The London School of Economics and Political Science

Essays on Spatial Economics

Kohei Takeda

A thesis submitted for for the degree of Doctor of Philosophy in Economics, May 2022

Declaration

I certify that the thesis I have presented for examination for the Ph.D. degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I confirm that Chapter 2 is jointly co-authored with Daniel Sturm and Anthony Venables.

I declare that my thesis consists of approximately 43,400 words excluding bibliography and appendices.

Kohei Takeda

Abstract

What are the causes and consequences of the spatial variation of economic activities both within and across cities? To contribute to our understanding of this question, the two chapters in this PhD thesis seek to advance two research agendas. The first is an understanding of the causes of the spatial variation in structural transformation in a country and how this affects spatial inequality and the upward income mobility of workers. The second is addressing the mechanisms that shape the internal structure of a city in a developing country. The approach to both is a mix of theory and empirics, leveraging the structure of the model for identification.

The first chapter develops a dynamic overlapping generations model of economic geography to explain variation in structural transformation across space and time. Despite the heterogeneity across locations, sectors, and time, the model remains tractable and is calibrated to match metropolitan area data for the U.S. economy from 1980 to 2010. The calibration allows us to back out measures of upward mobility and inequality, thereby providing theoretical underpinnings to the geographical variation of upward mobility and spatial inequality. The counterfactual analysis shows that structural transformation in the last decades has had substantial effects on mobility.

The second chapter studies how quantitative urban models can be calibrated in the data-sparse environments in developing countries using data from Dhaka. In particular, this paper shows how newly available satellite data on building heights can be used to estimate the housing supply elasticity. With the model parameters, we can also estimate the price of land and floor space in the city, which are prices that are usually difficult to observe for cities in developing countries directly. This paper also presents model counterfactuals to illustrate how essential it is to understand the general equilibrium impacts of the policy change.

(300 words)

Acknowledgements

It is no exaggeration to say that this thesis would not have been possible without Daniel Sturm. He has been a constant source of advice, encouragement, and thought-provoking discussions. I am truly grateful to Gianmarco Ottaviano for his advice, encouragement at various stages of the PhD and insightful comments on my papers. I really cannot say how grateful I am to them, and this thesis could not have been written without their help.

Many thanks to Oriana Bandiera, Robin Burgess and Steve Machin for their support toward the last stage of the PhD. The insightful discussion with them improved the first chapter of the thesis. Tony Venables co-authored the second chapter of the thesis. I have learned a lot from discussions with him. I will never be able to thank each of these remarkable people enough for all that they have taught me. A number of other individuals have kindly provided comments and suggestions on my papers. I am especially grateful to Swati Dhingra, Dennis Novy, Veronica Rappoport, Thomas Sampson and Catherine Thomas.

The financial support from the Funai Overseas Scholarship is very gratefully acknowledged. I have received additional logistic and financial support from the Economics department and the Centre For Economic Performance.

Finally, I am fortunate in having many close friends and family. Many thanks to all of them for their support and company over the years.

Contents

List of Figures

List of Tables

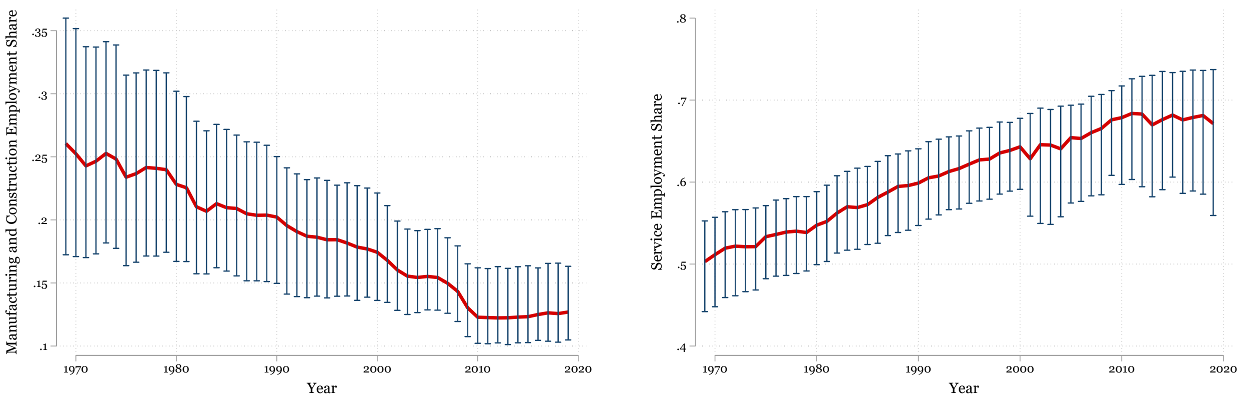
Chapter 1

The Geography of Structural Transformation: Effects on Inequality and Mobility

1 Introduction

The last half-century has seen a remarkable structural transformation of the world. While there has been sustained deindustrialization and a general shift towards the service sector in most developed countries, there is a significant variation in the extent of this structural transformation across geography within a given country. Figure 1.1 shows the change in employment shares for both the goods and services sectors in the U.S. cities over the last half-century. In the left-hand panel, both the median share of people employed in the manufacturing sector and its spatial variation represented by the inter-quartile bands have been declining in most cities since the 1970s, while in the right-hand panel, there has been concurrent growth in the share employed in the service sector. However, in contrast to the manufacturing sector, this continues to show large geographic variation. While

Figure 1.1: Change in Employment Share across the U.S. Cities (MSAs)



(a) Manufacturing and Construction **(b)** Services

Note: These figures show the change in employment share for the manufacturing and construction sector and the service sector (excluding public services) over 1969-2019. The red line shows the median across MSAs in the U.S. and the blue lines show inter-quartile ranges for any particular year. The data source is employment data from BEA.

the causes and consequences of structural transformation have been well documented at a national level, we know very little about what drives its variation across space within countries. And, importantly, the uneven impact of this structural transformation could explain both spatial inequality and geographical variation in the social mobility of workers.

This paper (i) shows how amenities and productivity spillovers are the main drivers of the geographical unevenness of structural transformation and (ii) use the model and fitted data to perform counterfactuals that allow us to trace out the consequences of this variation for inequality and mobility across cities in the U.S. To this end, we build a dynamic economic geography model that incorporates overlapping generations, multiple sectors and the frictional adjustment for workers who switch locations and industries. In their youth, workers' tastes for which industry to work in is a function the industries represented in their location of birth. Given their tastes for industry and locations, they choose cities and industries to work in later in their life, and this fuels the dynamics of labor allocation across industries. Incorporating overlapping generations of workers to characterize the evolution of labor allocation across space and industries is a novel extension of the economic geography model. Structural transformation in a given locality provides a tractable expression for understanding the key mechanisms that determine the spatial dynamics of total factor productivity (TFP), welfare, factor prices and intergenerational mobility. We then calibrate this model using data on the U.S. metropolitan areas (CBSAs) from 1980 to 2010 to obtain the amenity and productivity estimates that drive differential rates of structural transformation across locations and then trace out their effects on inequality and mobility.

The dynamic economic geography model proposed in this paper has three key components: (i) structural transformation caused by both non-homothetic preferences and differential productivity growth across sectors, (ii) a multi-location and multi-sector version of the gravity model, and (iii) barriers for workers to switch locations and industries. Conditional on the technological progress in fundamental productivity, the non-homothetic preferences of individuals between the manufacturing sector and services sector leads to a different slope of the Engel curve across workers in different locations and industries. We embed this mechanism of structural transformation in the multi-sector version of the gravity model and this enables us to consider the microstructure of spatial linkages in production and consumption. Firms are competitive and firms in each location benefit from other locations over time because they can exploit technology developed in other places through the in-migration of workers who bring knowledge with them. The different patterns of demand shifts by workers imply heterogeneous gains from trade by geography and sector, and disparity in real incomes leads to the localization and sector specialization of workers. These agglomerations are essential in the endogenous mechanisms creating the spatial variation of structural transformation and its relation to the spatial inequality in welfare.

Once we have defined the structure of demand, production and trade, we present an overlapping generation theory for workers' choice of local labor markets which drives the dynamics of labor allocation. Individuals live for two periods. In the first period, individuals choose the location and industry that will be the focus of the second period. Individual workers' decisions on where to supply labor depend on two probabilities: (i) their location choice is determined by amenities, real income and mobility costs; (ii) the choice of an industry that reflects the future expected return and exposure to the previous generation's sectors of employment in their home local labor market. Conditional on the choice of industry, lower migration costs increase the opportunity for labor mobility on geography, allowing workers to move where higher returns from work exist, leading to welfare gains. Turning to the industry choice of individuals in the first period, we introduce the simple microfoundation for the influence of the industrial composition in the previous generation on their choice. An individual receives information regarding jobs in an industry from the previous generation in the local labor market where they live in the first period. If there are a large number of workers in any particular industry among the previous generation, an individual in the next generation has more exposure to the industry and receives more information from it. This information leads to different taste values. An individual then decides on an industry that gives them the highest expected utility, taking into account their specific taste values. This, in turn, creates a path dependence in the local labor market over generations. Intuitively, an individual's choice of industry is affected by the degree of structural transformation in the local economy. This is consistent with a large body of sociological literature and empirical evidence from the study of the local labor market. In the model therefore, individuals' decisions feature two probability choices that take quite different roles in the transition of local labor markets. The former accounts for how local characteristics and spatial structure define labor supply, and the latter explains why the transition process of workers persists in some local economies.

Together with these key mechanisms which drive the geographical pattern of structural change, we provide a quantitatively oriented theory to study the consequences of the distributional effects of structural change on workers' inequality over space and time. The model allows us to characterize the local labor market dynamics with the Stolper-Samuelson effect and the Rybczynski theorem at work in the spatial economy. In equilibrium, the disparity of wages, consumption and sector-specific local agglomeration forces create cross-sectional inequality among workers. For upward mobility over generations, the two sets of workers' idiosyncratic preferences over locations and industries and the extent of structural transformation determine the equilibrium intergenerational income mobility. Therefore, our model speaks to the fundamental source of the variation of inequality and upward income mobility with a focus on the role of the geography of structural transformation.

After exploring the key qualitative and quantitative insights in the theoretical model, we calibrate the model with the data from the U.S. metropolitan areas and multiple industries. We consider 395 core based statistical areas (CBSAs) and 17 industries in the manufacturing sector and the services sector, and a construction sector. We first estimate some parameters by exploiting the structural equations in the model. We use gravity equations for internal trade and migration to estimate their elasticities, and we then estimate key parameters that determine workers' industry choice based on the data on wage and employment by industry and CBSAs, leveraging the model structure. Subsequently, we invert the model to recover the time-varying fundamental productivity and amenities by industry and CBSAs for different periods, 1980, 1990, 2000 and 2010. While we allow for high dimensions in locations, industries and time, we find that the model remains tractable and allows us to compute these fundamentals in the real economy. Based on the inverted fundamentals and computed workers' choice, we calculate the measured TFP, welfare and intergenerational inequality across space. The quantification highlights the quantitative importance of different margins in the model that determine the geographical variation of structural transformation and its impact on welfare and upward mobility.

Armed with the estimated parameters and inverted fundamentals in the economy, we perform two sets of counterfactual exercises varying (i) technological progress and (ii) local amenities. For the former, we start by quantifying the effect of fundamental technological progress on the geographical pattern of structural transformation, welfare and upward mobility. To do this, we conduct a counterfactual exercise where the evolution of fundamental productivity in the service sector shows different patterns to the baseline. We also look at what happens if information technology (IT) intensive services had not experienced technological advancement over time. Namely, we compute the counterfactual equilibrium when the fundamental productivity of communication services and finance, insurance and real estate (FIRE) was fixed after a negative shock to the baseline economy in 1990. In addition, we look at a counterfactual to assess the role of technological progress in the manufacturing sector due to the adoption of robots. We find that such fundamental productivity growth drives spatial variation in structural change via differential productivity spillovers and demand shifts. Technological progress, on average, lowers the upward mobility of workers and we find pronounced geographical variation in this effect. For the latter, we carry out a set of counterfactuals where we vary amenities across localities. In the model, fundamental amenities for workers are location and industry specific, and they include location-specific migration barriers and sector-specific taste shifters. To assess the importance of labor mobility, we first suppose that migration barriers are low. Further, we assume that the geographical variation of amenities becomes uniform so that every worker in any particular industry enjoys the same benefit from amenities across space. In these model counterfactuals, we find that the persistent variation of fundamental amenities is crucial for explaining the regional disparity in TFP changes and workers' mobility. This leads to the disparity in welfare and intergenerational income mobility among workers across CBSAs observed in the U.S. We also find that lower migration barriers yield higher geographical and income mobility for workers.

The power of the framework developed in this paper is that it is tractable and is capable of performing various counterfactual exercises to study policy interventions and their consequences of inequality among workers from both cross-sectional and intergenerational perspectives. It is applicable to a whole range of settings beyond that examined in this paper. The key finding is that interplay between structural transformation in the aggregate and local economies is critical for understanding spatial inequality and worker mobility. The dynamic nature of our spatial model allows us to study phenomena that have received limited scrutiny but which are of fundamental interest in a country which is increasingly riven by growing inequality and barriers to upward mobility. This paper addresses how the structure of the spatial economy - through trade and migration, local labor market exposures and agglomeration - shapes individual outcomes. We begin to understand why citizens in different cities in the same country have such different outcomes. Why some remain mired in the Rust Belt with limited prospects whilst others reside in the most dynamic cities on earth. We also begin to glimpse why rising inequality might constrain upward mobility thus providing microfoundations for the Great Gatsby Curve that the late Alan Krueger originally pointed to. These issues of inequality and limited mobility are perhaps the most important facing not just the U.S. but a whole range of countries across the world. This paper contributes by opening the black box of how the structure of economy can influence patterns of inequality and mobility in different locations.

This paper is related to the explanation of the structural transformation in macroeconomy (Matsuyama 1992, Caselli and Coleman II 2001, Ngai and Pissarides 2007, Matsuyama 2009, Buera and Kaboski 2012, Herrendorf et al. 2014, Matsuyama 2019, Comin et al. 2020) and the neoclassical analysis of regional disparity (Barro et al. 1991, Barro and Sala-I-Martin 1992). In the context of the spatial economy, there is a line of discussions about the sources of the diversity of spatial development: input-output linkages (Puga and Venables 1996), innovation and entrepreneurship (Brezis and Krugman 1997, Duranton and Puga 2001, citealtentrepreneurship2015), trade costs (Redding and Venables 2004, Duranton and Turner 2012, Allen and Arkolakis 2014), spatial spillover of technology (Desmet and Rossi-Hansberg 2009, Desmet and Rossi-Hansberg 2014), and amenities (Rappaport 2007, Glaeser et al. 2016). Our model integrates them to make these ideas quantitatively precise, and we propose the structural approach relating to the recent empirical findings of Hornbeck and Moretti (2020).

Theory adopts the recent modeling of non-homothetic preferences (Matsuyama 2019, Comin et al. 2020) to consider the role of heterogeneous Engel curves across local labor markets in the spatial pattern of structural transformation and inequality. We adopt the non-homothetic constant elasticity of substitution (CES) demand system for keeping the tractability of the model compared to previous works using different types of preferences. The modeling approach of dynamics is similar to that of Allen and Donaldson (2019). However, this study has different motivations. The extension of their framework to multiple sectors and the introduction of linkages between generations in labor supply add new insights into spatial inequality and worker mobility. By its nature, the model encompasses the interaction of comparative advantages and labor mobility that is focused on in Pellegrina and Sotelo (2021).

There is a list of papers that analyze the theory of dynamic equilibrium in economic geography (Krugman 1991, Matsuyama 1991, Ottaviano 1999, Baldwin 2001). At the expense of forward-looking choices, our approach provides tractability to isolate the importance of migration barriers, local labor market exposure, structural transformation and externalities in the workers' response to any particular shock. This is also our attitude toward the recent advancement in the formulation of the spatial economy with perfect foresight infinitely lived workers (Artuç et al. 2010, Dix-Carneiro 2014, Dix-Carneiro and Kovak 2017, Caliendo et al. 2018, Caliendo et al. 2019, Caliendo and Parro 2021, Kleinman et al. 2021). For labor mobility, our approach is also related to Porcher (2020) on the role of information friction in internal migration, although we underscore the past industry distribution in the decision to work within the local labor market. Among others, Michaels et al. (2012) provided the static model that studies the link between urbanization and the shift of labor from agriculture to manufacturing. Eckert and Peters (2018) considered the dynamic version of the structural transformation of the early U.S. economy. Fan et al. (2021) account for the regional difference in the service-led growth in India. Compared to them, we explicitly consider the role of spatial linkages and frictions that are abstract in their model because the spatial structure matters in assessing the fundamentals across locations.

