Mechanics of Spatial Growth*

Sheng Cai
City University of Hong Kong

Lorenzo Caliendo
Yale University, NBER

Fernando Parro
University of Rochester, NBER

Wei Xiang
University of Michigan

January 31, 2025

Abstract

This paper examines how internal migration and trade openness shape spatial and aggregate economic growth through knowledge diffusion. Using data from China's rapid growth period, we provide causal evidence that regions attracting migrants, especially from more productive regions, and those more exposed to international trade, experience faster knowledge accumulation. We develop a dynamic spatial model in which idea flows, mediated by trade and migration, drive forward-looking factor accumulation (labor and capital) and endogenous productivity growth. Our quantitative analysis highlights the significant impact of initial spatial conditions on China's long-term growth path and reveals that factor accumulation and idea diffusion vary in importance across different phases of its transition. Furthermore, our quantitative analysis highlights the crucial role of idea diffusion in explaining spatial growth heterogeneity.

*We thank Pol Antras, David Atkin, Arnaud Costinot, Jonathan Eaton, Elhanan Helpman, Oleg Itskhoki, Samuel Kortum, Marc Melitz, Ezra Oberfield, Esteban Rossi-Hansberg, Ana Maria Santacreu, Chris Tonetti, Fabrizio Zilibotti and many seminar and conference participants for their useful conversations and comments. We thank Tingyu Feng, Matthew Murillo, Xilin Yuan, and Yu Zhao for excellent research assistance. Correspondence by e-mail: shengcai@cityu.edu.hk, lorenzo.caliendo@yale.edu, fernando.parro@rochester.edu, xwe@umich.edu.

1 Introduction

Understanding economic growth requires understanding how countries accumulate factors of production and increase the productivity of these factors. Furthermore, aggregate economic growth is shaped by the process of development across space within a country—namely, the dynamics of the distribution of economic activity, 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 study empirically, theoretically, and quantitatively how factor accumulation and knowledge diffusion through internal migration and international trade drive spatial growth and, consequently, aggregate economic growth.

Empirically, we provide causal evidence showing that diffusion through migration and trade contributes to local knowledge growth. Theoretically, we develop a dynamic spatial growth model that incorporates knowledge diffusion through migration and trade, featuring forward-looking migrants and capital investors. We identify structural parameters that govern knowledge diffusion by estimating the law of motion of the local knowledge stock derived from the model. Using the estimated model, we quantitatively analyze how knowledge diffusion and factor accumulation have shaped spatial and aggregate growth in China from the 1990s to the 2010s.

We begin by documenting that internal migration and international trade positively affect local knowledge growth. We focus on China because it experienced significant increases in GDP, trade openness, and internal migration during the 1990s and 2000s, making it a highly relevant context for our research. We construct gross migration flows across provinces in China using population census data, and measure changes in the local knowledge stock through the stock and flow of patents. Building on Card (2001), we develop an identification strategy to estimate the causal impact of migration on the local stock of knowledge. We establish that provinces receiving more migrants experience greater knowledge growth. More importantly, our analysis reveals that migrants from more developed regions (defined as areas with 1990 total factor productivity above the median province) contribute more to the local knowledge stock than those from less developed regions. Additionally, we find that provinces more open to imports (instrumented by historical degrees of openness) experience higher growth. These findings hold across various outcome measures and specifications that control for year and province effects. Our results remain robust when using an alternative set of migration instruments. Specifically, we find consistent empirical results following Burchardi et al. (2020), extending Card (2001)'s approach to construct an instrument based on the pre-existing distribution of migrants in a given location. This constitutes our first contribution: to provide evidence that both migration and trade contribute to the local stock of knowledge and demonstrate that the origin of migrants plays a central role.

Our second contribution is theoretical. To interpret our empirical findings and quantify the role of migration and trade in shaping both spatial and aggregate economic growth, we develop a new dynamic spatial growth model that features endogenous local productivity growth driven by innovation and knowledge diffusion through trade and migration, as well as forward-looking agents making migration and capital investment decisions. Given the significant spatial heterogeneity within China, the model also accounts for spatial variation in international trade exposure and local factor accumulation driven by capital investment and internal migration. In particular, we consider a world economy with multiple countries and multiple locations within a country. In each location, growth is shaped by a process of factor accumulation and the endogenous evolution of local productivity. Workers supply labor in their location and make forward-looking migration decisions, which determine labor supply across locations. Capital accumulation at each location is determined by forward-looking landlords who make investment decisions in local capital. The evolution of local productivity is modeled as a stochastic process driven by the diffusion of global and local ideas. Global ideas are embedded in imported intermediate goods and diffuse more effectively to locations with greater exposure to international trade. Workers learn about local ideas and carry insights with them when they migrate, and the quality of these insights depends on the origin of the migrants. Consequently, the evolution of the local stock of knowledge depends on the degree of international trade openness and migration patterns, consistent with our empirical evidence described above.

Our third contribution is quantitative. We show how to bring the model to spatial data and demonstrate that the initial conditions significantly influence the transition dynamics of the economy. These conditions include the equilibrium spatial distribution of economic activity in 1990, which reflects prior policies and changes in economic fundamentals, such as the establishment of special economic zones. By extending the dynamic-hat algebra method of Caliendo et al. (2019) to a growth model, we do not need to assume that the economy is on a balanced growth path and instead perform counterfactual analysis along the transition path. This approach is essential for analyzing how initial conditions influence China's economic transition in the 1990s and 2000s, relative to the impact of changing trade and migration frictions during this period. Since comprehensive data to identify these earlier changes in fundamentals are unavailable, analyzing the economic transition driven by initial conditions offers novel insights into their role in subsequent growth in China.

Before performing the quantitative analysis, we estimate the elasticities that govern the strength of idea diffusion. We derive from the model a structural equation that connects trade and migration shares to changes in the local knowledge stock. This non-linear equation is the key equation for our estimates, though we face the challenge of having endogenous trade and migration flows in the non-linear estimation. To address this endogeneity, we construct instruments for those flows in a consistent way as in our empirical section. We then use instrumented non-linear least squares (NLLS) to estimate structural parameters that govern the rate of idea flows through trade and migration, as well as the rate of innovation. With data from 1990 and these estimates in hand, we proceed to our quantitative assessment.

We first examine how initial conditions shaped spatial and aggregate growth in China during the 1990s and 2000s. In particular, we ask: How would China have developed given the initial distribution of fundamentals in 1990 but assuming no changes in trade and migration costs thereafter? We find that initial conditions played a significant role in China's subsequent growth. In particular, the set of economic fundamentals in 1990, reflected in the initial conditions, placed the economy on a transition path characterized by high economic growth in the following decades. During the 1990s, idea diffusion and capital accumulation contributed roughly equally to this aggregate growth. However, while the contribution of capital accumulation to growth remained stable in the 2000s, the relative importance of idea diffusion increased significantly. Initial trade openness and worker mobility facilitated the spread of ideas and increased China's stock of knowledge. As knowledge diffused and locations expanded their knowledge stocks, individuals contributed better insights to their local areas and to other regions through migration. We also find that reforms related to changes in trade costs and migration frictions after 1990 played a relatively small role in growth compared to the impact of initial conditions.

The processes of factor accumulation and idea diffusion through trade and migration in our framework account for most of the observed spatial heterogeneity in both GDP levels and growth rates. Eastern provinces experienced higher growth rate due to capital accumulation and ideas diffusion from international trade relative to western and other provinces. Growth in the western and northern regions was primarily supported by idea diffusion from migrants coming from highly productive locations in eastern China. In the eastern provinces, growth during the 1990s was driven more by capital accumulation than by ideas diffusion. As noted earlier, the initial conditions in 1990 reflect earlier economic changes, including the establishment of special economic zones, most of which were located in eastern China and fostered capital accumulation and access to the world economy. By contrast, in the 2000s, the relative contribution of idea diffusion as a source of growth increased in eastern locations. Migrants from other regions brought better insights compared to the previous decade, enhancing the stock of knowledge in these areas. The growth effects of idea diffusion through trade and migration are crucial for understanding spatial growth during China's economic transition in the 1990s and 2000s. Removing knowledge diffusion from the model causes the correlation between the model-implied spatial growth and the observed growth in the data to decline by almost half.

Our empirical evidence on how migration affects knowledge flows is closely aligned with recent findings by Burchardi et al. (2020) on the impact of immigrants on ideas, innovation, and growth in the United States. We find evidence that internal migration enhances knowledge growth in China. Moreover, we present causal evidence demonstrating that migrants from more developed regions contribute more significantly to knowledge growth in a given destination area compared to those from less developed regions. These findings are closely related to the work of Pellegrina and Sotelo (2024), which examine the role of migration in disseminating knowledge across regions in Brazil. In particular, the authors present reduced-form evidence showing that average labor productivity in a specific region-crop group is higher when migrants to that region originate from areas with high productivity in the same crop. Moreover, we contribute to this literature by providing causal evidence on the impact of trade openness on local knowledge growth.[1]

The knowledge diffusion from trade in our model builds on quantitative frameworks of knowledge diffusion and growth, as in Eaton and Kortum (1999), Eaton and Kortum (2001) and Cai et al. (2022). More specifically, our approach of relating knowledge diffusion to trade is grounded in Buera and Oberfield (2020).[2] In a different context, our paper also relates to Cai and Xiang (2022), who study global growth and technology diffusion through multinational production across countries. Unlike those frameworks, we propose a dynamic spatial model that incorporates forward-looking factor accumulation and endogenous local productivity growth through the diffusion of global and local ideas to study aggregate and spatial growth.

The knowledge diffusion process in our framework also relates to spatial growth models involving the local diffusion of technology (e.g., Desmet and Rossi-Hansberg (2014) and Desmet et al. (2018)), and frameworks that incorporate frictional idea diffusion across space (e.g., Berkes et al. (2022)). Our framework shares some aspects with these papers, such as the spatial heterogeneity in fundamentals and the geographic aspect of local knowledge diffusion. However, in our framework, local knowledge growth is mediated by endogenous economic mechanisms—trade and migration—instead of being dictated by exogenous geographical distance or technological frictions. While distance is likely correlated with trade and migration due to gravity forces, our framework enables us to study how policy interventions or changes in economic fundamentals influence growth through equilibrium adjustments in trade and migration patterns. More broadly, this distinction sets our framework apart from studies focusing on other types of exogenous spatial spillovers or externalities. Additionally, our framework distinguishes itself from these papers by incorporating forward-looking migration and capital accumulation decisions, which are essential for capturing the impact of agents' anticipations on growth throughout China's transition dynamics.

The factor supply side of our framework is based on forward-looking migration decisions, as in Caliendo et al. (2019), while locations trade goods with the rest of the world, as in Eaton and Kortum (2002). As previously described, capital accumulation in our framework involves

[1] More generally, our connection of knowledge diffusion to migration is also related to recent empirical evidence on knowledge flows resulting from interactions among people (e.g., Atkin et al. (2022), Buzard et al. (2020)), as well as studies on the impact of immigrants on ideas, innovation, and growth in the United States and 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), Prato (2021)). There is also recent evidence on how internal migration impacts productivity and other related outcomes in destination regions in countries that have experienced significant internal migration episodes (e.g., Facchini et al. (2019), Imbert et al. (2022)).

[2] The model in Buera and Oberfield (2020) also relates to Kortum (1997) when there is no idea diffusion, and to Jones (1995) and Atkeson and Burstein (2019) which has intertemporal knowledge spillovers that are not modeled explicitly as a function of insights. Our paper also relates to and builds on studies that find that the diffusion of ideas across agents can generate sustained growth, as seen in Alvarez et al. (2013), Lucas and Moll (2014), and Perla and Tonetti (2014). 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), Sampson (2016), and Perla et al. (2021)). For a broad review of semi-endogenous growth models, see Jones (2022).

forward-looking atomistic landlords who make investment decisions in local capital to maximize intertemporal utility, following the structure outlined in Kleinman et al. (2023).[3] Our main departure is that, in our framework, aggregate growth arises from both factor accumulation and endogenous productivity growth at each location. This aspect is crucial for connecting spatial growth with the traditional sources of aggregate growth emphasized in macroeconomic literature.

This paper contributes to the literature on China's aggregate economic growth, which investigate factors such as financial frictions, innovation, hierarchical structures, and institutional reforms (e.g., Song et al. (2011), König et al. (2022), Song and Xiong (2023), Fernández-Villaverde et al. (2023), Brandt et al. (2023), Chen et al. (2024), Cheremukhin et al. (2024)).[4] In addition, Zhu (2024) highlights that knowledge diffusion, alongside the previously mentioned factors, is a primary driver of China's aggregate productivity growth. Relatedly, König et al. (2022) develop an endogenous growth model to study how diffusion through random interactions among firms and innovation contributed to TFP growth in China from 2007 to 2012, while Chen et al. (2024) empirically examine how policy diffusion across space impacts growth in China. Our work advances this strand of literature in three key dimensions. First, we provide novel causal evidence on knowledge diffusion across space, facilitated by the surge in international trade and internal migration during the 1990s and 2000s. Our empirical evidence offers insights into how each location acts as a source of ideas that diffuse spatially through migration and through trade, each making distinct contributions to growth in the destination. Second, we develop a new dynamic spatial model to incorporate this important diffusion mechanism and show that knowledge diffusion is crucial for rationalizing the observed heterogeneity in spatial growth in China. Third, we highlight the role of initial conditions in 1990, reflecting the influence of earlier policies and economic fundamentals on China's growth path post-1990. The feature of our quantitative framework, combined with our empirical evidence from the 1980s and 1990s—a pivotal period during which the Chinese economy began its rapid ascent—contrasts with existing literature, which primarily focuses on periods after the year 2000.

This paper also contributes to the literature on the role of China's internal migration in its development (Tombe and Zhu (2019), Fan (2019), Hao et al. (2020), Guo et al. (2022), Imbert et al. (2022), and Egger et al. (2024)). This strand of literature examines how internal migration in China impacts welfare, structural change, technology adoption, and TFP, typically through the lens of

[3]The distinction between landlords and workers is also related 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 key difference that labor is assumed to be immobile across countries in that strand of literature.

[4]Comprehensive reviews of the drivers behind China's rapid growth have been provided by Xu (2011), Zhu (2012), Storesletten and Zilibotti (2014), Zilibotti (2017), and Zhu (2024). A recent paper by Fernández-Villaverde et al. (2023) employs a Ramsey-Cass-Koopmans one-sector growth model augmented with a parsimonious TFP catch-up process to explain China's growth trajectory from 1995 to 2019, while remaining agnostic about the specific mechanisms driving TFP catch-up in China. We differ from their work in three key ways. First, our study provides explicit mechanisms that influence TFP to explain China's growth. Second, we explore how regional growth contributes to aggregate growth. Third, we account for growth in the 1990s and the role of initial conditions reflecting earlier economic fundamentals.

static models. Our study advances this literature by introducing a dynamic spatial framework with semi-endogenous growth to analyze how diffusion through internal migration influences both regional and aggregate growth during transitional dynamics—a defining characteristic of China's development in the 1990s.

The rest of the paper is structured as follows. In Section 2, we present stylized facts on trade, migration, and GDP in China during the 1990s, and provide casual evidence on how trade and migration affect local knowledge stock. In Section 3, we develop the dynamic spatial growth model. In Section 4, we estimate the model and describe the method for counterfactual analysis. Section 5 presents our quantitative results and Section 6 concludes. We relegate all proofs, theoretical derivations, and detailed data descriptions to the appendix.

2 Empirical Evidence

In this section, we document several salient characteristics of spatial development in China during the 1990s. To do this, we use the most comprehensive data available for China since 1987. With this data, we present facts on China's trade, growth, and migration. After that, we establish a casual relationship between trade, migration and idea flows. We now proceed to describe our dataset. Appendix A presents a more comprehensive description of all the data used in the empirical sections of this paper.

2.1 Data

The data we use include comprehensive information on inter-province migration flows in China over multiple decades. Specifically, it combines data from three Population Censuses conducted in 1990, 2000, and 2010, along with four Population Sample Surveys (1% representative sample) from 1987, 1995, 2005, and 2015. Additionally, the data we use includes macroeconomic indicators such as trade, production, GDP, employment, and capital formation. These are sourced primarily from the China Compendium of Statistics (1949–2008) and the National Bureau of Statistics online data (1995–2015). For data beyond China, we use resources such as the Penn World Tables and the World Development Indicators (WDI).

We access patent data through provincial-level statistics using National Statistical Yearbooks compiled by National Bureau of Statistics from 1985 to 2015, supplemented by Google patent data to represent global patent stocks. Google patent data serves as a publicly available counterpart of the European Patent Office's World Patent Statistical Database (PATSTAT) (Liu and Ma (2021)). With this dataset, we describe and analyze spatial growth in China and how it has been influenced by trade, and by migration dynamics in China.

2.2 Trade, Migration and Growth in China

Table 1 displays aggregate trends in trade, migration, and growth in China during the 1990s and 2000s. As we can see, during the first half of the 1990s, imports as well as exports increased by around 140%. Both figures continued to grow throughout the sample period presented in the table. Part of the growth in imports and exports during the beginning of our sample period can be attributed to the process of economic transformation and policies aimed at integrating the economy into the global economy. In fact, in the early 90s China had agreed to remove trade barriers and open its markets to foreign competition as it prepared to accede to the WTO by the year 2000. One policy in this direction was the implementation of place-based industrial policies for economic development, focusing on the establishment of special economic zones (see Lu et al. (2023) and Alder et al. (2016)). In the year 1990, special economic zones where in place in the provinces of Guangdong, Xinjiang, Fujian, and Hainan.

Table 1: Aggregate Trends

Year	Annual GDP Growth Rate	Change in Imports	Change in Exports	Number Migrants
1990-95	12.3%	148%	140%	10,440,127
1995-00	10.4%	70%	67%	34,264,210
2000-05	10.2%	193%	206%	38,597,107
2005-10	10.5%	112%	107%	57,400,107

Notes: This table shows the aggregate trends in GDP growth, changes in imports and exports, and inter-province migration. The data are sourced from the World Development Indicators, China Population Censuses, and population surveys conducted in various years.

Table 1 also shows the annual GDP growth in the 1990s and 2000s. As we can see, China experienced double-digit growth rates throughout this period. The annual growth rate in the first half of the 1990s was higher than that observed in the remainder of the period. This growth was very uneven across regions, with the eastern provinces experiencing the highest growth rates. Figure 1 presents the annualized growth rates across provinces over different time intervals.[5]

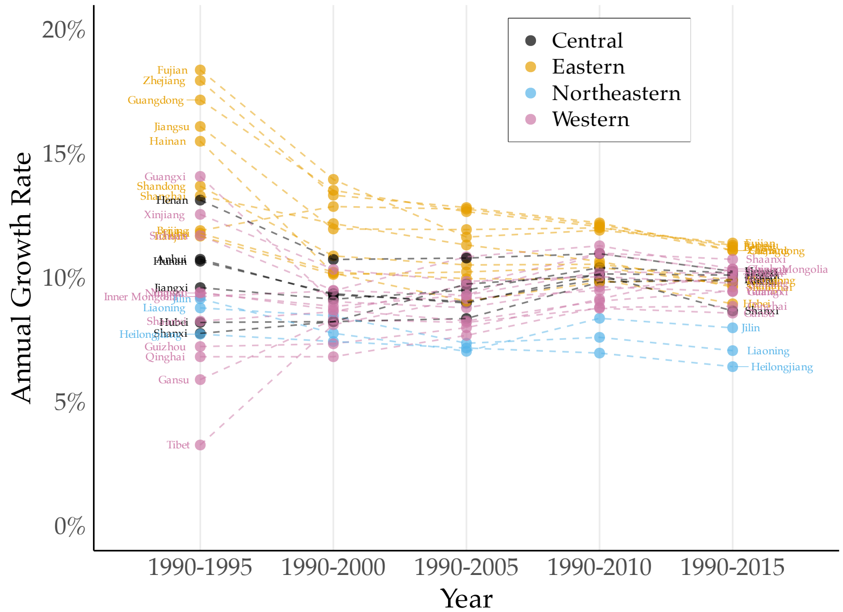
As we can see in Figure 1, eastern provinces experienced higher growth rates than those in the western and northern regions of China. In addition, growth rates were more dispersed across regions in the early 1990s compared to later periods, when they tended to converge. This heterogeneity in growth across space and time is a salient characteristics of China's growth path that we aim to explain with a spatial growth model. In Appendix A.2, Figure A.3 presents the levels of GDP for each province over time, indicating that the ranking of provinces in terms of real GDP levels of provinces over time remained stable despite the heterogeneity in spatial growth.

The last column in Table 1 shows inter-province migration flows. As we can see, the number of migrants has been rapidly increasing over time, starting from 10 million in the first half of the 1990s, and becoming more than 5 times large by the end of the sample period.[6]

[5] Appendix A.2 displays the annual growth rate and the level of GDP for each province.

[6] Note that this figure includes only inter-province migrants and excludes intra-province migrants, making it a

Figure 1: Real GDP Annual Growth Rates Across Provinces



Notes: The figure presents the annual GDP growth rate for each province in China for different time periods.

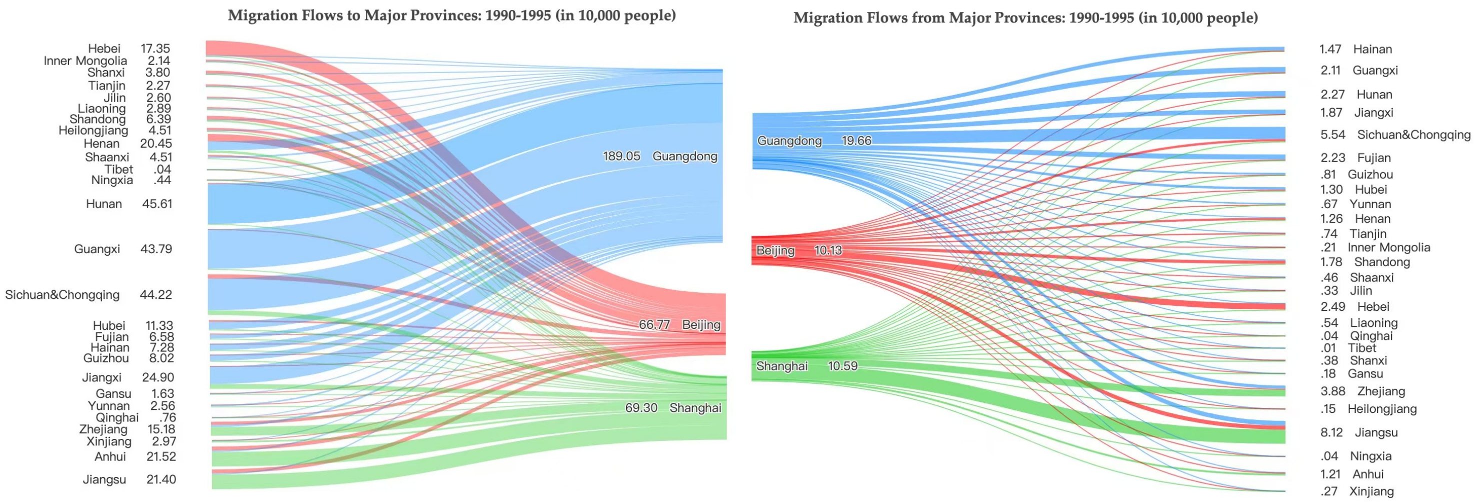
To depict mobility patterns across provinces in China, Figure 2 presents the five-year mobility flows across provinces that occurred from 1990 to 1995. As an illustration of the migration patterns (see Appendix A for more details on inter-province migration patterns), we show the mobility flows from other provinces to Beijing, Shanghai, and Guangdong (left panel) and from these three provinces to the rest of the provinces in China (right panel). Origin provinces are on the left axis, destination provinces are on the right axis, and a thicker line in the figure indicates a larger flow. These three provinces have higher initial measured productivity, and we can see how they receive migrants from all provinces in China. In the right panel, we also observe how migrants move from these high-productivity places to the rest of China, which indicates the importance of return migration in China, partly due to the Hukou restrictions, which impose extra costs for migrants to reside in provinces they were not born.[7] These return migration patterns in the data will be useful for studying the extent to which return migrants diffuse knowledge from high-productivity areas to lower ones.

All of these patterns of trade and migration in China influenced the development process during the 1990s and 2000s. We now proceed to establish a casual relationship between migration, trade, and knowledge diffusion.

lower bound on the total number of migrants.

[7]See Chan (2010) for a summary of the Hukou system.

Figure 2: Mobility Across Provinces in China (1990-1995)



Notes: The upper panel shows migration from other provinces to Beijing, Shanghai, and Guangdong between 1990 and 1995, measured in millions of people. The lower panel depicts migration from Beijing, Shanghai, and Guangdong to other provinces during the same period in the same units. Data are sourced from the 1995 mini population census.

