\theta\xi\gamma}},$$

and from equation (B.9) we also obtain that

$$1+g_k = (1+g_A)^{\frac{1}{\theta\xi\gamma}}.$$

C Existence and Uniqueness

Proposition 3. Existence and Uniqueness. *Given parameters and elasticities* $(\rho_\ell, \rho_m, \theta, \nu, \gamma, \xi, \beta)$, *and fundamentals* $\{\alpha_0, \bar{\kappa}_{in}, \bar{m}_{in}\}_{i=1,n=1}^{N,N}$, *there exists a unique (up to scale) solution given by* $\{\bar{w}_i, \bar{r}_i, \bar{L}_i, \bar{K}_i, \bar{V}_i, \bar{A}_i\}_{i=1}^N$ *that satisfies the equilibrium conditions of the detrended model at the balanced growth path.*

Proof. At the balanced growth path all the detrended variables are not growing, and as a result, the equilibrium variables of the detrended model reach a steady state. Hence, at the balanced growth path, $\tilde{y}_{t+1} = \tilde{y}_t = \bar{y}$, and it remains constant for all t. We use an upper bar to express the detrended equilibrium variables at the balanced growth path.

From the first-order condition of the firm's problem, we have that $\bar{w}_i\bar{L}_i = \frac{\xi}{1-\xi}\bar{r}_i\bar{K}_i$. It follows that $\bar{r}_i = \bar{w}_i\frac{\bar{L}_i(1-\xi)}{\bar{K}_i\xi}$ in the balanced growth path of the detrended model.

Hence, we have that

$$\frac{\bar{K}_i}{\bar{L}_i} = \frac{\bar{w}_i}{\bar{P}_i} \frac{(1-\xi)}{\xi} \frac{\beta}{(1+g_k) - \beta(1-\delta)}.$$

Now we use that $\bar{x}_i = B\left(\bar{w}_i^{\xi} \bar{r}_i^{1-\xi}\right)^{\gamma} \bar{P}_i^{1-\gamma}$ to obtain that

$$\begin{aligned}
\bar{x}_i &= B\left(\bar{w}_i^{\xi} \bar{r}_i^{1-\xi}\right)^{\gamma} \bar{P}_i^{1-\gamma} = B\left(\bar{w}_i^{\xi}\left(\bar{w}_i \frac{\bar{L}_i(1-\xi)}{\bar{K}_i \xi}\right)^{1-\xi}\right)^{\gamma} \bar{P}_i^{1-\gamma} \\
&= B\left(\frac{(1-\xi)}{\xi}\right)^{(1-\xi)\gamma} (\bar{w}_i)^{\gamma} \left(\frac{\bar{L}_i}{\bar{K}_i}\right)^{(1-\xi)\gamma} \bar{P}_i^{1-\gamma} \\
&= B\left(\frac{(1-\xi)}{\xi}\right)^{(1-\xi)\gamma} \left(\frac{((1+g_k) - \beta(1-\delta))\xi}{(1-\xi)\beta}\right)^{(1-\xi)\gamma} \left(\frac{\bar{P}_i}{\bar{w}_i}\right)^{-\xi\gamma} \bar{P}_i \\
&= \Psi (\bar{w}_i)^{\xi\gamma} \left(\bar{P}_i\right)^{1-\xi\gamma},
\end{aligned}$$

where $\Psi = B\left(\frac{(1-\xi)}{\xi}\right)^{(1-\xi)\gamma} \left(\frac{((1+g_k)-\beta(1-\delta))\xi}{(1-\xi)\beta}\right)^{(1-\xi)\gamma}$.

Hence, we can rewrite the labor market clearing condition as

$$\begin{aligned}
\bar{w}_i \bar{L}_i &= \sum_{n=1}^{N} \bar{A}_i \left(\frac{\bar{\kappa}_{ni} \bar{x}_i}{\bar{P}_n / T}\right)^{-\theta} \bar{w}_n \bar{L}_n \\
&= \sum_{n=1}^{N} \bar{A}_i \left(\frac{\bar{\kappa}_{ni} \Psi (\bar{w}_i)^{\xi\gamma} \left(\bar{P}_i\right)^{1-\xi\gamma}}{\bar{P}_n / T}\right)^{-\theta} \bar{w}_n \bar{L}_n
\end{aligned}$$

or

$$(\bar{w}_i)^{1+\xi\gamma\theta} \bar{L}_i \left(\bar{P}_i\right)^{\theta(1-\xi\gamma)} \left(\bar{A}_i\right)^{-1} = \sum_{n=1}^{N} (\bar{\kappa}_{ni} \Psi T)^{-\theta} \left(\bar{P}_n\right)^{\theta} \bar{w}_n \bar{L}_n.$$

Analogously, the price index can be written as

$$\begin{aligned}
\bar{P}_i^{-\theta} &= T^{-\theta} \sum_{n=1}^{N} \bar{A}_n (\bar{\kappa}_{in} \bar{x}_n)^{-\theta} \\
&= T^{-\theta} \sum_{n=1}^{N} \bar{A}_n \bar{\kappa}_{in}^{-\theta} \Psi^{-\theta} (\bar{w}_n)^{-\xi\gamma\theta} \left(\bar{P}_n\right)^{-\theta(1-\xi\gamma)}.
\end{aligned}$$