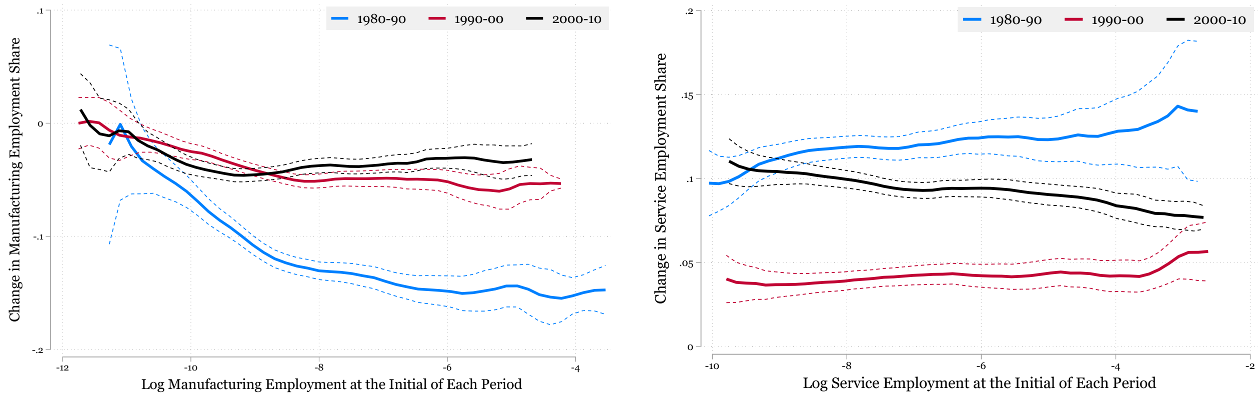
Finally, as an essential contribution, we take an approach focused on the structural mechanisms for the recent discussion on the dynamics of inequality by geography. In addition to the cross-sectional inequality across locations (Glaeser et al. 2009), this paper derives an implication for the heterogeneity in intergenerational mobility found in recent studies, including Ferrie (2005), Long and Ferrie (2013), Chetty et al. (2014), Feigenbaum (2015), Bütikofer et al. (2019), Fogli and Guerrieri (2019), and Boar and Lashkari (2021). This paper's approach and quantitative results complement their evidence, and we can obtain the absolute effects of structural change in the economy on the upward income mobility of individual workers.

The rest of this paper is structured as follows. Section 2 describes the spatial variation of structural changes and its relation to upward mobility in the U.S. Section 3 develops the model, and Section 4 describes the analytical results for accounting objectives in the model. The data and parameters for calibration and calibration procedure are described in Section 4. The results of calibration and quantitative analysis for the U.S. economy are discussed in Section 6. Armed with the data and parameters, Section 7 presents the results of the counterfactual analysis of the U.S. economy. Section 8 concludes. The Appendix contains technical details, including other results.

2 Spatial Variation of Structural Transformation in the U.S.

We start by documenting the spatial variation of structural transformation in the U.S. economy. Figure 2.2 displays the relationship between changes in employment share and initial employment level across CBSAs for the manufacturing sector and services sector over different periods, using the data on industry level employment from the county business pattern (CBP). In the left-hand panel, cities with large initial employment in the manufacturing sector showed a significant shift of workers to the services sector during 1980-1990. Although this pattern became less pronounced in the later periods, it shows that the deindustrialization of the U.S. economy has been led by cities where the size of the manufacturing sector was large. This implies that employment in the manufacturing sector has been dispersed across space over time. In the right-hand panel, the service sector exhibited a weak relationship between the change in the employment share of the services sector and the initial size of employment in the sector. This shows that the variation in the employment share of services across cities has not declined over time in contrast to the manufacturing sector. Another observation in these figures is that there is a variation in the change of employment composition for

Figure 2.2: Geography of Structural Transformation in the U.S.



(a) Manufacturing

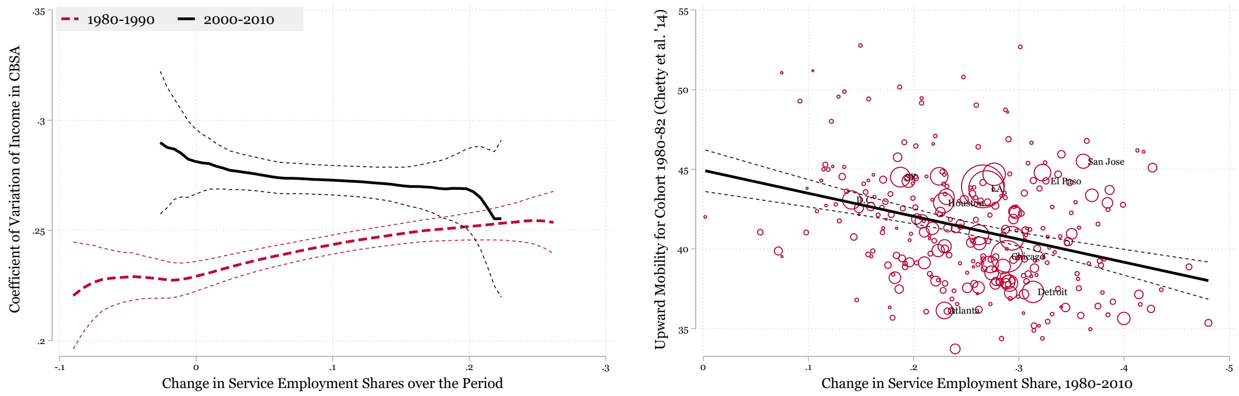
(b) Services

Note: These figures show the polynomial fitted line (local mean smoothing) for the change in employment share between different periods: 1980-1990, 1990-2000 and 2000-2010. Figure (a) shows that for the manufacturing sector, and Figure (b) shows that for the service sector. The sample includes 395 core based statistical areas (CBSAs) in the U.S. The dotted lines show 95% confidence intervals. Employment is normalized by the total employment in the economy.

specialized cities. The confidence intervals become large for cities with large size of employment in a particular sector. One logic that creates the spatial variation in this structural transformation is the differential productivity growth across space through fundamental technological differences and productivity spillovers across space. Therefore, our model allows spatial heterogeneity in fundamental productivity growth and spillovers. Another driver of this geographical unevenness in structural transformation is a significant difference in demand. In the U.S., while the average expenditure share of goods has declined and that for housing and services has increased in most places, there is considerable variation in expenditure share across cities. In addition, a larger expenditure share on services is associated with a large consumption expenditure and the relationship has been observed for different periods. See the Appendix for the expenditure share for some representative cities in the U.S. To reconcile this pattern, we incorporate the non-homothetic preferences of individuals in the model. This leads to the different slopes of the Engel curve of workers by their locations and industries. The Appendix also presents the changes in the U.S. economy in the average and standard deviation of the house price index and its relation to the structural change. The variations in the housing prices and the underlying local amenities are essential margins that account for the welfare disparity by place occurring in the structural transformation phase. Therefore, in the model, we introduce the different values of amenities for workers by location and sector and developers that supply residential stocks.

Next, Figure 2.3a shows the relationship between inequality in the local labor markets (CBSAs) and the change in service employment share. This confirms the significant income inequality in large cities where employment of services increased in the early period 1980-1990 but turned out less pronounced later. On intergenerational mobility, Figure 2.3b displays the variation in the measure of upward mobility of workers across metropolitan areas constructed by Chetty et al. (2014) and its relation to the change in the service employment share in the last decades. The measure of upward mobility represents the expected rank for children from families with below-median parents' income in the national distribution. They exploit residents born in 1980-82 and their income is evaluated in the years 2011-12, and related to the income of their parents back in 1996-2000. There is a large variation across U.S. cities in the upward mobility, and the structural transformation toward the service sector in the local labor market is associated with lower intergenerational income mobility for workers.

Figure 2.3: Inequality and Intergenerational Mobility across the U.S. Cities



(a) Urban Inequality

(b) Upward Income Mobility

Note: (a) Inequality within local labor market is measured by the coefficient of variation (CV) of income. The polynomial fitted line for CBSAs in the U.S. and dotted lines show 95% confidence intervals. (b) Measure of the absolute upward mobility comes from Chetty et al. (2014): the expected income rank for children from families with below-median parents' income in the national distribution.

Together with Figure 2.2, we notice that locations with a high employment share in manufacturing in the initial period show lower social mobility of workers, but we also find the variation of social mobility between cities, which show a similar pattern of the growth in the employment share of services(see Detroit and El Paso). In the next section, we develop a quantifiable model to consider the variation of upward mobility and its relation to structural transformation. Intuitively, more structural transformation to services inherently low productivity growth in the local economy and the lower degree of labor mobility together lead to lower upward mobility. Therefore, the current labor composition of the local economy and the pattern of structural transformation is important to create the variation of upward income mobility. Modeling with overlapping generations and workers' mobility speaks to the fundamental source of the variation of inequality and upward income mobility, focusing on the role of the geography of structural transformation.

The heterogeneity in structural transformation across space gives rise to the question of its redistributive effects across space and over generations. What are the underlying drivers that create the spatial variation of structural transformation? What is their quantitative importance in explaining the spatial inequality and upward mobility of workers in the U.S. economy? To address these questions, we build the quantifiable general equilibrium model that accommodates heterogeneous geography, fictional adjustment of workers across locations and industries, and structural transformation.

3 The Model

This section presents a model to understand the spatial heterogeneity of structural transformation and its consequence on workers over generations. The basic environment is the following. Time is discrete. A single country consists of a discrete number of locations, indexed by i, ℓ or $n \in \mathcal{N}$. We let $\mathcal{K}$ denote the set of $S+1$ industries. Among them, there are S tradable industries and a single sector providing the structure or housing services, which we refer sector 0. Each sector is indexed by j, k or s. Locations are different in fundamental productivity and amenities. Immobile landlords own the land and the total units of land are unchanged over time. At generic time t, the economy is inhabited by two overlapping generations of equal size $\bar{L}$: the old born at period $t-1$ and the young born at period t. Only the old work and consume with each of them supplying a unit of labor inelastically. Accordingly, at any time, $\bar{L}$ also represents the total number of consumers and workers in the economy. Each local labor market is characterized by a combination of location and industry. Young workers decide in which location to live and in which industry to work when old, thus potentially giving rise to intergenerational changes in employment across local labor markets. In this respect, the first period of individuals is the *formative years*. The Appendix A presents the details of each element in the model not included in the main text.

3.1 Demand, Land of Opportunity, and Exposure in Local Labor Market

We consider the individuals' decisions regarding the consumption, industry to work and location. At the initial of time $t-1$, people of generation t are homogeneous ex ante.[1] During the period $t-1$, individuals in location i observe the idiosyncratic taste shocks relating to the industry choice. They anticipate the wage and prices in the next period t and compute the expected payoff for the future. Given the expected payoff, they decide the industry, and we take that choice to be unchanged later. At the initial of period t, individuals draw and observe the taste shocks across locations and they decide location n where they live in period t. They move to the destination at the initial of period t subject to bilateral migration costs. In the location, they supply one unit of labor inelastically and decide consumption allocations. The lifetime utility of a worker ω of generation t who lived in i in period $t-1$ and works and consumes in location n and industry s in period t is:

$$\ln U^s_{ni,t}(\omega) = \ln B^s_{n,t} + \ln \mathbb{C}^s_{n,t}(\omega) - \ln D_{ni,t} + \ln z^s_{i,t}(\omega) + \ln v_{n,t}(\omega),$$

where $\mathbb{C}^s_{n,t}(\omega)$ is *subutility* function associated with consumption of individuals. The utility benefit from amenities, $B^s_{n,t}$, is common to sector s workers living in n, and migration from location i

[1] This can be easily extended to allow exogenous heterogeneity, including race and gender.

to n incurs the utility cost $D_{ni,t}$ that reflects any impediments that movers across locations face.[2] The idiosyncratic taste shocks from industry choice $z^s_{i,t}(\omega)$ depend on the origin of the worker. The second idiosyncratic shock of amenities related to location choices, $v_{n,t}(\omega)$, depends on the destination but is independent across i and s. We describe them in detail later.

For the demand system, our objective is to study the implication of demand heterogeneity across workers and locations along with the structural transformation in the economy. Therefore, we depart from the standard CES aggregation by introducing a heterogeneous income effect across sectors, keeping tractability in the substitution effect. In the baseline analysis, we adopt the implicitly additive separable consumption aggregator featuring non-homothetic CES demand discussed in Hanoch (1975) and recently Matsuyama (2019) and Comin et al. (2020). While there are alternative non-homothetic preferences used in the international trade and macroeconomics literature[3], as we discuss below in detail, the non-homothetic CES demand system has advantages: first, we keep non-homotheticity in the asymptotic; second, we easily accommodate multi-sectors; third, the elasticity of substitution between sectors is constant; and fourth, the elasticity of relative sectoral demand to aggregate demand is solely determined by parameter values. These gain tractability and entail the core mechanisms of demand shift.

Workers of generation t working in location n and sector s receive income $W^s_{n,t}$ which include labor earnings (wage) and surplus distributed among workers. We refer $\boldsymbol{p}_t = \{p^k_{n,t}\}$ to price of consumption of goods. The expenditure share of a worker with income $W^s_{n,t}$ is:

$$\psi^s_{k|n,t} = \alpha_k^{\sigma-1}\left(p^k_{n,t}/\mathcal{P}^s_{n,t}\right)^{1-\sigma}\left(W^s_{n,t}/\mathcal{P}^s_{n,t}\right)^{\theta_k-1}, \quad k \in \mathcal{K} \tag{1}$$

where $\boldsymbol{\alpha} = \{\alpha_k\}$, σ and $\boldsymbol{\theta} = \{\theta_k\}$ are exogenous preference parameters and we assume $(\theta_k - \sigma)/(1-\sigma) > 0$ for all industries.[4] $\mathcal{P}^s_{n,t}$ is the aggregate price index corresponding to the optimal consumption patterns for workers in sector s and location n that solves:

$$\mathcal{P}^s_{n,t} = \left(\sum_{k\in\mathcal{K}} \alpha_k^{\sigma-1}(p^k_{n,t})^{1-\sigma}(W^s_{n,t}/\mathcal{P}^s_{n,t})^{\theta_k-1}\right)^{1/(1-\sigma)} \tag{2}$$

Using the price index, we let $\mathcal{W}^s_{n,t}$ denote the real income for workers in location n and sector s: $\mathcal{W}^s_{n,t} \equiv W^s_{n,t}/\mathcal{P}^s_{n,t}$. We emphasize the three key elasticities for this demand system (1). First, the elasticity of substitution between sectors is constant, $1-\sigma$. Second, the elasticity of *relative* demand between two different sectors to the *aggregate demand* is specific to the pair of sectors and governed by θ_k by sector. Third, income elasticity varies across sectors and depends on expenditure patterns: individuals exhibit higher income elasticity of demand for the industry with a large θ_k. When ex-

[2]This conceptually includes moving costs between locations, the cost of job search in different locations, as well as the cost of searching for a place to live. In the quantification of the model, we allow $D_{ii,t}$ may differ across locations.

[3]The different types of non-homothetic preferences include Stone-Geary preference; price independent generalized linearity (PIGL) preference (Buera and Kaboski 2012, Eckert and Peters 2018); constant ratio of income elasticity (Fieler 2011, Caron et al. 2014); income specific elasticity of substitution between goods (Handbury 2019).

[4]This ensures the global monotonicity and quasi-concavity of the consumption aggregation.

penditure shifts to an industry with a large θ_k, the income elasticity of consumption becomes lower as the relative slopes of Engel decline for all sectors.

We let $Y^s_{n,t}$ be the aggregate income of workers in location n and sector s and $Y_{n,t} \equiv \sum_{s\in\mathcal{K}} Y^s_{n,t}$. We also define the income share of workers in any particular industry: $y^s_{n,t} \equiv Y^s_{n,t}/Y_{n,t}$. Then, the change of local expenditure share on sector k between time t and $t-1$ at the first order approximation becomes:

$$\mathrm{d}\ln\frac{E^k_{n,t}}{Y_{n,t}} = \sum_{s\in\mathcal{K}} \frac{\psi^s_{k|n,t-1} Y^s_{n,t-1}}{E^k_{n,t-1}} \left((1-\sigma)\mathrm{d}\ln\left(\frac{p^k_{n,t}}{\mathcal{P}^s_{n,t}}\right) + (\theta_k - 1)\mathrm{d}\ln \mathcal{W}^s_{n,t} + \mathrm{d}\ln y^s_{n,t} \right), \tag{3}$$

where $E^k_{n,t}$ is aggregate expenditure on sector k in location n. The local level Engel slope changes over time through substitution effect, real income change ($\mathcal{W}^s_{n,t}$) and the change of income distribution ($y^s_{n,t}$) given previous expenditure patterns.

Turning to the location choice of workers, we formally posit the followings for the stochastic factor:

ASSUMPTION 1 *An individual draws vector $\boldsymbol{v} = \{v_{i,t}(\omega)\}_{i\in\mathcal{N}}$ from the time invariant multivariate distribution: $G(\{v_{i,t}(\omega)\}) = \exp\left(-\sum_{i\in\mathcal{N}}(v_{i,t})^{-\varepsilon}\right)$. $v_{i,t}(\omega)$ and $v_{n,t}(\omega)$ are independent: $v_{i,t}(\omega) \perp v_{n,t}(\omega)$ for any $i \neq n$ conditional on industry choice.*

The shape parameter reflects the dispersion of the idiosyncratic utility. Low ε implies higher heterogeneity in taste across places to live, and $\varepsilon \to \infty$ implies that all individuals face the same order of locations in terms of the utility benefit. Under Assumption 1, the probability that a worker born in i at period $t-1$ ends up working in location n at period t conditional on choosing industry s equals:

$$\lambda^s_{ni,t} = \left(\frac{B^s_{n,t}\mathcal{W}^s_{n,t}}{D_{ni,t}\bar{U}^s_{i,t}}\right)^{\varepsilon} \quad \text{with} \quad \bar{U}^s_{i,t} = \left(\sum_{\ell\in\mathcal{N}} \left(B^s_{\ell,t}\mathcal{W}^s_{\ell,t}/D_{\ell i,t}\right)^{\varepsilon}\right)^{1/\varepsilon}, \tag{4}$$

where $\bar{U}^s_{i,t}$ is expected utility conditional on job choice s. By the law of large numbers across continuum of individuals, each element of matrix $\boldsymbol{\lambda}_{s,t} = \{\lambda^s_{ni,t}\}$ is the share of movers among individuals of generation t conditional on industry choice s. The share becomes large when the destination exhibits higher real income from consumption ($\mathcal{W}^s_{n,t}$) associated with the adjustment of amenity value ($B^s_{n,t}$) and discount of migration costs ($D_{ni,t}$). Therefore, $\bar{U}^s_{i,t}$ reflects the land of job opportunities for individuals born in i when working in industry s.