2.2.1 Trade, Migration and Knowledge Flows

In this section, we present evidence of idea flows through inter-province migration and international trade. We first document that provinces receiving more migrants experience larger increases in their stock of local knowledge. We then show that migrants from more productive origins contribute more significantly to the growth of knowledge in the destination provinces. Additionally, we document that provinces with higher import shares experience greater growth in their knowledge stock. We establish causality using instrumental variables for migration and import shares.

To establish that migration affects knowledge growth, we begin by estimating the following specification:

$$y_{n,t+1} - y_{n,t} = \beta * \text{Migration}_{n,t} + \gamma + \gamma_t + \gamma_n + \epsilon_{n,t}, \tag{1}$$

where $\text{Migration}_{n,t}$ measures the number of migrants flowing into destination province n between t and $t+1$, where each period is five years. We use migration data for the years t =1985, 1990, 1995, ..., and 2010. The outcome of interest $y_{n,t}$ is the number of patents per capita, hence $y_{n,t+1} - y_{n,t}$ is the change in the number of patents. We use several measures of $y_{n,t}$, including patents granted, patents filed, patent stocks, and patent flows. In particular, patent flows are defined as new patents in each period t, and patent stocks are defined as accumulated patents flows up until t.[8] This specification in changes controls for long-lasting differences between provinces,

[8]Our province-level patent data are obtained from the China Statistical Yearbook compiled by the National Bureau of Statistics, with the patents themselves sourced from China's State Intellectual Property Office (SIPO). In Appendix A.2, we show that patents are highly correlated with TFP at the provincial level. This is the same data source used by König et al. (2022), who also find a positive correlation between R&D and patenting.

Table 2: Migration, Trade, and Changes in Patents

	Granted		Filed	
	Flow	Stock	Flow	Stock
ΔPatent Per Capita	(1)	(2)	(3)	(4)
$\text{Migration}^H_{n,t}$	0.58***	0.85***	0.95***	1.55***
	(0.12)	(0.20)	(0.13)	(0.25)
$\text{Migration}^L_{n,t}$	-0.07**	-0.11**	-0.13***	-0.22***
	(0.03)	(0.05)	(0.04)	(0.06)
$(Import/GDP)_{n,t}$	0.23***	0.37***	0.34***	0.63**
	(0.04)	(0.13)	(0.10)	(0.24)
Observations	180	180	180	180
R^2	0.640	0.635	0.726	0.714
Number of provinces	30	30	30	30
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓

Notes: This table presents the correlation between changes in patents per capita and migration as well as the import-to-GDP ratio across provinces. Migrants are separated into those from high-TFP provinces ($\text{Migration}^H_{n,t}$) and low-TFP provinces ($\text{Migration}^L_{n,t}$). High-TFP provinces are defined as those with TFP levels above the national mean in 1990, while the remaining provinces are classified as low-TFP. $\text{Migration}^k_{n,t}$, where $k = H, L$ are measured in units of 10,000 people. $(Import/GDP)_{n,t}$ is measured in units of 1 percentage point. Patents per capita are expressed as the number of patents per 10,000 residents, including both migrants and stayers in 1990. The dependent variables in columns (1) to (4) are the flow per capita of granted patents, the stock per capita of granted patents, filed patent flow per capita, and filed patent stock per capita, respectively. The data covers years 1990, 1995, 2000, 2005, 2010, and 2015. All columns include controls for year and province fixed effects. Robust standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

mitigating concerns about the potential skewness of $y_{n,t}$. The terms γ_t and γ_n are time and province fixed effects, respectively, and γ is a constant. Since we control for province-specific trends in $y_{n,t}$ and also include a province fixed effect γ_n, our specification exploits only deviations from the province's average growth of patents flows and stocks over time. $\epsilon_{n,t}$ is the error term.

Table 3 presents the results. Panel A shows that, for all of our measures of knowledge, there is a positive correlation between the number of migrants a region receives and the change in local knowledge. Panel B shows that this correlation is stronger for migrants from more productive origins compared to those from less productive provinces. Specifically, $\text{Migration}^H_{n,t}$ and $\text{Migration}^L_{n,t}$ refer to migrants from H provinces and L provinces, where an H province is defined as one whose TFP was larger than the median TFP level across all provinces in 1990. Remarkably, the correlations with these two variables are stronger compared to the simple correlation with the total number of migrants.

We then expand the previous regressions to explore the role of international trade in knowledge diffusion by estimating the following specification:

Table 3: Effects of Migration on Knowledge Growth

	Granted		Filed	
	Flow	Stock	Flow	Stock
ΔPatents Per Capita	(1)	(2)	(3)	(4)
Panel A: Pooled Migrants				
$\text{Migration}_{n,t}$	0.07*	0.10**	0.10*	0.16*
	(0.03)	(0.05)	(0.05)	(0.08)
Observations	180	180	180	180
R^2	0.451	0.452	0.507	0.484
Number of provinces	30	30	30	30
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
Panel B: Migrants of Two Types				
$\text{Migration}^H_{n,t}$	0.57***	0.83***	0.93***	1.51***
	(0.07)	(0.11)	(0.13)	(0.18)
$\text{Migration}^L_{n,t}$	-0.06**	-0.09**	-0.11**	-0.19***
	(0.03)	(0.04)	(0.04)	(0.06)
Observations	180	180	180	180
R^2	0.596	0.587	0.683	0.659
Number of provinces	30	30	30	30
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓

Notes: This table presents the correlation between migration and changes in patents per capita across provinces. $\text{Migration}_{n,t}$ and $\text{Migration}^k_{n,t}$, where $k = H, L$ are measured in units of 10,000 people. Patents per capita are expressed as the number of patents per 10,000 residents, including both migrants and stayers in 1990. The dependent variables in columns (1) to (4) are the flow per capita of granted patents, the stock per capita of granted patents, filed patent flow per capita, and filed patent stock per capita, respectively. In Panel A, the main independent variable is the total number of migrants from other provinces ($\text{Migration}_{n,t}$). In Panel B, migrants are separated into those from high-TFP provinces ($\text{Migration}^H_{n,t}$) and low-TFP provinces ($\text{Migration}^L_{n,t}$). High-TFP provinces are defined as those with TFP levels above the national mean in 1990, while the remaining provinces are classified as low-TFP. All columns include controls for year and province fixed effects. Robust standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

$$y_{n,t+1} - y_{n,t} = \sum_{j \in \{H,L\}} \beta^j * \text{Migration}^j_{n,t} + \beta^{\text{IM}} * \left(\frac{Import}{GDP}\right)_{n,t} + \gamma + \gamma_t + \gamma_n + \epsilon_{n,t}. \tag{2}$$

The results in Table 2 show that, as before, knowledge growth is positively correlated with migrants from high-TFP provinces compared to low-TFP provinces. Additionally, the results show a positive correlation between knowledge growth and international trade, measured as import shares in GDP.

Even though we estimate (2) in differences and control for fixed effects that account for province-specific trends, a concern with the OLS regressions is that unobserved shocks may influence both knowledge growth and migration. We discuss potential sources of endogeneity and propose a solution using two sets of instrumental variables.

Suppose there is a positive shock to productivity and knowledge growth in provinces with high TFP (e.g., coastal provinces). This shock could lead to production expansion by incumbent firms, attracting cheap labor from low-TFP provinces and retaining local workers. Consequently GDP may increase, causing the import-to-GDP ratio to decline in those provinces. This scenario would introduce a downward bias in β^H and β^{IM} and an upward bias in β^L. Alternatively, consider a positive shock to productivity and knowledge growth in interior provinces with low TFP. Such shock could encourage the entry of new firms requiring high-skilled managers from high-TFP provinces, alongside local cheap labor. In this case, the bias would be upward in β^H and β^{IM}, and downward in β^L.

To address these endogeneity concerns and similar issues, we follow Card (2001) and Burchardi et al. (2020) in constructing instrumental variables for internal migration. Additionally, we use a five-year lag of import shares as an instrument for current import shares.

The Card IV is defined as follows: $I_{in,t}^{Card} = I_{i\cdot,t} \times \frac{I_{in,t-1}}{I_{i\cdot,t-1}}$, where $I_{i\cdot,t}$ is total number of out-migration from i in time t, and $n \neq i$. The IV is a combination of push factors $I_{i\cdot,t}$, and pull factors measured by the previous share of migrants from i and into n. Our IV is constructed as

$$I_{n,t}^{j,IV} = \sum_{i \in \mathcal{I}^j} I_{in,t}^{Card}, j = L, H,$$

where $\mathcal{I}^L$ stands for the set of origin provinces with TFP below the national median in 1990, and $\mathcal{I}^H$ is the set with TFP above the national median in 1990. The identification assumption for our Card IV is that $\epsilon_{n,t}$ is orthogonal to $I_{i\cdot,t} \times \frac{I_{in,t-1}}{I_{i\cdot,t-1}}$ for all i and t. This requires that any unobserved shocks $\epsilon_{n,t}$ that cause temporary increases in a given destination province's knowledge growth do not systematically correlate with migrants from an origin province to other provinces ($I_{i\cdot,t}$) interacted with the share of migrants in that destination five years ago ($\frac{I_{in,t-1}}{I_{i\cdot,t-1}}$).

Table 4 presents the results from the second stage, and Tables B.1 in Appendix B presents the first-stage results. The results show causal evidence that provinces that receive more migrants from more productive places experience a larger increase in their knowledge relative to those that receive fewer migrants from those places. The results also indicate, as reflected by the negative sign on $\text{Migration}_{n,t}^L$, that migration from less productive provinces contributed less to knowledge growth compared to migration from more productive places, where the results vary slightly across specifications. Finally, we also find that a larger import share contributed to higher growth in local knowledge.

One of the possible concerns regarding the Card IV is that both push and pull factors might still be endogenous. As a result, we also follow Burchardi et al. (2020) and construct an instrumental variable that accounts for the possible endogeneity of the push and pull factors. In particular, the approach relies on the past history of migrants to instrument the push and pull factors in

Table 4: IV Regressions: Migration, Trade, and Changes in Patents

	Granted		Filed	
	Flow	Stock	Flow	Stock
ΔPatent Per Capita	(1)	(2)	(3)	(4)
Panel A: Card (2001)'s IV				
$\text{Migration}^{H}_{n,t}$	0.67***	1.00***	0.88***	1.56***
	(0.23)	(0.28)	(0.13)	(0.28)
$\text{Migration}^{L}_{n,t}$	-0.08	-0.12*	-0.10**	-0.22***
	(0.06)	(0.07)	(0.05)	(0.07)
$(Import/GDP)_{n,t}$	0.47***	0.85***	0.81***	1.58***
	(0.16)	(0.27)	(0.25)	(0.40)
Observations	150	150	150	150
Number of provinces	30	30	30	30
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
First Stage F-statistic	16.41; 18.72; 48.26			
AR Wald F-test P-value	0.000	0.000	0.000	0.000
Panel B: Burchardi et al. (2020)'s IV				
$\text{Migration}^{H}_{n,t}$	1.77***	2.85**	2.21**	4.05**
	(0.68)	(1.15)	(0.88)	(1.59)
$\text{Migration}^{L}_{n,t}$	-0.40**	-0.64**	-0.49**	-0.91**
	(0.16)	(0.28)	(0.22)	(0.40)
$(Import/GDP)_{n,t}$	1.35**	2.29***	1.87***	3.52***
	(0.57)	(0.89)	(0.68)	(1.11)
Observations	150	150	150	150
Number of provinces	30	30	30	30
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
First Stage F-statistic	12.93; 11.11; 6.12			
AR Wald F-test P-value	0.015	0.005	0.016	0.006

Notes: This table shows the IV regression results. Migrants are separated into those from high-TFP provinces ($\text{Migration}^{H}_{n,t}$) and low-TFP provinces ($\text{Migration}^{L}_{n,t}$). High-TFP provinces are defined as those with TFP levels above the national mean in 1990, while the remaining provinces are classified as low-TFP. $\text{Migration}^{k}_{n,t}$, where $k = H, L$ are measured in units of 10,000 people. $(Import/GDP)_{n,t}$ is measured in units of 1 percentage point. Patents per capita are expressed as the number of patents per 10,000 residents, including both migrants and stayers in 1990. The dependent variables in columns (1) to (4) are the flow per capita of granted patents, the stock per capita of granted patents, filed patent flow per capita, and filed patent stock per capita, respectively. We instrument migration following Card (2001) in Panel A and Burchardi et al. (2020) in Panel B. We instrument $(Import/GDP)_{n,t}$ using $(Import/GDP)_{n,t-1}$. We report the Sanderson-Windmeijer F-statistic in the first stage to accommodate our specification with multiple endogenous variables. The Anderson-Rubin Wald F-test p-value is reported in each column. All columns include controls for year and province fixed effects. Robust standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

the Card IV. In Appendix B.2 we describe all the steps taken to construct the Burchardi et al. (2020)-style instrument. Panel B of Table 4 shows that we find the same casual evidence as we did using the Card IV; namely, that migrants from more productive provinces contribute more to local knowledge than migrants from less productive places, and that provinces more open to international trade also experience a larger increase in local knowledge. In Table B.2, we present the estimated first-stage regression using the Burchardi et al. (2020)'s IV.

Motivated by the empirical evidence, we now introduce a dynamic spatial growth model that both aligns with these findings and allows us to quantify the role of diffusion through trade and migration in shaping regional and aggregate growth.

3 Dynamic Spatial Growth Model

In this section, we develop the dynamic spatial growth model. We begin with a description of technology diffusion in a single economy in Section 3.1. In Section 3.2 we introduce locations in the framework and describe the trade and production structure of the model. After that, we specify the supply of factors in our framework. In Section 3.3 we describe the capital accumulation decisions made by local landlords, and in Section 3.4 we specify the dynamic labor supply decisions made by migrants. In Section 3.5 we endogenize the idea diffusion process, relate it to migration and trade, and derive the evolution of the stock of knowledge across space. We also define the balanced growth path equilibrium of the economy and detrended model.

3.1 Innovation and Idea Diffusion

In this section, we model idea diffusion as a stochastic process in a generic economy to derive the distribution of ides and law of motion of the stock of knowledge. Later, by incorporating the structure of our framework, this law of motion of knowledge across locations will be linked to trade and migration in a way that rationalizes our empirical findings from the previous section.

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 if the new idea's 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, 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. Producers generate new ideas originating from their internal source of ideas, drawn from their own distribution of original ideas $H(z)$. Diffusion is a component that is external to the producer and that allows her to be exposed to the ideas of other producers. 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 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.

Assumption 1

a) The 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 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 Appendix C.

Proposition 1 shows that the probability distribution of the best new idea is Fréchet with shape 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 arrive with potentially 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. *The frontier of knowledge, $F_t(\cdot)$, at any t given follows a* **Fréchet** *distribution 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 Appendix C.

Proposition 2 establishes two results that we use in subsequent sections. First, at each moment in time the frontier of knowledge follows a Fréchet distribution, which we use to specify the production and trade structure in our framework, as described in the next section. Second, we can see that both the arrival rate of new ideas α_t and the source distribution $G_t(\cdot)$ matter for the evolution of A_t. Later in the paper, after we describe the economic environment, we return to discuss how ideas diffuse over space and relate the source distribution $G_t(\cdot)$ to ideas from sellers and from migrants.[9]

3.2 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. 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

[9] In Appendix C.1, we provide two approaches to derive the endogenous emergence of the Fréchet distribution for frontier productivity. The first approach is closely related to that of Buera and Oberfield (2020). We characterize how the distribution of frontier knowledge at t, combined with the arrival of ideas between t and $t+1$, gives rise to the distribution in $t+1$. This is akin to Buera and Oberfield (2020)'s method in a continuous-time framework, where they derive the evolution of the frontier knowledge distribution from t to $t+\Delta$ as $\Delta \to 0$. In both cases, we assume the initial distribution follows a Fréchet distribution. However, with the second approach (which we refer to as the *Poisson approach*), we can relax this assumption. The characterization of productivity distributions using the Poisson approach began with Eaton et al. (2011) and has since been adopted by Eaton et al. (2025) in the context of firm-to-firm trade. Cai and Xiang (2022) use the Poisson approach to characterize the endogenous emergence of Fréchet-distributed frontier technology in the context of technology diffusion through multinational production. Xiang (2023) employs similar approach to demonstrate how the endogenous innovation of multinational firms results in a Fréchet-distributed frontier technology. Similarly, Lind and Ramondo (2023)'s characterization of Max-stable Fréchet productivity, emerging from Poisson innovation and independent diffusion, mirrors the underlying spirit of the Poisson approach.

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 labor, capital, and material inputs, respectively. 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, respectively. It follows from the cost minimization problem of the producers that 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, respectively, and where B is a constant.[10]

We now use the results from the previous section in which we derived the law of motion for 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 time 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 (see Appendix D.1 for the derivation), which is given by

$$\lambda_{in,t} = \frac{A_{n,t} \left(\kappa_{in,t} x_{n,t} \right)^{-\theta}}{\sum_{h=1}^{N} A_{h,t} \left(\kappa_{ih} x_{h,t} \right)^{-\theta}}. \tag{3}$$

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

$$P_{i,t} = T \left(\sum_{n=1}^{N} A_{n,t} \left(\kappa_{in,t} x_{n,t} \right)^{-\theta} \right)^{-1/\theta}, \tag{4}$$

where intermediate varieties are aggregated with a constant elasticity of substitution η, and T is related to the gamma function; $T \equiv \Gamma \left(1 + (1-\eta)/\theta\right)^{1/(1-\eta)}$. Given this environment, total expenditure in location i, which we denote by $X_{i,t}$, is given by

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

[10]In particular, $B = \left[\xi^{\xi} (1-\xi)^{1-\xi} \right]^{-\gamma} \gamma^{-\gamma} (1-\gamma)^{\gamma-1}$.

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). 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

$$w_{i,t}L_{i,t} = \sum_{n=1}^{N} \lambda_{ni,t}w_{n,t}L_{n,t}. \tag{5}$$

3.3 Capital Accumulation Across Locations

We now turn to the supply of factors of production in the model. We start by describing capital accumulation decisions across space. At each location, we assume that there are atomistic landowners who consume the local consumption bundle with logarithmic preferences over consumption goods and whose source of income is from renting capital structures.[11] 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. (2023) 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,

$$\begin{aligned} c_{i,t} &= (1-\beta)\left[r_{i,t}/P_{i,t} + (1-\delta)\right]K_{i,t}, \\ K_{i,t+1} &= \beta\left[r_{i,t}/P_{i,t} + (1-\delta)\right]K_{i,t}, \end{aligned} \tag{6}$$

which give rise to the law of motion of capital accumulation across locations. In Appendix D.3

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

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 across space. Similar to Kleinman et al. (2023), 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.[12]

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

3.4 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

$$v_{i,t} = log(w_{i,t}/P_{i,t}) + \max_{\{n\}_{n=1}^{N}} \{\beta E_t[v_{n,t+1}] - m_{in,t} + \nu\epsilon_{n,t}\}, \tag{7}$$

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 (7) over $\epsilon_{n,t}$. We then obtain the value of location i for a representative worker in that

[12] As long as landlords are immobile, 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).

location at time t, denoted by $V_{i,t} = E_t[v_{i,t}]$. The value of location i is given by

$$V_{i,t} = log(w_{i,t}/P_{i,t}) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta V_{n,t+1} - m_{in,t}\right)^{1/\nu}\right). \tag{8}$$

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

$$\mu_{in,t} = \frac{exp\left(\beta V_{n,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta V_{h,t+1} - m_{ih,t}\right)^{1/\nu}}. \tag{9}$$

This equilibrium condition determines the gross migration flows of workers across space (see Appendix D.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 locations n (including stayers in i) at time t. Therefore, the stock of workers at each location evolves according to

$$L_{i,t+1} = \sum_{n=1}^{N} \mu_{ni,t} L_{n,t}. \tag{10}$$

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.

3.5 Idea Diffusion with Trade and Migration

We now specify the innovation and diffusion process described in Section 3.1 to allow for migration and trade to contribute to the local pool of ideas. Producers in location n obtain new insights from two sources. First, they obtain insights from sellers; namely, ideas from producers in other locations are embedded in imported intermediate varieties; like blueprints. Second, migrants carry insights with them when they arrive in a new location. 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 trade. The parameters $\rho_\ell, \rho_m \in [0,1)$ capture the learning intensity from both types of insights (migration and trade) with $\rho_\ell + \rho_m < 1$.

After imposing Assumption 1 and following the same steps as in Section 3.1, 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

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

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 of 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,

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

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 Appendix C.2).

In the case of the source distribution of goods, we assume that diffusion opportunities are randomly drawn from the set of best practices across all goods sold to location n. In this way the source distribution $G_{n,t}^m(q_m)$ is given by the fraction of goods for which the lowest-cost provider of the good to location n is a producer with productivity less than or equal to q_m. Under these mechanisms for idea diffusion, we derive the law of motion of the stock of knowledge across locations with idea flows from people and goods (see Appendix C.3). We obtain that the difference equation that determines the evolution of the stock of knowledge at each location is given by

$$A_{n,t+1} - A_{n,t} = \alpha_t \Gamma_{\rho_\ell,\rho_m} \underbrace{\left[\sum_{i=1}^N s_{in,t} \left(A_{i,t}\right)^{\rho_\ell}\right]}_{migration} \underbrace{\left[\sum_{i=1}^N \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}\right]}_{trade}, \tag{11}$$

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.

Equilibrium condition (11) 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 term $A_{i,t}/\lambda_{ni,t}$ on the right-hand side reflects the selection effect in trade, where location n gets insights only from the best producers (i.e., lowest-cost suppliers) in i. We also emphasize that the diffusion of ideas from migrants and sellers is endogenous since both migration and trade patterns are equilibrium objects in our framework. Additionally, it is worth noting that ideas diffuse not only from migrants and foreign sellers but also from local active producers and from non-migrants, hence the stock of knowledge also grows even in locations that are closed to trade or migration, namely where $s_{ii,t} = 1$ or $\lambda_{ii,t} = 1$. As described previously, the relative strength of idea diffusion, governed by the diffusion parameters ρ_ℓ and ρ_m, determine the importance of learning from people and goods.

Importantly, we can see how the evolution of stock of knowledge at each location rationalize the causal evidence in Section 2.2.1 regarding how trade and migration impact the local stock of knowledge. In particular, the equilibrium condition (11) reflects how locations receiving migrants from places with a higher stock of knowledge experience a larger knowledge stock relative to locations receiving migrants from less productive places. Additionally, locations more open to trade experience higher knowledge growth through the selection in trade. Our empirical evidence found in Section 2.2.1 is consistent with both aspects of equilibrium condition (11).

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, that is*

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

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

Definition 1. Equilibrium of the Spatial Growth Model. *Given an initial distribution of the local stock of knowledge $\{A_{i,0}\}_{i=1}^N$ and factor endowments $\{L_{i,0}, K_{i,0}\}_{i=1}^N$, the 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, $\{V_{i,t}, w_{i,t}, r_{i,t}, P_{i,t}, L_{i,t}, K_{i,t}, A_{i,t}\}_{i=1,t=0}^{N,\infty}$, that satisfies the equilibrium conditions determined by the bilateral trade shares (3), the equilibrium location prices (4), the labor market clearing condition (5), the capital accumulation condition (6), the location value function (8), the worker gross flow condition (9), the law of motion of labor (10), and the evolution of the stock of knowledge (11).*

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).

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 $1 + g_A = (1 + g_\alpha)^{\frac{1}{(1-\rho_\ell-\rho_m)}}$, capital grows at a rate $1 + g_k = (1 + g_A)^{\frac{1}{\theta\xi\gamma}}$, and values grow at a rate $1 + g_v = (1 + g_A)^{\frac{1}{\theta\xi\gamma(1-\beta)}}$.*

Appendix E solves for the equilibrium long-run growth rates of all variables along the balanced growth path.

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

4 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. 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

$$V_{i,t}^H = log(\psi_i^H w_{i,t}/P_{i,t}) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta V_{n,t+1}^H - m_{in,t}^H\right)^{1/\nu}\right), \tag{12}$$

$$\mu_{in,t}^H = \frac{exp\left(\beta V_{n,t+1}^H - m_{in,t}^H\right)^{1/\nu}}{\sum_{g=1}^N exp\left(\beta V_{g,t+1}^H - m_{ig,t}^H\right)^{1/\nu}}, \tag{13}$$

$$L_{i,t+1} = \sum_H \sum_{n=1}^N \mu_{ni,t}^H L_{n,t}^H, \tag{14}$$

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.

4.1 Data for Quantification

To bring the model to the data, we use the data described in Section 2.1 and Appendix A across provinces in China and for the rest of the world to obtain 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 conditional on Hukou type $\mu_{in,t}^H$. We also obtain 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 $(\alpha_0, \rho_\ell, \rho_m)$.