Turning to the value functions, we use the following change of variables:

$$\widetilde{m}_{in} \equiv exp\left(\bar{m}_{in}\right)^{-1/\nu},$$
$$\bar{\phi}_i = \sum_{n=1}^{N} exp\left(\beta \bar{V}_n - \bar{m}_{in}\right)^{1/\nu}.$$

Using these conditions, we express

$$\exp\left(\frac{\beta}{\nu}\bar{V}_i\right) = \left(\zeta \bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}} \left(\sum_{n=1}^{N} \widetilde{m}_{in} \exp\left(\frac{\beta}{\nu}\bar{V}_n\right)\right)^{\beta},$$

$$\exp\left(\frac{\beta}{\nu}\bar{V}_i\right) = \left(\zeta \bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}} \bar{\phi}_i^{\beta},$$

with $\zeta = (1+g_v)^{\beta}$. Hence,

$$\bar{\phi}_i = \sum_{n=1}^{N} \widetilde{m}_{in} \left(\zeta \bar{w}_n/\bar{P}_n\right)^{\frac{\beta}{\nu}} \bar{\phi}_n^{\beta}.$$

We also re-express the gross flows equation as

$$\bar{\mu}_{ni} = \frac{exp\left(\beta \bar{V}_i - \bar{m}_{ni}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta \bar{V}_h - \bar{m}_{nh}\right)^{1/\nu}} = \frac{\widetilde{m}_{ni} \left(\zeta \bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}} \bar{\phi}_i^{\beta}}{\bar{\phi}_n}.$$

Hence, we express the law of motion of labor as

$$\begin{aligned}
\bar{L}_i &= \sum_{n=1}^{N} \bar{\mu}_{ni} \bar{L}_n, \\
\bar{L}_i &= \sum_{n=1}^{N} \frac{\widetilde{m}_{ni} \left(\zeta \bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}} \bar{\phi}_i^{\beta}}{\bar{\phi}_n} \bar{L}_n, \\
\bar{w}_i^{-\frac{\beta}{\nu}} \bar{P}_i^{\frac{\beta}{\nu}} \bar{\phi}_i^{-\beta} \bar{L}_i \quad \zeta^{-\frac{\beta}{\nu}} &= \sum_{n=1}^{N} \widetilde{m}_{ni} \bar{\phi}_n^{-1} \bar{L}_n.
\end{aligned}$$

Finally, the evolution of technology is given by

$$\bar{A}_n = \frac{\alpha_0 \Gamma_\rho}{g_A} \sum_{i=1}^{N} \bar{s}_{in} \left(\bar{A}_i\right)^{\rho_l} \sum_{i=1}^{N} \bar{\lambda}_{ni} \left(\frac{\bar{A}_i}{\bar{\lambda}_{ni}}\right)^{\rho_m}$$

and using $\bar{\lambda}_{in} = \frac{\bar{A}_n(\bar{\kappa}_{in}\bar{x}_n)^{-\theta}}{\left(\bar{P}_i/T\right)^{-\theta}}$, $\bar{\mu}_{ni} = \frac{\widetilde{m}_{ni}\left(\zeta\bar{w}_i/\bar{P}_i\right)^{\frac{\beta}{\nu}}\bar{\phi}_i^{\beta}}{\bar{\phi}_n}$, and $\bar{s}_{in} = \frac{\bar{\mu}_{in}\bar{L}_i}{\sum_{h=1}^{N}\bar{\mu}_{hn}\bar{L}_h}$, we obtain

$$
\begin{aligned}
\bar{A}_n &= \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\bar{s}_{in}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\left(\bar{\lambda}_{ni}\right)^{1-\rho_m}\left(\bar{A}_i\right)^{\rho_m} \\
&= \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\frac{\bar{\mu}_{in}\bar{L}_i}{\bar{L}_n}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\left(\frac{\left(\bar{\kappa}_{ni}\bar{x}_i\right)^{-\theta}}{\left(\bar{P}_n/T\right)^{-\theta}}\right)^{1-\rho_m}\bar{A}_i \\
&= \frac{\alpha_0\Gamma_\rho}{g_A}\sum_{i=1}^{N}\frac{L_i}{L_n}\frac{\widetilde{m}_{in}\left(\zeta\bar{w}_n/\bar{P}_n\right)^{\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}}{\bar{\phi}_i}\left(\bar{A}_i\right)^{\rho_l}\sum_{i=1}^{N}\left(\frac{\left(\bar{\kappa}_{ni}\Psi\left(\bar{w}_i\right)^{\xi\gamma}\left(\bar{P}_i\right)^{1-\xi\gamma}\right)^{-\theta}}{\left(\bar{P}_n/T\right)^{-\theta}}\right)^{1-\rho_m}\bar{A}_i,
\end{aligned}
$$