We turn to the distribution of idiosyncratic taste shocks relating to the choice of industry, $\boldsymbol{z}_t = \{z^s_{i,t}(\omega)\}$. An individual of generation $t+1$ in location i receives the discrete number of taste shocks for each sector from previous generation t during the formative period, t. An individual of generation $t+1$ (*young*) spends an entire time for job choice during period t. An individual acquires information containing taste shock from existing workers in the local labor market. An individual split one unit of time into T time spans with intervals Δ. Let $J^s_{i,t}$ refer to the probability that she receives the valuable information during Δ. Within each time span, an individual decides time allocation across different industries to maximize the logit of probabilities of receiving the valuable shocks. We let $\mathcal{O}(J^s_{i,t}, L^s_{i,t})$ denote the time required to achieve the probability $J^s_{i,t}$. This is increasing in $J^s_{i,t}$ and decreasing in $L^s_{i,t}$. Intuitively, marginal time needed for obtaining valuable information becomes small if there is a large pool of existing workers. For the objective function, an individual maximizes the average of odds that captures the chance of receiving valuable taste shocks relative to valueless ones regarding industries, with minimizing the average coefficient of variation for the binomial distribution of receiving the number of valuable taste shocks in a unit time.[5] Specifically, during time span Δ, an individual of generation $t+1$ in location i solves the following problem:

$$\max_{j^s_i \in (0,1)} \left\{ \sum_{k\in\mathcal{K}} \ln \frac{j^k_i}{1-j^k_i} \text{ s.t. } \sum_{k\in\mathcal{K}} \mathcal{O}(j^k_i, L^k_{i,t}) \leq \Delta, \text{ and } \mathcal{O}(j^k_i, L^k_{i,t}) \equiv \frac{1}{\zeta_{k,t}} \ln\left(\frac{1}{1-j^k_i}\right) (L^k_{i,t})^{-\eta} \right\}, \tag{5}$$

and we let $J^s_{i,t}$ refer to its solution. The first constraint is time constraint. In the specification for $\mathcal{O}(J^s_{i,t}, L^s_{i,t})$, $\zeta_{s,t}$ and η are strictly positive constants. $\zeta_{s,t}$ is a scale shifter and η quantifies how much an individual can save time when there are more existing workers in the local labor market. Taking the limit $\Delta \to 0$, the problem above can lead to the number of shocks an individual of generation $t+1$ receives during a unit of time following Poisson distribution with arrival rate $J^s_{i,t}$. Further, to gain the tractability, the value of each shock is supposed to be following Pareto distribution with the shape parameter ϕ and shocks are independent. A small value of ϕ implies fat tail distribution for the size of shocks. Intuitively, if ϕ becomes small, an individual is more likely to receive a higher value of shock in job choice, leading to more idiosyncrasy in the industry choice. Summarizing the assumptions about the taste shocks that an individual of cohort $t+1$ receives:

ASSUMPTION 2 *An individual of cohort $t+1$ solves (5) and consider the limit case $\Delta \to 0$ to characterize the distribution for the number of arrival shocks. The value of each taste shock follows independent Pareto distribution with shape parameter $\phi > 1$.*

Intuitively, this assumption argues that individuals face the *consideration set* when deciding future industry and location of work, and the set is influenced by workers' exposure to the historical employment composition. Given the set, individuals make their decisions following subjective expectations about future returns. Let $m^s_{i,t}(\omega)$ be the number of shocks an individual receives from location i and industry s. An individual decides industry s to work in if and only if:

$$s \in \left\{ k \ : \max_{m\in\{1,2,\cdots,m^k_{i,t}(\omega)\}} \bar{U}^k_{i,t} z^{k\ (m)}_{i,t} \geq \max_{s'\in\mathcal{K}} \max_{m\in\{1,2,\cdots,m^{s'}_{i,t}\}} \bar{U}^{s'}_{i,t} z^{s'\ (m)}_{i,t} \right\}.$$

Under Assumption 2, the share of cohorts $t+1$ in location i that choose industry s becomes:

$$\varsigma^s_{i,t+1} = \zeta_{s,t}(L^s_{i,t})^\eta \left(\frac{\bar{U}^s_{i,t+1}}{V_{i,t+1}}\right)^\phi \quad \text{with} \quad V_{i,t+1} \equiv \left(\sum_{k\in\mathcal{K}} \zeta_{k,t}\left(L^k_{i,t}\right)^\eta \left(\bar{U}^k_{i,t+1}\right)^\phi\right)^{1/\phi}. \tag{6}$$

[5] The coefficient of variation captures the relative variation of the number of valuable information over the average number of valuable information. Minimizing such variation is isomorphic to maximizing the logit.

The matrix $\varsigma_t = \{\varsigma_{i,t}^s\}$ closes the individuals' decision process. The share of individuals, $\varsigma_{i,t+1}^s$, depends on three components. The shifter $\zeta_{s,t}$ translates the macro effect in the industry choice that is common across locations. The large probability of choosing sector s is associated with large size of employment in the previous generation ($L_{i,t}^s$): more existing workers in the local labor market can save the marginal cost of information acquisition and it turns to be a large expected number of shocks that arrive to young generation *ceteris paribus*. Intuitively, the more people you meet who work there, the more likely you meet someone who prefers it and transmits to you the love for the profession. This result can be interpreted as a path dependence in job choices in the local labor market over generations. Lastly, individuals of cohort $t+1$ choose sector s with high probability when conditional expected utility ($\bar{U}_{i,t}^s$) is large since it determines the advantage of industry s in terms of net gain for their future.

This formulation under Assumption 2 is related to empirical evidences of intergenerational linkage in job choices and work behavior in labor economics[6]. In particular, the specification may capture the path dependence in the local labor market through education. Some U.S. manufacturing cities, including Buffalo, Cincinnati and Youngstown (Ohio), have underdeveloped the infrastructure to educate young generations for a long time, and the number of high schools and college graduates has been low in these cities. For these cities, the industrial specialization leads to the underinvestment into education: workers of steelmaking or paper-pulping tied to specialized industries did not have any motivations for higher education or education for the new technology in services. Therefore, specialization of the industry has a long-term effect over generations through the accumulation of schooling.[7] The specification of workers' idiosyncratic taste shocks also reflects the recent literature in the intergenerational transmission of preference apart from the endogenous creation of human capital or productivity.[8]

3.2 Technology and Trade

The production side builds on the multi-sector and multi-location Ricardian model embedded with input-output linkages and externalities from agglomeration. In each sector, there are final good pro-

[6]The intergenerational linkage in the job choice found in the literature is one potential feature behind the recent trend of intergenerational mobility, as discussed in Corak (2013). Loury (2006) showed that around half of jobs are found in the network among relatives and friends in the U.S., and the highest wage was paid to workers who found the job through male relatives in the prior generation, and Kramarz and Skans (2014) showed that young workers find the first stable job in a parent's firm, and the effect is more substantial for low skilled jobs. Corak and Piraino (2011) found the direct evidence on intergenerational transmission of employers in Canada;

[7]To consider the movement of people for education, we extend the baseline model to include the additional choice of individuals for education. See subsection 3.5 for further discussion.

[8]The relationship between generations in the job choice can be explained by the (unobserved) transmission of taste or preference through formal or informal social interactions (Manski (2000)) instead of investment of education or financial assets. Doepke and Zilibotti (2008) highlighted the impact of the previous economic environment on the formation of the preference among the future generation. Dohmen et al. (2012) found that the risk attitudes in preference are transmitted from parents to children and there is a neighborhood effect in the transmission. Fernández et al. (2004) and Fernandez and Fogli (2009) suggested that there are significant effects of female labor participation in the previous generation on the work and fertility behavior among the second generation, and this reflects the persistence of the formation of preference between generations.

ducers and intermediate good producers. In each location, final good producers supply consumption goods and materials in a competitive fashion that are consumed locally. They use sector-specific intermediate goods, and their technology is constant elasticity of substitution. The time span of each period is not too short, and final goods are produced and used as inputs simultaneously in each period.

Intermediate goods' as well as the factors' markets are perfectly competitive. Intermediate goods are produced using labor and materials exploiting a Cobb-Douglas function. Firms face location and sector specific productivity $\boldsymbol{Z} = \{Z^s_{i,t}\}$ and firm specific productivity that is drawn from Fréchet distribution with shape parameter $\kappa_s > 1$ in the wake of Eaton and Kortum (2002). Intermediate goods can be traded incurring a sector-specific iceberg trade cost, so that delivering one unit of an intermediate good from n to i requires $\tau^s_{in,t} \geq 1$ units, with $\tau^s_{ii,t} = 1$. The probability that final producers of sector s in location i source intermediate goods from location n is:

$$\pi^s_{in,t} = \frac{(\tau^s_{in,t}\Xi^s_{n,t}/Z^s_{n,t})^{-\kappa_s}}{\sum_{\ell\in\mathcal{N}}(\tau^s_{i\ell,t}\Xi^s_{\ell,t}/Z^s_{\ell,t})^{-\kappa_s}} \quad \text{with} \quad \Xi^s_{n,t} = (w^s_{n,t})^{\beta_s} \prod_{j\in\mathcal{K}\backslash 0} (p^j_{n,t})^{\beta_{sj}} \tag{7}$$

In turn, price of final good in location i for consumers is:

$$p^s_{i,t} = \Gamma_s \left(\sum_{\ell\in\mathcal{N}} \left(\tau^s_{i\ell,t}\Xi^s_{\ell,t}/Z^s_{\ell,t}\right)^{-\kappa_s}\right)^{-1/\kappa_s} \tag{8}$$

where $\Gamma_s \equiv \Gamma\left(1 - \frac{\tilde{\kappa}-1}{\kappa_s}\right)^{1/(1-\tilde{\kappa})}$ is constant. The gravity structure of regional trade characterized by (7) and (11) summarize the spatial linkage of goods.

The aggregate productivity in the local production place is increasing in employment size and evolves through the spatial spillovers. We make the following assumption:

ASSUMPTION 3

$$Z^s_{i,t} = A^s_{i,t}\left(\sum_{n\in\mathcal{N}} L^s_{in,t} Z^s_{n,t-1}\right)^{\rho}\left(L^s_{i,t}\right)^{\gamma_s}$$

for all $i \in \mathcal{N}$ and $s \in \mathcal{K}\backslash 0$.

The fundamental productivity $A^s_{i,t}$ changes over time to reflect the technological change in sector s in the local economy. Suppose that $\rho = 0$. Then, productivity increases in the size of local workers to power $\gamma_s > 0$, which naturally arises when economies of scale exist. Suppose that $\rho > 0$. Each location benefits from other locations through workers (including stayers) who have ideas of sector s. Then, the formulation of productivity spillover in Assumption 3 captures has two features. First, the "technology" is embodied with workers in tacit form (Polanyi 1958), and it moves across locations over generations. Intuitively, a large inflow of workers from productive places enhances local productivity. This is microfounded by the movement of workers who produce ideas based on the knowledge accumulated in the previous places. Second, technology spillover across space hinges on the local economic conditions. Intuitively, the inflow of workers ($L_{in,t}$) reflects the current economic condition in location i. Therefore, gains from the productivity spillover are high in the location with high real income. This is in line with the classical study of technology diffusion across space (Griliches 1957). The exogenous environment may create a random difference in productivity across space through $A^s_{i,t}$, while employment growth and flow of "ideas" create the self-organizing technological advancement across space that is related to labor mobility and demand-led growth.[9]

3.3 Development of Residential Stocks

Sector 0 denotes the residential structure. The structures are produced by a competitive developer sector who can convert structures over the residential land $\boldsymbol{T} = \{T_i\}$. We let $h_{i,t}$ refer to the stock of structure per unit of land in period t and $\bar{h}_i$ refer to the constant depreciation rate. The production technology of a developer sector exhibits constant return to scale. Letting $l^0_{i,t}$ be the employment per unit of land for the development sector, the technology of developers is:

$$h_{i,t} = \nu_i (l^0_{i,t})^{\chi} ((1 - \bar{h}_i) h_{i,t-1})^{1-\chi} \tag{9}$$

Therefore, we think of development as the process of adding structure to the previous stocks by exploiting labor. The share of labor in construction is χ and the location specific productivity ν_i is unchanged over time.[10]

We consider the bidding process for developers to obtain the right to develop the place by paying rent to landlords. Letting $r_{i,t}$ be the bidding price per unit of land, the aggregate surplus extracted from developers in location i through bidding becomes:

$$R_{i,t} = r_{i,t} T_i = (1 - \chi) \nu_i p^0_{i,t} (L^0_{i,t})^{\chi} ((1 - \bar{h}_i) H_{i,t-1})^{1-\chi} \tag{10}$$

Landlords in each location collect the surplus $R_{i,t}$ and the total land rent is equal to the share of land in the total cost of production. Given the fixed amount of land, the bidding price for a unit of land is determined endogenously to balance the total endowment of land and the surplus from development of land.

Lastly, we make an assumption about the division rule of the surplus among the population to take the general equilibrium effects into account.

ASSUMPTION 4 *In each location, individuals hold a portfolio of land that is proportional to their labor earnings share.*

[9]When $\rho = 0$ and $\gamma_s = 1/\kappa_s$, this specification is isomorphic to the new economic geography model in which the mass of firms is proportional to the mass of labor due to the fixed cost of entry and monopolistic competition. Nevertheless, in the present model, the agglomeration forces work as externalities in production, but not through love of variety or extensive margins. Hence, the results of quantification are different. See discussion in the Appendix A.2.

[10]This is in line with Davis and Heathcote (2005) that show almost no change in productivity in the U.S. construction sector.

On top of the tractability, Assumption 4 does not distort the income distribution at the location since income is proportional to wage.[11] We refer $\widetilde{W}_{it}$ to the total labor earnings of generation t in location i and we let $\mu_{i,t} = 1 + R_{i,t}/\widetilde{W}_{i,t}$. Then, income of each individual of sector s in location i becomes $W_{i,t}^s = \mu_{i,t} w_{i,t}^s$.

3.4 Equilibrium and Aggregate Dynamics

This subsection describes the aggregation in the economy and defines the equilibrium. Combining individuals' choices in self-selection in (6) and the gravity structure of migration in (4) determine the spatial allocation of labor and its dynamics:

$$L_{i,t}^s = \sum_{n \in \mathcal{N}} \lambda_{ni,t}^s \varsigma_{n,t}^s L_{n,t-1}, \tag{11}$$

where $L_{n,t-1}$ is the total population of generation $t-1$ that choose location n. This equilibrium conditions supposes that the *ex ante* indirect utility of generation t born in i is equalized and the value of the outside option for generation t born in i becomes equal to $V_{i,t}$ to preserve the total population over generations.[12]

The market clearing conditions for final goods imply that the total value of production of sector s is:

$$X_{i,t}^s = \sum_{j \in \mathcal{K} \setminus 0} \beta_{js} \sum_{n \in \mathcal{N}} \pi_{ni,t}^j X_{n,t}^j + \sum_{k \in \mathcal{K}} \psi_{s|i,t}^k W_{i,t}^k L_{i,t}^k, \tag{12}$$

where, on the right-hand side, the first term is demand from intermediate producers in location i for the use of materials, and the second term is aggregate demand from individuals consumption.[13] Analogously, the market clearing condition for residential stocks is:

$$p_{i,t}^0 H_{i,t} = \sum_{k \in \mathcal{K}} \psi_{0|i,t}^k W_{i,t}^k L_{i,t}^k. \tag{13}$$

The right-hand side is the total expenditure on housing of workers in location i and $\psi_{0|i,t}^k$ captures the different expenditure pattern of workers by their sector. The labor market of industry s in

[11] Another way of distribution rule is that the total land rent is divided among people with equal share. Then, the income becomes $w_{i,t}^s + R_{i,t}/L_{i,t}$. The drawback of this specification is that the income ratio between workers in different sectors is not preserved. This feature is not convenient in the analysis of the inequality among workers. However, the definition of competitive equilibrium is not largely different from this assumption. In Caliendo et al. (2019), land is owned by a national investment fund to which all workers participate with shares taken from the data. In the present model, land is locally owned by local workers. Hence, in their case land prices do not affect the location decision, while in ours they do.