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 $(1-\delta) = 0.89^5$, which corresponds to an annual depreciation rate of 10.96 percent following Shan (2008), consistent with our capital stock estimation (see Appendix A.1 for more details.).[13] 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 $\eta = 2$

[13]Our choice capital depreciation rate is also consistent with that of Bai et al. (2006).

in the gamma function in equation (4). Our quantitative analysis will center on the role of local idea diffusion through internal migration and global idea diffusion through international trade.

4.2 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

$$Real\ GDP_{n,0} = \frac{w_{n,0}L_{n,0} + r_{n,0}K_{n,0}}{P_{n,0}} = \left(A_{n,0}/\left(\lambda_{nn,0}\Upsilon\right)\right)^{\frac{1}{\gamma\theta}} \left(K_{n,0}\right)^{(1-\xi)} \left(L_{n,0}\right)^{\xi}, \tag{15}$$

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 $\left(A_{n,0}/\left(\lambda_{nn,0}\Upsilon\right)\right)^{\frac{1}{\gamma\theta}}$. It has two main components: average 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 average productivity $A_{n,o}$ (scaled by a constant power coefficient), which is the average efficiency of the set of all goods produced and consumed in n. In an open economy, due to selection, 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.

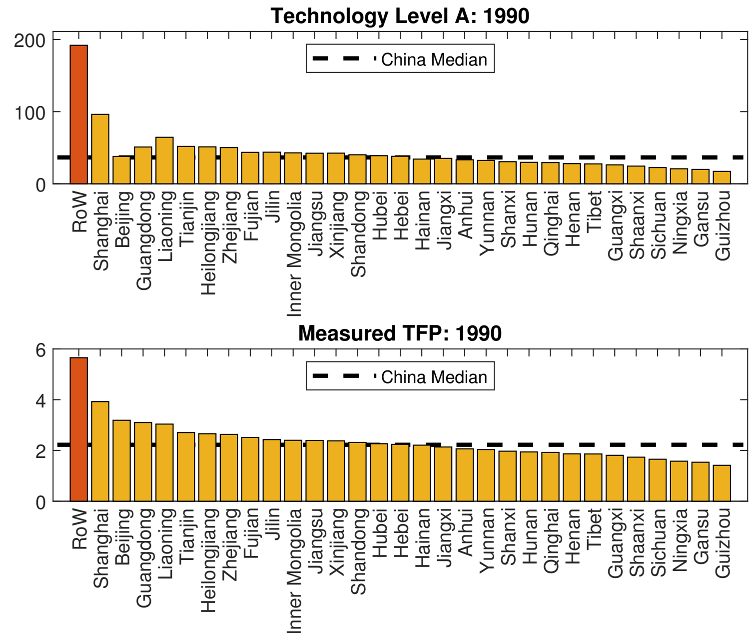
Inverting equation (15), and solving for fundamental productivity $A_{n,0}$, we obtain

$$A_{n,0} = \Upsilon \left(\frac{Real\ GDP_{n,0}}{\left(K_{n,0}\right)^{1-\xi} \left(L_{n,0}\right)^{\xi}} \right)^{\gamma\theta} \lambda_{nn,0}. \tag{16}$$

Using our data, we compute the initial stock of knowledge across provinces in China as well as for the rest of the world. Figure 3 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 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.

[14]See Appendix A.1 for the details of this derivation.

Figure 3: Initial Stock of Knowledge and Measured Productivity Across Locations (1990)



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

4.3 Estimation of Idea Diffusion from Trade and Migration

In our dynamic spatial growth model, three parameters govern productivity growth and idea diffusion across locations: the strength of idea diffusion through trade in goods ρ_m, the strength of idea diffusion through migration ρ_ℓ, and the arrival rate of insights α_0. To discipline these parameters in the our dynamic spatial model, we proceed as follows.

The equilibrium condition (11) provides a structural relation between the evolution of knowledge stocks, the diffusion elasticities, the arrival rate of insights, migration and trade. We use this structural relation to estimate the diffusion parameters. To do so we need to confront several issues. The first issue is the measure of knowledge stock to estimate these parameters. For example, one possibility is to use our model-consistent inverted A's (from the previous subsection), but one could argue that this involves using data affected by migration and trade, and as a result, it might bias our estimates. Another possibility is to use our patent measures as proxies of knowledge stock; while using this measure could help identify $\rho's$, it might get the units of α wrong (note that the magnitude of α is directly related to the magnitude of the change in A on the left-hand side of equation (11), which depends on the units of A's). To address this issue, we proceed with an iterative procedure. We first estimate $\rho's$ using the cumulative number of patents granted as proxies for A's. After we estimate the $\rho's$, we then use the estimated coefficients along with our inverted $A's$ into (11) to estimate α. A second issue we need to address is the fact that the observed evolution of knowledge stock, or TFP, could be partly influenced by determinants outside of our model.

To address this, we assume that the arrival rate of ideas is subject to *i.i.d.* location-specific shocks resulting in unobserved residuals in the TFP evolution that are not predictable by migrants and therefore are not part of their migration decisions. Consequently, our empirical strategy allows for an unobserved residual that captures the effects of factors influencing TFP besides idea diffusion. Finally, there is a possible concern that the migration shares $s_{in,t}$ may be correlated with the error term, as in the model people make forward-looking migration decisions. Similarly, trade shares $\lambda_{ni,t}$ might also be correlated with the error term. To address these concerns, we use instruments for both, migration shares (based on Card IV) and trade shares (based on lagged trade shares). Therefore, we use instrumented non-linear least squares (NLLS) from (11) and estimate ρ_m, ρ_ℓ, and α_0.

We first define the left-had side of (11) as y_{nt} and the right hand side as $g(\mathbf{s}_{nt}, \boldsymbol{\lambda}_{nt}, \mathbf{A}_t, t, \boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})$,with $\tilde{\alpha}_t \equiv \alpha_t \Gamma_{\rho_\ell, \rho_m}$. For notational brevity, we use bold letters to represent vectors. Our NLLS estimator is defined as the solution to

$$\min_{\boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}}} \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} \left(y_{nt} - g(\mathbf{s}_{nt}, \boldsymbol{\lambda}_{nt}, \mathbf{A}_t, t, \boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})\right)^2,$$

where $\mathbf{A}_t$ are the cumulative number of patents granted. Equivalently, the NLLS corresponds to the following moment condition

$$\mathbb{E}\left[u_{nt} \cdot D_{(\boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})} g(\mathbf{s}_{nt}, \boldsymbol{\lambda}_{nt}, \mathbf{A}_t, t, \boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})\right] = 0, \tag{17}$$

where $D_{(\boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})} g(\mathbf{s}_{nt}, \boldsymbol{\lambda}_{nt}, \mathbf{A}_t, t, \boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})$ is the gradient of the $g(\cdot)$ function, and $u_{nt} \equiv y_{nt} - g(\mathbf{s}_{nt}, \boldsymbol{\lambda}_{nt}, \mathbf{A}_t, t, \boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})$. Denote by $\mathbf{s}_{nt}^{iv}$, and $\boldsymbol{\lambda}_{nt}^{iv}$ to the instrumented migration and trade flows such that $\mathbb{E}\left[u_{nt} | \mathbf{s}_{nt}^{iv}, \boldsymbol{\lambda}_{nt}^{iv}, \mathbf{A}_t\right] = 0$. The NLLS IV estimator that we construct is given by

$$\mathbb{E}\left[u_{nt} \cdot D_{(\boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})} g(\mathbf{s}_{nt}^{iv}, \boldsymbol{\lambda}_{nt}^{iv}, \mathbf{A}_t, t, \boldsymbol{\rho}, \tilde{\boldsymbol{\alpha}})\right] = 0. \tag{18}$$

Table 5 presents the results of our NLLS estimation with and without IV. As we mentioned before, unlike $\hat{\rho}$'s, the magnitude of the estimate of α_0 is closely linked to the units used in the measure of knowledge stock.

Table 5: Non-linear Least Squares Estimation Results

	NLLS	With IV:
Diffusion from people ($\hat{\rho}_\ell$)	0.363 ***	0.353 ***
	(0.063)	(0.062)
Diffusion from trade ($\hat{\rho}_m$)	0.247 **	0.265 **
	(0.107)	(0.113)

Notes: This table shows the Non-linear Least Squares estimation results. Standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

We proceed to estimate α_0 given $\hat{\rho}_\ell$ and $\hat{\rho}_m$ from the previous step. In particular, we run the following specification to estimate α_0,

$$y_{nt} = \alpha_0 m_{nt} + u_{nt},$$

where

$$m_{nt} \equiv (1+g_\alpha)^{t-1}\Gamma(1-\hat{\rho}_\ell)\Gamma(1-\hat{\rho}_m)\left[\sum_{i=1}^{N} s_{in,t}\left(A_{i,t}\right)^{\hat{\rho}_\ell}\right]\left[\sum_{i=1}^{N}\lambda_{ni,t}\left(\frac{A_{i,t}}{\lambda_{ni,t}}\right)^{\hat{\rho}_m}\right].$$

In this equation, $1+g_\alpha = (1+g_{TFP})^{\gamma\theta(1-\hat{\rho}_\ell-\hat{\rho}_m)}$, we use our model-inverted $A_{i,t}$, and the values of $g_{TFP} = 0.03974$, $\gamma = 0.38$, and $\theta = 4.55$.[15] We then run a pooled linear regression without intercept to estimate α_0. We obtain $\hat{\alpha}_0 = 0.67(0.10)$. With this set of estimates, $\hat{\rho}_\ell$, $\hat{\rho}_m$, $\hat{\alpha}_0$, we proceed to conduct our quantitative analysis.

4.4 Computing Counterfactuals

To compute the dynamic spatial growth model, we extend the dynamic-hat algebra technique 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. Proposition 4 establishes this result. As a previous step, we first define formally the detrended economy. Appendix E 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.*

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.

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

[15] The U.S. TFP data in the 1990s is extracted from https://fred.stlouisfed.org/series/RTFPNAUSA632NRUG. We use data between 1990 and 2000 to calculate the five-year growth rate of TFP.

$(\rho_\ell, \rho_m, \theta, \nu, \gamma, \xi, \beta)$, the 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 Appendix F.

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 Appendix G. 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 parameters 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.

Applying dynamic-hat algebra to compute the spatial growth model without assuming that the economy is initially on the balanced growth path requires conditioning on the factual allocations in the initial year, which, in our case, is 1990. There are two important aspects of conditioning on initial observable allocations. First, these initial factual allocations contain all the relevant information on initial fundamentals (e.g., productivities, trade costs, migration costs) and also dictate how far the economy is in its transition to a balanced growth path. Second, one way to interpret the initial allocations is as summarizing the equilibrium or state of the economy due to past changes in economic fundamentals. Specifically, changes in trade costs, migration costs, and other policies that affect productivity across provinces in China—such as the special economic zones implemented by 1990—resulted in an equilibrium of the economy that is reflected in the distribution of economic activity, the level of internal mobility, and trade openness across space, as shown by the data in 1990. These two features of our approach allow us to use our framework to study the importance of initial conditions on spatial and aggregate growth in China, which, as mentioned above, resulted from economic changes prior to 1990. This is a key contribution of our quantitative empirical analysis which we now proceed to describe.

5 Mechanics of Spatial Growth in China

Table 6 highlights the importance of initial conditions for aggregate growth in China during the 1990s and 2000s. The first row shows that changes in economic fundamentals prior to 1990, as reflected in the 1990 distribution of economic activity in China, set the economy on a transition path marked by high economic growth in the following decades. The subsequent rows illustrate the relative contributions of different sources of growth within this transitional dynamic.

In the 1990s, idea diffusion and capital accumulation contributed roughly equally to aggregate growth. The second and third rows of Table 6 show that aggregate growth would have been halved without either mechanism during this period. However, while the contribution of capital

Table 6: Annual GDP Growth Rates

	90-95	90-00	90-05	90-10	90-15
With fundamentals in 1990	10.23%	9.86%	9.51%	9.17%	8.86%
W/o capital accumulation	5.66%	5.46%	5.28%	5.12%	4.98%
W/o idea diffusion ($\rho_\ell = 0$, $\rho_m = 0$)	5.22%	4.31%	3.65%	3.15%	2.76%
W/o ideas from people ($\rho_\ell = 0$)	6.03%	5.21%	4.60%	4.14%	3.78%
W/o ideas from goods ($\rho_m = 0$)	6.49%	5.73%	5.16%	4.72%	4.38%
W/o technology growth ($\alpha_0 = 0$)	5.09%	4.16%	3.48%	2.98%	2.58%

Notes: This table shows the annual GDP growth rate in different cases with 1990 fundamentals. GDP growth with 1990 fundamentals is computed by solving the dynamic spatial growth model with constant fundamentals. The second row presents the aggregate GDP growth in the absence of capital accumulation. The third row presents the aggregate growth with capital accumulation and no idea diffusion, obtained by computing the model with $\rho_\ell = 0$ and $\rho_m = 0$. The growth rate without idea flows from people is obtained by computing the model with $\rho_\ell = 0$, and the growth rate without idea flows from goods is obtained by computing the model with $\rho_m = 0$. The last row presents the case with no technology growth.

accumulation to transitional growth remained stable during the 2000s, the relative importance of idea diffusion increased significantly. Initial trade openness and worker mobility facilitated the spread of ideas and enriched China's stock of knowledge across locations. As knowledge diffused and locations expanded their knowledge stocks, individuals contributed better insights to their local areas and to other regions through migration.

Provinces more open to international trade also benefited disproportionately more from global insights as the global stock of knowledge grew over time. As we will describe below, this dynamic is also important for rationalizing the observed heterogeneity in spatial growth in China during this period. Among the two sources of idea diffusion, rows four and five of Table 6 show that diffusion through people contributed slightly more to aggregate growth than diffusion through goods. On the one hand, international trade spread good ideas to all provinces in China, particularly as the global stock of knowledge in 1990 far exceeded China's domestic stock of knowledge. On the other hand, idea flows through people had more nuanced effects on growth. For example, return migration from high-productivity areas stimulated growth in the destination province's knowledge stock, while receiving migrants from low-productivity locations slowed knowledge accumulation. Nonetheless, the best insights from migrants contributed more to a province's stock of knowledge (due to a higher diffusion elasticity) than the best insights from goods.

The last row of Table 6 presents the annual growth rates that would occur in the absence of any technology growth. In this scenario, GDP growth is driven by capital accumulation and internal migration. The small difference between the last and third rows of the table suggest that exogenous technology growth alone doesn't play an important role in aggregate growth in this period.

In the 1990s and 2000s, China implemented reforms related to changes in trade costs and migration frictions that may have also influenced aggregate growth. Specifically, when China joined the World Trade Organization, provinces more exposed to trade may have experienced greater relative development compared to less exposed provinces. Similarly, Hukou reforms likely facilitated the flow of ideas by increasing mobility across provinces.

Using the framework, we quantitatively assess the impact of these reforms on spatial growth in China. We capture the changes in trade costs between China and the rest of the world over the period 1990-2010 using the time variation in bilateral trade shares relative to domestic expenditure shares across provinces, given by $\frac{\hat{\lambda}_{in,t}\hat{\lambda}_{ni,t}}{\hat{\lambda}_{ii,t}\hat{\lambda}_{nn,t}} = \left(\hat{\kappa}_{in,t}\hat{\kappa}_{ni,t}\right)^{-\theta}$.[16] We also explore migration reforms in a simple and similar 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 2010-2015 as $\frac{\hat{\mu}_{in,t}\hat{\mu}_{ni,t}}{\hat{\mu}_{ii,t}\hat{\mu}_{nn,t}} = \left(\hat{m}_{in,t}\hat{m}_{ni,t}\right)^{-\frac{1}{v}}$.[17]

Table 7: Annual GDP Growth Rates: Changes in Fundamentals

	90-95	90-00	90-05	90-10	90-15
With fundamentals in 1990	10.23%	9.86%	9.51%	9.17%	8.86%
Fund. in 1990 & change in trade cost	10.32%	10.04%	9.85%	9.57%	9.26%
Fund. 1990 & change in mig. restrictions	10.23%	9.90%	9.58%	9.26%	8.96%
Fund. in 1990 & change in fundamentals	10.32%	10.07%	9.92%	9.66%	9.36%

Notes: The first row of the table reproduces the growth rate with 1990 fundamentals displayed in Table 6. The second row presents the annual growth rate with 1990 fundamentals and changes in international trade costs. The third row presents the annual growth rate with 1990 fundamentals and changes in migration restrictions. The fourth row presents the annual growth rate with 1990 fundamentals and changes in international trade costs and migration restrictions.

Table 7 presents aggregate growth in China attributable to both the 1990 initial conditions and the changes in international trade costs and migration restrictions during the 1990s and 2000s. The first row of the table reproduces the growth rate based solely on the 1990 initial conditions, while the subsequent three rows incorporate the growth effects of changes in trade costs and migration restrictions. As shown in the table, our second key quantitative finding is that, relative to the initial conditions, changes in trade costs and mobility restrictions during the transition period contributed relatively little to growth—less than one percentage point annually. Another interpretation of these empirical findings is that the reforms that resulted in changes in economic fundamentals (e.g., productivity, trade costs, migration costs) before 1990 were much more significant in setting the economy on a high growth path, relative to those captured by changes in trade and migration costs in the 1990s and 2000s.

We then turn to analyze the implications of our quantitative analysis for the spatial distribution of economic activity and spatial growth in China. In particular, in Table 8 we describe how factor accumulation and idea diffusion influence the spatial distribution of economic activity, measured by real GDP across provinces in China. This analysis considers both the 1990 initial conditions

[16] 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.

[17] 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. Also note that most of the Hukou migration reforms occurred after 2015 (see Fan (2019)).

and changes in trade and migration costs and compares them with the actual spatial distribution of real GDP at different points during the economy's transition in the 1990s and 2000s. In 1990, the correlation is exactly one by construction because, as previously explained, the dynamic-hat algebra aligns the model with the data for that year. Over the remainder of the 1990s and into the 2000s, the table shows that the processes of factor accumulation and idea diffusion in our framework account for most of the spatial heterogeneity in real GDP, with correlations ranging between 0.95 and 0.99.

Table 8: Correlation of Real GDP Across Space

	1990	1995	2000	2005	2010	2015
Fund. in 1990 & change in fundamentals	1	0.994	0.991	0.984	0.972	0.954

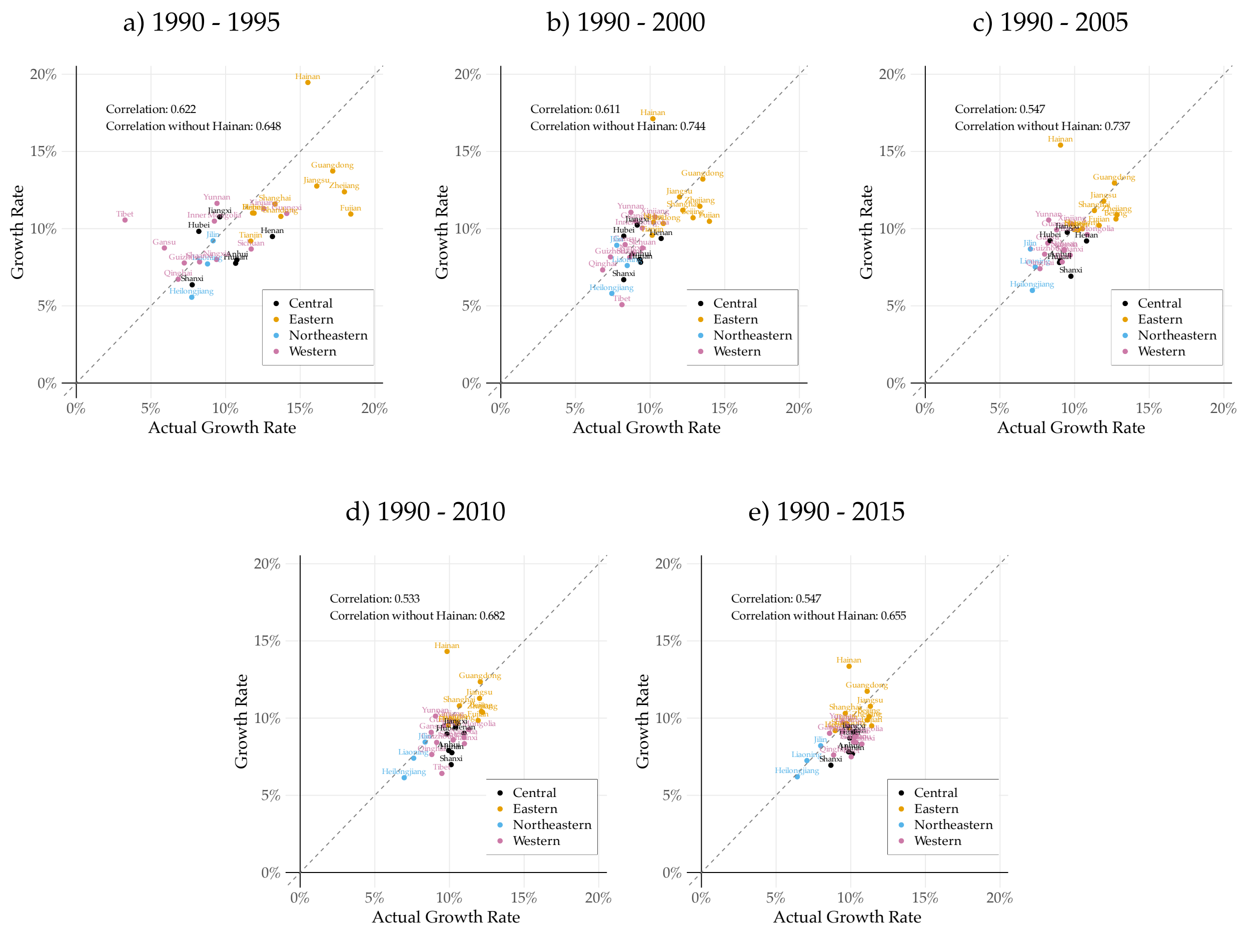
Notes: The table shows the correlations between observed real GDP across provinces in China and the spatial real GDP based on 1990 initial conditions and changes in economic fundamentals thereafter.

In terms of spatial growth, Figure 4 illustrates how factor accumulation and idea diffusion—stemming from the 1990 initial conditions and subsequent changes in fundamentals—shape growth across provinces, and how these results compare with actual growth rates during the economy's transition in the 1990s and 2000s.

Each dot in Figure 4 represents a province, and provinces are color-coded by region. As we can see, capital accumulation and idea diffusion lead to higher growth rates in eastern provinces such as Guangdong, Jiangsu, Beijing, Shanghai, Fujian, and Zhejiang, aligning with the higher observed growth rates in these provinces during this period. Conversely, in western and north-eastern provinces such as Heilongjiang, Liaoning, Gansu, and Tibet, the mechanisms of the spatial model result in lower growth rates, which are also consistent with the relatively lower observed growth rates in these provinces. The only notable outlier is the southern island province of Hainan, where capital accumulation and idea diffusion in our framework predict much higher growth rates than the observed factual rates. Excluding Hainan, the spatial correlation between growth rates derived from factor accumulation and idea diffusion in our framework and the actual growth rates during the economy's transition in the 1990s and 2000s is approximately 0.7.

To gain insights how the previously described spatial growth in China materialized, Figure 5 illustrates the relative importance of idea diffusion versus capital accumulation for growth in each province in China during the 1990s and 2000s. Specifically, the dashed line at one in the figure represents an equal contribution from both margins in a given province. Values above one indicate a relatively greater contribution from idea diffusion, while values below one signify a relatively larger contribution from capital accumulation.