Finally, rearranging this expression, we obtain

$$
\begin{aligned}
&\bar{w}_n^{-\frac{\beta}{\nu}}\bar{P}_n^{\frac{\beta}{\nu}-(1-\rho_m)\theta}\bar{\phi}_n^{-\beta}\bar{A}_n\bar{L}_n \\
&= \frac{\alpha_0\Gamma_\rho\left(\Psi T\right)^{-(1-\rho_m)\theta}}{\zeta^{-\frac{\beta}{\nu}}g_A}\sum_{i=1}^{N}\widetilde{m}_{in}\bar{\phi}_i^{-1}\bar{A}_i^{\rho_l}\bar{L}_i\sum_{i=1}^{N}\bar{\kappa}_{ni}^{-(1-\rho_m)\theta}\bar{w}_i^{-(1-\rho_m)\theta\xi\gamma}\bar{P}_i^{-(1-\rho_m)\theta(1-\xi\gamma)}\bar{A}_i.
\end{aligned}
$$

We end up with the following system of equations to solve for the equilibrium variables at the balanced growth path:

$$
\bar{w}_i^{1+\xi\gamma\theta}\bar{P}_i^{\theta(1-\xi\gamma)}\bar{L}_i\bar{A}_i^{-1} = \left(\Psi T\right)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{ni}^{-\theta}\bar{w}_n\bar{P}_n^{\theta}\bar{L}_n, \tag{C.1}
$$

$$
\bar{P}_i^{-\theta} = \left(\Psi T\right)^{-\theta}\sum_{n=1}^{N}\bar{\kappa}_{in}^{-\theta}\bar{w}_n^{-\theta\xi\gamma}\bar{P}_n^{-\theta(1-\xi\gamma)}\bar{A}_n, \tag{C.2}
$$

$$
\bar{w}_i^{-\frac{\beta}{\nu}}\bar{P}_i^{\frac{\beta}{\nu}}\bar{L}_i\bar{\phi}_i^{-\beta}\zeta^{-\frac{\beta}{\nu}} = \sum_{n=1}^{N}\widetilde{m}_{ni}\bar{L}_n\bar{\phi}_n^{-1}, \tag{C.3}
$$

$$
\bar{\phi}_i = \sum_{n=1}^{N}\widetilde{m}_{in}\zeta^{\frac{\beta}{\nu}}\bar{w}_n^{\frac{\beta}{\nu}}\bar{P}_n^{-\frac{\beta}{\nu}}\bar{\phi}_n^{\beta}, \tag{C.4}
$$

$$
\begin{aligned}
&\bar{w}_n^{-\frac{\beta}{\nu}}\bar{P}_n^{\frac{\beta}{\nu}-(1-\rho_m)\theta}\bar{\phi}_n^{-\beta}\bar{A}_n\bar{L}_n \\
&= \frac{\zeta^{\frac{\beta}{\nu}}\alpha_0\Gamma_\rho}{g_A(\Psi T)^{(1-\rho_m)\theta}}\sum_{i=1}^{N}\widetilde{m}_{in}\bar{\phi}_i^{-1}\bar{A}_i^{\rho_l}\bar{L}_i\sum_{i=1}^{N}\kappa_{ni}^{-(1-\rho_m)\theta}\bar{w}_i^{-(1-\rho_m)\theta\xi\gamma}\bar{P}_i^{-(1-\rho_m)\theta(1-\xi\gamma)}\bar{A}_i. \qquad \text{(C.5)}
\end{aligned}
$$

To prove the existence and uniqueness of the balanced growth path equilibrium in the detrended model, we start with the following change of variables:

$$\bar{A}_n = \bar{A}^l_n \bar{A}^m_n,$$

$$\bar{A}^l_n = \sum_{i=1}^{N} \frac{\bar{\mu}_{in} \bar{L}_i}{\bar{L}_n} \left(\bar{A}_i\right)^{\rho_l},$$

$$\bar{A}^m_n = \frac{\alpha_0 \Gamma_\rho}{g_A} \sum_{i=1}^{N} \bar{\lambda}_{ni} \left(\frac{\bar{A}_i}{\bar{\lambda}_{ni}}\right)^{\rho_m}.$$

Using this change of variables in the equilibrium system, we can write the matrices Λ and Γ representing the exponents of $\{\bar{w}_i, \bar{P}_i, \bar{L}_i, \bar{\phi}_i, \bar{A}^l_i, \bar{A}^m_i\}$ on the left-hand side and right-hand side of the system of equations, respectively. These matrices are given by

$$\Lambda = \begin{pmatrix} 1+\theta\xi\gamma & \theta\left(1-\xi\gamma\right) & 1 & 0 & -1 & -1 \\ 0 & -\theta & 0 & 0 & 0 & 0 \\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{\beta}{\nu} & \frac{\beta}{\nu} & 1 & -\beta & 1 & 0 \\ 0 & -\left(1-\rho_m\right)\theta & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\Gamma = \begin{pmatrix} 1 & \theta & 1 & 0 & 0 & 0 \\ -\theta\xi\gamma & -\theta\left(1-\xi\gamma\right) & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ \dfrac{\beta}{\nu} & -\dfrac{\beta}{\nu} & 0 & \beta & 0 & 0 \\ 0 & 0 & 1 & -1 & \rho_l & \rho_l \\ -\left(1-\rho_m\right)\theta\xi\gamma & -\left(1-\rho_m\right)\theta\left(1-\xi\gamma\right) & 0 & 0 & 1 & 1 \end{pmatrix}.$$