[12] Let $\mathbb{V}_{i,t}$ be the value of outside option for workers of generation t born in location i. If $V_{i,t} < \mathbb{V}_{i,t}$, people move to outside option and total population of generation t is strictly lower than $L_{i,t-1}$. If $V_{i,t} = \mathbb{V}_{i,t}$, we suppose that people stay in the economy and total population of generation t is equal to $L_{i,t-1}$. When $V_{i,t} > \mathbb{V}_{i,t}$, potentially people in outside option enter into the economy, therefore total population of generation t is equal to or more than $L_{i,t-1}$. The baseline analysis supposes that $V_{i,t} = \mathbb{V}_{i,t}$ in equilibrium to equalize the total population of generation t to $L_{i,t-1}$, and $\mathbb{V}_{i,t}$ is determined endogenously according to (6).

[13] To simplify the discussion, the baseline analysis does not include the net export to the international market although it is straightforward to include the exogenous term of the net export.

location i clear at each point of time:

$$\begin{aligned} w_{i,t}^{s}L_{i,t}^{s} &= \beta_s \sum_{n\in\mathcal{N}} \pi_{ni,t}^{s} X_{n,t}^{s}, \\ w_{i,t}^{0}L_{i,t}^{0} &= \chi\nu_i p_{i,t}^{0}(L_{i,t}^{0})^{\chi}[(1-\bar{h}_i)H_{i,t-1}]^{1-\chi} \end{aligned} \tag{14}$$

where β_s is the labor share of sector s in production of intermediate goods and χ is the labor share in development of residential stocks. To close the description of the aggregate economy, $\sum_{i\in\mathcal{N}} L_{i,t} = \bar{L}$ for all period t. This implies that the total population size is fixed at the national level.

We now define the equilibrium in the economy. The notations are following: $\boldsymbol{\mathcal{F}}_t$ denotes the set of time-varying fundamentals including migration costs between locations ($D_{ni,t}$), trade costs ($\tau_{in,t}^{s}$), exogenous productivity growth ($A_{i,t}^{s}$), amenities ($B_{i,t}^{s}$) and exogenous shifter of macroeconomy taste ($\zeta_{s,t}$), and $\overline{\boldsymbol{\mathcal{F}}}$ denotes the set of time-invariant fundamentals that consist of efficiency in development of housing (ν_i), re-structuring parameter ($\bar{h}_i$) and endowment of land (T_i). The initial state, $\boldsymbol{\mathcal{G}}_0$ includes the initial population distribution in the economy, the initial productivity ($Z_{i,0}^{s}$) and the initial endowment of residential structure (i.e., housing). $\boldsymbol{\Omega}$ denotes the set of parameters associated with demand system, choice of individuals, migration elasticity, production technology, trade elasticities, and productivity spillover. Then, variables of interest are dynamics of $(\boldsymbol{\psi}_t, \boldsymbol{\lambda}_t, \boldsymbol{\varsigma}_t, \boldsymbol{\pi}_t, \mathbf{p}_t, \mathbf{w}_t, \boldsymbol{H}_t, \boldsymbol{r}_t)$: expenditure patterns, location choice of workers, sector choice of workers, the pattern of trade, price of consumption goods and housing, wage, amount of residential structure and land rent.

DEFINITION 1 *Given* $(\boldsymbol{\mathcal{F}}_t,\ \overline{\boldsymbol{\mathcal{F}}},\ \boldsymbol{\mathcal{G}}_0,\ \boldsymbol{\Omega})$, *the dynamic equilibrium of the economy is characterized by endogenous sequences of:* $\boldsymbol{\psi}_t$ *solving utility maximization,* $\boldsymbol{\lambda}_t$ *determined by (4),* $\boldsymbol{\varsigma}_t$ *determined by (6),* $\boldsymbol{\pi}_t$ *determined by (7),* $\mathbf{p}_t$ *that solve market clearing conditions (12) and (13),* $\mathbf{w}_t$ *that solves labor market clearing condition (14), and* $\boldsymbol{H}_t$ *and* $\boldsymbol{r}_t$ *solving profit maximization of developers (9) and (13).*

The dynamic equilibrium describes the full transition of economic activities over time and space. The Appendix B.1 presents the forward solution in which model is solved given the pre-period state. To guarantee the uniqueness of the forward solutions, we need assumptions on parameters of (i) variation of idiosyncratic shocks, (ii) trade elasticity, (iii) non-homotheticity of demand system, and (iv) externalities in productivity. Intuitively, larger variation in labor mobility (ε and ϕ) and trade (κ_s) and difference in expenditure patterns (θ_s and σ) across workers are related to more labor mobility in the equilibrium, while lower agglomeration forces (γ_s) prevents the concentration of workers as in *black hole.* For the concrete discussion, we consider the special case in which $\rho = 0$ and $\chi = 1$. This implies that the externalities in productivity are purely local economies of scale and the supply of residential stocks is elastic. In this case, the dynamic equilibrium conditional on the initial state is unique when $\gamma_s \leq \frac{\theta_s-\sigma}{\kappa_s+(1-\sigma)}\left(1+\frac{1}{\varepsilon}\right)$. This condition is intuitive. When $\varepsilon \to \infty$, the idiosyncratic shocks for migration are homogeneous across workers, and it requires a small value of γ_s to avoid generating multiple equilibria. If θ_s becomes large, the condition becomes slack as large heterogeneity in consumption across workers of different incomes leads to more dispersion. This condition in the special case ($\rho = 0$ and $\chi = 1$) is conservative since productivity spillover is purely local agglomeration and congestion force from land is small. Therefore, we consider the condition as a bound for the unique dynamic equilibrium conditional on the initial state. The Appendix B.1 displays the analytics for the special cases. We can also compute the dynamic path recursively from any state of the economy, and this is the tractable way to analyze the spatial dynamics featuring structural change, regional specialization and inequality. The Appendix B.2 describes the system of equations and solution methods. While the main aim of the model is a characterization of the transition process, the level of the spatial distribution of economic activities in the (very) long run is characterized by the *stationary steady-state equilibrium* in which all aggregate variables are constant given that the exogenous time-varying factors ($\boldsymbol{\mathcal{F}}_t$) are constant ($\boldsymbol{\mathcal{F}}^*$). The following statement gives sufficient conditions for the uniqueness of the steady-state equilibrium in this economy.

PROPOSITION 1 *Suppose that there exists a sequence of fundamentals such that $\boldsymbol{\mathcal{F}}_t \to \boldsymbol{\mathcal{F}}^*$. Then, the stationary steady-state equilibrium exists. The steady-state is unique under the regularity conditions:*

$$\boldsymbol{\Upsilon} \geq 0, \quad \max_{s \in \mathcal{K} \setminus 0} \max_{(i,n) \in \mathcal{N} \times \mathcal{N}} \left| \frac{\partial \ln X_n^s}{\partial \ln L_i^s} \right| < 1,$$

$$\sup \left| \mathcal{E}_{ii}^s + \sum_{n \in \mathcal{N}} \frac{L_{in}^s}{L_i^s} \frac{\partial \ln \mathcal{E}_{in}^s}{\partial \ln L_i^s} \right| < 1, \quad \left| \sum_{(n,k) \neq (i,s)} \mathcal{E}_{in}^s \frac{\partial L_n^k}{\partial w_i^s} \right| \geq \max_{(\ell,s') \neq (i,s)} \left| \sum_{(n,k) \neq (i,s)} \mathcal{E}_{in}^s \frac{\partial L_n^k}{\partial w_\ell^{s'}} \right|$$

where $\boldsymbol{\Upsilon}$ is matrix in which each entry is elasticity of export from local market (i, s) to wage of other local market (ℓ, s'). $\mathcal{E}_{in}^s \equiv \lambda_{in}^s \cdot \varsigma_n^s$ is the transition probability for workers sorting into sector s and moving from n to i.

The set of conditions argues the following. The first and second condition implies that the linkage between the local labor market through trade shows the regularity conditions. The third condition argues that labor mobility across space is large enough not to be clustered in one location, and the last condition is about the regularity condition for the linkage in local labor markets. The Appendix B.3 shows the manipulation of the system of equations for the steady-state equilibrium and discusses its uniqueness.

3.5 Discussion of the assumptions and possible generalization

Efficient labor. The taste shock in the industry choice in the model is crucial to characterize the aggregate equilibrium straightforwardly. It is not isomorphic to the model where an individual worker draws a vector of idiosyncratic labor efficiency she can supply. With non-homothetic preference, its realization determines a worker's real income that is not linear in the labor efficiency. Therefore, the choice probabilities of workers become different and depend on the realization of labor efficiency. This leads to complications in the characterization of the aggregate equilibrium conditions.

Education. The baseline model abstracts any endogenous mechanisms that generate heterogeneity of labor supply and productivity among workers. The framework can be extended to include an explicit education choice. Workers are supposed to differ in terms of not just sector and location but also education level. Consider two different education levels, for instance, graduate or non-graduate. During the first period, an individual decides whether to obtain graduate education and do so in the city of birth or other cities. Assume that she can only leave the city of birth to obtain graduate education in the junior period. Other choices are the same as in the baseline model. Introducing additional idiosyncratic factors in the net return of education can formulate the probabilities of education choice by similar representations. See the Appendix A.5 for details.

Infinitely lived workers with perfect foresight. Individuals work only in the second period of their life. Other approaches to seeing the dynamics entail infinitely lived workers with perfect foresight (McLaren 2017, Caliendo et al. 2019, Caliendo and Parro 2021, Kleinman et al. 2021). Infinitely lived workers and households determine the future path of mobility in a forward-looking way, taking into account future shocks. Their choice of a given location is based on current real income but also an option value associated with that location. Comparing such an approach and the present approach, *forward* solutions of the model upon the transitory shocks are different, and therefore different transitions arise. At the expense of forward-looking choices, the present approach provides tractability to isolate the importance of migration barrier, local labor market exposure, structural transformation and externalities over space in the workers' long-run response to the common shocks. With such externalities and lower costs of labor mobility, there may exist the potential issue of self-fulfilling prophecy and multiplicity of transitions that hinges on expectations rather than the past, and it is challenging to characterize the option values by sector and geography and discuss the intergenerational link (Krugman 1991, Matsuyama 1991, Ottaviano 1999, Baldwin 2001). It is also noted that there is the equivalence between the two approaches when considering the *backward* solution to back out the past fundamentals in the economy from the steady state. The Appendix A.6 presents details.

4 Dynamics of Spatial Economy and Inequality

This section derives positive and normative analytical results regarding how changes in exogenous fundamentals shape the spatial disparity of productivity, welfare and inequality along with the transition. Throughout this section, the fundamental amenities, sector-specific taste parameters and migration costs are assumed to be unchanged. First, subsection 4.1 considers the transition dynamics for the total factor productivity (TFP) in local economy and its spatial variation, then discuss welfare gains and losses in the transition. Next, subsection 4.2 derives the model's implications for the dynamics of the equilibrium prices in the local labor market. Lastly, the model's simple framework speaks to the spatial difference in the degree of intergenerational mobility in subsection 4.3. In the Appendix, Section C provides details associated with these analytical results and other results and Section D presents the numerical results of equilibrium to explore the pattern of structural transformation and inequality for the simple economy.

4.1 Measured Local TFP and Welfare Dynamics

The first objective is to see how exogenous shocks in the economy change the local level TFP differently by geography. Intuitively, the remoteness of the production places in the regional trade network, the pattern of migration and local labor exposure in the sectoral choice together define the geographical variation of local TFP change. Let $\delta^s_{i,t}$ denote the local TFP of sector s in location i, the following proposition summarizes them:

PROPOSITION 2 *Suppose that there is a common shock in the fundamental productivity in period t. The change in measured TFP in the local economy is:*

$$\frac{\mathrm{d}\ln\delta^s_{i,t}}{\mathrm{d}\ln A^s_{i,t}} = 1 - \frac{1}{\kappa_s}\frac{\mathrm{d}\ln\pi^s_{ii,t}}{\mathrm{d}\ln A^s_{i,t}} + \sum_n \left(\rho\widetilde{z}^s_{in,t} + \gamma_s\widetilde{l}^s_{in,t}\right)\left(\frac{\mathrm{d}\ln\lambda^s_{in,t}}{\mathrm{d}\ln A^s_{i,t}} + \frac{\mathrm{d}\ln\varsigma^s_{n,t}}{\mathrm{d}\ln A^s_{i,t}}\right) \tag{15}$$

where $\widetilde{z}^s_{in,t} \equiv \frac{L^s_{in,t}Z^s_{n,t-1}}{\sum_\ell L^s_{i\ell,t}Z^s_{\ell,t-1}}$ is the contribution of location n in the baseline equilibrium; and $\widetilde{l}^s_{in,t} \equiv \frac{L^s_{in,t}}{\sum_\ell L^s_{i\ell,t}}$ is the share of workers' inflow. In the steady state, the local TFP converges to:

$$\ln\delta^s_i = -\frac{1}{\kappa_s}\ln\pi^s_{ii} + \sum_n K^s_{in}\Big(\ln A^s_n + (\gamma_s+\rho)\ln L^s_n + \rho\Delta^s_n\Big) \tag{16}$$

where K^s_{in} is (i,n)-th element of the matrix $\mathbf{K}^s \equiv \sum_{m=0}^{\infty}\rho^m\left\{\lambda^s_{in}\varsigma^s_n\widetilde{l}^s_{in}\right\}^m$ and Δ^s_n is a small positive constant.

To a common shock to the technology of sector s in the economy at period t, the second term in (15) reflects the gains from trade: an increase in local TFP is associated with more export to other locations. A small trade elasticity (κ_s) leads to a large variation of local TFP gains *ceteris paribus*. The third term in (15) conflates the scale effect and spillover from the in-migration of workers. A large value of scale economies (γ_s) and spillover effect (ρ) are associated with the significant variation of local TFP gains *ceteris paribus*. An increase in sectoral productivity leads workers away from the sector, and its reallocation differs by location according to the industrial specialization. Therefore, higher mobility of labor and a higher degree of industrial specialization leads to a large variation of local TFP gains. In the steady state, the first term in (16) captures the comparative advantage in trade, and the matrix $\{K_{in}\}$ is the matrix summarizing the linkages between productivity in other locations and the local labor market. See the Appendix C.1 for derivation.

Next, we consider the welfare dynamics in the transition of the economy. Our interests are the spatial difference in welfare change and its decomposition into margins in the model. To this end, $V_{i,t}$ in (6) is a measure of welfare for individuals of generation t who have origin i. Then, the welfare change of individuals between two consecutive generations of workers who have the same origin is given by the following proposition:

PROPOSITION 3 *In the dynamic equilibrium, change of welfare measure over generations* $\dot{V}_{i,t}$ *is proportional to:*

$$\prod_{s\in\mathcal{K}\backslash 0}\left(\dot{\lambda}^{s}_{ii,t}\right)^{-1/\varepsilon}\left(\dot{\varsigma}^{s}_{i,t}\left(\dot{L}^{s}_{i,t-1}\right)^{-\eta}\right)^{-1/\phi}\left(\left(\dot{e}^{s}_{s|i,t}\right)^{\frac{1}{1-\sigma}}\left(\prod_{j}\left(\frac{\dot{w}^{j}_{i,t}}{\dot{\delta}^{j}_{i,t}}\right)^{\beta_s}\left(\dot{\pi}^{j}_{ii,t}\right)^{-\frac{1-\beta_s}{\kappa_j}}\right)^{-\widetilde{\beta}_{sj}}\right)^{\widetilde{\theta}_s} \tag{17}$$

where $e^{s}_{s|i,t}$ *is expenditure on sector* s *by workers in sector* s *and location* i*,* $\widetilde{\beta}_{sj}$ *is an element of matrix* $(\mathbf{I}-\widetilde{\mathbf{B}})^{-1}$ *with* $\widetilde{\mathbf{B}}\equiv\{\beta_{sk}\}$*, and* $\widetilde{\theta}_s\equiv(1-\sigma)/(\theta_s-\sigma)$*.*

The Appendix C.2 presents details of derivation. The first term is the change in non-migration probability with elasticity $-1/\varepsilon$. Conditional on the sector choice, $\dot{\lambda}^{s}_{ii,t}$ is expected to be declining as migration frictions are smaller, *ceteris paribus.* This term depends on the responses of labor mobility across all local labor markets to arbitrary changes in the environment and summarizes the degree of the land of opportunity for workers. When $\varepsilon\to\infty$, idiosyncratic shocks in location choice are homogeneous, and gains from migration become zero. The second term captures how flexibly workers move across sectors or how labor is specific to the sector. Greater job opportunity for workers in location i is associated with less labor specificity to the sectors in their origin. Instead, a huge distortion in the sector choice (ς^{s}_{i}) implies a lower opportunity for the future location choice, and it turns out to lower welfare gain in dynamics. Given these endogenous responses, the large heterogeneity in the taste shocks across industries (small ϕ) leads to greater welfare changes as it allows the variety of industry choices during the young for workers or less labor specificity. The local labor market externalities lead to further job opportunities for sector s when the sector exhibits employment growth in the previous period.

Apart from these choice probabilities of individuals, the last part in the welfare dynamics stands for the change of real income from the consumption of tradable goods. With a non-homothetic demand system, change in demand for sector s is decomposed into the change in expenditure patterns, change of purchasing power in the local market and change in terms of trade. Comparing the non-homothetic demand and homothetic demand ($\widetilde{\theta}_s=1$), the welfare growth to the local price change depends on the curvature of the local Engel curve. If the local Engel curve shows a relatively high slope (i.e., $\widetilde{\theta}_s>1$), the size of welfare change and its spatial variation becomes large.