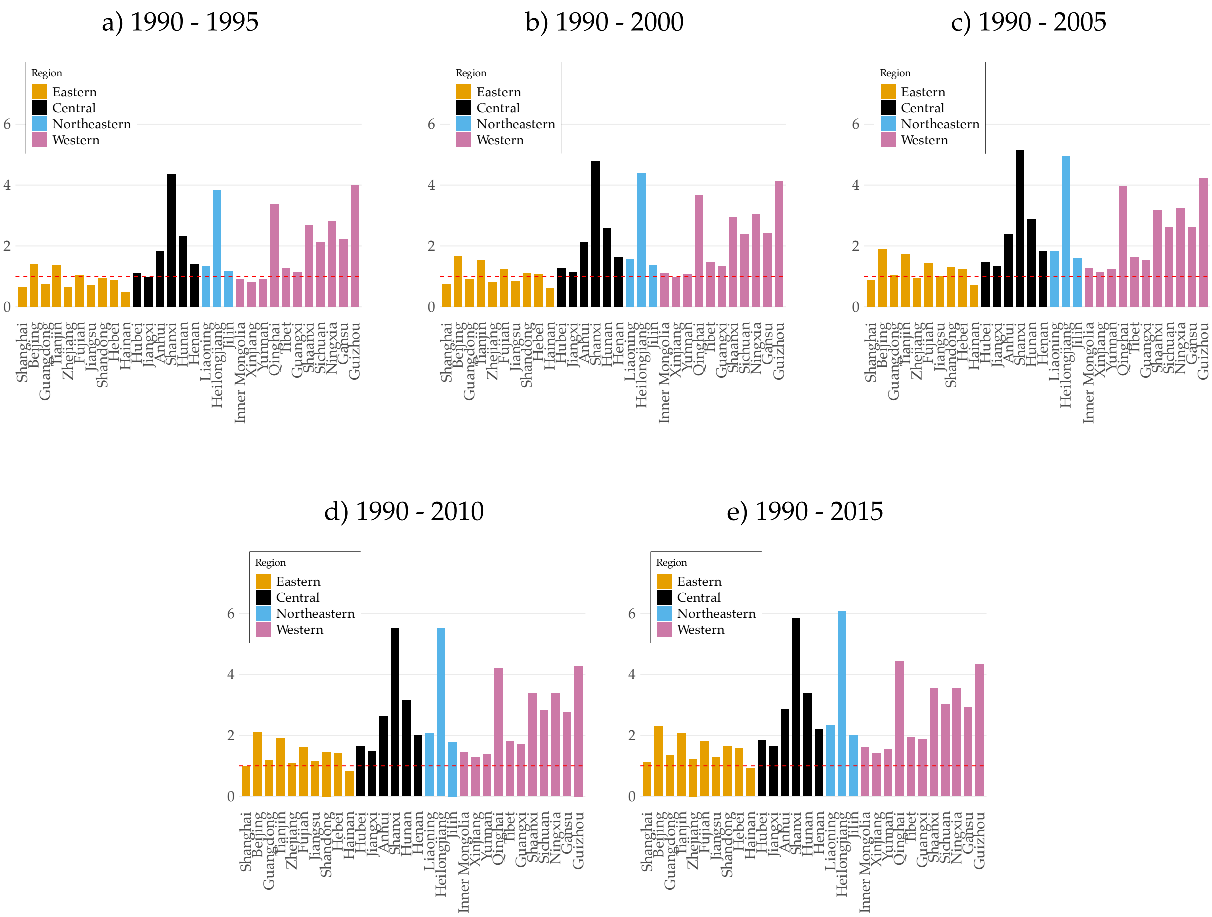
Figure 4: Growth Across Space (1990 Conditions and Changes in Trade and Migration Costs)



Notes: The graphs plot the annual growth rates (since 1990) in model against the observed ones. Annual growth rates in model is obtained by computing the baseline model with 1990 initial conditions and subsequent changes in trade and migration costs. The dashed line represents a 45-degree line.

Interestingly, in the eastern provinces, where growth rates were relatively higher, growth during the 1990s was primarily driven by capital accumulation. As noted earlier, the 1990 initial conditions reflect earlier economic changes, including the establishment of special economic zones, most of which were located in eastern China. These zones spurred capital investment and triggered rapid capital accumulation during the 1990s. By contrast, in the 2000s, the relative contribution of idea diffusion as a source of growth increased. Migrants from other regions brought better insights compared to the previous decade, reflecting the enhanced stock of knowledge in the western and northern regions resulting from idea diffusion during the 1990s. In fact, in the western provinces, where growth rates were relatively lower, growth was primarily driven by idea diffusion, as migrants from coastal, high-productivity areas carried good-quality insights and spurred the local stock of knowledge in the western provinces. This knowledge diffusion benefited the eastern provinces receiving migrants in the 2000s.

Figure 5: Diffusion versus Capital Accumulation



Notes: The graphs show the ratio of annual growth rates (since 1990) in two scenarios: with only idea diffusion versus with only capital accumulation. A ratio above 1 (dashed red line) indicates that idea diffusion is more important than capital accumulation in the province. △ indicates values exceeding 9, ▽ indicates values below 0. With only capital accumulation, the model predicts a negative growth rate for figure (b) (-0.99%) and figure (d) (-0.17%) for Tibet.

Importantly, these growth effects of idea diffusion through trade and migration are crucial for understanding spatial growth during China's economic transition in the 1990s and 2000s. To see this, we consider a version of our model with capital accumulation and exogenous technological growth but with no idea diffusion. We examine the spatial heterogeneity in growth rates predicted by this alternative model against data, as shown in Figure 6. The correlation in the first part of the 1990s is similar to that in Figure 4, as growth in high-growth areas was primarily driven by capital accumulation, as described previously. However, during the transition, without the contribution of idea diffusion, the correlation between spatial growth and actual growth drops to approximately 0.36, as idea diffusion became a more important contributor of growth across provinces, and especially in high-growth areas. As a result, over the period 1990 to 2015, removing knowledge diffusion from the model causes the correlation between the model-implied spatial growth and the observed growth in the data to decline by almost half. This finding reinforces our previous message about the importance of both factor accumulation and idea flows through trade and migration as essential mechanisms for understanding the heterogeneity in spatial growth in China during its transition in the 1990s and 2000s. As discussed above, we find that the conditions of 1990 are the primary drivers of China's growth path compared to changes in trade and migration costs. Therefore, we relegate to Appendix H the results on spatial growth driven solely by initial conditions, as they closely resemble those resulting from initial conditions and changes in the fundamentals described in this section.

Figure 6: Annual Growth Rates Without Idea Flows

Notes: The graphs plot the annual growth rates (since 1990) in model against data. Annual growth rates in model is obtained by computing the model with 1990 initial conditions and subsequent changes in trade and migration costs, but we shut down the diffusion channel by setting ($\rho_\ell = \rho_m = 0$). In panel (b), the data point for Tibet (8.1%, -1.0%) is out of range and not shown. In panel (d), the data point for Tibet (9.5%, -0.2%) is out of range and not shown. The dashed line represents a 45-degree line.

6 Concluding Remarks

Using various instruments and province-level data from China, we provide causal evidence that migrants from more developed regions contribute more significantly to the local knowledge stock than those from less developed regions. Additionally, we find that provinces with greater openness to imports (instrumented by historical levels of openness) experience higher growth. Motivated by this empirical evidence, we develop a dynamic spatial growth model to study, understand, and quantify the impact of spatial growth on aggregate economic activity. In our model, internal migration and trade drive spatial growth, as producers and migrants share ideas with one another. The flow of ideas across space and time serves as the primary mechanism fueling spatial growth. We show how to take the model to data, enabling quantitative analysis without assuming that the economy is initially on its balanced growth path, and how to perform counterfactuals along the transition path.

As an application, we study the importance of trade and migration as engines of growth for the Chinese economy after 1990. Initial conditions, idea diffusion through international trade and internal migration, and capital accumulation, play a crucial role in shaping spatial development and aggregate growth during the 1990s and 2000s in China. Changes in economic fundamentals driven by trade openness and migration restrictions contribute less to aggregate growth and the observed heterogeneous spatial development across the country.

The framework can be used to explore a wide range of questions related to spatial and aggregate growth. This opens up interesting avenues for research, such as the study of optimal placed-based, migration, industrial, and other policies. For instance, the model assumes that firms and workers do not consider the impact of their insights on the stock of knowledge across locations, leading to externalities. In such a context, there might be a role for regional policies to correct these externalities. Furthermore, the framework can also be applied to understand the growth patterns of cities and towns, and gain a better understanding of issues such as optimal infrastructure investment across different regions, and the aggregate growth implications of place-based policies. Also, in the study we have abstracted from a sectoral analysis, where insights from producers and people in the same industry might be better than those from other industries. The framework is tractable enough to accommodate this extension.

References

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

Alder, S., L. Shao, and F. Zilibotti (2016): "Economic reforms and industrial policy in a panel of Chinese cities," *Journal of Economic Growth*, 21, 305–349.

Alvarez, F. (2017): "Capital Accumulation and International Trade," *Journal of Monetary Economics*, 91, 1–18.

Alvarez, F. E., F. J. Buera, and J. Robert E. Lucas (2013): "Idea Flows, Economic Growth, and Trade," NBER Working Papers 19667, National Bureau of Economic Research, Inc.

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.

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.

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.

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.

Brandt, L., G. Kambourov, and K. Storesletten (2023): "Barriers to entry and regional economic growth in China," .

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," *National Bureau of Economic Research*.

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.

CARD, D. (2001): "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 19, 22–64.

CHAN, K. (2010): "The household registration system and migrant labor in China: notes on a debate," *Population and development review*, 36, 357–364.

CHEN, H., B. LI, AND X. ZHU (2024): "Bottom-Up Institutional Change and Growth in China," .

CHEN, W., X. CHEN, C.-T. HSIEH, AND Z. SONG (2019): "A Forensic Examination of China's National Accounts," NBER Working Papers 25754, NBER.

CHEREMUKHIN, A., M. GOLOSOV, S. GURIEV, AND A. TSYVINSKI (2024): "The Political Development Cycle: The Right and the Left in People's Republic of China from 1953," *American Economic Review*, 114, 1107–1139.

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

——— (2001): "Technology, Trade, and Growth: A Unified Framework," *European Economic Review*, 45, 742–755.

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

EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An anatomy of international trade: Evidence from French firms," *Econometrica*, 79, 1453–1498.

EATON, J., S. KORTUM, B. NEIMAN, AND J. ROMALIS (2016): "Trade and the Global Recession," *American Economic Review*, 106, 3401–38.

EATON, J., S. S. KORTUM, AND F. KRAMARZ (2025): "Firm-to-firm trade: Imports, exports, and the labor market," .

EGGER, D., B. FABER, M. LI, AND W. LIN (2024): "Rural-Urban Migration and Market Integration," *Working Paper*.

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.

FERNÁNDEZ-VILLAVERDE, J., L. E. OHANIAN, AND W. YAO (2023): "The neoclassical growth of China," *National Bureau of Economic Research*.

GUO, Q. G. S. N., B. L. H. Q. SHI, AND X. Z. HKU (2022): "Migration Costs, Sorting, and the Agricultural Productivity Gap," .

HAO, T., R. SUN, T. TOMBE, AND X. ZHU (2020): "The effect of migration policy on growth, structural change, and regional inequality in China," *Journal of Monetary Economics*, 113, 112–134.

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.

JONES, C. I. (2022): "The Past and Future of Economic Growth: A Semi-Endogenous Perspective," *Annual Review of Economics*, 14, 125–152.

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 (2023): "Dynamic Spatial General Equilibrium," *Econometrica*, 91, 385–424.

KÖNIG, M., K. STORESLETTEN, Z. SONG, AND F. ZILIBOTTI (2022): "From imitation to innovation: Where is all that Chinese R&D going?" *Econometrica*, 90, 1615–1654.

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.

LIND, N. AND N. RAMONDO (2023): "Global innovation and knowledge diffusion," *American Economic Review: Insights*, 5, 494–510.

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

LU, F., W. SUN, AND J. WU (2023): "Special Economic Zones and Human Capital Investment: 30 Years of Evidence from China," *American Economic Journal: Economic Policy*, 15, 35–64.

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

LUCAS, R. E. AND B. MOLL (2014): "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 122, 1–51.

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.

PELLEGRINA, H. S. AND S. SOTELO (2024): "Migration, Specialization, and Trade: Evidence from Brazil's March to the West," *Journal of Political Economy, forthcoming*.

PERLA, J. AND C. TONETTI (2014): "Equilibrium Imitation and Growth," *Journal of Political Economy*, 122, 52–76.

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

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.

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.

SONG, Z., K. STORESLETTEN, AND F. ZILIBOTTI (2011): "Growing Like China," *American Economic Review*, 101, 196–233.

SONG, Z. AND W. XIONG (2023): "The Mandarin model of growth," *National Bureau of Economic Research*.

STORESLETTEN, K. AND F. ZILIBOTTI (2014): "China's great convergence and beyond," *Annu. Rev. Econ.*, 6, 333–362.

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

XIANG, W. (2023): "Clean Growth and Environmental Policies in the Global Economy," *Available at SSRN*.

XU, C. (2011): "The fundamental institutions of China's reforms and development," *Journal of economic literature*, 49, 1076–1151.

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.

ZHU, X. (2012): "Understanding China's growth: Past, present, and future," *Journal of Economic Perspectives*, 26, 103–124.

——— (2024): *China's Productivity Challenge*, University of Toronto, Department of Economics.

ZILIBOTTI, F. (2017): "Growing and slowing down like China," *Journal of the European Economic Association*, 15, 943–988.

Appendix: Mechanics of Spatial Growth

The appendix includes detailed theoretical derivations and proofs, additional quantitative results, and detailed data descriptions described in the paper.

A Data Sources and Empirical Moments

In this section of the appendix, we provide a detailed description of the data sources and construction used in the empirical and quantitative analysis.

A.1 Data and Measurements

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.

Province-Level Data and National Accounts. We obtain GDP, employment, export, and import data from the China Compendium of Statistics 1949–2008[1], National Bureau of Statistics online data 1995–2015, as well as Provincial Statistical Yearbooks 2009–2015. In particular, we obtain GDP, export, and import data after 1995 from the National Bureau of Statistics online database, while data prior to 1995 come from the Compendium. We obtain employment data after 2008 from Provincial Statistical Yearbooks and data prior to 2008 from the Compendium. We obtain patent data from National Statistical Yearbooks 1985–2015.[2] The China Compendium of Statistics consists of three main parts.The first part contains national-level data compiled by the National Bureau of Statistics. The second part includes provincial, regional, and municipal data compiled by local statistical bureaus for provinces, autonomous regions, and municipalities under the jurisdiction of the central government. The third part provides data for the Special Administrative Regions of Hong Kong and Macao, 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.

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

[1] The digitized data is available at https://data.oversea.cnki.net/yearBook/single?id=N2010042091&pinyinCode=YXZLL

[2] The digitized data can be extracted from the China statistical yearbooks available at https://www.stats.gov.cn/english/Statisticaldata/yearbook/ and https://data.stats.gov.cn/easyquery.htm?cn=C01

(2006), and Chen, Chen, Hsieh, and Song (2019)). Given the incentive for local governments to overstate local GDP figures and other measurement discrepancies, the National Bureau of Statistics adjusts data reported by local governments to calculate national GDP using independent data sources. As a result, the reported aggregate national GDP is generally lower than the sum of the province-level GDPs. To address this discrepancy, we scale down the province-level GDP uniformly across all provinces to align with the reported national GDP. We apply the same adjustment strategy to province-level employment, export, and import data to match their respective national aggregates.

Second, we account for the evolving status of Chongqing. Prior to 1997, Chongqing was not designated as a municipality directly under the central government. Since this paper focuses on China's economic rise during the 1990s—a period when Chongqing remained part of Sichuan—we treat Chongqing and Sichuan as a single integrated region, referred to as Sichuan-Chongqing, or simply Sichuan in quantitative result when the context is clear, throughout our analysis. Relevant variables for the two regions are aggregated accordingly.

Third, for some provinces, measurement units are inconsistent with those used for national aggregates. For instance, Guangdong Province's export and import data are inaccurately reported by the local statistical bureau in units of 100 million Chinese Yuan, despite being labeled as 10,000 Chinese Yuan. We carefully identify and correct such inconsistencies in the data.

International Trade Data. We obtain export and import data between Chinese provinces and the rest of the world from the China Compendium of Statistics, 1949–2008 and the National Bureau of Statistics online data 1990–2015. This data includes export and import values based on firm locations for each province. However, the National Bureau of Statistics online database does not provide province-level trade data for the early 1990s. Therefore, we rely on the China Compendium of Statistics, 1949–2008 to obtain province-level international trade data for the 1980s and 1990s.

One potential concern about the international trade data from the 1990s is that two government agencies reported export and import data in China over the last four decades: the Ministry of Trade and Customs. The names of the Ministry of Trade's provincial counterparts varied; for instance, in Shanghai, it was known as the Foreign Trade and Economic Cooperation Commission, which oversaw export and import data collection in Shanghai before 1999. Historically, during the 1980s and 1990s, trade data for many provinces were reported by the Ministry of Trade, whereas, in the past two decades, nearly all provinces have had their trade data reported by Customs.

The province-level international trade data from the China Compendium of Statistics are consistent with those from National Bureau of Statistics for the year 2000 onward. For these years, we rely on the Yearbook data, as it explicitly states the definitions and methodologies used to report province-level international trade flows between provinces and the rest of the world. However, for the year 1995, due to mixed sources of province-level export and import data, the trade data for some provinces in the Compendium show discrepancies compared to the National Bureau of Statistics. Specifically, for 1995, the Compendium data for provinces such as Beijing, Fujian, Guangdong, Guizhou, Hebei, Heilongjiang, Henan, Hunan, Jiangxi, Liaoning, Shaanxi, Shandong, Shanxi, Sichuan, and Zhejiang align with the Yearbook data, with discrepancies smaller than 3%, while the discrepancies are larger for the remaining provinces.

We use the deviations between the Compendium and National Bureau of Statistics data in 1995 to adjust the data for 1985 and 1990. A potential concern is that the data sources for 1985, 1990, and 1995 in the Compendium might differ, making this adjustment less reliable. Reassuringly, the international trade data for all provinces except Gansu are reported by the same government agency across these years.[3]

GDP Data. We use the GDP deflator from the World Development Indicators compiled by the World Bank, to compute the real GDP of each province at 1990 prices. We rely on the Penn World Table 10.0 (PWT 10.0) to construct data for the rest of the world. The PWT 10.0 reports real GDP at constant 2017 national prices (rgdpna) and employment (emp). 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. We express GDP, exports, and imports in 100 million USD, while employment is measured in units of 10,000 people.

Capital Stock. We follow Shan (2008) to estimate province-level capital stock from 1952 to 2015. 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 and National Bureau 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, 2000, 2005, 2010, and 2015 to match the average gross capital formation (percentage of GDP) in 1985-1990, 1990-1995, 1995-2000, 2000-2005, 2005-2010, and 2010-2015 respectively. Afterward, by excluding the capital stock of China, we obtain the capital stock for the rest of the world.

[3]The international trade data for Gansu in 1985 and 1990 were reported by the Ministry of Foreign Trade, while the data for 1995 were sourced from Customs Statistics. Gansu represents a very small fraction of China's international trade; in 1990, its exports accounted for 0.24% of the national total, while its imports made up 0.18% of the national total.

Input 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.

Gross Migration Flows. The People's Republic of China (PRC) conducted its first population census in 1953, with subsequent censuses held in 1964, 1982, 1990, 2000, 2010, and 2020. Between these full censuses, inter-census surveys, or mini-censuses, were conducted in 1987, 1995, 2005, and 2015, each sampling 1% of the total population. Migration data was first collected during the 1987 mini-census. This study uses migration data from Chinese population censuses and surveys conducted in 1987, 1990, 1995, 2000, 2005, 2010, and 2015. Data from 1987, 1995, 2005, and 2015 are drawn from 1% sample surveys, while data from 1990, 2000, and 2010 are obtained from full population censuses.

We focus on inter-province migration flows for several reasons. First, all population censuses and mini-censuses report inter-province migration flow matrices, providing a consistent and comparable measure of spatial mobility across time from 1990 to 2015. Second, obtaining migration flows at a more granular geographic level typically requires using micro-samples from population or mini-censuses. However, inter-city migration flow data is unavailable in the micro-sample of the 1990 population census, and there is no access to the micro-sample of the 1995 mini-census. While the micro-sample of the 2000 population census includes province-to-city migration information, the 1990s represent a crucial period of analysis for this paper. Therefore, we focus exclusively on inter-province migration flows.

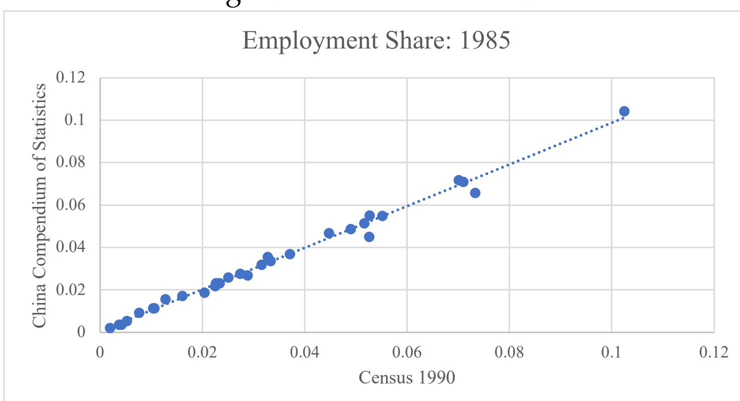
For inter-provincial migration between 1995 and 1990, we directly obtain the data from the 1990 Population Census, where respondents reported both their current province of residence and the province they lived in five years earlier. In the 2000 and 2010 Population Censuses, there were both short and long forms. All individuals completed the short form, while 10% of the population filled out the long form. The long form included migration-related questions for respondents aged five years and older. We extracted the inter-provincial migration flow matrix from responses on the long form, then scaled the migration flow and stayers values proportionally to match the total population aged five and older. For the 1987, 1995, 2005, and 2015 surveys, we obtain inter-provincial migration matrices and similarly scale the migration flows and stayers values to align with the total population aged five and older.

Hukou Status We obtain the total number of people holding Hukou from different provinces within a given province (Hukou stocks) using data from the Chinese population censuses and surveys conducted in 1987, 1990, 1995, 2000, 2005, 2010, and 2015. For the Hukou stocks in 1990, the data is directly sourced from the 1990 Population Census. For the 2000 and 2010 censuses, the data is derived from the short-form respondents. For the years 1987, 1995, 2005, and 2015, we obtain the Hukou stocks matrices and scale the values to align with the total population.

Migration Flows Conditional on Hukou Status In the quantitative exercise, we use inter-provincial migration flows conditional on Hukou status between 1985 and 1990. To do so, we use IPUMS data to construct a migration flow matrix by Hukou status. Specifically, we rely on the 1% sample from the 1990 Population Census provided by IPUMS to calculate migration flows for the period 1985-1990.

To construct the migration flows for 1985-1990, we proceed as follows. Using the 1% sample from the 1990 Census, we focus on the working-age population (15-64) and retain only those who were actively employed in 1990. We apply a weight to each province to ensure that the provincial employment shares in our sample match the shares reported in the China Compendium of Statistics.

Figure A.1: Data validation



Notes: This figure shows each province's employment share in the national total employment. The horizontal axis represents the employment share based on the 1% micro sample from the 1990 population census, while the vertical axis represents the employment share based on the 1985 data from the China Compendium of Statistics 1949–2008.

For each individual, we determine their Hukou location as follows. In the 1990 Census, respondents were asked about the status and nature of their Hukou registration. If a respondent chose "(1) residing and registered here," we assign their 1990 location as their Hukou place. If they chose "(2) residing here for over 1 year, but registered elsewhere," "(3) living here for less than 1 year but absent from the registration place for over 1 year," or "(4) living here with unsettled registration," we assign their Hukou place as the province they lived in during 1985.

For respondents whose Hukou status falls under categories (2)-(4) but who remained in the same province in both 1985 and 1990 (i.e., "stayers"), we assign their Hukou location based on the following procedure. We first identify a sample of migrants who changed their residence province between 1985 and 1990, as recorded in the data. For each destination province, we calculate the share of migrants coming from each origin province, then assign the Hukou place to the stayer according to this distribution.[4] For each Hukou location (at the provincial level), we construct a

[4] 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

five-year migration flow matrix from origin to destination provinces. By combining this migration matrix with the 1990 Census data, we can verify whether the employment share of each province, relative to the nationwide total, aligns with the figures reported in the China Compendium of Statistics (see Figure A.1).

Measures for Knowledge We measure knowledge using patent flow and accumulated patents (patent stock) per capita. Province-level data on patents applications and patents granted were obtained from the National Statistical Yearbooks compiled by National Bureau of Statistics for the years 1985 to 2015. There are three types of patents: innovation, utility, and design. For each type, the yearbook reports the number of patent applications and patents granted in a given year. To proxy the measure of knowledge stock, we calculate the cumulative number of patents granted of all three types at the provincial level for each year, starting from 1985. Using this data, we compile province-level knowledge stock for the years 1985, 1990, 1995, 2000, 2005, 2010, and 2015, enabling us to calculate changes in knowledge stock every five years from 1985 to 2015. For data on patents granted in the rest of the world, we use information from Google Patent spanning 1985–2015, following the methodology of Liu and Ma (2021). To represent global patent stocks, we use data from Google Patents data, which serves as a reliable counterpart to PATSTAT.

Our province-level patent data are obtained from the China Statistical Yearbooks compiled by the National Bureau of Statistics, with the patents themselves sourced from China's State Intellectual Property Office (SIPO). Figure A.2 shows that patents are highly correlated with measured TFP and knowledge stock (A) at the provincial level. This finding is related to the one presented in König et al. (2022); where using the same data source find a positive correlation between TFP, R&D, and patenting.