We then define the matrix $\Omega = \Gamma\Lambda^{-1}$. Following Kleinman et al. (2021) and Allen et al. (2020), we show that if the spectral radius of Ω is equal to one ($\rho\left(\Omega\right) = 1$) and if Ω is invertible, then the solution must be unique up to scale. Evaluating the eigenvalues of Ω, we have

$$\begin{bmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \frac{-b+\sqrt{b^2-4ac}}{2a} \\ \frac{-b-\sqrt{b^2-4ac}}{2a} \\ 0 \\ \rho_m + \rho_l \end{bmatrix},$$

where

$$\begin{aligned} a &= \beta+\nu+\theta\gamma\nu\xi, \\ b &= -\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right), \\ c &= \beta\left(\nu-1-\gamma\xi\nu\left(1+\theta\right)\right). \end{aligned}$$

We proceed to show that $|\zeta_3| < 1$ and $|\zeta_4| < 1$.

First, $\left|-b+\sqrt{b^2-4ac}\right| < 2a$. If $-b+\sqrt{b^2-4ac} > 0$, then this is equivalent to show that $b^2-4ac < (2a+b)^2$ or that $b^2-4ac < (2a+b)^2 = 4a^2+4ab+b^2$, or equivalently that $0 < a+b+c$ which holds since $a+b+c = \gamma\nu\xi(1-\beta)(1+2\theta) > 0$. Otherwise, if $-b+\sqrt{b^2-4ac} < 0$, we need to show that $b-\sqrt{b^2-4ac} < 2a$, or that $-\sqrt{b^2-4ac} < 2a-b$ but note that $2a-b = 2\left(\beta+\nu+\theta\gamma\nu\xi\right)+\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right) = 2\beta+3\nu+\beta\nu+\left(\theta-1\right)\gamma\nu\xi+\theta\beta\gamma\nu\xi > 0$ so it holds since $\theta \geq 0$. Note that at $\theta = 0$, $2\beta+3\nu+\beta\nu-\gamma\nu\xi > 0$.

Second, $\left|-b-\sqrt{b^2-4ac}\right| < 2a$. If $-b-\sqrt{b^2-4ac} < 0$ then this is equivalent to show that $b+\sqrt{b^2-4ac} < 2a$, or that $\sqrt{b^2-4ac} < 2a-b$, and given that $2a-b = 2\left(\beta+\nu+\theta\gamma\nu\xi\right)+\nu\left(1+\beta-\gamma\xi\left(1+\theta\left(1-\beta\right)\right)\right) = 2\beta+3\nu+\beta\nu+\left(\theta-1\right)\gamma\xi\nu+\theta\beta\gamma\nu\xi > 0$ then this is equivalent to show that $b^2-4ac < (2a-b)^2$, $b^2-4ac < (2a-b)^2 = 4a^2+b^2-4ab$ or $-c < a-b$, or $0 < a-b+c = \nu(1+\beta)(2-\gamma\xi) > 0$. If $-b-\sqrt{b^2-4ac} > 0$ then this is equivalent to show that $-b-\sqrt{b^2-4ac} < 2a$, or that $0 < 2a+b$, but since we know that $a+b = \beta(1-\nu)+\gamma\nu\xi+(2-\beta)\theta\gamma\nu\xi > 0$, then $0 < 2a+b$.

The two additional eigenvalues are zero and $\rho_m+\rho_l < 1$.

It follows that the balanced growth path of the detrended equilibrium with idea flows from migration and trade is unique up to scale.

In Online Supplement H, we prove the existence and uniqueness of the steady state equilibrium of the detrended model at the balanced growth path in different versions of the model with and without idea diffusion from trade or migration.

D Dynamic-Hat Algebra

Proposition 4. Dynamic-Hat Algebra. Define $\hat{y}_{t+1}$ as the time difference in the detrended endogenous variable $\tilde{y}$; namely, $\hat{y}_{t+1} = (\tilde{y}_{t+1}/\tilde{y}_t)$. *Given an initial observed allocation* $\left\{\{\lambda_{in,0}\}_{i=1,n=1}^{N,N}, \{\mu_{in,0}\}_{i=1,n=1}^{N,N}, \{w_{i,0}L_{i,0}\}_{i=1}^{N}, \{K_{i,0}\}_{i=1}^{N}, \{L_{i,0}\}_{i=1}^{N}\right\}$, *parameters and elasticities* $(\rho_\ell, \rho_m, \theta, \nu, \gamma, \xi, \beta)$, *initial rate and growth rate in the arrival of ideas* (α_0, g_α) *and a convergent sequence of future changes in fundamentals under perfect foresight* $\{\hat{\kappa}_{in,t}, \hat{m}_{in,t}\}_{i=1,n=1,t=1}^{N,N,\infty}$, *the solution for the sequence of changes in the model's endogenous variables in the detrended model* $\{\hat{y}_{t+1}\}_{t=1}^{\infty}$ *does not require information on the level of fundamentals (trade and migration costs).*