These welfare dynamics relate to the key mechanisms of reallocation of workers along with the structural transformation in the model. Large migration opportunities, job opportunities, and consumption opportunities provide an incentive for workers to move to the local labor market, and production relocates to the place in response to the productivity changes and demand shift. The spatial linkages between local labor markets determine the distributional effects of TFP change and welfare change over time.

4.2 Measures of Local Market Dynamics

In the model, the key measures for the local labor market are twofold. First, employment distribution is given by the share of employment across industries, $f^s_{i,t} \equiv L^s_{i,t}/\sum_{k\in\mathcal{K}} L^k_{i,t}$ and its variation across space represents the geography of structural transformation. Another one is the fraction of income for different industries $y^s_{i,t} \equiv w^s_{i,t}L^s_{i,t}/\sum_{k\in\mathcal{K}} w^k_{i,t}L^k_{i,t}$ characterize the income distribution. These two measures are sufficient statistics of the local labor markets. By definition, the change of employment ($\mathrm{d}\ln \mathbf{L}$) and change of income ($\mathrm{d}\ln \mathbf{Y} = \mathrm{d}\ln \mathbf{w} + \mathrm{d}\ln \mathbf{L}$) characterize the change of the fraction of employment ($\boldsymbol{f} = \{f^s_{it}\}$) and income ($\boldsymbol{y} = \{y^s_{i,t}\}$) conditional on previous equilibrium state.

The employment evolves by location and industry choice of workers (11), and labor market clearing condition for each industry and location pins down income distribution that is consistent with the employment growth in the local labor market. For the transition dynamics of wages, the closed-form representation is given by the following proposition:

PROPOSITION 4 *Suppose that $\mu_{i,t} = 1$. Then, wage growth for generation t satisfies:*

$$\mathrm{d}\ln w^s_{i,t} = \sum_j \varrho^{sj}_{i,t-1}\left(-\frac{\mathrm{d}\ln\bar{\psi}^j_{i,t}}{1-\sigma} + \frac{1}{\varepsilon}\frac{\bar{\theta}_{i,t-1}-\sigma}{1-\sigma}\sum_n \widetilde{\lambda}^j_{in,t-1}\mathrm{d}\ln\lambda^j_{nn,t}\right) - \varrho^s_{i,t-1}\left(\sum_j \widetilde{\Psi}^j_{i,t-1}\mathrm{d}\ln\delta^j_{i,t}\right) \tag{18}$$

where we use the following notations: $\bar{\psi}^s_{i,t} \equiv \prod_j(\psi^s_{j|i,t})^{\Psi^j_{i,t-1}}$ is the weighted geometric mean of expenditure share for each type of worker with expenditure share on s among tradable goods in previous period $\Psi^s_{i,t-1}$, $\bar{\theta}_{i,t-1} = \sum_j \Psi^j_{i,t-1}\theta_j$ is the weighted average of Engel slope, $\widetilde{\lambda}^j_{in,t-1}$ is an element of the matrix $(\mathbf{I} - \{\lambda^j_{ni,t-1}\}^\top_{n,i})^{-1}$, $\widetilde{\Psi}^j_{i,t-1} \equiv \sum_s \Psi^s_{i,t-1}\widetilde{\beta}_{sj}$, $\varrho^{sj}_{i,t-1}$ is the element of the matrix $(\mathbf{I} - \{\beta_j\widetilde{\Psi}^j_{i,t-1}\}_{j,j})^{-1}$ and $\varrho^s_{i,t-1} = \sum_j \varrho^{sj}_{i,t-1}$.

To keep the discussion clear, we assume that the revenue from land goes to landlords absent in the economy. See the Appendix C.3 for details and further general discussion. This proposition states how wage evolves in the dynamic equilibrium. We consider a change of productivity over time but keep labor mobility costs fixed and assume land development revenue is distributed to the absentee landlords. Then, the wage dynamics combine the Rybczinski derivatives and the Stolper-Samuelson derivatives in the spatial economics framework. In the first term on the right-hand side in (18), the pre-determined elements $\{\varrho^{sj}_{i,t-1}\}$ summarize the substitution of labor between sector s and j. Within the parenthesis, the first term is about the workers' heterogeneity in consumption. Analytically, $\{\bar{\psi}^j_{i,t}\}$ evaluates the distortion in expenditure patterns relative to the uniform expenditure share. The second term factors the change of real income and the slope of the local Engel curve. When considering homothetic demand system, $\bar{\theta}_{i,t-1} = 1$ and this term is reduced to pure change of real income, $\mathrm{d}\ln\mathcal{W}^j_{i,t}$. These two terms together determine the expansion of labor in industry j in location i, and its impact on industry s depends on the labor intensity in production. Hence, this term is about the Rybczinski derivatives.

The second parenthesis on the right-hand side in (18) states the relationship between TFP changes and wage growth. As we discussed in Proposition 2, change in the import penetration and productivity contribute to the change in TFP. Therefore, the matrix $\{\varrho^{s}_{i,t-1}\}$ gives information of the Stolper-Samuelson derivatives that summarize how the change in trade patterns affect the wage. The input-output linkages and expenditure patterns together characterize the derivative. These two derivatives determine wage changes to the common shocks in the economy. The differences in elements $\{\varrho^{s,j}_{i,t-1}\}$ govern the difference in wage growth between industries, and the spatial variation of wage growth results from the variation in the probabilities of labor reallocation (i.e., migration) and trade conditional on the local expenditure patterns.

4.3 Upward Mobility

We are now in a position to discuss income mobility. We aim to understanding the relationship between spatial structural transformation and intergenerational mobility of workers – how does the next generation climb up the income ladder compared to the previous generation? The model abstracts the exact linkage between individual pairs of parents and children, and therefore there is no explicit inter-generational link between specific pair of parents among generation $t-1$ and children among generation t. Nevertheless, the model emphasizes the importance of location choices and sector choices in shaping the geography of intergenerational mobility. In particular, for each location, the model allows us to characterize (i) income distribution of generation t (i.e., parents) working there, and (ii) income distribution of generation $t+1$ (i.e., children) who have the origin there. Therefore, we assess the general equilibrium relationship of income distribution between parents and children in each location.[14] We start with the discussion of the measure. We let $\mathcal{R}^{\mathbf{o}}_{i,t}$ be the average percentile in the national income distribution for generation t working in location i, and $\mathcal{R}^{\mathbf{y}}_{i,t+1}$ refers to the expected percentile in the national income distribution for the next generation who are born in location i. Using these percentiles, the baseline index of intergenerational mobility for individuals in location i is:

$$\mathcal{M}_{i,t+1} = \mathcal{R}^{\mathbf{y}}_{i,t+1}/\mathcal{R}^{\mathbf{o}}_{i,t} \tag{19}$$

The ratio between the expected income percentile of generation $t+1$ and the average income percentile of generation t, $\mathcal{M}_{i,t+1}$, shows the expected climb up on the income ladder for individuals who have origin in location i. When location i exhibits greater land of opportunity in terms of upward income mobility for the future, $\mathcal{M}_{i,t+1}$ returns a high value. The measure (19) becomes large when workers of generation $t+1$ sort into the industry with high wage growth and move to the location with relatively high wages and a large surplus from land. The relationship between the measure and the equilibrium of the model is summarized in the following proposition:

PROPOSITION 5 *Define the income distribution in the whole economy $\mathcal{Q}_t$ such that:*

$$\mathcal{Q}_t(W^s_{i,t}) = \sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{K}} f^j_{n,t}\mathbf{1}[W^j_{n,t} \leq W^s_{i,t}]\frac{L_{n,t}}{\bar{L}} \equiv \mathcal{Q}^s_{i,t}$$

[14] Note that the income distribution in the model is characterized by the probability mass function across different income levels. Income levels take $N \times (S+1)$ different values.

The upward income mobility measure for generation $t+1$ *in terms of average rank is:*

$$\mathcal{M}_{i,t+1} = \sum_{s\in\mathcal{K}} \varsigma^{s}_{i,t+1} \left(\sum_{n\in\mathcal{N}} \lambda^{s}_{ni,t+1} \frac{\mathcal{Q}_{t+1}(W^{s}_{n,t+1})}{\sum_{j\in\mathcal{K}} f^{j}_{i,t} \mathcal{Q}_t(W^{j}_{i,t})} \right), \tag{20}$$

The measure (20) is intuitive. It is useful to see the decomposition of this measure into the different margins in the model:

$$\mathcal{M}_{i,t+1} = \sum_{j\in\mathcal{K}} \underbrace{\varsigma^{j}_{i,t+1}}_{\text{Job Opportunity}} \underbrace{\frac{\mathcal{Q}^{j}_{i,t}}{\sum_{s\in\mathcal{K}} f^{s}_{i,t} \mathcal{Q}^{s}_{i,t}}}_{\text{Local Inequality}} \underbrace{\frac{\mathcal{Q}^{j}_{i,t+1}}{\mathcal{Q}^{j}_{i,t}}}_{\text{Local Growth}} \underbrace{\left(\sum_{n\in\mathcal{N}} \lambda^{j}_{ni,t+1} \frac{\mathcal{Q}^{j}_{n,t+1}}{\mathcal{Q}^{j}_{i,t+1}} \right)}_{\text{Spatial Mobility}}.$$

The first term of sector choice probability reflects the job opportunity in location i for generation $t+1$. The second term is about the local income inequality for generation t as it is the relative position of workers in sector s to the local average in terms of income. The third term is the growth of the local labor market over generations represented by the change of positions in national income distribution between two generations for each industry. The last term in parenthesis captures gains from the geography of labor mobility for generation $t+1$. Thus, the variation of intergenerational mobility in geography is the consequence of the different extent of structural change and evaluates the importance of spatial economy regarding how further the young generation can climb up the income ladder.

To understand this measure more concretely, we consider the special cases. The first case supposes no geographical mobility of workers and two different sectors. Then, this measure is reduced to:

$$\mathcal{M}_{i,t+1}|_{D_{ij,t}\to\infty} = \sum_{j\in\mathcal{K}} \varsigma^{j}_{i,t+1} \frac{\mathcal{Q}^{j}_{i,t}}{\sum_{s\in\mathcal{K}} f^{s}_{i,t} \mathcal{Q}^{s}_{i,t}} \frac{\mathcal{Q}^{j}_{i,t+1}}{\mathcal{Q}^{j}_{i,t}} \tag{21}$$

Assume that sector j is sufficiently productive compared to k and becomes more productive in the next period. Then, $Q^{j}_{i,t} > Q^{k}_{i,t}$ and $\frac{\mathcal{Q}^{j}_{i,t+1}}{\mathcal{Q}^{j}_{i,t}} > \frac{\mathcal{Q}^{k}_{i,t+1}}{\mathcal{Q}^{k}_{i,t}}$. Then, if location i shows large inequality if $f^{j}_{i,t} < f^{k}_{i,t}$. Suppose workers are sorting more into industry k in the next period due to persistence in their job choices. In that case, we see lower social mobility in location i compared to the case in which labor is fully adjusted to the expected growth of an industry. This is one important mechanism that relates inequality in the local economy to low social mobility there. Another extreme case is the economy, where industries are differentiated by their locations, a lá Armington model. Then, the measure of intergenerational mobility is:

$$\mathcal{M}_{i,t+1} = \sum_{n\in\mathcal{N}} \lambda_{ni,t+1} \frac{\mathcal{Q}_{n,t+1}}{\mathcal{Q}_{i,t}} \tag{22}$$

This is the weighted average of the relative expected income ranking ($\mathcal{Q}_{n,t+1}/\mathcal{Q}_{i,t}$) with migration patterns ($\lambda_{ni,t+1}$). Therefore, intergenerational mobility becomes low when the origin shows low productivity growth or the high migration costs to the growing regions. This relationship relates the geography of industrial growth to the degree of intergenerational mobility. In our model, these mechanisms work together to define how the location shows high social mobility along with a geographical variation of structural transformations.

We can define alternative measures for intergenerational income mobility. One index is related to absolute upward mobility: what is the likelihood of earning more than parents? We can represent them by the probability of earning higher than a worker at α-th quantile in the previous generation with any particular value of α. Another measure captures the upward mobility from the bottom to the top by comparing the workers at the bottom of the quantile and the top of the quantile. Intuitively, a large value of such index in particular place i implies that the top income individuals arise from the cohort of generation $t+1$ born in i where workers in the previous generation are relatively lower income group at the national level. Therefore, this can be seen as the "American Dream." The Appendix C.6 discusses these measures. As a baseline, however, we use (20) since it is robust and shows continuity over time compared to other measures.

5 Model's Calibration

The goal is to quantitatively assess the extent of spatial structural change and its impact on individual consequences of welfare and inequality. To this end, we use data and model structure to estimate parameters and obtain the fundamentals of the real economy.

The model is mapped into the U.S. economy. The spatial unit of locations is the core based statistical area (CBSA). The time range is from 1980 to 2010 when there have been a considerable decline in the relative price of goods to services and an increase in real housing prices in the macroeconomy. The set of industries in the model is mapped into 18 industries. Among them, we consider the construction sector, 9 manufacturing industries, and 8 service industries. The construction sector corresponds to sector index 0 that develops the residential stock in the model. All of the sectors classified in the manufacturing sector are tradable, while one sector among service sectors, retail, is non-tradable. For CBSAs and sectors, data of employment and industry wage are from the County Business Pattern (CBP), the American Community Survey (ACS) and decennial censuses. Through the analysis, we focus on 395 CBSAs where we are able to construct these data for different periods. For each pair of CBSAs, geographical distance is computed between the reference points for the pair of most populated counties.

Model calibration proceeds in two parts. Subsection 5.1 discusses the parameters in the model. First, we set parameters in the demand system ($\alpha_s, \theta_s, \sigma$), production technology for manufacturing and service sectors (β_{sj}) and residential stocks (χ). Second, we exploit the gravity structure for manufacturing sectors and tradable services and determine the trade elasticity (κ_s). Third, we combine the structure of the model and parameter value of migration elasticity (ε) from the literature to obtain the industry choice parameters (η, ϕ). Fourth, we discuss the parameter choice for the economies of scale (γ_s) and productivity spillover (ρ).

In subsection 5.2, we leverage the structure of the model to back out the fundamentals in the U.S. economy. This procedure is sequential, and therefore we discuss it step by step. We assume that the economy is in the stationary steady state equilibrium in the last period of 2010. Then, the structural relationships allow us to derive the fundamentals in the development of residential stocks, amenities and productivity that are consistent with the distribution of workers to be the steady state equilibrium. Then, the inversion of the equilibrium conditions leads to fundamentals in past periods. The details of data construction and technical details are in the Appendix E.

5.1 Parameters

We explain the parameters in the baseline analysis. Table 5.1 reports the summary.

Demand and production. The demand system has three parameters. We set the elasticity of substitution between different industries $\sigma = 0.40$ which ensures their complementarity. We assign the slope of the Engel curve based on the estimation from Comin et al. (2020) and set different values between two large categories of manufacturing sector and service sector. Namely, θ_s is normalized to 1.0 for the construction sector and manufacturing sector. For service sector, we set θ_s such that $(\theta_s - \sigma)/(1 - \sigma) = 1.75$, which is in the middle of estimates from Table I in Comin et al. (2020). Therefore, the expenditure share on manufacturing sectors is independent of real income, while that on service sectors increases in real income. For the rest of the parameters in the demand system, the parameters of demand shift (α_s) are chosen to match the year 2010 expenditure shares in the manufacturing and service sector.

We need input share for each industry. Using the US Bureau of Economic Analysis (BEA) table of input-output accounting, we compute these shares to match the average values during 2011-15. Since we do not consider the international trade of intermediate goods as a baseline, we need to adjust the input-output identity, and the labor share is computed as the share of labor compensation in the adjusted total production values based on the identity. The parameters of input-output linkages are equal to the share of input purchases from other industries. On the development of residential stock, the production technology exhibits the labor share equal to χ. The input-output accounting from BEA gives $\chi = 0.35$ for labor share in the construction sector on average.

Trade elasticity. The regional trade in the model is the gravity fashion. We parametrize the impediment of trade such that trade costs between different locations are elastic function of geographical distance with elasticity $\bar{d}$. Then, the restricted gravity equation for the value of export from n to i is:

$$\ln X^s_{in,t} = \mathbb{D}^s_{i,t} + \mathbb{O}^s_{n,t} - (\kappa_s \bar{d}) \ln \text{dist}_{in} + \varepsilon^s_{in,t} \tag{23}$$

where $\mathbb{D}^s_{i,t}$ factors destination characteristics and $\mathbb{O}^s_{n,t}$ factors characteristics of source locations, respectively. We estimate $\kappa_s \times \bar{d}$ for manufacturing sector by using U.S. Commodity Flow Survey (CFS) in 2012. After the estimation of the gravity equation, to decompose the trade elasticity of each industry (κ_s) and trade cost elasticity ($\bar{d}$), we assume that $\bar{d} = 0.125 = 1/8$. The value is close to the estimates in Eaton and Kortum (2002) and lower than trade cost elasticities estimated for international trade. This gives the inferred different trade elasticities (κ_s) by manufacturing industries that are in the range of estimates in the literature of international trade (Head and Mayer 2014, Simonovska and Waugh 2014) as well as domestic trade (Gervais and Jensen 2019). Turning to the service sector, we cannot directly observe the trade flows and we rely on the estimation by Anderson et al. (2014). Their estimates can be directly used in our definition of service sectors to pin down the trade elasticity of services. We assume the same value of trade cost elasticity as manufacturing sectors and obtain the different trade elasticity by service industries. They are within the range of estimates in Gervais and Jensen (2019). See Table E.3 in the Appendix for numbers.