Stock of Knowledge in the Model We derive the model inversion used to estimate the initial local stock of knowledge. We start from the domestic expenditure $\lambda_{nn,0} = A_{n,0}\left(\frac{x_{n,0}}{P_{n,0}/T}\right)^{-\theta}$. Using this equation, we obtain

$$A_{n,0} = \left(\frac{B\left(w_{n,0}^{\xi} r_{n,0}^{1-\xi}\right)^{\gamma} P_{n,0}^{1-\gamma}}{P_{n,0}/T}\right)^{\theta} \lambda_{nn,0}.$$

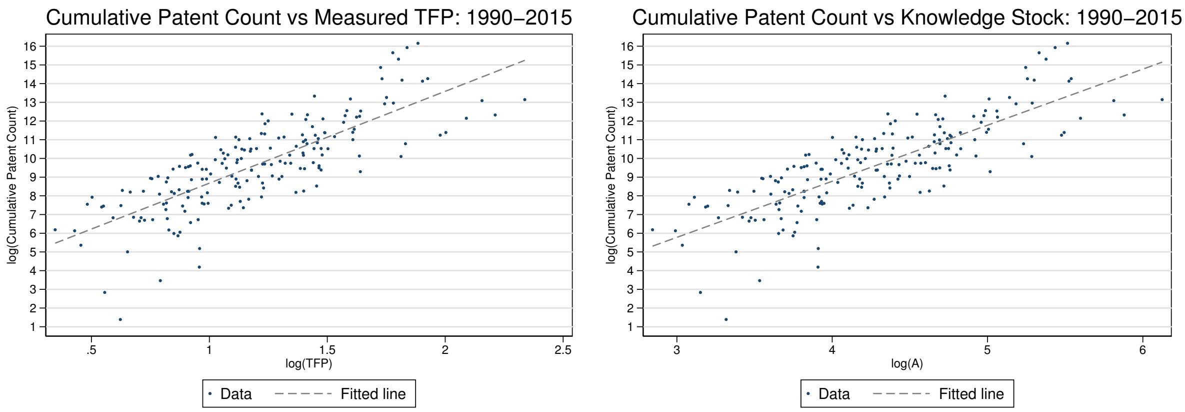
Using the first-order condition of the firm's problem, $\frac{w_{n0}L_{n0}}{r_{n0}K_{n0}} = \frac{\xi}{1-\xi}$, we obtain

$$A_{n,0} = (BT)^{\theta}\left(\frac{1-\xi}{\xi}\right)^{(1-\xi)\gamma\theta}\left(\frac{\frac{w_{n,0}L_{n0}}{P_{n,0}}}{(K_{n,0})^{1-\xi}(L_{n,0})^{\xi}}\right)^{\gamma\theta} \lambda_{nn,0}.$$

Finally, using the fact that $w_{n,0}L_{n0} = \xi\left(w_{n,0}L_{n,0} + r_{n,0}K_{n,0}\right)$, we find that the initial stock of

the period 1985 to 1990.

Figure A.2: Patent Count, Measured TFP, and Knowledge Stock



Notes: The left panel shows the correlation between cumulative patent counts and measured Total Factor Productivity (TFP), while the right panel shows the correlation between cumulative patent counts and knowledge stock. Each dot represents a province in a specific year. The patent data are sourced from the National Statistical Yearbooks, compiled by the National Bureau of Statistics. Measured TFP and knowledge stock are calculated by the authors.

knowledge across locations is given by

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

where $\Upsilon = (BT)^{\theta} (1-\xi)^{(1-\xi)\gamma\theta} (\xi)^{\xi\gamma\theta} = \Gamma \left(1 - \frac{\eta-1}{\theta}\right)^{\frac{\theta}{1-\eta}} \left[\gamma^{-\gamma} (1-\gamma)^{\gamma-1}\right]^{\theta}$.

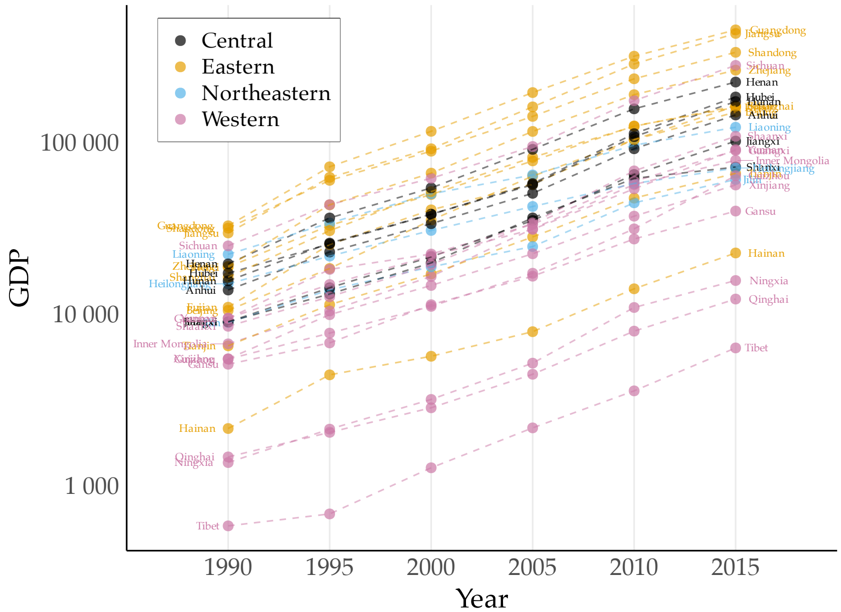
A.2 Descriptive Patterns

We begin by categorizing provinces into four regions—eastern, central, western, and northeastern—based on classifications from the National Bureau of Statistics and examine inter-regional migration flows. Next, we present the level of real GDP across provinces and the annual growth rates of real GDP for each province. One takeaway is that provinces experienced uneven economic growth in this period. We then focus on migration patterns from major origin provinces and to key destination provinces, showing that both origins and destinations exhibit significant geographic diversity.

Real GDP Levels and Growth

Figure A.3 shows the dispersion in levels of real GDP across provinces in China for different time periods.

Figure A.3: GDP by Provinces



Notes: The figure presents the level of real GDP for each province in China over different time periods.

Table A.1 shows the annual growth rates of real GDP in each province in data.

Cross-region Migration

People migrated across all regions and not necessarily to the geographically less distant region. In addition, not all migration is to coastal areas (high-growth places), and coastal areas received migrants from all locations, not only from the geographically closer provinces. To illustrate these migration patterns, we categorize provinces into four regions as described above, and exam inter-regional migration flows. We focus on migration patterns from major origin provinces and to key destination provinces, showing that both origins and destinations exhibit significant geographic diversity. Figure A.4 illustrates cross-region migration flows in China. The figure highlights that a significant number of migrants move from the central, eastern, and western regions to the eastern region. While the eastern region is a major destination, it also serves as an important source of migrants relocating to the central, eastern, and western regions.

In-migration

Inter-province migration gained momentum in the 1980s, with a significant surge in the 1990s. All provinces receive significant number of migrants. The primary destination provinces were Guangdong, Jiangsu, Beijing, Shanghai, and, later, Zhejiang (Figure A.5).

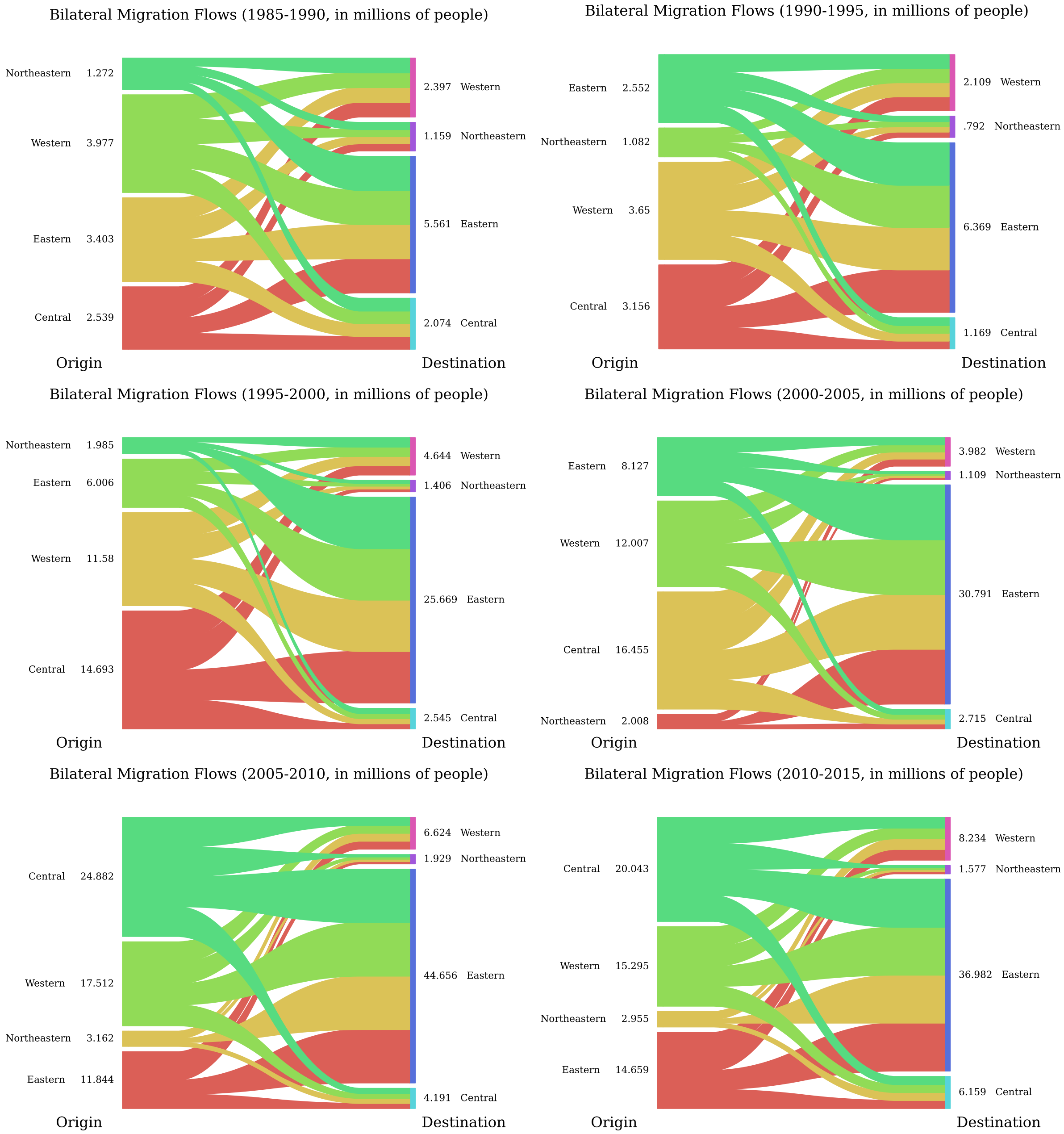
Among migrants to Guangdong, the top destination province, the sources of origin span a wide geographic range, from South to North and East to West. In the late 1980s and early 1990s,

Table A.1: Annual Growth Rates of Real GDP

Province	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015
Beijing	11.91%	13.86%	12.56%	10.28%	7.71%
Tianjin	11.68%	8.61%	10.45%	11.00%	6.87%
Hebei	11.80%	8.68%	9.44%	9.84%	5.12%
Shanxi	7.76%	8.71%	12.85%	11.21%	3.08%
Inner Mongolia	9.25%	9.71%	13.65%	12.64%	6.69%
Liaoning	8.79%	8.15%	5.19%	8.32%	4.95%
Jilin	9.16%	6.39%	5.60%	12.46%	6.47%
Heilongjiang	7.73%	7.14%	6.65%	6.36%	4.26%
Shanghai	13.31%	11.08%	9.62%	8.70%	5.61%
Jiangsu	16.10%	7.99%	11.90%	12.22%	8.55%
Zhejiang	17.96%	8.89%	11.85%	10.38%	6.81%
Anhui	10.75%	7.90%	8.51%	12.71%	9.50%
Fujian	18.38%	9.72%	7.14%	12.75%	9.37%
Jiangxi	9.61%	8.67%	10.26%	13.16%	9.38%
Shandong	13.70%	8.14%	9.79%	10.63%	7.37%
Henan	13.14%	8.40%	10.92%	11.50%	7.46%
Hubei	8.20%	8.27%	8.61%	14.34%	10.36%
Hunan	10.67%	8.07%	8.24%	13.76%	9.91%
Guangdong	17.17%	10.00%	10.97%	10.27%	7.32%
Guangxi	14.09%	4.21%	8.30%	12.23%	8.66%
Hainan	15.51%	5.12%	6.83%	12.21%	10.10%
Sichuan	11.71%	7.33%	8.92%	13.11%	9.93%
Guizhou	7.23%	7.46%	9.30%	12.66%	15.35%
Yunnan	9.42%	8.01%	7.37%	11.50%	11.10%
Tibet	3.26%	13.20%	11.32%	10.45%	12.23%
Shaanxi	8.25%	9.02%	11.87%	14.98%	9.74%
Gansu	5.90%	10.81%	7.96%	10.51%	7.79%
Qinghai	6.82%	6.82%	9.42%	12.29%	9.00%
Ningxia	9.39%	8.31%	10.24%	16.13%	7.52%
Xinjiang	12.56%	8.17%	8.89%	10.63%	8.73%

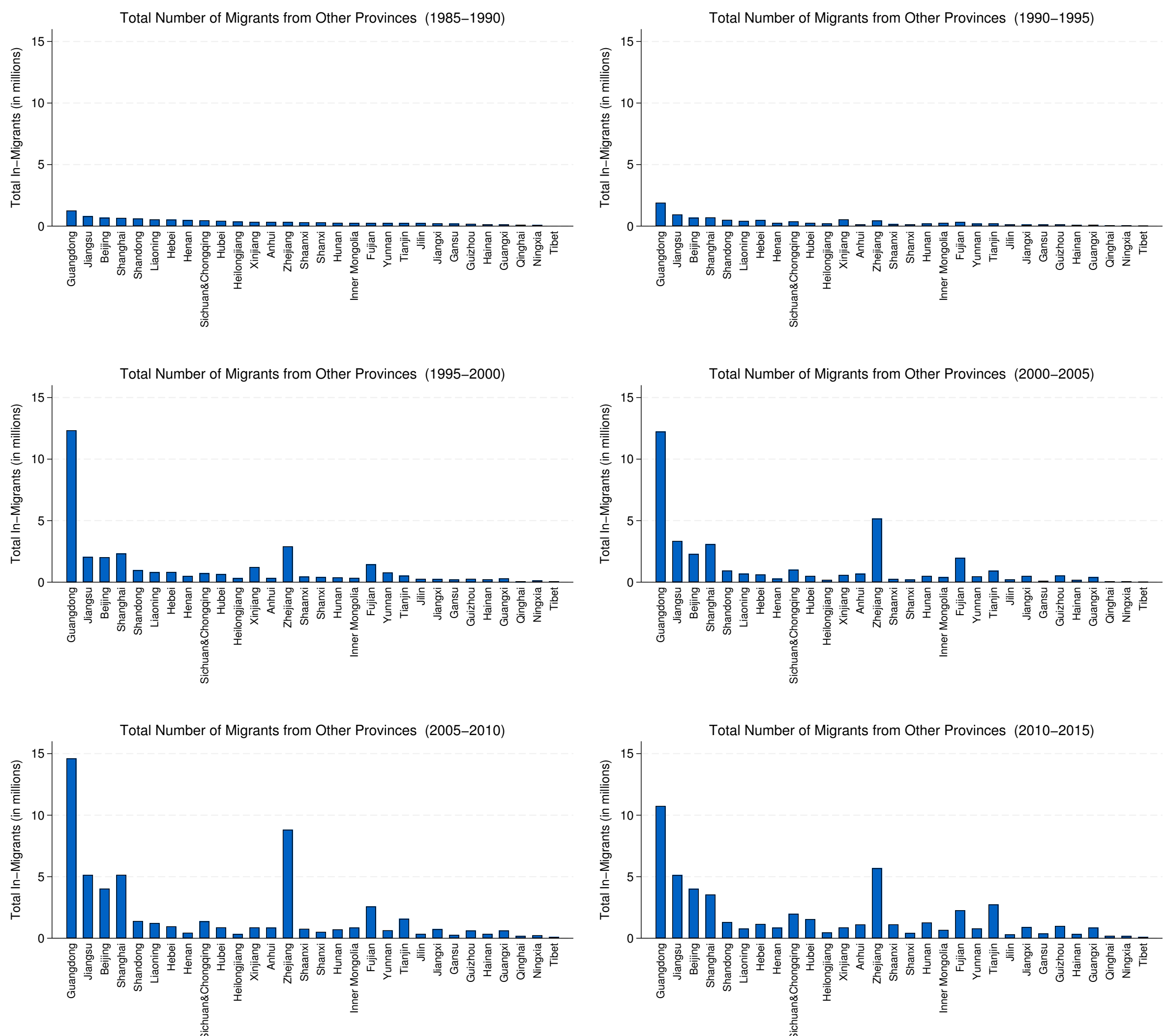
Notes: GDP growth in data. GDP after 1995 are from the National Bureau of Statistics online database, while data prior to 1995 come from the Cempendium.

Figure A.4: Cross-region Bilateral Migration Flows



Notes: These figures illustrate cross-region migration flows within the country every five years from 1985 to 2015, measured in millions of people. The Eastern regions include Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The Central regions consist of Shanxi, Anhui, Jiangxi, Henan, Hubei, and Hunan. The Western regions encompass Inner Mongolia, Guangxi, Sichuan, Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang. Lastly, the Northeastern regions include Liaoning, Jilin, and Heilongjiang.The data are sourced from population censuses and mini-censuses conducted in various years.

Figure A.5: Distribution of In-migration across Provinces



Notes: These figures show the distribution of in-migration across provinces every five years from 1985 to 2015, measured in millions of people. The data are sourced from population censuses and mini-censuses conducted in various years.

Sichuan, Chongqing, Guangxi, and Hunan were the primary sources. By the late 1990s and early 2000s, Henan, Hubei, and Jiangxi also emerged as major contributors, each sending over 800,000 people to Guangdong within a five-year period (Figure A.6).

Notably, Guangdong later became significant sources of migrants relocating to other regions, as we will illustrate next.

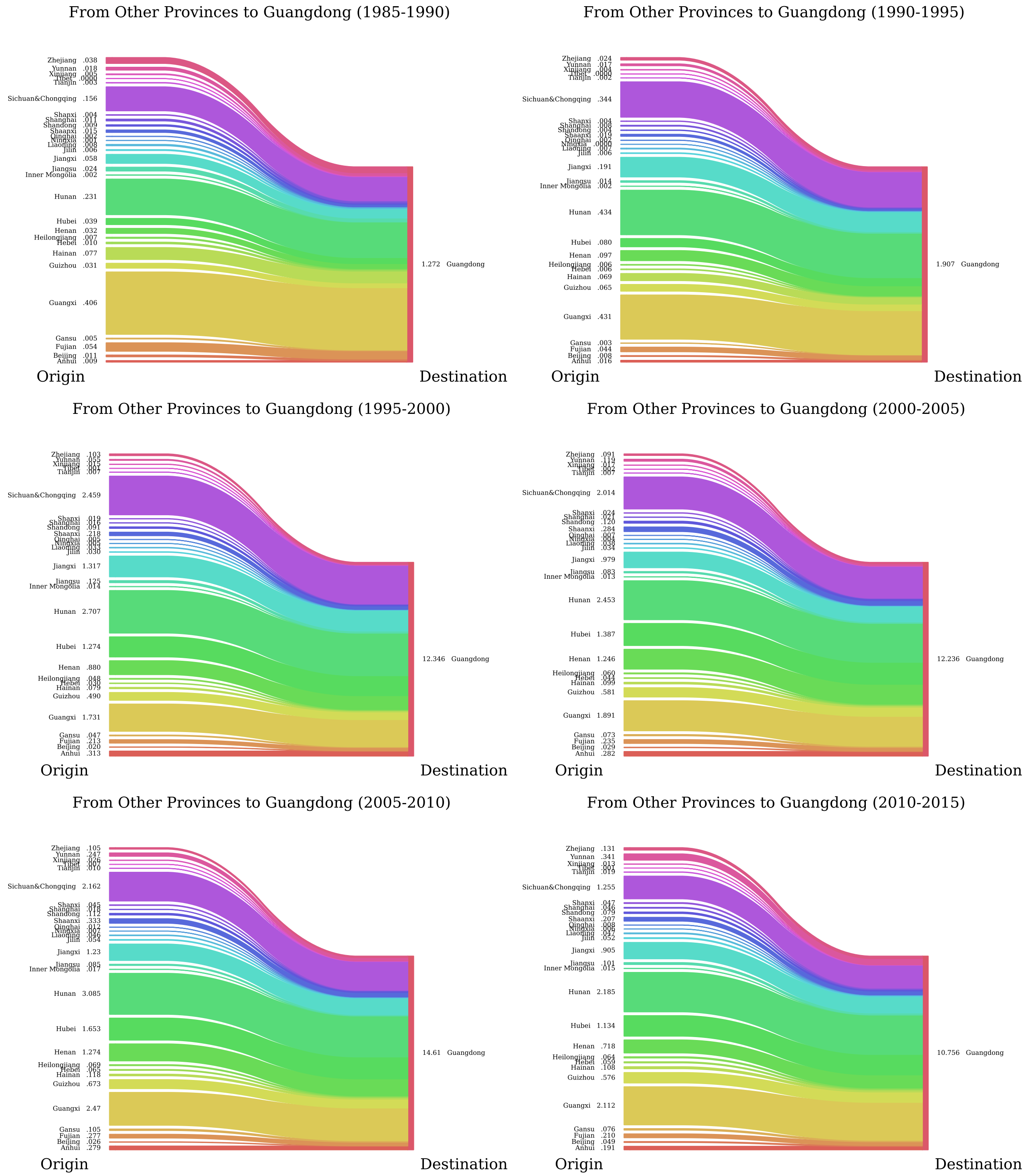
Out-migration

All provinces experience a notable number of outmigrants (Figure A.7). Leading the list are Sichuan, Chongqing, Henan, Hunan, and Anhui. Figures A.8 and A.9 illustrate that the destination provinces for migrants from these top origin provinces are also geographically diverse.

Figure A.8 illustrates that the destination provinces of migrants from the top origin provinces are also geographically diverse. Sichuan and Chongqing, located in the Midwest of mainland China, have seen the majority of their migrants move to Guangdong in the Southwest. By the late 1990s, Zhejiang in the East emerged as the second-largest destination, followed by Fujian in the Southeast, Xinjiang in the Northwest, and Yunnan in the South.

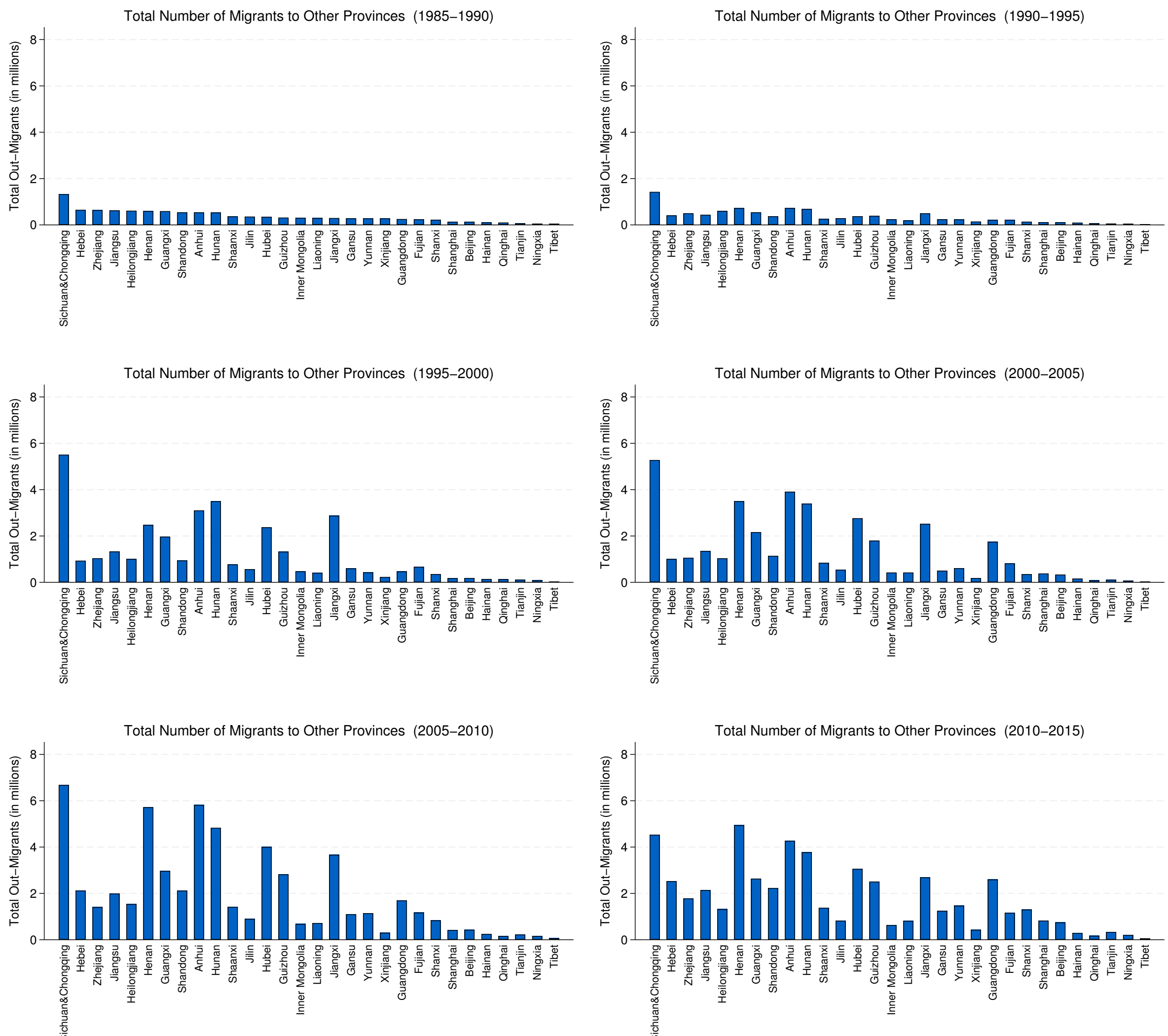
Since we combine Sichuan and Chongqing into one geographic unit, the fact that they send out the most number of outmigrants might be driven by the size of population. Hence, we also report the the Sankey diagram in Figure A.9 for origin provinces that send the second largest number of migrants to other provinces. Notably, Henan surpassed Sichuang and Chongqing combined and became the largest origin province between 2010 and 2015.