Proof. Let us define $\hat{y}_{t+1}$ as the time difference in the detrended variable $\tilde{y}$; namely, $\hat{y}_{t+1} = (\tilde{y}_{t+1}/\tilde{y}_t)$. The equilibrium conditions in time differences of the detrended system are given by

$$\log(\hat{u}_{i,t+1}) = \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \mu_{in,t} \left(\hat{u}_{n,t+2}\right)^{\beta/\nu} \left(\hat{m}_{in,t+1}\right)^{-1/\nu}\right), \quad \text{(D.1)}$$

$$\mu_{in,t+1} = \frac{\mu_{in,t} \left(\hat{u}_{n,t+2}\right)^{\beta/\nu} \left(\hat{m}_{in,t+1}\right)^{-1/\nu}}{\sum_{h=1}^{N} \mu_{ih,t} \left(\hat{u}_{h,t+2}\right)^{\beta/\nu} \left(\hat{m}_{ih,t+1}\right)^{-1/\nu}}, \quad \text{(D.2)}$$

$$L_{i,t+1} = \sum_{n=1}^{N} \mu_{ni,t} L_{n,t}, \quad \text{(D.3)}$$

$$\hat{x}_{i,t} = \left(\hat{w}_{i,t}^{\xi} \hat{r}_{i,t}^{1-\xi}\right)^{\gamma} \hat{P}_{i,t}^{1-\gamma}, \quad \text{(D.4)}$$

$$\hat{P}_{i,t+1} = \left(\sum_{n=1}^{N} \lambda_{in,t} \hat{A}_{n,t+1} \left(\hat{\kappa}_{in,t+1} \hat{x}_{n,t+1}\right)^{-\theta}\right)^{-1/\theta}, \quad \text{(D.5)}$$

$$\lambda_{in,t+1} = \lambda_{in,t} \hat{A}_{n,t+1} \left(\frac{\hat{\kappa}_{in,t+1} \hat{x}_{n,t+1}}{\hat{P}_{i,t+1}}\right)^{-\theta}, \quad \text{(D.6)}$$

$$\hat{w}_{i,t+1} \hat{L}_{i,t+1} = \frac{1}{\tilde{w}_{i,t} L_{i,t}} \sum_{n=1}^{N} \lambda_{ni,t+1} \hat{w}_{n,t+1} \hat{L}_{n,t+1} \tilde{w}_{n,t} L_{n,t}, \quad \text{(D.7)}$$

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)} \tilde{R}_{i,t} \tilde{K}_{i,t}, \quad \text{(D.8)}$$

$$\tilde{R}_{i,t+1} = 1 - \delta + \frac{\hat{w}_{i,t+1} \hat{L}_{i,t+1}}{\hat{P}_{i,t+1} \hat{K}_{i,t+1}} \left[\tilde{R}_{i,t} - (1-\delta)\right], \quad \text{(D.9)}$$

$$\hat{A}_{n,t+1} = \frac{1}{(1+g_A)} + \frac{\alpha_0 \Gamma_\rho}{\tilde{A}_{n,t}(1+g_A)} \sum_{i=1}^{N} s_{in,t} \left(\tilde{A}_{i,t}\right)^{\rho_\ell} \left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}\right], \quad \text{(D.10)}$$

where $\hat{u}_{i,t+1} = \exp(\tilde{V}_{i,t+1} - \tilde{V}_{i,t})$, $\hat{m}_{in,t+1} = \exp(m_{in,t+1} - m_{in,t})$, $\tilde{R}_{i,t} = \tilde{r}_{i,t}/\tilde{P}_{i,t} + (1-\delta)$. Note we use the fact that $L_{n,t} = \tilde{L}_{n,t}$,$\mu_{ni,t} = \tilde{\mu}_{ni,t}$ and $\lambda_{in,t} = \tilde{\lambda}_{in,t}$.

In Online Supplement I we provide the algebra to arrive in the system of equilibrium conditions in changes. As the system of equations in time differences shows, solving the model in relative time differences requires conditioning the model on the initial observable allocations $\lambda_{in,0}$, $\tilde{w}_{i,0} L_{i,0} + \tilde{r}_{i,t} \tilde{K}_{i,0}$, $L_{i,0}$, $\mu_{in,0}$, and $\tilde{K}_{i,0}$, and elasticities θ, ν, β, δ, ρ_ℓ, ρ_m, and α_0, which contains information on the initial level of fundamentals as the model inversion shows.

E Data Sources and Empirical Moments

In this section of the appendix, we describe in more detail the data sources and construction used in the quantitative analysis.