Migration costs and elasticities in labor supply. There are three parameters in the choice of workers and also need to characterize the migration frictions. The first parameter is the shape parameter of Fréchet distribution of the idiosyncratic shocks in location choice, which captures the elasticity of labor allocation across different locations with respect to real income: $\varepsilon = \frac{\partial \ln\left(L^s_{i,t}/L^s_{n,t}\right)}{\partial \ln\left(\mathcal{W}^s_{i,t}/\mathcal{W}^s_{n,t}\right)}$. Following Fajgelbaum et al. (2019), we set $\varepsilon = 1.5$. Next, we consider the migration costs. Suppose that the bilateral migration cost is decomposed into an elastic function of bilateral distance and destination characteristics. In particular, we parametrize $\ln D_{in,t} = \widetilde{d} \ln \text{dist}_{in} + \ln M_{i,t}$ for migration cost from n to i. $\widetilde{d}$ is positive constant and $M_{i,t}$ is destination characteristics that can very over time. Under this parametrization, the model derives the gravity equation of labor mobility across space conditional on sector choice:

$$\ln L^s_{in,t} = \mathbb{W}^s_{i,t} - (\varepsilon \widetilde{d}) \ln \text{dist}_{in} + \mathbb{H}^s_{n,t} \tag{24}$$

where $\mathbb{W}^s_{i,t}$ and $\mathbb{H}^s_{n,t}$ contain source location and industry characteristics, and destination and industry characteristics, respectively. To estimate $\widetilde{d}$, we use American Community Survey (ACS) 5 year sample data between 2006-10 and 2011-15. In their sample, the ACS data allows us to identify the current county, previous county and industry of the worker. We extract workers in sectors of our analysis and map their locations to the CBSA level and focus on workers who moved between different CBSAs in the sample to estimate the gravity equations. Based on the estimates during different sample periods, 2006-10, 2011-15 and 2006-10, we set $\widetilde{d} = 0.50$.[15]

Once we have the migration elasticity and bilateral term in migration cost, we leverage the structural equations for labor mobility to calibrate the other two key parameters in the choice of individuals (η and ϕ). Given the parameters (ε, $\widetilde{d}$), the model allows us to characterize the mobility of workers in equilibrium. For each pair of values (ϕ, η), exploiting the equilibrium condition for the labor allocation (11), we uniquely determine the set of endogenous characteristics that rationalize the observed change in the distribution of workers. In turn, we can compute predicted migration

[15] The estimates of the gravity equation are similar to the findings for intra-national migration elasticity to distance in the literature (e.g., Bryan and Morten 2019). Compared to Allen and Arkolakis (2018), estimates are small. This difference may arise from the difference in periods. For the old period, it would be large because of the higher moving cost per unit of distance.

flow in equilibrium, $\widehat{L}_{in,t}$, between any particular pair of CBSAs. Therefore, we can define the moment conditions that argue the differences between the observed pattern of labor mobility ($L_{in,t}^{Data}$) between CBSAs and the predicted one in the model ($\widehat{L}_{in,t}$) are not systematically correlated to the bilateral distances between source and destination within the same range of distances. As an observation of labor mobility, we exploit Internal Revenue Service (IRS) county-to-county migration data and aggregate them to the CBSA pairs for two time periods, 1990-2000 and 2000-2010. Comparing the pattern of labor mobility between data and prediction in the moment conditions, we obtain the estimated value of two parameters: $\phi = 2.50$ and $\eta = 0.80$.

Productivity spillovers. We assign the value of parameters in agglomeration economies (γ_s) based on the discussion in theory. The one condition imposed on the parameter argues that the dynamic equilibrium converging to the stationary steady state equilibrium is unique when γ_s is not too large to avoid the degenerate equilibria. Since the long-run equilibrium in history does not show such a degenerate equilibrium, we use the condition to set γ_s. As we discussed in Section 3, one of the condition that is related to the uniqueness of the dynamic equilibrium conditional on the initial state is given by $\gamma_s \leq \frac{\theta_s - \sigma}{\kappa_s + (1-\sigma)} \left(1 + \frac{1}{\varepsilon}\right)$. This condition gives the upper bound of the parameter when allowing labor mobility across locations, productivity spillover happens only locally ($\rho = 0$), and the supply of residential stocks is perfectly elastic ($\chi = 1$). In the quantification, however, we allow $\rho \neq 0$. The spillover in productivity across space through migration of workers leads to further agglomeration forces in the steady state since favorable locations attract workers while the remote places lose. Therefore, we take the conservative values that satisfy the condition with additional restriction $\varepsilon \to \infty$. This assures that the dynamic equilibrium is unique when idiosyncratic shocks for location choices are even homogeneous. This gives us the parameter values by industry such that $\gamma_s = \frac{\theta_s - \sigma}{\kappa_s + (1-\sigma)}$. For comparison to the existing values in empirical studies, we also refer values in Combes et al. (2012) and Bartelme et al. (2021) in the Appendix E.2. It is worth emphasizing that we assume that $\chi = 1$ to derive the condition. If $\chi < 1$, the supply of residential stocks becomes less elastic and congestion force arises. Therefore, setting $\chi = 1$ keeps the conservative value for the purpose. Lastly, for the parameter of spatial spillover (ρ), we discuss it in the next subsection along with the inversion of productivity.

5.2 Calibration of Fundamentals

Next, we solve the model for the fundamentals of the economy conditional on the information about the local labor markets. The inversion of the fundamental productivity ($A_{i,t}^s$), fundamental location characteristics ($B_{i,t}^s$, $M_{i,t}$, ν_i and $\bar{h}_i$) and fundamental sectoral parameter ($\zeta_{s,t}$) proceeds in two steps. In the first step, we assume that the economy is in the steady state level in 2010 and compute the time-invariant location characteristics in the development of residential stock (ν_i and $\bar{h}_i$) by using the information on the price of housing. Then, we use the system of equations in the equilibrium to back out the overall productivity (Z_i^s) and attractiveness of locations that combine amenities (B_i^s) and migration friction (M_i). Once we obtain the overall productivity, we compute the

Table 5.1: Parameters

Parameter	Source and Comments
1. **Demand**: $\sum_j (\alpha_j c_j)^{(\sigma-1)/\sigma} \mathbb{U}^s_{n,t}(\omega)^{(\theta_j-\sigma)/\sigma} = 1$	
$\sigma = 0.40$	Elasticity of substitution between industries
$\theta_M = 1.00, \quad \theta_S = 1.375$	Sector specific non-homotheticity; Comin et al. (2020)
	θ_M for all industries in manufacturing; θ_S for all industries in service
$\alpha_M = 3.00, \ \alpha_S = 0.1, \quad \alpha_0 = 0.25$	Sector level demand shifter
	Matched to expenditure share ratio between service and manufacturing sector
2. **Production**: $\Xi^s_{nt} = (w^s_{nt})^{\beta_s} \prod_{j\in\mathcal{K}\setminus 0} (p^j_{nt})^{\beta_{sj}}$, $h_{i,t} = \nu_i (l^0_{i,t})^{\chi} ((1-\bar{h}_i) h_{i,t-1})^{1-\chi}$	
β_s, β_{sj}	Input share; BEA table of input-output accounting
$\chi = 0.35$	Labor share in housing construction; BEA table for construction sector
3. **Productivity**: $F_s(\varphi) = \exp(-\varphi^{-\kappa_s})$, $Z^s_{i,t} = A^s_{i,t} \Big(\sum_n L^s_{in,t} Z^s_{n,t-1} \Big)^{\rho} (L^s_{i,t})^{\gamma_s}$	
κ_s	Trade elasticity from gravity estimates using CFS for manufacturing;
	Anderson et al. (2014) for services; see Table E.1 in the Appendix
$\rho = 0.0284$	Spillover in productivity across space; internal estimation
γ_s	Local externalities; upper bound of uniqueness condition
	see Table E.3 in the Appendix
4. **Workers' Choice**: $G(v) = \exp(-\sum_n v_n^{-\varepsilon})$, $\mathcal{O}(j^k_i, L^k_{i,t}) = \frac{1}{\zeta_k} \ln \Big(\frac{1}{1-j^k_i} \Big) (L^k_{i,t})^{-\eta}$, $F(z) = 1 - z^{-\phi}$	
$\varepsilon = 1.50$	Migration elasticity; Fajgelbaum et al. (2019)
$\phi = 2.50$	Variation of taste shocks; estimation from migration pattern
$\eta = 0.80$	Local labor market exposure effect; estimation from migration pattern
5. **Spatial frictions**: $D_{in,t} = (\text{dist}_{in})^{\tilde{d}} M_{i,t}, \ \ \tau_{in,t} = (\text{dist}_{in})^{\bar{d}}$	
$\bar{d} = 0.125$	Trade cost elasticity; Eaton and Kortum (2002)
$\widetilde{d} = 0.50$	Migration cost elasticity from gravity estimates

Note: This table reports parameters in calibration and quantitative analysis.

exogenous part of productivity (A_i^s) and estimate parameter ρ. In the next step, we solve the model for the time-varying fundamentals. We match the dynamic equilibrium and observation of wage and employment for the inversion of the path of exogenous part of productivity ($A_{i,t}^s$), amenities ($B_{i,t}^s$), migration frictions ($M_{i,t}$) and sectoral shifter ($\zeta_{s,t}$). The whole process is sequential, so we explain the procedure by step.

Efficiency of development of residential stock. In the steady state, (9) implies that

$$\widetilde{\nu}_i \equiv \nu_i(1-\bar{h}_i)^{1-\chi} = \exp\big(\chi(-\ln\chi + \ln w_i^0 - \ln p_i^0)\big) \tag{25}$$

where we have equilibrium wage (w_i^0) and housing price (p_i^0) from Federal Housing Finance Agency (FHFA). We exploit the Housing Price Index (HPI) of all-transactions index across CBSAs for 2010. In the Appendix E.3, we show the spatial distribution of efficiency of development across CBSAs inverted in the model. Intuitively, CBSAs with large values $\widetilde{\nu}_i$ implies that there is a persistence of residential stock and higher efficiency of construction. In equilibrium, the geographical differences in $\{\widetilde{\nu}_i\}$ affect the spatial variation of housing supply and price changes.

Productivity. The zero profit condition of producers (11) implies:

$$\Gamma_s^{\kappa_s}(p_i^s)^{-\kappa_s} = \sum_n \left((\widehat{\tau}_{in}^s)^{-\kappa_s}(Z_n^s)^{\kappa_s} \left((w_n^s)^{\beta_s} \prod_j (p_n^j)^{\beta_{sj}} \right)^{-\kappa_s} \right) \tag{26}$$

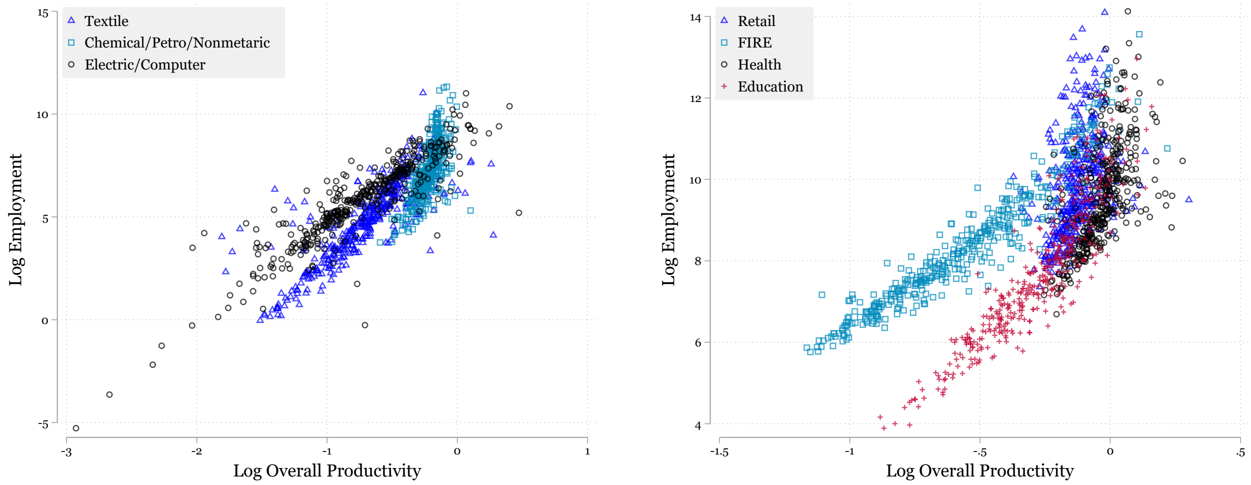
where $\widehat{\tau}_{in}^s \equiv (\mathrm{dist}_{in})^{\bar{d}}$ is bilateral trade costs. Given productivity, wages and the set of parameters (κ_s, $\bar{d}$, β_s, β_{sj}), solving this allows us to characterize the equilibrium prices. Turning to the demand system, we solve (2) for the non-homothetic aggregate price index:

$$(\mathcal{P}_i^s)^{1-\sigma} = \sum_j \alpha_j^{\sigma-1}(p_i^j)^{1-\sigma} \left(\frac{W_i^s}{\mathcal{P}_i^s} \right)^{\theta_j - 1} \tag{27}$$

where income of workers (W_i^s) is constructed from the wage and employment. Given the parameters of preference ($\sigma, \theta_s, \alpha_s$) and $\theta_s \geq \sigma$ and $\sigma \in (0,1)$, we obtain the unique matrix of price index $\{\mathcal{P}_i^s\}$ solving (27).

Combining (26), (27) and the labor market clearing condition (14) allows inversion of productivity $\{Z_i^s\}$ that is consistent with the observation to be the equilibrium. Figure 5.4 shows the relationship between overall productivity (Z_i^s) implied by the model and employment (L_i^s) in 2010. Each mark shows CBSA in the analysis. The left-hand panel 5.4a shows the relationship for three industries in the manufacturing sector: textile, chemical/petro/coal products and electric/computer industry. We find a larger variation in overall productivity for the electric/computer industry than for the other two industries. This is consistent with the dispersion of workers in the industry over space. The right-hand panel 5.4b presents four industries in the service sector: retail, finance, insurance and real estate (FIRE), health and education/legal services. FIRE and education services exhibit

Figure 5.4: Productivity and Employment for CBSAs



(a) Manufacturing industries **(b)** Service industries

Note: These figures show log scale of overall productivity (Z_i^s) and log scale of employment for particular industries in 2010.

relatively large variation across CBSAs in overall productivity compared to retail and health.

Amenities and industry taste. As a next step, we use the model structure for the inversion of amenities and location characteristics in the migration frictions – location attractiveness. In the steady state, amenities ($B_{i,t}^s$) and migration barriers ($M_{i,t}$) are constant. These two location fundamentals decide the exogenous gains for workers who choose the destination, and we are unable to isolate them. In addition, there are another fundamental constant parameters $\{\zeta_s\}$ that capture the tastes of workers by sector. Therefore, we let $\Omega_i^s \equiv (B_i^s/M_i)\zeta_s^{1/\phi}$ conflate these fundamentals. (4) implies that the adjusted average utility of a worker in location n in sector s is:

$$\bar{\mathbb{U}}_n^s = \zeta_s^{1/\phi}\bar{U}_n^s = \left(\sum_i \left(\Omega_i^s \widehat{D}_{in}\frac{W_i^s}{\mathcal{P}_i^s}\right)^{\varepsilon}\right)^{1/\varepsilon} \tag{28}$$

where $\widehat{D}_{in} \equiv \text{dist}_{in}^{-\widetilde{d}}$ is the inverse of bilateral migration frictions. Using this, we compute the inferred probabilities of location choice for workers in sector s ($\{\widehat{\lambda}_{in}^s\}$) and the probabilities of industry choice ($\{\widehat{\varsigma}_n^s\}$). Then, we use the labor mobility condition (11) for computing the attractiveness of location i and sector s with implementing the aggregate price index derived in (27). In this step of inversion, we also compute the predicted labor mobility ($\{\widehat{L}_{in}^s\}$) in the steady state and use this information in the next step.

Fundamental productivity. We proceed to the inversion of fundamental productivity and calibration of parameter ρ. In the steady state, the exogenous fundamental productivity satisfies:

$$\ln A_i^s = \ln Z_i^s - \rho \ln\left(\sum_n \widehat{L}_{in}^s Z_n^s\right) - \gamma_s \ln L_i^s \tag{29}$$

We implement the productivity ($\{Z_i^s\}$) that we have derived in the previous step and the labor mobility ($\{\widehat{L}_{in}^s\}$) into this and given the employment data ($\{L_i^s\}$) and parameter values ($\{\gamma_s\}$), we are able to compute the fundamental productivity for any particular value of parameter ρ. To estimate ρ, we consider the following moment conditions:

$$\mathbb{E}\left[\left(\ln A_i^s - \frac{1}{N}\sum_n \ln A_n^s - \frac{1}{S}\sum_k \ln A_i^k\right) \times \mathbb{I}_g\right] = 0, \quad g \in \mathbb{G}_0, \mathbb{G}_1, \cdots, \mathbb{G}_P \tag{30}$$

where $\mathbb{I}_g$ is an indicator that the location i and sector s is in the group of g. The group is defined by the labor market potential for each location and sector. Namely, for location i and sector s, we compute the measure $\sum_{n\neq i}(\text{dist}_{ni})^{-\varepsilon\widetilde{d}} L_n^s$, and we order locations and sectors by this measure. Based on the order, we use 20 groups defined by 5 percentile of the measure. Therefore, the moment condition assumes that the location and industry specific fundamental productivity after eliminating the sector-level and location-level averages is not systematically related to the labor market access since the spatial dependence of productivity is captured by the second term in (29). We use (30) and search parameter ρ that minimizes the distances of the moment conditions. We obtain $\widehat{\rho} = 0.0284$ that is reported in Table 5.1.