Figure A.6: Distribution of Source Provinces Among Migrants to Guangdong



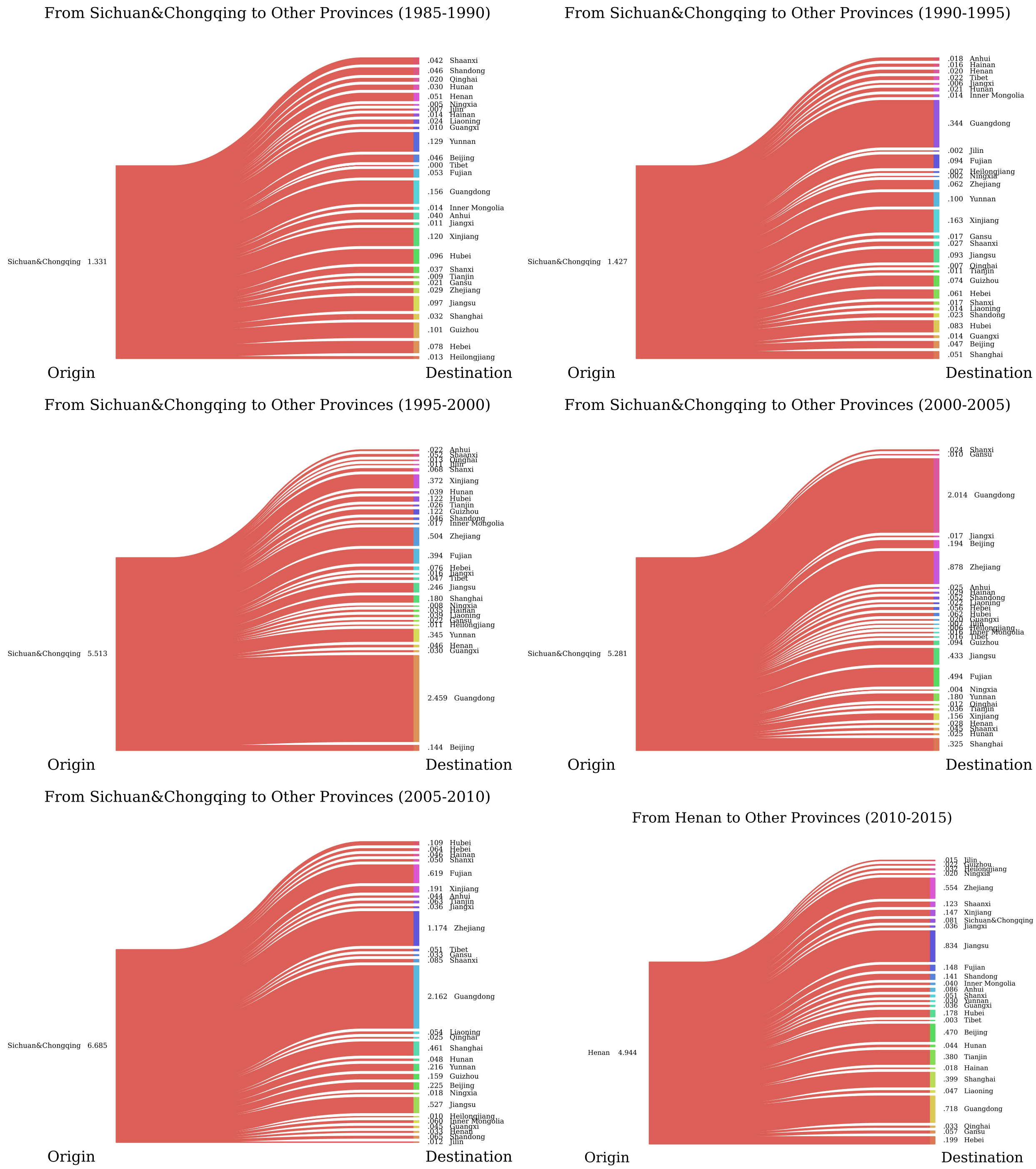
Notes: These figures show migration flows from other provinces to Guangdong, the largest migration destination, every five years from 1985 to 2015, measured in millions of people. The data are sourced from population censuses and mini-censuses conducted in various years.

Figure A.7: Distribution of Out-migration across Provinces



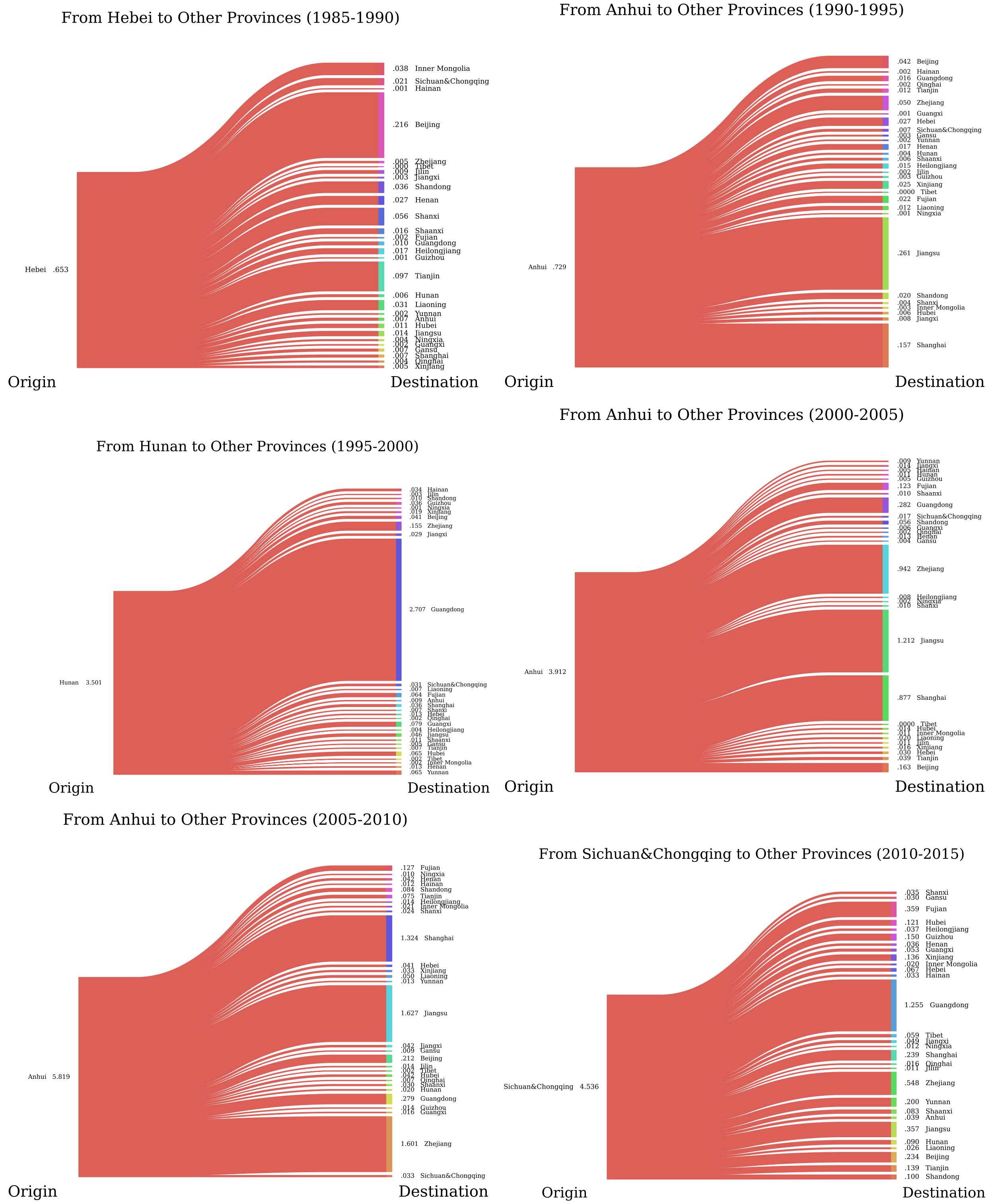
Notes: These figures show the distribution of out-migration across provinces every five years from 1985 to 2015, measured in millions of people. The data are sourced from population censuses and mini-censuses conducted in various years.

Figure A.8: Distribution of Out-migration across Provinces



Notes: These figures show migration flows from the largest migration origin to other provinces every five years from 1985 to 2015, measured in millions of people. The data are sourced from population censuses and mini-censuses conducted in various years.

Figure A.9: Distribution of Out-migration across Provinces



Notes: These figures show migration flows from the second largest migration origin to other provinces every five years from 1985 to 2015, measured in millions of people. The data are sourced from population censuses and mini-censuses conducted in various years.

B Identification Strategy – Migration, Trade, Idea Flows

B.1 Regressions using Card IV

We present the first-stage results of the IV regression using Card (2001)'s instrument in Table B.1. Results suggest a strong correlation in the first stage.

Table B.1: First Stage Results for Card IV

	$\text{Migration}^{H}_{n,t}$ (1)	$\text{Migration}^{L}_{n,t}$ (2)	$(\text{Import/GDP})_{n,t}$ (3)
$\text{Migration}^{H,Card\ \text{IV}}_{n,t}$	0.598***	-0.333	0.482***
	(0.15)	(0.46)	(0.10)
$\text{Migration}^{L,Card\ \text{IV}}_{n,t}$	0.031	0.819***	-0.093***
	(0.03)	(0.09)	(0.02)
$(\text{Import/GDP})_{n,t-1}$	-0.090	-0.008	0.452***
	(0.10)	(0.18)	(0.04)
Observations	150	150	150
R^2	0.793	0.810	0.751
Number of provinces	30	30	30
Year FE	✓	✓	✓
Province FE	✓	✓	✓
F-statistic	16.41	18.72	48.26

Notes: This table shows the first-stage results of the IV regression using Card (2001)'s instrument. Migrants are separated into those from high-TFP provinces ($\text{Migration}^{H}_{n,t}$) and low-TFP provinces ($\text{Migration}^{L}_{n,t}$). High-TFP provinces are defined as those with TFP levels above the national mean in 1990, while the remaining provinces are classified as low-TFP. $\text{Migration}^{k}_{n,t}$, where $k = H, L$, are measured in units of 10,000 people. is measured in units of 1 percentage point. Dependent variables in column (1)-(3) are$\text{Migration}^{H}_{n,t}$, $\text{Migration}^{L}_{n,t}$, and $(\text{Import/GDP})_{n,t}$, respectively. $\text{Migration}^{k,Card\ \text{IV}}_{n,t}$ is the instrument for $\text{Migration}^{k}_{n,t}$, where $k = H, L$, following Card (2001), and $(Import/GDP)_{n,t-1}$ is the instrument for $(Import/GDP)_{n,t}$. We report the Sanderson-Windmeijer F-statistic in the first stage. All columns include controls for year and province fixed effects. Robust standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

B.2 Regressions using Burchardi IV

We also construct instruments proposed by Burchardi et al. (2020). We first estimate, for each t,

$$M_{in,t} = \delta_{n,t} + \delta_{i,t} + \sum_{\tau=1982}^{t} a_{\tau,t} I_{i,-n,\tau} \frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}} + \nu_{in,t}, \tag{1}$$

where $I_{i,-n,\tau}$ is the total number of migrants flowing from i at τ who arrive in provinces outside of the province n, which captures a "push" shock from i. $\frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}}$ is the share of migrants flowing from provinces other than i who arrive in n at τ, capturing a "pull" shock to n. $M_{in,t}$ is the stock of migrants in province n from province i, measured by the number of residents in province n that have their Hukou from province i. The terms $\delta_{i,t}$ and $\delta_{n,t}$ are origin and destination province fixed effects, respectively. We estimate (1) separately for each $t = 1995, 2000, 2005, 2010, 2015$. We then derive the predicted stock of migrants

$$\hat{M}_{in,t} = \sum_{\tau=1982}^{t} \hat{a}_{\tau,t} \left(I_{i,-n,\tau} \frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}} \right)^{\perp},$$

where $\hat{a}_{\tau,t}$ are the coefficients estimated from (1), and $\left(I_{i,-n,\tau} \frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}} \right)^{\perp}$ is the obtained from residualizing $I_{i,-n,\tau} \frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}}$ with respect to $\delta_{n,t}$ and $\delta_{i,t}$ in (1). Having predicted migration stocks in hands, we apply the shift-share approach by interacting the predicted migration stock in a province with its migration flow from the same origin province,

$$I_{in,t} = \delta_{n,t} + \delta_{i,t} + \delta_t + b_t \left[\hat{M}_{in,t} \times \left(I_{i,-n,t} \frac{I_{-i,n,t}}{I_{-i,-n,t}} \right) \right] + u_{in,t}$$

where $\delta_{i,t}$, $\delta_{n,t}$, and δ_t are origin province, destination province, and time fixed effects, respectively, and $I_{i,-n,t} \frac{I_{-i,n,t}}{I_{-i,-n,t}}$ is a scaled push shock from i. $u_{in,t}$is an error term.

We finally construct our instrument for the total number of migrating arriving in province n in period t, $\text{Immigration}_{n,t}$, in (1),

$$\hat{I}_{in,t} = \hat{b}_t \left[\hat{M}_{in,t} \times \left(I_{i,-n,t} \frac{I_{-i,n,t}}{I_{-i,-n,t}} \right) \right].$$

Finally, we construct our IV as

$$I_{n,t}^{j,\text{BCHTT IV}} = \sum_{i \in \mathcal{I}^j} \hat{I}_{in,t}, j = L, H.$$

Identifying assumption. The identification assumption for our IV, constructed following Card (2001), is that $\epsilon_{n,t}$ is orthogonal to $I_{i\cdot,t} \times \frac{I_{in,t-1}}{I_{i\cdot,t-1}}$ for all i and t. This requires that any unobserved shocks $\epsilon_{n,t}$ that cause temporary increases in a given destination province's knowledge growth do not systematically correlate with migrants from an origin province to other provinces ($I_{i\cdot,t}$) interacted with the share of migrants in that destination five years ago ($\frac{I_{in,t-1}}{I_{i\cdot,t-1}}$).

The identification assumption of our IV, constructed following Burchardi et al. (2020), is that $\hat{M}_{in,t}$ is exogenous. The corresponding exclusion restriction is that $\epsilon_{n,t}$ is orthogonal to $I_{i,-n,\tau} \times \frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}}$ for all i and $\tau \leq t$. This requires that any unobserved shocks $\epsilon_{n,t}$ affecting knowledge growth in a given destination province do not systematically correlate with migration from an origin province to other provinces, excluding the given destination province ($I_{i,-n,\tau}$), interacted with the share of migrants in that destination among those not flowing from the given origin province ($\frac{I_{-i,n,\tau}}{I_{-i,\cdot,\tau}}$).

Table B.2: First-stage Results for Burchardi et al. IV

	$\text{Migration}^{H}_{n,t}$ (1)	$\text{Migration}^{L}_{n,t}$ (2)	$(\text{Import/GDP})_{n,t}$ (3)
$\text{Migration}^{H,BCHT\ \text{IV}}_{n,t}$	-2.809**	-4.711	-1.468
	(1.11)	(2.87)	(1.09)
$\text{Migration}^{L,BCHT\ \text{IV}}_{n,t}$	0.779***	3.470***	0.123
	(0.15)	(0.75)	(0.09)
$(\text{Import/GDP})_{n,t-1}$	-0.197	0.405*	0.314**
	(0.21)	(0.21)	(0.14)
Observations	150	150	150
R^2	0.649	0.725	0.530
Number of provinces	30	30	30
Year FE	✓	✓	✓
Province FE	✓	✓	✓
F-statistic	12.93	11.11	6.12

Notes: This table shows the first-stage results of the IV regression using Burchardi et al. (2020)'s instrument. Migrants are separated into those from high-TFP provinces ($\text{Migration}^{H}_{n,t}$) and low-TFP provinces ($\text{Migration}^{L}_{n,t}$). High-TFP provinces are defined as those with TFP levels above the national mean in 1990, while the remaining provinces are classified as low-TFP. $\text{Migration}^{k}_{n,t}$, where $k = H, L$ are measured in units of 10,000 people. is measured in units of 1 percentage point. Dependent variables in column (1)-(3) are $\text{Migration}^{H}_{n,t}$, $\text{Migration}^{L}_{n,t}$, and $(\text{Import/GDP})_{n,t}$, respectively. $\text{Migration}^{k,BCHT\ \text{IV}}_{n,t}$ is the instrument for $\text{Migration}^{k}_{n,t}$, where $k = H, L$, following Burchardi et al. (2020), and $(Import/GDP)_{n,t-1}$ is the instrument for $(Import/GDP)_{n,t}$. We report the Sanderson-Windmeijer F-statistic in the first stage. All columns include controls for year and province fixed effects. Robust standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

As our migration stocks measure started in 1990, we lost one cross section of data. To make coefficients comparable, we run our OLS specification using data after 1990. We present our results in Table B.3.

C 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.

C.1 Law of Motion of the Stock of Knowledge

We present two distinct approaches to derive the endogenous emergence of the Fréchet distribution for frontier productivity and subsequently link them to the existing literature.

Table B.3: Migration, Trade, and Changes in Patents

	Granted		Filed	
	Flow	Stock	Flow	Stock
ΔPatent Per Capita	(1)	(2)	(3)	(4)
$\text{Immigration}^H_{n,t}$	0.46***	0.62***	0.82***	1.20***
	(0.10)	(0.21)	(0.12)	(0.28)
$\text{Immigration}^L_{n,t}$	-0.06*	-0.08	-0.11***	-0.18***
	(0.03)	(0.05)	(0.04)	(0.06)
$(Import/GDP)_{n,t}$	0.62***	1.04***	0.77***	1.62***
	(0.17)	(0.14)	(0.18)	(0.21)
Observations	150	150	150	150
R^2	0.698	0.712	0.756	0.779
Number of provinces	30	30	30	30
Year FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓

Notes: This table presents the correlation between changes in patents per capita and migration as well as the import-to-GDP ratio across provinces. Migrants are separated into those from high-TFP provinces ($\text{Migration}^H_{n,t}$) and low-TFP provinces ($\text{Migration}^L_{n,t}$). High-TFP provinces are defined as those with TFP levels above the national mean in 1990, while the remaining provinces are classified as low-TFP. $\text{Migration}^k_{n,t}$, where $k = H, L$ are measured in units of 10,000 people. is measured in units of 1 percentage point. Patents per capita are expressed as the number of patents per 10,000 residents, including both migrants and stayers in 1990. The dependent variables in columns (1) to (4) are the flow per capita of granted patents, the stock per capita of granted patents, filed patent flow per capita, and filed patent stock per capita, respectively. The data covers years 1995, 2000, 2005, 2010, and 2015. All columns include controls for year and province fixed effects. Robust standard errors, clustered at the province level, are reported in parentheses, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

C.1.1 Approach 1: Distribution in Time Difference (Limit)

The arrival of new ideas between t and $t+1$drives changes in the distribution of frontier knowledge. Our focus is on the evolution of the entire distribution over each unit time interval.

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

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

where the fourth equality follows from the fact that $\Pr\left[Z \leq \frac{q}{Q'^{\rho}} \middle| Q' > (q/\bar{z})^{1/\rho}\right] = \Pr\left[Z \leq \bar{z}\right] = 0$.

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

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

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. *. The frontier of knowledge, $F_t(\cdot)$, at any t given follows a* Fréchet *distribution 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. We prove this result following two different approaches. The first approach relies on the assumption that the initial distribution at time 0 follows a Fréchet distribution. The second approach, which we label as the Poisson Approach, does not impose this assumption.

First note that 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 $t = 0$.

Approach 1: 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.

Related literature. This approach shares similarities with Buera and Oberfield (2020). We describe how the distribution of frontier knowledge at time t, combined with the arrival of ideas between t and $t+1$, shapes the distribution at $t+1$. This is akin to Buera and Oberfield (2020) who use a continuous-time framework to trace the evolution of the frontier knowledge distribution from t to $t + \Delta$ as $\Delta \to 0$. Both approaches assume an initial Fréchet distribution; however, the second method, which we introduce next, allows for relaxing this assumption.

Approach 2: Poisson Approach

We propose an alternative approach to deriving the Fréchet-distributed frontier technology. At any given moment, the distribution of frontier knowledge reflects the cumulative arrival of ideas up to that point. Rather than examining the evolution of the distribution over a unit time interval, we focus on how the accumulation of ideas from the initial time period to the present shapes the resulting distribution.

With $H(z) = 1 - (z/\bar{z})^{-\theta}$ and $\Lambda_t = \alpha_t \bar{z}^{-\theta}$, the number of newly arrived ideas with an original component's productivity $Z \geq z$ between t and $t+1$ follows a Poisson distribution with mean $R_t(z) = (1 - H(z))\Lambda_t = \alpha_t z^{-\theta}$. The support of $R_t(z)$ is $(0, \infty]$ when $\bar{z} \to 0$.

The number of new ideas with productivity greater than or equal to q between t and $t+1$ follows a Poisson distribution with mean

$$\begin{aligned}\tilde{R}_t(q) &= \int_0^\infty R_t\left(q/q^{'\rho}\right) dG_t\left(q^{'}\right) \\ &= q^{-\theta} \int_0^\infty \alpha_t \left(q^{'}\right)^{\rho\theta} dG_t\left(q^{'}\right) \\ &= q^{-\theta}\alpha_t \int_0^\infty x^{\rho\theta} dG_t(x).\end{aligned}$$

It follows from that given $Q^{'} = q^{'}$, an idea $Z \geq q/q^{'\rho}$ implies $ZQ^{'\rho} \geq q$. Now we denote $\Delta A_t \equiv \alpha_t \int_0^\infty x^{\rho\theta} dG_t(x)$. The number of new ideas with productivity greater than or equal to q between 0 and t follows a Poisson distribution with mean

$$\begin{aligned}\mu_t(q) &= \sum_{\tau=0}^{t-1} \tilde{R}_\tau(q) \\ &= q^{-\theta} \sum_{\tau=0}^{t-1} \alpha_\tau \int_0^\infty x^{\rho\theta} dG_\tau(x) \\ &= q^{-\theta} \sum_{\tau=0}^{t-1} \Delta A_\tau \\ &= q^{-\theta} A_t,\end{aligned}$$

where we denote $A_t \equiv \sum_{\tau=-\infty}^{t-1} \Delta A_\tau$, which we refer to as knowledge stock. The probability that no idea with productivity greater than or equal to q follows

$$F_t(q) = \exp\left[-\mu_t(q)\right] = \exp\left[-q^{-\theta} A_t\right],$$

where

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

Note that if the knowledge stock starts at zero, the random arrival of new ideas with varying productivity levels immediately results in a Fréchet-distributed frontier technology from the consequent period onwards. In this case, we no longer need to assume that the initial distribution of frontier technology has to be Fréchet.

Related literature. Eaton et al. (2011) first introduced the Poisson approach, which has since been adopted by Eaton et al. (2025) to study firm-to-firm trade. Cai and Xiang (2022) and Xiang (2023) use the Poisson approach to characterize the endogenous emergence of multivariate Fréchet-distributed frontier technology in the context of multinational production. Similarly, Lind and Ramondo (2023)'s characterization of Max-stable Fréchet productivity, resulting from Poisson innovation and independent diffusion, aligns with the spirit of the Poisson approach.