We obtain GDP, employment, export, and import data from the China Compendium of Statistics, 1949-2008.[1] The book consists of three main parts. The first part contains data at the national level compiled by National Bureau of Statistics. The second part presents data from provinces, autonomous regions, and municipalities under the direct jurisdiction of the central government; the data are complied by local statistical bureaus. The third part provides data from the Special Administrative Regions (SARs) of Hong Kong and Macao that have been edited by the National Bureau of Statistics. The national GDP, employment, and trade data do not include those of the Hong Kong SAR, Macao SAR, or Taiwan Province.

E.0.1 List of Provinces

The geographic units used in the quantitative analysis are Chinese provinces and the rest of the world. Strictly speaking, the province-level administrative divisions in China include provinces, autonomous regions, and municipalities under the direct jurisdiction of the central government. For simplicity, we call provinces to these highest-level administrative divisions of China. These provinces are Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Hainan, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

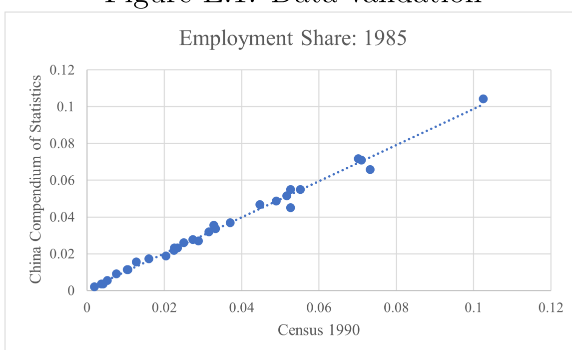
Gross Migration Flows. We use the Chinese census data from IPUMS to construct the migration flow matrix. Our constructed migration flow matrix matches the employment share of each province reported in the China Compendium of Statistics. We leverage the 1% samples of the 1990 and 2000 censuses from IPUMS as our data source to calculate migration flows for 1985-1990, 1990-1995, and 1995-2000.

To construct the migration flows for 1985-1990, we proceed as follows. With the 1% sample of the 1990 census, we include the working-age (15-64) population in our sample. Furthermore, we keep any census respondent who is actively employed in 1990. We put a weight on each province to match exactly the provincial employment share shown in the census with that of each province reported by the China Compendium of Statistics. For each individual, we determine the Hukou location as follows. For the 1990 census, the status and nature of registration was asked. If the person chose "(1) residing and registered here", we use the person's location in 1990 as the registration location; if the person chose "(2) residing here over 1 year, but registered elsewhere", "(3) living here less than 1 year absent from registration place over 1 year" or "(4) living here with registration unsettled", we use the person's location in 1985 as the registration location. For a person whose Hukou registration is

[1] The digitized data can be extracted from the China statistical yearbooks available at https://data.cnki.net/area/yearbook/Single/N2010042091?dcode=D03.

in category (2)-(4) but who lived in the same province (stayer) in 1985 and 1990, we assign a Hukou place to them as follows. We first construct a sample of migrants who switched their habitant province between 1985 and 1990, as measured in the data. Then, for each destination province, we compute the share of migrants coming from different origin provinces. We assign the Hukou place to the aforementioned stayer according to this share.[2] For each Hukou location (province-level), we construct a five-year migration flow matrix from origin province to destination province. Combining the migration matrix and the data in 1990, we can check whether the employment share of each province out of the nationwide total employment is consistent with the data from the China Compendium of Statistics (see Figure E.1).

Figure E.1: Data validation



Using the same method, we calculate migration flow between 1995 and 2000 using the 1% sample of the 2000 census from IPUMS. To the best of our knowledge, there are no publicly available micro-level people census data from 1995. When necessary, we thus use as proxy the migration flows from 1985 to 1990 for the flows from 1990 to 1995.

Trade and Production Data. To obtain the trade and production data for China and the rest of the world, we proceed as follows. To take the model to the data, in addition to the data described in the previous section of this appendix, the bilateral trade shares $\lambda_{in,t}$, total expenditure $X_{i,t}$,value added $w_{i,t}L_{i,t} + r_{i,t}K_{i,t}$, and the initial capital stock $K_{i,0}$ must be obtained. We also need to compute the share of value added in gross output, γ, and the share of labor in value added, ξ.

GDP, Employment, Export, and Import Data. We obtain GDP, employment, export, and import data from the China Compendium of Statistics, 1949-2008.

We make several adjustments to the data. First, the Chinese national accounts are based on data provided by local governments to the National Bureau of Statistics (Bai, Hsieh, and Qian (2006), and Chen, Chen, Hsieh, and Song (2019)). Given the

[2] A potential concern is step migration, i.e., a person does not directly migrate from her registration location to the current location. We cannot check this using the 1990 census. Imbert, Seror, Zhang, and Zylberberg (2022) uses 2005 mini census data to show that step migration was negligible in 2000-2005. We do not expect this to be any different for the period 1985 to 1990.

incentive of local governments to overstate the local GDP and other measurement discrepancies, the National Bureau of Statistics adjusts the data reported from the local governments to calculate the national-level GDP using independent data sources. Consequently, the reported aggregate GDP is generally lower than the sum of reported province-level GDPs. We address this issue by scaling down province-level GDP by the same proportion for each Chinese province to match the reported GDP at the national level. We follow the same strategy to adjust province-level employment, export, and import data to match their reported national aggregates.