Dynamics of Fundamentals. Once we have characterized the steady state fundamentals, we compute the change in fundamental productivity and location attractiveness for 2000-2010, 1990-2000, and 1980-1990. We suppose that the economy reached the steady state equilibrium in 2010 and compute the change of these fundamentals in the past. We follow the steps close to the previous procedure for the steady state equilibrium.

First, we compute the residential stock and their prices in the past conditional on the current observations. The production function of developers (9) implies the previous residential stocks:

$$(1-\chi)\ln H_{i,t-1} = \ln H_{i,t} - \ln \widetilde{\nu}_i - \chi \ln L_{i,t}^0 \tag{31}$$

and market clearing condition implies the price of residential stocks:

$$\ln p_{i,t-1}^0 = -\ln \chi + \ln w_{i,t-1}^0 + \ln L_{i,t-1}^0 - \ln H_{i,t-1} \tag{32}$$

Implementing parameters in production technology of residential structure (χ), observed wage and employment in the construction sector and location fundamentals ($\{\widetilde{\nu}_i\}$), it gives the path of ($\{p_{i,t-1}^0\}$, $\{H_{i,t-1}\}$, $\{R_{i,t-1}\}$) in the dynamic equilibrium that are not directly observable. The HPI have limited data of prices for CBSAs in 1990 and 1980 that can gauge the model specification in (31) and (32). The Appendix E.3 shows the comparison between prices across CBSAs predicted by the model and the limited data for 1980 and 1990.

Second, we compute the change in productivity over periods $\mathrm{d}\ln Z_{i,t}^s$, such that wage and employment in the past are consistent with the dynamic equilibrium. We guess the productivity in the past ($\mathrm{d}\ln Z^s_{i,t}$) and compute the change in prices and trade patterns. We solve the static equation (27) for the aggregate price index and use the market clearing condition to update the productivity change.

Third, we use the forward equations in the model for computing the path of location attractiveness. The labor mobility condition (11) implies that the adjusted attractiveness for workers in location i and sector s satisfies:

$$\Omega^s_{i,t} = \left(\frac{1}{L^s_{i,t}} \sum_{n \in \mathcal{N}} \left(\frac{\widehat{D}_{in}}{\mathbb{U}^s_{n,t}} \frac{W^s_{i,t}}{\mathcal{P}^s_{i,t}} \right)^{\varepsilon} \frac{(L^s_{n,t-1})^{\eta} (\mathbb{U}^s_{n,t})^{\phi}}{\sum_j (L^j_{n,t-1})^{\eta} (\mathbb{U}^j_{n,t})^{\phi}} L_{n,t-1} \right)^{-1/\varepsilon} \tag{33}$$

where $\{\mathbb{U}^s_{n,t}\}$ are derived for each generation as in (28). Conditional on the observation about employment ($\{L^s_{i,t}\}$) and income and aggregate price index constructed by the model, we obtain the location and sector specific adjusted amenities ($\{\Omega^s_{i,t}\}$) in each period and we are able to compute the two probabilities of workers' choice ($\{\lambda^s_{in,t}\}$, $\{\varsigma^s_{n,t}\}$) predicted by the model. Lastly, we compute the development of fundamental productivity in the analogous way to (29). The overall productivity of two consecutive periods ($\{Z^s_{i,t}\}$, $\{Z^s_{i,t-1}\}$), employment ($\{L^s_{i,t}\}$) and labor mobility ($\{L^s_{in,t}\}$) give fundamental productivity ($\{A^s_{i,t}\}$) that is consistent with the dynamic equilibrium. For the initial period, we set $A^s_{i,t} = Z^s_{i,t}$ in 1980.

6 Quantitative Analysis

Having an inversion of the model to obtain the fundamentals in the economy and estimated parameters, we assess the role of these fundamentals and analyze the dynamics of TFP, welfare and upward mobility across CBSAs in the U.S. economy that we have discussed in Section 4. We first see the role of industry and location specific amenities for workers' distribution along with the structural transformation. Then, we derive the measured TFP predicted by the model and discuss the industry level and aggregate TFP at the sector level: manufacturing and service. Next, we discuss the welfare differences of individuals between two generations and how they are different across locations. Following theoretical results in Proposition 3, we present the different margins in welfare dynamics. Lastly, we explain how the measure of intergenerational income mobility derived in the model shows spatial variation and investigate its relationship to the underlying mechanisms in the general equilibrium.

Amenities. As we discussed in the previous section, the amenities in each local labor market, $\{\Omega^s_{i,t}\}$ is obtained by exploiting the model structure. Figure 6.5 shows the relationship between the real income of workers and employment share for different industries in 2010. The vertical axis is the real income of workers in any particular industry in each CBSA ($\{\mathcal{W}^s_{i,t}\}$) derived using the calibrated income and nonhomothetic price index in the equilibrium relationship. The horizontal axis shows the de-meaned employment share of each industry in CBSA.

Figure 6.5: Real income and Employment

(a) Manufacturing

(b) Service

Note: The employment share of each industry is de-meaned by the average employment share across 395 CBSAs. Each circle represents the size of total employment in 2010 for CBSAs. The real income of workers is computed in the model.

The left-hand panel 6.5a displays three industries in the manufacturing sector. The employment share exhibits large variation relative to real income, and the pattern is different across industries. This confirms that there exist industry-specific amenities for workers. The right-hand panel 6.5b shows two distinctive industries – finance, insurance and real estate (FIRE) and retail among the services sector. For FIRE, a large employment share is associated with higher real income for workers. In contrast, the retail industry exhibits the importance of amenities to explain the spatial variation of workers. These results are consistent with industry and location specific amenities for workers' location choices and such amenities are crucial to explaining the spatial variation of employment shifts. In the Appendix D.2, we also confirm the relationship between the average level of amenities and the size of employment in CBSA. The relationship is stable over time, suggesting the importance of location fundamentals for the persistence in the aggregate size of employment in CBSA. Furthermore, in the Appendix F.1, we report the geographical distribution of the average value of amenities, CBSAs with the highest average amenities and the correlation between average amenities and some observed characteristics in CBSA.

Productivity. As we discussed in Section 4 and proposition 2, we are able to compute measured TFP for each CBSA and industry given overall productivity ($\{Z^s_{i,t}\}$) and trade probabilities ($\{\pi^s_{ii,t}\}$). We find distinctive dynamics of the spatial distribution of measured TFP by industry. These spatial distributions are in the Appendix F.1. For instance, we can identify an increase in the spatial variation of TFP for the electric and computer industry, along with the development of clusters in California and large metropolitan areas on the East coast. We also see the different evolution of TFP for FIRE. Over time, there has been a remarkable increase in its level and variation. The industry has seen a significant development on the East coast (New York metropolitan area) and in large cities that are the hub of the financial market in each region (Chicago, Dallas, Atlanta and Nashville) from

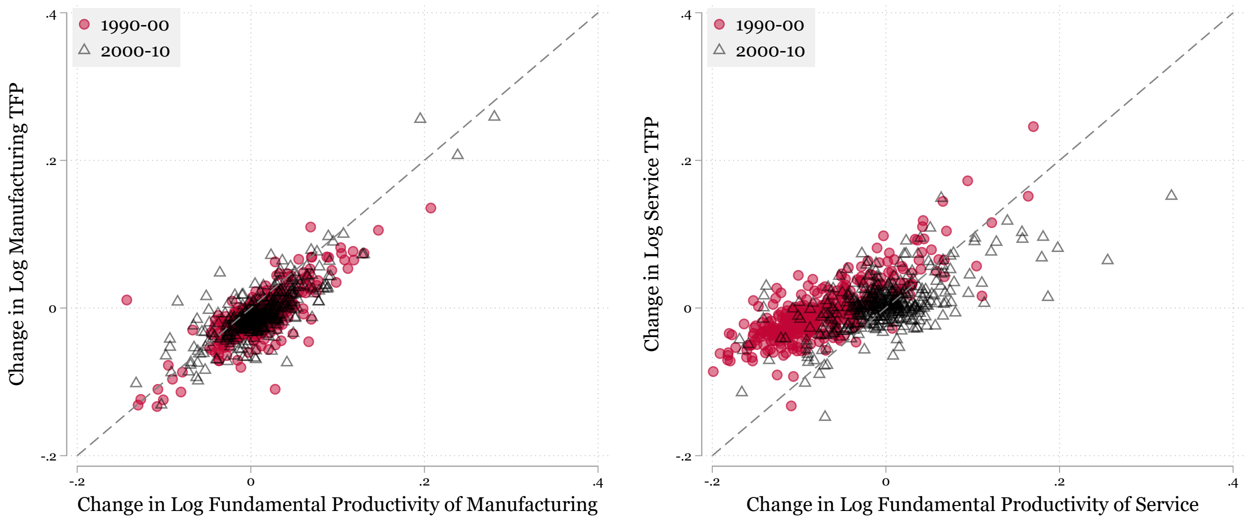
1980 to 1990. Then, these clusters show persistent development over time, while some other inland cities also have seen a rise in FIRE. The developments of TFP over time are the combination of exogenous productivity change and endogenous spillovers in theory. We find a significant variation in the fundamental productivity for the service sector.

Having measured TFP of each industry, we compute the aggregate TFP for the aggregate sectors: manufacturing and services sector. Namely, for an aggregate sector, we compute:

$$\delta_{i,t}^{\mathcal{S}} = \sum_{j \in \mathcal{S}} \frac{X_{i,t}^{j}}{\sum_{k \in \mathcal{S}} X_{i,t}^{k}} \delta_{i,t}^{j} \tag{34}$$

where $\delta_{i,t}^{j}$ is measured TFP of industry j in location i for period t and $X_{i,t}^{j}$ is the value of output of the industry. $\mathcal{S}$ is the set of industries in aggregated sectors, manufacturing or service sector. In an analogous way, we can compute aggregated fundamental productivity. Figure 6.6 shows the relationship between change in aggregate sector level TFP and fundamental productivity for the manufacturing and services sector. This corresponds to the implication in Proposition 2. In the left-hand panel, changes in TFP and fundamental productivity for the manufacturing sector exhibit a similar pattern. In contrast, changes in TFP of the services sector show large values relative to the fundamental changes. This implies that TFP growth in the services sector over these periods is driven by the endogenous mechanisms of labor reallocation and productivity spillovers.

Figure 6.6: Change in TFP and Fundamental Productivity for Manufacturing and Service



(a) Manufacturing **(b)** Service

Note: These figures show the change in log of fundamental productivity for aggregate sector ($\mathrm{d} \ln A_{i,t}^{K}$) and the change in log of TFP for aggregate sector ($\mathrm{d} \ln \delta_{i,t}^{K}$).

Welfare. Next, we quantitatively evaluate the welfare dynamics discussed in Proposition 3. There is a large variety of welfare effects across CBSAs.[16] In sum, the welfare gains between generations

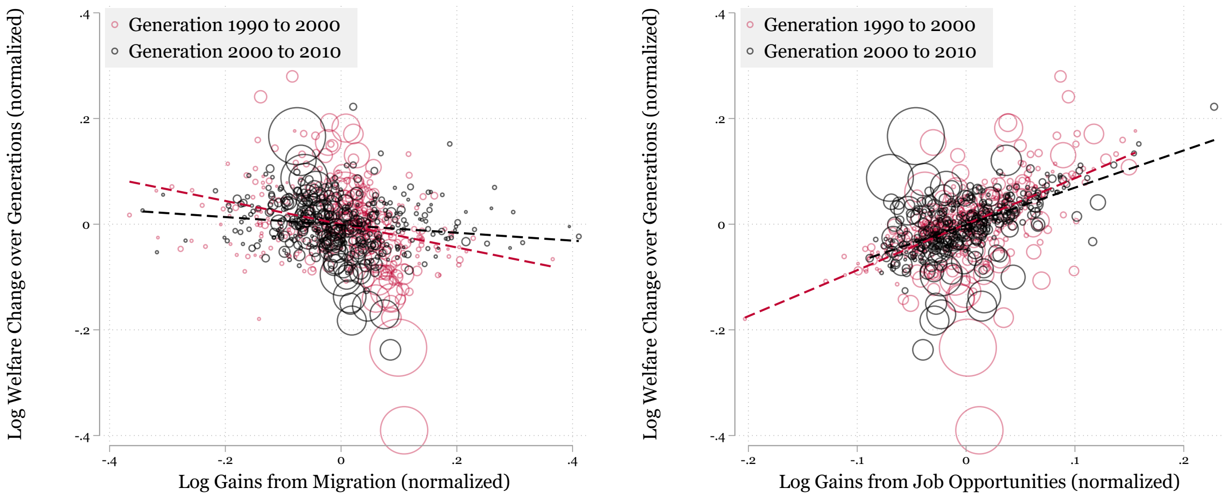
[16]See the Appendix D.2 for the map of its spatial pattern.

t and $t-1$ can be decomposed into three different terms:

$$\dot{V}_{i,t} = \underbrace{GM_{i,t}}_{\prod_s (\dot{\lambda}^s_{ii,t})^{-1/\varepsilon}} \times \underbrace{GJO_{i,t}}_{\prod_s (\dot{\varsigma}^s_{i,t})^{-1/\phi} (\dot{L}^s_{i,t-1})^{\eta/\phi}} \times GC_{i,t} \tag{35}$$

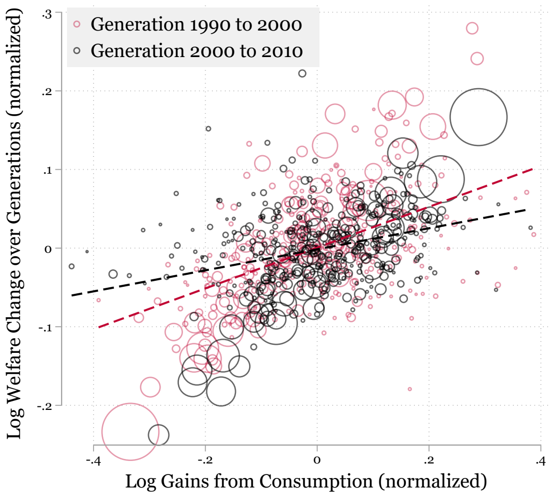
where $GM_{i,t}$ represents gains from labor mobility across space, $GJO_{i,t}$ summarizes gains from job opportunities in the local labor market and $GC_{i,t}$ stands for local gains from consumption and amenities. Figure 6.7 presents this relationship for U.S. CBSAs.

Figure 6.7: Welfare Differences



(a) Migration

(b) Job opportunity



(c) Consumption

Note: These figures show the different margins in welfare differences for CBSAs. Each circle represents the employment size of CBSA in the baseline year (i.e., employment size in 1990 for the blue circle). All variables are normalized by their means.

In the first panel (a), higher gains from migration are associated with small welfare differences. The logic is clear. Conditional on industry choice and growth of real income, an increase in the probability of staying in the original location requires higher welfare gains for individuals who stay in the local labor market. Comparing the two periods, the elasticity of welfare difference to gains of migration becomes small. This is consistent with the recent decline of the migration rate in the

U.S. economy. The second panel (b) shows the positive relationship between job opportunities in the local labor market and welfare. Individuals gain from the labor specificity in relatively small local labor markets. In these CBSAs, the specialization of workers into a particular industry in a growing sector leads to significant welfare differences over generations. The positive relationship is steady for these two periods. The third panel (c) exhibits the positive relationship between the change in real income adjusted with amenities and welfare differences. The change in average real income shows large variation and the role of real income disparity in the welfare change is large in the early period. The smaller elasticity of welfare differences to gains of migration and growth in the real income account for a decline of welfare differences over periods, while the gains of job opportunities account for the persistence in local labor market adjustment. These three margins are quantitatively consistent with the theoretical implications.

Inequality. The final objectives in this section are worker inequality and upward income mobility. For the inequality, we use the coefficient of variation in income within CBSA as a measure of income inequality. For the intergenerational income mobility, we compute normalized measure based on $\mathcal{M}_{i,t+1}$ defined in Proposition 5. Specifically, we let $\widetilde{\mathcal{M}}_{i,t+1} = (\mathcal{M}_{i,t+1}/\bar{\mathcal{M}}_{t+1}) \times 25$ where $\bar{\mathcal{M}}_{t+1}$ is average of $\mathcal{M}_{i,t+1}$ in the economy. Intuitively, this measure gives an expected rank of individuals in CBSA i when their previous generations are in the 25 percentile in the income distribution in the economy.[17] Figure 6.8 display the measure for different generations.