C.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

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

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

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

C.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 $\bar{z}_n \to 0$, $\alpha_t \lim_{n\to\infty} \bar{z}_n^{-\theta} \left[1 - \int\int_{B(\bar{z}_n)} dG_t^\ell(q_\ell) dG_t^m(q_m)\right] = 0$, where $B(\bar{z}) := \left\{(x_1, x_2) : \bar{z} x_1^{\rho_l} x_2^{\rho_m} < q\right\} \subset \mathbb{R}^2$. In addition, the integral $\int\int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t^\ell(q_\ell) dG_t^m(q_m)$ 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*

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

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

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

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

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

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

$$\lim_{n\to\infty} \alpha_t \bar{z}_n^{-\theta} (1 - F_t^{new}(q)) = \alpha_t \int\int \left(\frac{q}{q_\ell^{\rho_\ell} q_m^{\rho_m}}\right)^{-\theta} dG_t^\ell(q_\ell) dG_t^m(q_m).$$

The integral exists under Assumption 1′ d).

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

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

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

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

so

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

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

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

or, using the Riemann integral,

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

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

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

Then the law of motion becomes

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

The first integral,

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

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,

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

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,

$$
\begin{aligned}
G^m_{n,t}(x) &= \sum_i \int_0^x f_{i,t}(q) \left(\prod_{j \neq i} F_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} q \right) \right) dq \\
&= \sum_i \int_0^x A_{i,t}\theta q^{-\theta-1} \exp\left(-A_{i,t}q^{-\theta}\right) \exp\left(-\sum_{j\neq i} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q^{-\theta} \right) dq \\
&= \sum_i \int_0^x A_{i,t}\theta q^{-\theta-1} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q^{-\theta} \right) dq \\
&= \sum_i \frac{A_{i,t}\left(w_{i,t}\kappa_{ni,t}\right)^{-\theta}}{\sum_j A_{j,t}\left(w_{j,t}\kappa_{nj,t}\right)^{-\theta}} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} x^{-\theta} \right) \\
&= \sum_i \lambda_{ni,t} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} x^{-\theta} \right).
\end{aligned}
$$

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

$$
\begin{aligned}
\int_0^\infty q_m^{\theta\rho_m} dG^m_{n,t}(q_m) &= \int_0^\infty q_m^{\theta\rho_m} d \sum_i \lambda_{ni,t} \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q_m^{-\theta} \right) \\
&= \sum_i \lambda_{ni,t} \int_0^\infty q_m^{\theta\rho_m} d \exp\left(-\sum_{j} A_{j,t} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta} q_m^{-\theta} \right)
\end{aligned}
$$

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

$$
\begin{aligned}
\int_0^\infty q_m^{\theta\rho_m} dG^m_{n,t}(q_m) &= \sum_i \lambda_{ni,t} \int_0^\infty \sum_j A_{j,t}^{\rho_m} \left(\frac{w_{j,t}\kappa_{nj,t}}{w_{i,t}\kappa_{ni,t}} \right)^{-\theta\rho_m} x^{-\rho_m} d \exp\left(-x\right) \\
&= \Gamma(1-\rho_m) \sum_i \lambda_{ni,t} \left(\frac{A_{i,t}}{\lambda_{ni,t}} \right)^{\rho_m}.
\end{aligned}
$$

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

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

D Additional Derivations

In this section, we provide detailed derivations of the trade shares, migration shares, and the solution to landowner consumption and investment decisions.

D.1 Derivation of the Trade Shares

Let Ω be the variety space and intermediate variety $\omega \in \Omega$. Let $p_{in,t}(\omega)$ be the price that firms in location i pay for good ω purchased from location n at time t. Then perfect competition implies

$$p_{in,t}(\omega) = \frac{\kappa_{in,t} x_{n,t}}{q(\omega)},$$

where $x_{n.t}$ is the unit cost of inputs to produce in location n. Since $\{q(\omega)\}_{\omega\in\Omega}$ are i.i.d., for all $\omega \in \Omega$, they have the same distribution. Let $H_{in,t}(p)$ be the cumulative distribution of prices, i.e., $H_{in,t}(p) = \mathbb{P}\left[p_{in,t}(\omega) \leq p\right]$. Then

$$\begin{aligned}
H_{in,t}(p) &= \mathbb{P}[p_{in,t}(\omega) \leq p] \\
&= \mathbb{P}\left[\frac{\kappa_{in,t} x_{n,t}}{q(\omega)} \leq p\right] \\
&= \mathbb{P}\left[q(\omega) \geq \frac{\kappa_{in,t} x_{n,t}}{p}\right] \\
&= 1 - \mathbb{P}\left[q(\omega) \leq \frac{\kappa_{in} x_{n,t}}{p}\right] \\
&= 1 - F_{n,t}\left(\frac{\kappa_{in,t} x_{n,t}}{p}\right) \\
&= 1 - \exp\left\{-A_{n,t}\left(\frac{\kappa_{in,t} x_{n,t}}{p}\right)^{-\theta}\right\},
\end{aligned} \tag{D.1}$$

where $F_{n,t}(\cdot)$ denotes the Fréchet distribution with scale parameter $A_{n,t}$ and shape parameter θ.

Let $\lambda_{in,t}$ be the fraction of goods purchased by location i from n. For location i to buy good ω from n, n must be the lowest-cost supplier among all locations. By the law of large numbers, we have

$$\begin{aligned}
\lambda_{in,t} &= \mathbb{P}\left[p_{in,t}(\omega) \leq \min_{h\in S\setminus\{i\}} p_{ih,t}(\omega)\right] \\
&= \int_0^\infty \mathbb{P}\left[\min_{h\in S\setminus\{i\}} p_{ih,t}(\omega) \geq p\right] dH_{in,t}(p) \\
&= \int_0^\infty \mathbb{P}\left[\bigcap_{h\in S\setminus\{i\}} \{p_{ih,t}(\omega) \geq p\}\right] dH_{in,t}(p) \\
&= \int_0^\infty \prod_{h\in S\setminus\{i\}} \mathbb{P}\left[p_{ih,t}(\omega) \geq p\right] dH_{in,t}(p) \\
&= \int_0^\infty \prod_{h\in S\setminus\{i\}} \left[1 - H_{ih,t}(p)\right] dH_{in,t}(p),
\end{aligned}$$

where the law of iterated expectation is used for the second equality and independence is used for the fourth equality.

Using the expression of price distribution derived in (D.1), we have

$$\lambda_{in,t} = \int_0^\infty \prod_{h\in S\setminus\{i\}} \exp\left\{-A_{h,t}\left(\frac{\kappa_{ih,t}x_{h,t}}{p}\right)^{-\theta}\right\} \exp\left\{-A_{n,t}\left(\frac{\kappa_{in,t}x_{n,t}}{p}\right)^{-\theta}\right\} A_{n,t}(\kappa_{in}x_{in,t})^{-\theta}dp^\theta$$
$$= A_{n,t}(\kappa_{in,t}x_{n,t})^{-\theta}\int_0^\infty \exp\left\{-\sum_{h=1}^N A_{h,t}(\kappa_{ih,t}x_{h,t})^{-\theta}p^\theta\right\}dp^\theta$$
$$= \frac{A_{n,t}(\kappa_{in,t}x_{n,t})^{-\theta}}{\sum_{h=1}^N A_{h,t}(\kappa_{ih,t}x_{h,t})^{-\theta}}.$$

D.2 Derivation of Gross Flows Equation

Let $\mu_{in,t}$ be the fraction of individuals who relocate from location i to location n at time t. By definition, we have

$$\mu_{in,t} = \mathbb{P}\left[\frac{\beta V_{n,t+1}-m_{in,t}}{\nu}+\epsilon_{n,t} \geq \max_{l\neq n}\left\{\frac{\beta V_{l,t+1}-m_{il,t}}{\nu}+\epsilon_{l,t}\right\}\right]$$
$$= \int_\infty^\infty \mathbb{P}\left[\frac{\beta V_{n,t+1}-m_{in,t}}{\nu}+x \geq \max_{l\neq n}\left\{\frac{\beta V_{l,t+1}-m_{il,t}}{\nu}+\epsilon_{l,t}\right\}\right]dM(x)$$
$$= \int_{-\infty}^\infty \mathbb{P}\left[\bigcap_{l\neq n}\left\{\frac{\beta V_{n,t+1}-m_{in,t}}{\nu}+x \geq \frac{\beta V_{l,t+1}-m_{il,t}}{\nu}+\epsilon_{l,t}\right\}\right]dM(x)$$
$$= \int_{-\infty}^\infty \prod_{l\neq n}\mathbb{P}\left[\frac{\beta V_{n,t+1}-m_{in,t}}{\nu}+x \geq \frac{\beta V_{l,t+1}-m_{il,t}}{\nu}+\epsilon_{l,t}\right]dM(x)$$

$$= \int_{-\infty}^\infty \prod_{l\neq n}\mathbb{P}\left[\epsilon_{l,t} \leq \frac{\beta(V_{n,t+1}-V_{l,t+1})-(m_{in,t}-m_{il,t})}{\nu}+x\right]dM(x)$$
$$= \int_{-\infty}^\infty \prod_{l\neq n} M\left(\frac{\beta(V_{n,t+1}-V_{l,t+1})-(m_{in,t}-m_{il,t})}{\nu}+x\right)dM(x),$$

where $M(\cdot)$ denotes the cumulative distribution function of a Gumbel Type I distribution.

Define $\bar{\epsilon}_{ln,t} = \frac{\beta(V_{n,t+1}-V_{l,t+1})-(m_{in,t}-m_{il,t})}{\nu}$ with $\bar{\epsilon}_{nn,t}=0$. Using this notation and the expression of $M(\cdot)$, we have

$$\mu_{in,t} = \int_{-\infty}^\infty \exp\{-e^{-x-\gamma}\}e^{-x-\gamma}\exp\{-e^{-x-\gamma}\sum_{l\neq n}e^{-\bar{\epsilon}_{ln,t}}\}dx$$
$$= \int_{-\infty}^\infty e^{-x-\gamma}\exp\{-e^{-x-\gamma}\sum_{l=1}^N e^{-\bar{\epsilon}_{ln,t}}\}dx.$$

Define $\Xi_{in} = \log(\sum_{l=1}^{N} e^{-\bar{\epsilon}_{ln,t}})$ and $y = x + \gamma - \Xi_{in}$. Then

$$\begin{aligned}\mu_{in,t} &= \int_{-\infty}^{\infty} e^{-y-\Xi_{in}} e^{-e^{-y}} dy \\ &= e^{\Xi_{in}}.\end{aligned}$$

Finally, plugging in the expression of Ξ_{in}, we have

$$\begin{aligned}\mu_{in,t} &= \frac{1}{\sum_{l=1}^{N} \exp\left\{\frac{\beta(V_{l,t+1}-V_{n,t+1})-m_{il,t}+m_{in,t}}{\nu}\right\}} \\ &= \frac{\exp(\beta V_{n,t+1} - m_{in,t})^{\frac{1}{\nu}}}{\sum_{l=1}^{N} \exp(\beta V_{l,t+1} - m_{il,t})^{\frac{1}{\nu}}}.\end{aligned}$$

D.3 Landlord's Problem

The landlord's problem is defined as

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

where δ is the depreciation rate and $K_{i,0}$ is taken as given. Set up the Lagrangian equation,

$$\mathcal{L} = \left[\beta^t \{\log(c_{i,t}) + \lambda_t [r_{i,t} K_{i,t} - P_{i,t}(c_{i,t} + K_{i,t+1} - (1-\delta) K_{i,t})]\}\right], \tag{D.2}$$

where λ_t is the Lagrangian multiplier for the constraint in period t.

The first-order conditions for the problem are

$$\begin{aligned}\frac{1}{c_{i,t}} &= \lambda_t P_{i,t} \\ \lambda_t P_{i,t} &= \beta\left[\lambda_{t+1}[r_{i,t+1} + P_{i,t+1}(1-\delta)]\right].\end{aligned}$$

Define $R_{i,t} = 1 - \delta + \frac{r_{i,t}}{P_{i,t}}$. Then eliminating λ_t yields the Euler equation,

$$\frac{1}{c_{i,t}} = \beta\left[R_{i,t+1} \frac{1}{c_{i,t+1}}\right], \tag{D.3}$$

together with the budget constraint

$$R_{i,t} K_{i,t} = c_{i,t} + K_{i,t+1}. \tag{D.4}$$

To solve this problem, we use the guess-and-verify strategy. We guess that $c_{i,t} = \varsigma R_{i,t} K_{i,t}$, where ς is a constant to be determined. Plugging in (D.4), we have

$$K_{i,t+1} = (1-\varsigma) R_{i,t} K_{i,t}. \tag{D.5}$$

Combining equations (D.5) and (D.3), we have

$$\frac{1}{\varsigma R_{i,t} K_{i,t}} = \beta \left[R_{i,t+1} \frac{1}{\varsigma R_{i,t+1} (1-\varsigma) R_{i,t} K_{i,t}} \right].$$

The undetermined coefficient method implies that $\varsigma = 1 - \beta$. Hence, the consumption and saving policy functions are as follows:

$$c_{i,t} = (1-\beta)[r_{i,t}/P_{i,t} + (1-\delta)]K_{i,t},$$
$$K_{i,t+1} = \beta[r_{i,t}/P_{i,t} + (1-\delta)]K_{i,t}.$$

E 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

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

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

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

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

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

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

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

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

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

or

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

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

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

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 express the labor market clearing condition in a detrended form as

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

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

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

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

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

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

Using equation (E.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

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

which immediately implies that

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

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

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

or

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

We then require that

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

and in this way, the detrended capital accumulation equation becomes

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

From equation (E.9) we obtain that

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

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

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

F Dynamic-Hat Algebra

Proposition 4. Dynamic-Hat Algebra. Define the term $\hat{y}_{t+1}$ as the relative 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* $\left\{\left\{\lambda_{in,0}\right\}_{i=1,n=1}^{N,N},\left\{\mu_{in,0}\right\}_{i=1,n=1}^{N,N},\left\{w_{i,0}L_{i,0}\right\}_{i=1}^{N},\left\{K_{i,0}\right\}_{i=1}^{N},\left\{L_{i,0}\right\}_{i=1}^{N}\right\}$*, the parameters and elasticities* $\left(\rho_\ell,\rho_m,\theta,\nu,\gamma,\xi,\beta\right)$*, the initial rate and growth rate in the arrival of ideas* $\left(\alpha_0,g_\alpha\right)$ *and a convergent sequence of future changes in fundamentals under perfect foresight* $\left\{\hat{\kappa}_{in,t},\hat{m}_{in,t}\right\}_{i=1,n=1,t=1}^{N,N,\infty}$*, the solution for the sequence of changes in the model's endogenous variables in the detrended model* $\left\{\hat{y}_{t+1}\right\}_{t=1}^{\infty}$ *does not* *require information on the level of fundamentals (trade and migration costs).*

Proof. Let us define the variable $\hat{y}_{t+1}$ 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

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

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

$$L_{i,t+1} = \sum_{n=1}^{N}\mu_{ni,t}L_{n,t}, \tag{F.3}$$

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

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

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

$$\hat{w}_{i,t+1}\hat{L}_{i,t+1} = \frac{1}{\tilde{w}_{i,t}L_{i,t}}\sum_{n=1}^{N}\lambda_{ni,t+1}\hat{w}_{n,t+1}\hat{L}_{n,t+1}\tilde{w}_{n,t}L_{n,t}, \tag{F.7}$$

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

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

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

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 what follows 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.

To derive the system of equations in time differences, we first reproduce the equilibrium conditions of the detrended model derived in Appendix E,

$$\tilde{V}_{i,t} = \beta log\left(1+g_v\right) + log\left(\frac{\tilde{w}_{i,t}}{\tilde{P}_{i,t}}\right) + \nu log\left(\sum_{n=1}^{N} exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}\right), \tag{F.11}$$

$$\tilde{P}_{i,t} = T\left(\sum_{n=1}^{N} \tilde{A}_{n,t}\left(\kappa_{in,t}\tilde{x}_{n,t}\right)^{-\theta}\right)^{-1/\theta}, \tag{F.12}$$

$$\tilde{w}_{i,t}L_{i,t} = \sum_{n=1}^{N} \tilde{A}_{i,t}\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}}{\tilde{P}_{n,t}/T}\right)^{-\theta} \tilde{w}_{n,t}L_{n,t}, \tag{F.13}$$

$$\tilde{r}_{i,t}\tilde{K}_{i,t} = \sum_{n=1}^{N} \tilde{A}_{i,t}\left(\frac{\kappa_{ni,t}\tilde{x}_{i,t}}{\tilde{P}_{n,t}/T}\right)^{-\theta} \tilde{r}_{n,t}\tilde{K}_{n,t}, \tag{F.14}$$

$$L_{i,t+1} = \sum_{n=1}^{N} \mu_{ni,t}L_{n,t}, \tag{F.15}$$

$$\tilde{K}_{i,t+1} = \frac{\beta}{(1+g_k)}\left(\tilde{r}_{i,t}/\tilde{P}_{i,t} + (1-\delta)\right)\tilde{K}_{i,t}, \tag{F.16}$$

$$\tilde{A}_{n,t+1} - \frac{\tilde{A}_{n,t}}{(1+g_A)} = \frac{\alpha_0\Gamma_{\rho_\ell,\rho_m}}{(1+g_A)}\sum_{i=1}^{N} s_{in,t}\left(\tilde{A}_{i,t}\right)^{\rho_l}\sum_{i=1}^{N}\lambda_{ni,t}\left(\frac{\tilde{A}_{i,t}}{\lambda_{ni,t}}\right)^{\rho_m}, \tag{F.17}$$

Recall first that the share of workers moving from location i to n at time $t+1$ is given by

$$\mu_{in,t+1} = \frac{exp\left(\beta V_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta V_{h,t+2} - m_{ih,t+1}\right)^{1/\nu}} = \frac{exp\left(\beta \tilde{V}_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\sum_{h=1}^{N} exp\left(\beta \tilde{V}_{h,t+2} - m_{ih,t+1}\right)^{1/\nu}},$$

where for the second equality we use the definition $e^{V_{i,t}} = e^{\tilde{V}_{i,t}}(1+g_v)^t$ for all i and t.

By multiplying and dividing $\mu_{in,t}$ in the numerator and $\mu_{ih,t}$ for each term in the summation in the denominator, we have

$$\begin{aligned}\mu_{in,t+1} &= \frac{\mu_{in,t}exp\left(\beta \tilde{V}_{n,t+2} - \beta \tilde{V}_{n,t+1} + m_{in,t+1} - m_{in,t}\right)^{1/\nu}}{\sum_{h=1}^{N} \mu_{ih,t}exp\left(\beta \tilde{V}_{h,t+2} - \beta \tilde{V}_{h,t+1} + m_{ih,t+1} - m_{ih,t}\right)^{1/\nu}} \\ &= \frac{\mu_{in,t}\left(\hat{u}_{n,t+2}\right)^{\beta/\nu}\left(\hat{m}_{in,t+1}\right)^{-1/\nu}}{\sum_{h=1}^{N} \mu_{ih,t}\left(\hat{u}_{h,t+2}\right)^{\beta/\nu}\left(\hat{m}_{ih,t+1}\right)^{-1/\nu}},\end{aligned}$$

which is equation (F.2).

To obtain equilibrium condition (F.1), we take the time difference using (F.11). We obtain

$$
\begin{aligned}
\log(\hat{u}_{i,t+1}) &= \tilde{V}_{i,t+1} - \tilde{V}_{i,t} \\
&= \log\left(\frac{\tilde{w}_{i,t+1}/\tilde{P}_{i,t+1}}{\tilde{w}_{i,t}/\tilde{P}_{i,t}}\right) + \nu \log\left(\sum_{n=1}^{N} \frac{\exp\left(\beta \tilde{V}_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\sum_{h=1}^{N} \exp\left(\beta \tilde{V}_{h,t+1} - m_{ih,t}\right)^{1/\nu}}\right) \\
&= \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \frac{\exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu} \frac{\exp\left(\beta \tilde{V}_{n,t+2} - m_{in,t+1}\right)^{1/\nu}}{\exp\left(\beta \tilde{V}_{n,t+1} - m_{in,t}\right)^{1/\nu}}}{\sum_{h=1}^{N} \exp\left(\beta \tilde{V}_{h,t+1} - m_{ih,t}\right)^{1/\nu}}\right) \\
&= \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \mu_{in,t} \exp\left(\tilde{V}_{n,t+2} - \tilde{V}_{n,t+1}\right)^{\beta/\nu} \exp\left(m_{in,t+1} - m_{in,t}\right)^{-1/\nu}\right) \\
&= \log\left(\hat{w}_{i,t+1}/\hat{P}_{i,t+1}\right) + \nu \log\left(\sum_{n=1}^{N} \mu_{in,t} \left(\hat{u}_{n,t+2}\right)^{\beta/\nu} \left(\hat{m}_{in,t+1}\right)^{-1/\nu}\right),
\end{aligned}
$$

where for the third equality we use the expression of $\mu_{in,t}$ previously derived.

Since labor in each location is constant in the long run, we immediately obtain (F.3) from the law of motion (F.15).

To obtain equation (F.4), note that

$$
\tilde{x}_{i,t} = B\left(\tilde{w}_{i,t}^{\xi} \tilde{r}_{i,t}^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t}^{1-\gamma}.
$$

Taking the time difference yields

$$
\hat{x}_{i,t} \equiv \frac{\tilde{x}_{i,t+1}}{\tilde{x}_{i,t}} = \frac{\left(\tilde{w}_{i,t+1}^{\xi} \tilde{r}_{i,t+1}^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t+1}^{1-\gamma}}{\left(\tilde{w}_{i,t}^{\xi} \tilde{r}_{i,t}^{1-\xi}\right)^{\gamma} \tilde{P}_{i,t}^{1-\gamma}} = \left(\hat{w}_{i,t}^{\xi} \hat{r}_{i,t}^{1-\xi}\right)^{\gamma} \hat{P}_{i,t}^{1-\gamma}.
$$

Recall that in the detrended version of the model, the trade flow share from location n to location i at time t is

$$
\lambda_{in,t} = \frac{T^{-\theta} \tilde{A}_{n,t} \left(\kappa_{in,t} \tilde{x}_{n,t}\right)^{-\theta}}{\tilde{P}_{i,t}^{-\theta}},
$$

where T is some constant. Taking the time difference yields

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

which leads to equilibrium condition (F.6).

Note that the detrended price index in location i is

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

Taking the time difference, we have

$$\hat{P}_{i,t+1} = \left(\sum_{n=1}^{N} \frac{\tilde{A}_{n,t+1} \left(\kappa_{in,t+1} \tilde{x}_{n,t+1}\right)^{-\theta}}{\sum_{h=1}^{N} \tilde{A}_{h,t} \left(\kappa_{ih,t} \tilde{x}_{h,t}\right)^{-\theta}} \right)^{-1/\theta}$$

$$= \left(\sum_{n=1}^{N} \frac{\tilde{A}_{n,t} \left(\kappa_{in} \tilde{x}_{n,t}\right)^{-\theta} \tilde{A}_{n,t+1} \left(\kappa_{in,t+1} \tilde{x}_{n,t+1}\right)^{-\theta} / \tilde{A}_{n,t} \left(\kappa_{in,t} \tilde{x}_{n,t}\right)^{-\theta}}{\sum_{h=1}^{N} \tilde{A}_{h,t} \left(\kappa_{ih,t} \tilde{x}_{h,t}\right)^{-\theta}} \right)^{-1/\theta}$$

$$= \left(\sum_{n=1}^{N} \lambda_{in,t} \hat{A}_{n,t+1} \left(\hat{\kappa}_{in,t+1} \hat{x}_{n,t+1}\right)^{-\theta} \right)^{-1/\theta},$$

where we use $\lambda_{in,t} = \frac{\tilde{A}_{n,t}(\kappa_{in}\tilde{x}_{n,t})^{-\theta}}{\sum_{h=1}^{N} \tilde{A}_{h,t}(\kappa_{ih}\tilde{x}_{h,t})^{-\theta}}$ and which gives equilibrium condition (F.5).

To obtain equilibrium condition (F.7), we use labor market clearing condition (F.13),

$$\tilde{w}_{i,t+1} L_{i,t+1} = \sum_{n=1}^{N} \lambda_{ni,t+1} \tilde{w}_{n,t+1} L_{n,t+1},$$

and divide by $\tilde{w}_{i,t} L_{i,t}$ on both sides, to obtain

$$\hat{w}_{i,t+1} \hat{L}_{i,t+1} = \frac{1}{\tilde{w}_{i,t} L_{i,t}} \sum_{n=1}^{N} \lambda_{ni,t+1} \tilde{w}_{n,t+1} L_{n,t+1}$$

$$= \frac{1}{\tilde{w}_{i,t+1} L_{i,t+1}} \sum_{n=1}^{N} \lambda_{ni,t+1} \hat{w}_{n,t+1} \hat{L}_{n,t+1} \tilde{w}_{n,t} L_{n,t},$$

where as before we use $\tilde{L}_{n,t} = L_{n,t}$.