Second, we account for the changing status of Chongqing. Before 1997, Chongqing was not considered a municipality under the direct jurisdiction of the central government. One of the focuses of this paper is to understand the rise of China in 1990s. For most of this period, Chongqing was still part of Sichuan, we thus treat Chongqing and Sichuan as an integrated province, Sichuan-Chongqing, throughout our paper. We aggregate relevant variables for the two regions.

Third, for some provinces, the measurement units are not consistent with the those of the national aggregates. For example, the export and import data of Guangdong Province are inaccurately reported by the local statistical bureau in units of 100 million Chinese Yuan, although the indicated unit is still 10,000 Chinese Yuan. We carefully checked and addressed this type of issues in the data.

We use the GDP deflator from the World Development Indicators compiled by the World Bank, to compute the real GDP at province level at 1990 prices.

To construct the production data for the rest of the world, the PWT 10.0 reports real GDP at constant 2017 national prices (rgdpna) and employment (emp). We rely on the Penn World Table 10.0 (PWT 10.0) to construct data for the rest of the world. We first keep all countries but China. Second, we drop countries with missing data for either GDP or employment. We aggregate all countries in our sample to obtain GDP and employment for the rest of the world. The World Development Indicators database reports the world GDP deflator from 1985 to 2017. Combining the two data sources, we compute GDP for the rest of the world at current year prices and real GDP at 1990 prices.

Capital stock. We follow Shan (2008) to estimate province-level capital stock from 1952 to 2010. We use the perpetual inventory method to estimate the time series of capital stock. For capital stock at the base year, we follow Young (2003), using 10 percent of the gross capital formation in 1952. As Young (2003) and Bai et al. (2006) argue, the most appropriate measure of investment in China is fixed capital formation. We obtain this measure from the China Compendium of Statistics. The investment price deflator is constructed by Shan (2008) based on official statistics. We follow Shan (2008) to choose the value for the depreciation rate.

For the rest of the world, we obtain capital stock at constant 2017 national prices from the PWT. We deflate country-level capital stock to reflect 1990 national prices using the GDP deflator. We further adjust the capital stock of the rest of the world by matching the percentage gross fixed capital formation in GDP compiled by the World

Bank. We start from the aggregate capital stock of all countries (including China) in 1985 according to the PWT. We adjust for the aggregate capital stock in the years 1990, 1995, and 2000 to match the average gross capital formation (percentage of GDP) in 1985-1990, 1990-1995, and 1995-2000, respectively. Afterward, by excluding the capital stock of China, we obtain the capital stock for the rest of the world.

Shares. We compute the values of $\gamma = 0.38$ and $\xi = 0.54$, which correspond to the parameter values for the year 1990 from world's aggregates in the Eora multi-region input-output table.

To unify the units of measure, GDP, exports, and imports are measured in 100 million USD, while employment is measured in units of 10,000 people.

E.1 Empirical Moments

Table E.1 presents the empirical moment conditions targeted to discipline the elasticities that govern innovation and idea diffusion $(\alpha_0, \rho_l, \rho_m)$, and the model-implied moments predicted by the evolution of fundamental productivity using equation (9).

Table E.1: Moment Conditions

Moment	Data	Model
Moment 1	164.5	114.5
Moment 2	0.84	0.60
Moment 3	18.5	2.8
Moment 4	5.3	7.1
Moment 5	-8.0	-9.1

Note: Moment 1 is the average change in fundamental productivity levels across locations. Moment 2 is the average growth rate in fundamental productivities. Moment 3 is the variance in the time changes in fundamental productivity levels (in thousands). Moment 4 is the covariance between the initial fundamental productivities and the change in fundamental productivity levels (in thousands). Moment 5 is the covariance between the initial fundamental productivities and the growth rate in fundamental productivities.

F Empirical Evidence of Idea Diffusion

In this section of the appendix, we provide further empirical evidence related to the idea diffusion mechanism in our model. In particular, we use province-level patent data, along with trade and migration data, to support the role played by trade and migration in the diffusion of ideas.

We obtain province-level patent data from the China Statistics Yearbooks. There are three types of patents: innovation, utility, and design. For each type of patent, the yearbook reports the number of applications and number of approved patents in a given year. To proxy the measure of knowledge stock, $A_{n,t}$, we calculate the cumulative approved patents of all three types at the province level for each year, starting from 1985. We then compile the province-level knowledge stock in 1985, 1990, 1995, 2000, 2005, and 2010, with which we calculate the change in the knowledge stock every five years from 1985 to 2010. For the approved patents in the rest of the world, we obtain data from Google Patent from 1985-2010 following Liu and Ma (2021).[3]

In the next section, we empirically test the model-implied law of motion of the knowledge stock (equation (9)) by implementing an instrumental variable strategy. The section complements Section 5 in the main text that provides reduced-form evidence of the contribution of idea diffusion from trade and migration to local knowledge.