Equation (F.8) is exactly the detrended law of motion of capital as in equation (F.16). To obtain equation (F.9), we use the equilibrium condition:

$$\frac{\tilde{w}_{i,t} \tilde{L}_{i,t}}{\left[\tilde{R}_{i,t} - (1-\delta)\right] \tilde{P}_{i,t} \tilde{K}_{i,t}} = \frac{\xi}{1-\xi}.$$

Taking the time difference and rearranging this expression yields the desired result.

Finally, to obtain the law of motion of knowledge in relative time changes (F.10), note that equation (F.17) gives the detrended law of motion of knowledge:

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

Divided by $\tilde{A}_{n,t}$ on both sides, we have

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

G Solution Algorithm

In this section we describe the algorithm used to compute the dynamic spatial growth model.

G.1 Algorithm to Solve for the Sequential Equilibrium Given Initial Conditions

In what follows, we describe the algorithm to solve the detrended model given an initial allocation of the economy, $\left(\{L_{i,0}\}_{i=1}^{N}, \{\tilde{K}_{i,0}\}_{i=1}^{N}, \{\lambda_{in,0}\}_{i,n=1}^{N}, \{\mu_{in,-1}\}_{i,n=1}^{N}, \{\tilde{A}_{i,0}\}_{i=1}^{N}\right)$, and given an unanticipated convergent sequence of changes in fundamentals, $\left\{\{\hat{m}_{in,t}\}_{i,n=1}^{N}, \{\hat{\kappa}_{in,t}\}_{i,n=1}^{n}\right\}_{t=1}^{\infty}$. We first describe the algorithm to solve the model under the given initial conditions and constant fundamentals going forward; namely, with $\left\{\{\hat{m}_{in,t}=1\}_{i,n=1}^{N}, \{\hat{\kappa}_{in,t}=1\}_{i,n=1}^{n}\right\}_{t=1}^{\infty}$. We then describe how to solve the model under a change in fundamentals.

1. Initiate the algorithm at $t=0$ with a guess for the path of $\left\{\hat{u}_{i,t+1}^{(0)}\right\}_{t=0}^{T}$, where the superscript (0) indicates that it is a guess. The path should converge to $\hat{u}_{i,T+1}^{(0)}=1$ for sufficiently large T.

2. For all $t \geq 0$, use $\left\{\hat{u}_{i,t+1}^{(0)}\right\}_{t=0}^{T}$ and $\{\mu_{in,-1}\}_{i,n=1}^{N}$ to solve for the path of $\left\{\{\mu_{in,t}\}_{i,n=1}^{N}\right\}_{t=0}^{T}$ using equation (F.2).

3. Use the path for $\left\{\{\mu_{in,t}\}_{i,n=1}^{N}\right\}_{t=0}^{T}$ and $\{L_{i,0}\}_{i=1}^{N}$ to obtain the path for $\left\{\{L_{i,t+1}\}_{i=1}^{N}\right\}_{t=0}^{T}$ using equation (F.3).

4. Solve for the trade equilibrium:

   (a) For each $t \geq 0$, given $\hat{L}_{i,t+1}$, define the term $\hat{\omega}_{i,t} = \tilde{w}_{i,t}^{\xi}\tilde{r}_{i,t}^{1-\xi}$. Guess a value for $\hat{\omega}_{i,t+1}$.

   (b) Obtain $\hat{x}_{i,t+1}$, $\hat{P}_{i,t+1}$, $\lambda_{in,t+1}$, $\tilde{R}_{i,t+1}$, $\tilde{K}_{i,t+1}$, and $\hat{A}_{i,t+1}$ using equations (F.4), (F.5), (F.6), (F.8) and (F.9). Use the fact that $\hat{r}_{i,t+1} = \hat{w}_{i,t+1}\hat{L}_{i,t+1}/\hat{K}_{i,t+1}$ and $\hat{w}_{i,t+1} = \hat{\omega}_{i,t+1}\left(\hat{K}_{i,t+1}/\hat{L}_{i,t+1}\right)^{1-\xi}$.

   (c) Check if the market clearing condition (F.7) holds using $\hat{w}_{i,t+1} = \hat{\omega}_{i,t+1}\left(\hat{K}_{i,t+1}/\hat{L}_{i,t+1}\right)^{1-\xi}$. If it does not, go back to step (a) and adjust the initial guess for $\hat{\omega}_{i,t+1}$ until labor markets clear.

   (d) Repeat steps (a) through (d) for each period t and obtain paths for $\{\hat{w}_{i,t+1}, \hat{P}_{i,t+1}\}_{t=0}^{T}$ for all i.

5. For each t, use $\mu_{in,t}$, $\hat{w}_{i,t+1}$, $\hat{P}_{i,t+1}$, and $\hat{u}_{n,t+2}^{(0)}$ to solve backwards for $\hat{u}_{i,t+1}^{(1)}$ using equation (F.1). This solution delivers a new path for $\left\{\left\{\hat{u}_{i,t+1}^{(1)}\right\}_{i=1}^{N}\right\}_{t=0}^{T}$, where the superscript 1 indicates an updated value for $\hat{u}$.

6. Check whether $\left\{\left\{\hat{u}_{i,t+1}^{(1)}\right\}_{i=1}^{N}\right\}_{t=0}^{T} \approx \left\{\left\{\hat{u}_{i,t+1}^{(0)}\right\}_{i=1}^{N}\right\}_{t=0}^{T}$. If it does not, go back to step 1 and update the initial guess with $\left\{\left\{\hat{u}_{i,t+1}^{(1)}\right\}_{i=1}^{N}\right\}_{t=0}^{T}$.

G.2 Solving for Counterfactual Changes in Fundamentals

We now describe how to solve the dynamic spatial growth model given an unanticipated convergent sequence of changes in fundamentals, $\hat{\Theta}_{t+1} = \left\{ \{\hat{m}_{in,t}\}_{i,n=1}^{N}, \{\hat{\kappa}_{in,t}\}_{i,n=1}^{n} \right\}_{t=1}^{\infty}$.

The algorithm used to solve for a change in fundamentals follows the same steps described in the previous section, but the sequence of changes in fundamentals is fed into the model. The main difference from the previous section is that we now consider the fact that agents are surprised in the first period by the changes in fundamentals. The surprise in the changes in fundamentals is captured in the initial gross flow equation. That is, in the first period we now use the following equilibrium condition:

$$\mu_{in,1}(\hat{\Theta}) = \frac{\vartheta_{in,0} \left(\hat{u}_{n,2}(\hat{\Theta})\right)^{\beta/\nu} \left(\hat{m}_{in,1}\right)^{-1/\nu}}{\sum_{i=1}^{N} \vartheta_{ih,0} \left(\hat{u}_{h,2}(\hat{\Theta})\right)^{\beta/\nu} \left(\hat{m}_{ih,1}\right)^{-1/\nu}},$$

where $\vartheta_{in,0} = \mu_{in,0} \exp\left(V_{n,1}(\hat{\Theta}) - V_{n,1}\right)^{\beta/\nu}$. Therefore, we also use the equilibrium condition,

$$\log(\hat{u}_{i,1}) = \log\left(\hat{w}_{i,1}/\hat{P}_{i,1}\right) + \nu \log\left(\sum_{n=1}^{N} \vartheta_{in,0} \left(\hat{u}_{n,2}\right)^{\beta/\nu} \left(\hat{m}_{in,1}\right)^{-1/\nu}\right).$$

H Additional Quantitative Results

In this section of the appendix we describe additional results from our quantitative analysis.

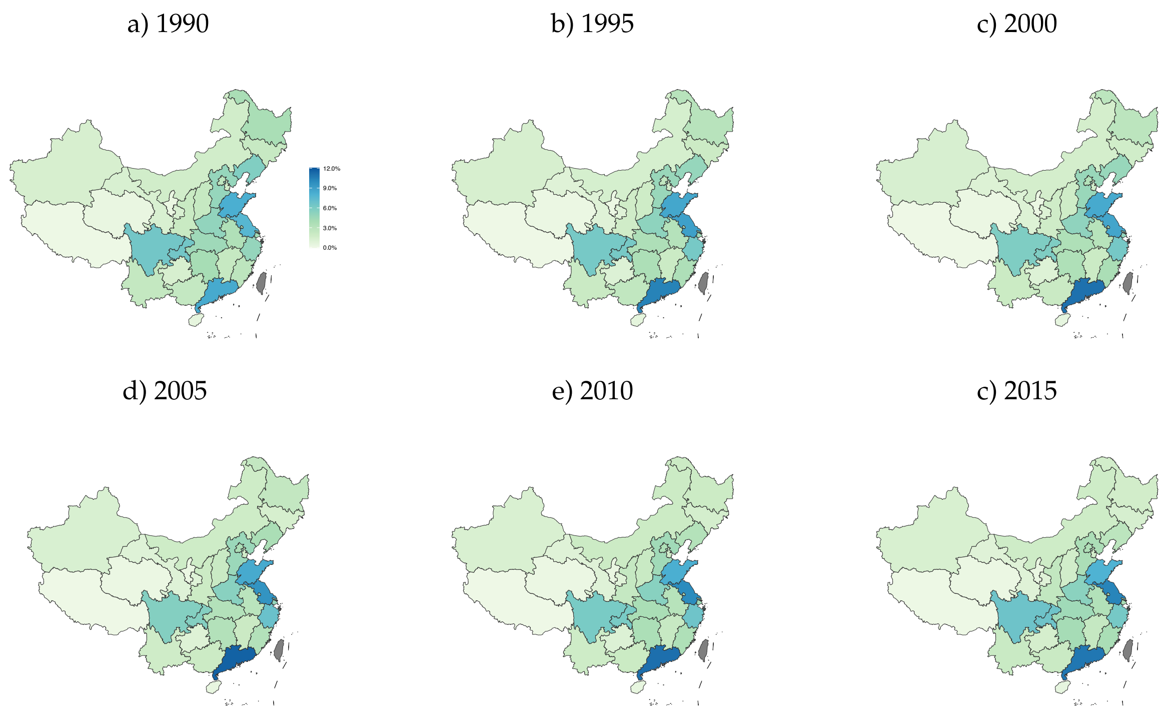
H.1 Calculating Migration Costs

We capture the changes in migration frictions across provinces in China using the cross-variation in five-year mobility rates from 1985-1990 to 2010-2015 as $\frac{\hat{\mu}_{in,t}\hat{\mu}_{ni,t}}{\hat{\mu}_{ii,t}\hat{\mu}_{nn,t}} = \left(\hat{m}_{in,t}\hat{m}_{ni,t}\right)^{-\frac{1}{v}}$. We apply the same change in migration costs to all Hukou types. We calculate mobility rates $\mu_{in,t}$ from the raw migration flows. Sometimes, the raw migration flows contain 0. In this case, we add a small value (one-hundredth of the smallest non-zero migration flow in the same year) first before we calculate mobility rates. We restrict our focus on $\hat{m}_{in,t}$ that reflect reductions of migration costs.

H.2 Regional Distribution of Economic Activity

Figures H.1 and H.2 display the evolution of actual GDP shares in China and their evolution under initial 1990 conditions. The figure presents the GDP shares across provinces in China every five years during the period 1990-2015.

Figure H.1: GDP Shares



Notes: The figures show the distribution of economic activity across provinces in China, measured as GDP shares, in the data.

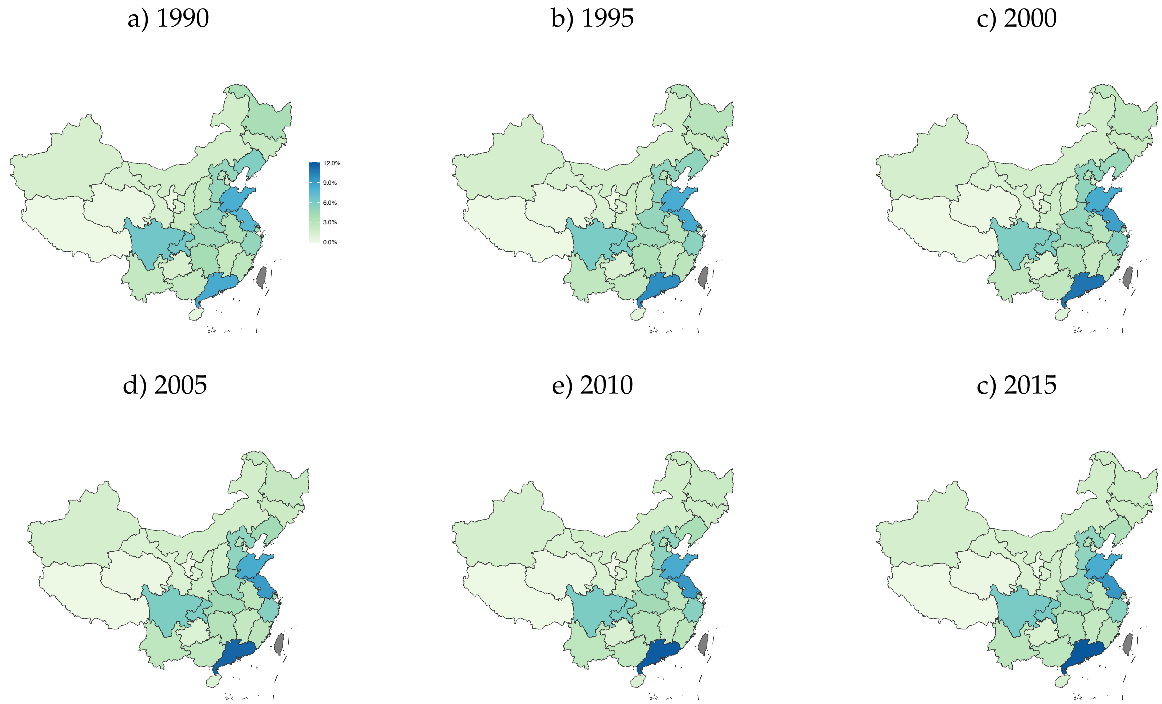
H.3 Initial Conditions and Spatial Growth

Figure H.3 plots initial distribution of fundamentals and spatial development. Figure H.4 presents the growth rate in each province under initial conditions.

H.4 Additional Tables

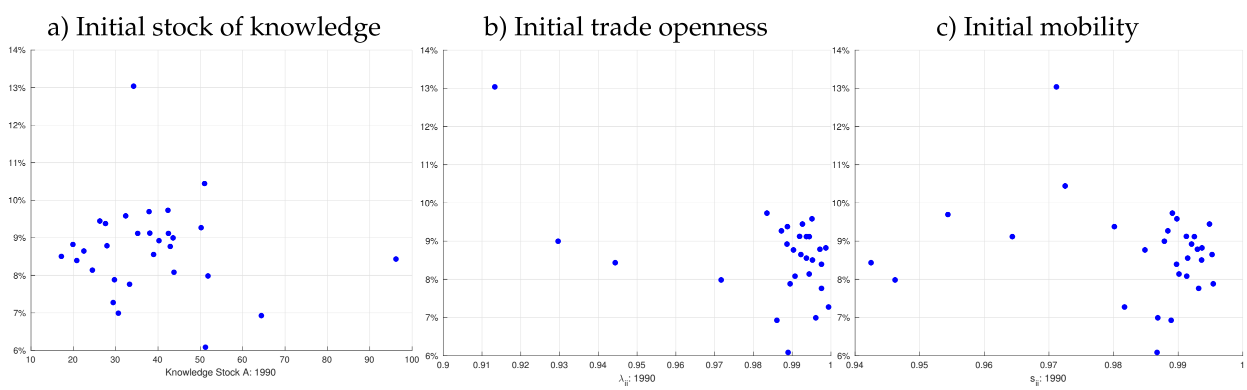
Table H.1 presents the results with fundamentals in 1990. Table H.2 presents the results with fundamentals, incorporating changes in both trade and migration costs.

Figure H.2: GDP shares (1990 conditions)



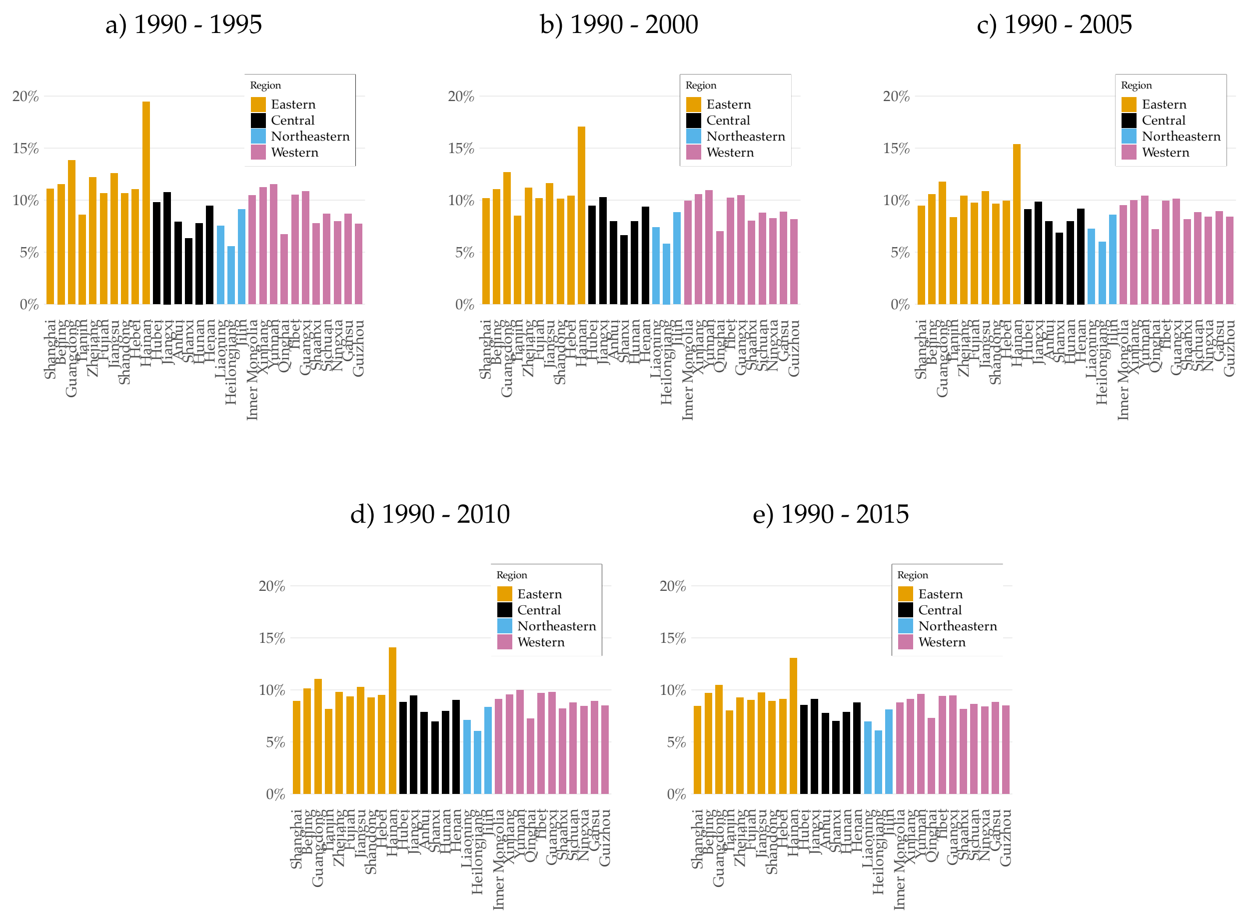
Notes: The figures show the distribution of economic activity across provinces in China, measured as GDP shares, in the model with 1990 fundamentals.

Figure H.3: Real GDP Growth versus Initial Conditions



Notes: 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 domestic expenditure share, λ_{ii}, in Panel (b) (two outlier provinces were trimmed), and initial share of stayers, s_{ii}, in Panel (c).

Figure H.4: Spatial Growth (Annual, Percent)



Notes: The figures show the annual real GDP growth across provinces in China in different time frames over the period 1990-2015. Spatial growth in each figure is computed in the model under the initial 1990 conditions.

Table H.1: With Fundamentals in 1990

Province	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015
Beijing	11.54%	10.51%	9.59%	8.78%	8.08%
Tianjin	8.61%	8.39%	8.04%	7.65%	7.24%
Hebei	11.02%	9.85%	8.94%	8.22%	7.62%
Shanxi	6.31%	6.96%	7.26%	7.29%	7.15%
Inner Mongolia	10.48%	9.44%	8.62%	7.95%	7.39%
Liaoning	7.55%	7.26%	6.94%	6.61%	6.29%
Jilin	9.11%	8.57%	8.05%	7.57%	7.12%
Heilongjiang	5.57%	6.06%	6.28%	6.31%	6.21%
Shanghai	11.09%	9.27%	8.07%	7.22%	6.58%
Jiangsu	12.56%	10.66%	9.37%	8.43%	7.71%
Zhejiang	12.18%	10.17%	8.86%	7.94%	7.25%
Anhui	7.92%	8.02%	7.90%	7.64%	7.33%
Fujian	10.67%	9.70%	8.88%	8.18%	7.58%
Jiangxi	10.77%	9.80%	8.99%	8.31%	7.73%
Shandong	10.64%	9.61%	8.78%	8.10%	7.52%
Henan	9.47%	9.24%	8.86%	8.42%	7.96%
Hubei	9.78%	9.12%	8.50%	7.95%	7.45%
Hunan	7.79%	8.09%	8.08%	7.88%	7.57%
Guangdong	13.81%	11.53%	10.00%	8.90%	8.07%
Guangxi	10.86%	10.12%	9.40%	8.74%	8.14%
Hainan	19.44%	14.73%	12.01%	10.26%	9.05%
Sichuan	8.67%	8.94%	8.86%	8.58%	8.20%
Guizhou	7.73%	8.54%	8.86%	8.82%	8.57%
Yunnan	11.52%	10.36%	9.42%	8.65%	8.00%
Tibet	10.49%	9.97%	9.39%	8.80%	8.24%
Shaanxi	7.79%	8.28%	8.39%	8.25%	7.97%
Gansu	8.69%	9.07%	9.07%	8.83%	8.46%
Qinghai	6.70%	7.29%	7.53%	7.51%	7.34%
Ningxia	7.96%	8.53%	8.68%	8.55%	8.26%
Xinjiang	11.22%	9.89%	8.90%	8.12%	7.49%

Notes: GDP growth with 1990 fundamentals is computed by solving the dynamic spatial growth model with constant fundamentals.

Table H.2: Changes in Fundamentals

Province	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015
Beijing	11.00%	10.41%	10.47%	9.86%	8.77%
Tianjin	9.20%	9.95%	10.59%	9.34%	9.23%
Hebei	11.02%	9.82%	9.01%	8.36%	7.77%
Shanxi	6.36%	7.03%	7.36%	7.19%	6.79%
Inner Mongolia	10.49%	9.59%	8.78%	8.18%	7.55%
Liaoning	7.72%	7.50%	7.34%	7.06%	6.63%
Jilin	9.22%	8.63%	8.20%	7.75%	7.28%
Heilongjiang	5.56%	6.05%	6.39%	6.56%	6.44%
Shanghai	11.60%	10.79%	11.13%	9.67%	8.37%
Jiangsu	12.76%	11.35%	11.24%	9.78%	8.74%
Zhejiang	12.39%	10.52%	9.78%	8.85%	7.91%
Anhui	7.97%	8.03%	7.87%	7.73%	7.49%
Fujian	10.95%	10.01%	9.68%	8.78%	8.09%
Jiangxi	10.76%	9.71%	8.83%	8.42%	7.88%
Shandong	10.79%	9.88%	9.31%	8.74%	7.99%
Henan	9.49%	9.24%	8.87%	8.48%	8.00%
Hubei	9.82%	9.21%	8.64%	8.21%	7.66%
Hunan	7.76%	7.89%	7.77%	7.63%	7.39%
Guangdong	13.73%	12.69%	12.45%	10.57%	9.31%
Guangxi	10.98%	9.93%	8.87%	8.27%	7.65%
Hainan	19.47%	14.82%	12.07%	11.10%	9.59%
Sichuan	8.68%	8.81%	8.45%	8.36%	7.96%
Guizhou	7.78%	8.56%	8.71%	8.66%	8.37%
Yunnan	11.64%	10.49%	9.57%	8.84%	8.17%
Tibet	10.56%	-0.13%	13.60%	2.25%	11.92%
Shaanxi	7.85%	8.43%	8.57%	8.53%	8.24%
Gansu	8.75%	9.19%	9.26%	9.13%	8.75%
Qinghai	6.72%	7.94%	7.57%	8.35%	7.48%
Ningxia	8.00%	8.75%	8.99%	9.22%	8.81%
Xinjiang	11.29%	10.23%	9.38%	8.70%	8.01%

Notes: This table presents the annual growth rate with 1990 fundamentals and changes in international trade costs and migration restrictions.