F.1 Instrumental Variable Regressions

In this section of the appendix, we use the structure of our model to provide further evidence of our spatial mechanisms using our patent, production, and migration data. Recall that the evolution of the stock of knowledge in our model, according to equation (9), is given by

$$A_{n,t+1} - A_{n,t} = \alpha_t \Gamma_{\rho_\ell,\rho_m} \left[\sum_{i=1}^{N} s_{in,t} \left(A_{i,t} \right)^{\rho_\ell} \right] \left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}} \right)^{\rho_m} \right].$$

Taking logs on both sides, we obtain the following estimating equation

$$\log \left(A_{n,t+1} - A_{n,t} \right) = \beta_m \log \left(goods_{n,t} \right) + \beta_l \log \left(people_{n,t} \right) + \tau_t + \tau + \epsilon_{n,t},$$

where we define $\log \left(goods_{n,t} \right) = \log \left[\sum_{i=1}^{N} \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}} \right)^{\rho_m} \right]$ and we define $\log \left(people_{n,t} \right) = \log \left[\sum_{i=1}^{N} s_{in,t} \left(A_{i,t} \right)^{\rho_\ell} \right]$, τ_t controls for the time fixed effect, which captures the term $\log \alpha_t$; τ is a constant that captures $\log \Gamma_{\rho_\ell,\rho_m}$; and $\epsilon_{n,t}$ follows an $i.i.d.$ standard normal distribution.

It is important to highlight that our model allows for two-way causality in this structural equation due to general equilibrium effects. That being said, in what follows, we still try to establish causality between the growth in knowledge stock and the diffusion through goods and people in order to provide further evidence of our idea diffusion mechanisms. Hence, we try to address potential endogeneity issues in the estimating equation.

As described before, locations more exposed to international trade benefit more from the global diffusion of ideas and experience faster growth in the stock of knowledge. At the same time, fast-growing locations may build up their comparative advantages, impacting international trade as a result. To address potential endogeneity issues, we need an instrument for $\log \left(goods_{n,t} \right)$. We instrument $\lambda_{ni,t}$ by $\lambda_{ni,1985}$, as the growth prospect can hardly affect the trade pattern 15-25 years before.

For the variable $\log \left(people_{n,t} \right)$, the reverse causality concern also holds. Locations where the stock of knowledge grows faster might experience more immigration and less

[3] We are grateful to Song Ma for sharing the Google Patent data.

outmigration, which affects the knowledge diffusion through people in those locations. Furthermore, a higher share of immigration might lead to faster or slower growth in the stock of knowledge depending on the relative knowledge level and insights of the locals and the immigrants. Hence, the endogeneity issues might lead to either upward or downward estimation bias. To address this, we need an instrumental variable for $\log(people_{n,t})$. We instrument $s_{in,t}$ by $s_{in,1985}$. The rationale is similar to the case of trade; the growth of the stock of knowledge stock can hardly be anticipated by people 15-25 years before, so $s_{in,1985}$ is exogenous from the perspective of location location n in year $t \geq 2000$. In short, the instrument for $\log(goods_{n,t})$ is defined as $\log\left[\sum_{i=1}^{N} \lambda_{ni,1985} \left(\frac{A_{i,t}}{\lambda_{ni,1985}}\right)^{\rho_m}\right]$, and the instrument for $\log(people_{n,t})$ is defined as $\log\left[\sum_{i=1}^{N} s_{in,1985} \left(A_{i,t}\right)^{\rho_\ell}\right]$, where $\rho_m = 0.61$ and $\rho_l = 0.2$ are taken from our GMM estimation. The IV regression results are reported in Table F.1. In Column (1), we directly test our model-implied law of motion of the change in knowledge stock, and we do not control for endogeneity. The positive and significant coefficients of the two diffusion variables are consistent with our spatial mechanisms and in line with the reduced-form evidence presented in Section 5. As the patent stock is a proxy for the knowledge stock in the model and is likely a function of knowledge stock and other factors, the magnitudes of the coefficients do not need to be constrained to be close to one. In Columns (2), we report the instrumental variable regression results.[4] The positive effects of idea diffusion through international trade and migration on the growth in knowledge stock are still salient.

Table F.1: Estimates of the effects of knowledge diffusion through trade and migration

Dept. Var: $\log(A_{n,t+1} - A_{n,t})$	OLS	IV
	(1)	(2)
$\log(goods_{n,t})$	0.470***	0.390**
	(0.127)	(0.162)
$\log(people_{n,t})$	4.996***	5.055***
	(0.238)	(0.300)
Constant	-3.317***	-4.157***
	(0.835)	(1.034)
Kleibergen-Paap rk Wald F statistic		87.54
Observations	90	87
R-squared	0.935	0.935
Year FE	✓	✓

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[4] In 1985 Hainan had not been elevated to the status of a province; therefore, Hainan is dropped from the sample in the instrumental variable regressions.