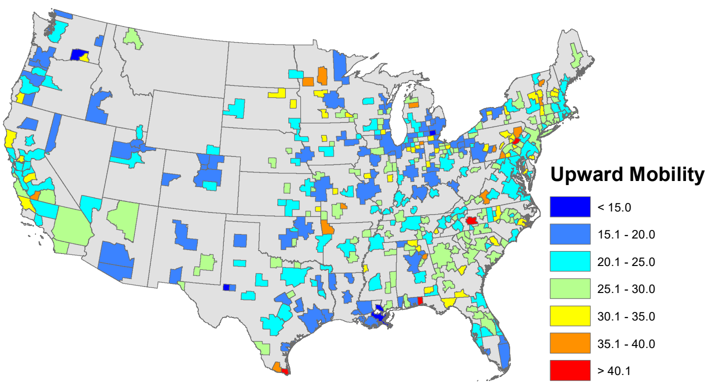
We find a considerable variation in upward mobility. For the first generation who worked in 1990, central cities in the region show relatively higher upward mobility. In later periods, upward mobility becomes lower on average. Given this spatial variation, we consider the relation of upward mobility to the underlying mechanisms in equilibrium. Following the discussion in Section 4.3, the measure of upward mobility can be written as:

$$\widetilde{M}_{i,t+1} \propto \sum_{s\in\mathcal{K}} \underbrace{LL^s_{i,t+1}}_{\frac{\mathcal{Q}^s_{i,t}}{\sum_k f^k_{i,t}\mathcal{Q}^k_{i,t}}\frac{\mathcal{Q}^s_{i,t+1}}{\mathcal{Q}^s_{i,t}}} \times \underbrace{ISM^s_{i,t+1}}_{\varsigma^s_{i,t+1}} \times \underbrace{GLM^s_{i,t+1}}_{\sum_n \lambda^s_{ni,t+1}\mathcal{Q}^s_{n,t+1}/\mathcal{Q}^s_{i,t+1}} \tag{36}$$

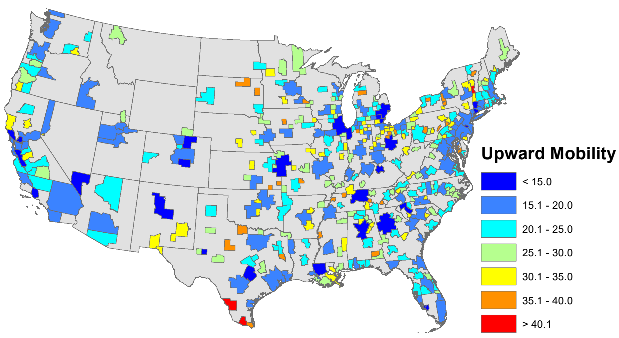
where $LL^s_{i,t+1}$ captures the inequality in the local labor market in period t and local economic growth, $ISM_{i,t+1}$ is patterns of industry choice and $GLM^s_{i,t+1}$ is the geographical labor mobility. Figure 6.9 present these margins.

[17] See the Appendix D.2 for further discussion about the measure and relation to measures in the literature.

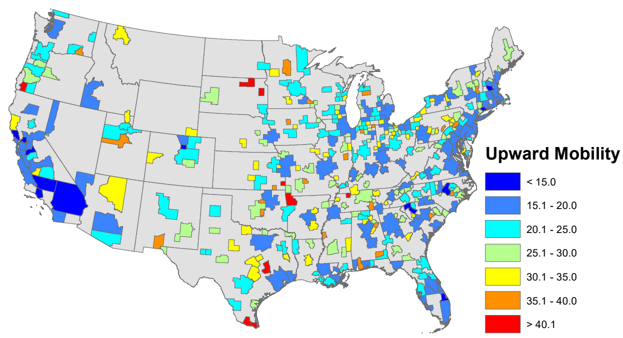
Figure 6.8: Geography of Upward Mobility



(a) Generation 1990

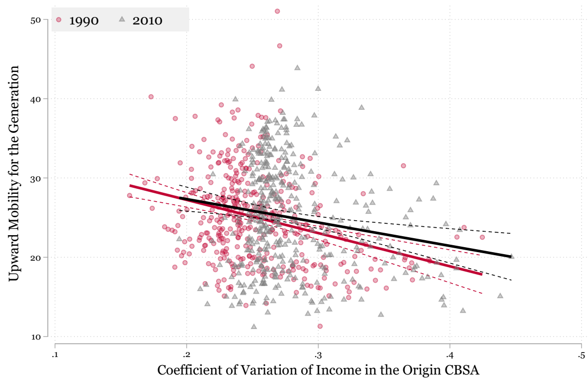


(b) Generation 2000

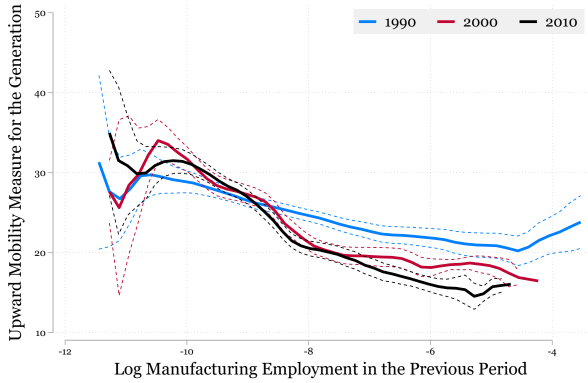


(c) Generation 2010

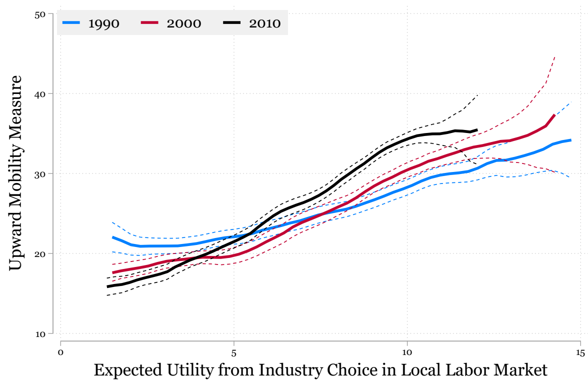
Figure 6.9: Intergenerational Income Mobility



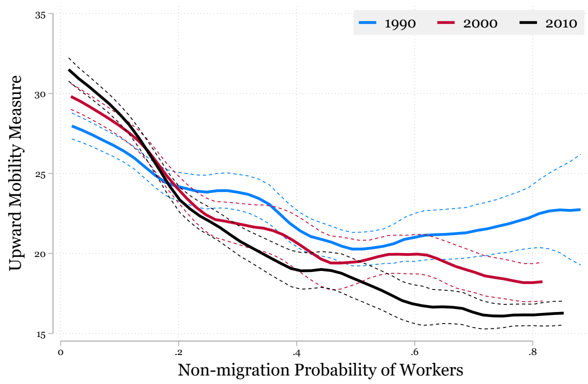
(a) Local inequality



(b) Employment



(c) Opportunity of Industry Choice



(d) Geographical Mobility

Note: These figures show the relationship between the measure of intergenerational income mobility and relevant measures. Panel (a) shows the relationship between the measure of upward income mobility for generation 1990 and 2000 to the inequality in the local labor market in period 1980 and 1990 measured by the coefficient of variation. Each marker shows CBSA and red (black) solid line is fitted line, and dash lines are 95% confidence intervals. Panels (b), (c) and (d) report the polynomial fitted line for CBSAs. Dash lines are 95% confidence intervals. Panel (b) shows the relationship between upward income mobility of three different generations (1990, 2000 and 2010) to the log of manufacturing employment in 1980, 1990 and 2000, respectively. Panel (c) shows the relationship between upward income mobility and individuals' expected utility from industry choice in each CBSA for generations 1990, 2000 and 2010. Panel (d) displays the relationship between upward income mobility to the probability that individuals stay in the CBSA for three generations, 1990, 2000 and 2010.

The first panel (a) displays the relationship between upward mobility and local inequality for two generations. The vertical axis is the upward mobility measures for generations 1990 and 2010, and the horizontal axis is inequality in CBSA in 1980 and 2000, respectively. We find a negative relationship: individuals from CBSAs with large income inequality among workers are likely to experience lower upward mobility. This is related to the *Great Gatsby curve* in the U.S., showing the negative relationships between local inequality and upward mobility. In my model, this arises from the specialization and wage disparity in the local labor market, leading to less opportunity in the choice of industry for the next generation. The second panel (b) shows the structural transformation and upward mobility. As seen in Figure 2.2, CBSAs of large employment size in the manufacturing sector exhibited structural transformation to services. Therefore, panel (b) implies that structural transformation lowers the upward mobility of individuals. This accounts for the part of $LL^s_{i,t+1}$. In the third panel (c), we consider the land of opportunities for individuals that are related to intersectoral mobility ($ISM^s_{i,t+1}$). The horizontal axis is an expected utility from industry choice for individuals in CBSAs: large values correspond to the land of opportunities for the future. Therefore, such CBSAs exhibit high upward mobility. Over generations, the relationship becomes more robust. This confirms that the disparity in the land of opportunities drives an increase in the spatial variation of upward mobility. The last panel (d) describes the role of labor mobility across CBSAs that is related to $GLM^s_{i,t+1}$. Intuitively, the low mobility of workers in geography predicts less possibility of climbing up the location ladder *ceteris paribus*. This panel shows the probability of non-migration for individuals from the CBSAs on the horizontal axis. As predicted in theory, a high probability of staying in origin is associated with low intergenerational income mobility. This is consistent with the decline of upward mobility along with a lower migration rate in the U.S. economy during the last decades.

This section has presented the calibrated results for the U.S. CBSAs and quantitative analysis for the general equilibrium implications for measured TFP, workers' welfare, inequality and intergenerational income inequality. The quantitative analysis reveals the underlying mechanisms to create their spatial variation and dynamics. In the next section, we perform counterfactual exercises to understand the contribution of these mechanisms when there are shocks in the fundamentals.

7 Counterfactual Experiments

Armed with the data and parameters calibrated above, we undertake counterfactual experiments to understand the quantitative impact of the development of fundamental productivity and amenities in shaping the variation of structural transformation, welfare and intergenerational income mobility. The calibration yields the trajectory of fundamental productivity ($A^s_{i,t}$) and fundamental amenities ($\Omega^s_{i,t}$) that conflates location and industry specific amenities for workers ($B^s_{i,t}$), migration barrier to CBSA ($M_{i,t}$) and sector specific taste parameters (ζ_s). The objective of undertaking counterfactuals is to understand their quantitative importance of these fundamentals to explain the spatial heterogeneity of workers' location choices and industry choices and finds that the changes in workers' mobility across locations and industries determine their welfare gains relative to the previous generation and their position on the income ladder. Shocks to fundamental productivity are salient in shaping individual consequences. Consider the sector level negative shock to the productivity in any particular period. The standard mechanisms are the following: it directly lowers TFP and overall productivity and lowers labor demand and wage, and workers are less likely to sort into the sector, and lower labor supply counteracts the negative impact on wages. However, the present model has additional channels to amplify the general equilibrium effects. First, change in income leads to demand shift of workers due to non-homothetic preference, therefore feedback loop in the goods market: lower income leads to less demand for the services sector. Second, there are frictions in the workers' adjustment due to the exposure effect in the local labor market. These effects play out across space, leading to different consequences on their mobility.

We undertake counterfactuals for the fundamental productivity and amenities, respectively. In the first subsection, we consider the impact of productivity shocks on structural change, welfare and intergenerational income mobility of workers. This allows us to study the importance of technological progress in shaping individual-level consequences in the last decades. In the second subsection, we undertake the counterfactuals when there are shocks in the fundamental amenities. Specifically, we consider the lower barrier for migration across CBSAs. The counterfactual experiment reveals the importance of the location specific environment in explaining the spatial pattern of structural change and labor mobility. In addition, for amenities, we undertake the counterfactuals where amenities become uniform across space. The motivation for this counterfactual is to understand the persistent role of differences in fundamentals to explain the equilibrium allocation in the later period. In the counterfactual experiments, the economy starts from the actual equilibrium observed in the data in 1980 and we implement the changes in the fundamentals to solve the counterfactual equilibrium.[18]

7.1 Productivity Changes

The first set of counterfactuals assumes that there are negative shocks to the fundamental productivity. We undertake the first counterfactual where the fundamental productivity of the services sector ($\{A^j_{i,t}\}_{j\in\text{Service}}$) is dropped by 10 percent in 1990 relative to the observed level and fixed at the level for the later periods, 2000 and 2010. Therefore, the fundamental productivity of all service industries is fixed at 90 percent of the level of 1990 throughout time. The second exercise focuses on particular industries. As discussed in the previous section, the IT-intensive industry such as FIRE has shown a significant increase in TFP in CBSAs in the U.S. economy. To see the role of such a rise in IT-intensive industries in the service sector, we assume that there was no such positive technological progress in FIRE and communication services. We set their productivity at 90 percent

[18] The uniqueness of the dynamic equilibrium is not guaranteed in the presence of spillovers in productivity and intersectoral linkages in production. Therefore, we compute the counterfactual equilibrium with the observed equilibrium as a starting point and run the model with a small perturbation of the initial equilibrium to assess the local uniqueness of the counterfactual equilibrium.

of the level in 1990 throughout time. Further, we also consider the technological progress in the manufacturing sector for the comparison of their impacts. In particular, we focus on industries that use robots intensively, the electric and computer industry and transport equipment.[19] We set lower productivity for these two industries in the same manner as services.

The Panel A of Table 7.2 reports the results for these three counterfactuals about the TFP changes and structural changes. The rows for the first counterfactual experiment (i) shows the negative impact on services sector TFP ($\delta_{i,t}^{\text{Service}}$) defined in (34). In 1990, it shows 8.5 percent lower than the baseline economy. Since the TFP is determined by both fundamental productivity and endogenous mechanisms through labor mobility (Proposition 2), the absolute effect is less than 10 percent, and it implies that the workers' adjustment mitigates the negative shock on average. More interestingly, the negative effect becomes smaller over time. This implies that the negative impact of fundamental productivity shocks in the initial period can be faded out through workers' mobility over generations. We also find an increase in the variation of the negative effect over time, implying large heterogeneity in adopting the negative shocks across CBSAs. Row 2 in Panel A shows the difference in the employment share of services to the baseline economy. When turning off the technological progress in the service sectors, we see a significant drop in the employment share of services. This happens for two reasons. The first channel is the traditional effect of factor mobility across sectors. The second channel is an additional impact of the demand-driven structural changes. When we abstract the exogenous fundamental productivity growth, the real income of workers becomes low and the expenditure shift from goods (manufacturing and housing) to services is slowed down. Therefore, it further prevents the labor shift to services. This mechanism through the demand side is fundamental as we see its effect in the counterfactual (iii) in the table. The counterfactual (iii) does not introduce the direct negative shocks to the services sector, but we see a low employment share compared to the baseline.

Table 7.3 reports the results for the impact on welfare. For each counterfactual, the first row shows the average percentage change of the welfare dynamics ($\mathrm{d}\ln V_{i,t}$), and other rows show those for different margins in (35). The average of changes in welfare dynamics is small over periods, but there is a significant variation. The standard deviation of the changes shows 3.2 for generation 2000 and 1990 and 3.1 for generation 2010 and 2000. The pattern is the same for other counterfactuals (ii) and (iii). When we investigate the margins, the main contribution to the change of welfare dynamics is the gains from consumption, and the gains exhibit large variation. This confirms that the adverse productivity shocks create a large variation in real income across CBSAs. Part of this result is due to the fluctuation of housing prices, as a large variation in real income is associated with a large variation in housing prices. Among other mechanisms, counterfactual migration gains show large values compared to the job opportunity gains. We find substantial differences in these two margins across CBSAs. Their inter-quantile ranges are similar to those of changes in welfare dynamics.

Figure 7.10 presents the welfare effects and change in intergenerational income mobility across 395 CBSAs when undertaking the counterfactual for the low productivity in all service industries.

[19] These industries show the highest penetration rate of robots in the U.S. economy (Acemoglu and Restrepo 2020).

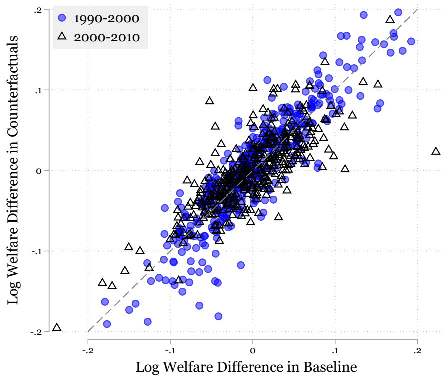
Panel (a) displays the variation of welfare effects across CBSAs. For the welfare difference between generation 1990 and 2000, the CBSAs with welfare losses in the baseline show further welfare losses in the counterfactual. The technological progress in services and structural transformation benefit these CBSAs in the baseline economy. Panel (b), (c) and (d) displays the distribution of the upward income mobility across CBSAs for different generations. An important takeaway from the first generation, in Panel (b), is that the negative impact on the productivity of services leads to a large variation of upward income mobility for the generation 1990. Once the productivity of services is fixed after 1990, Panels (c) and (d) show less variation of upward mobility. Table 7.4 reports the impact on intergenerational income mobility. Row 1 to 3 show the results for the first set of counterfactuals. For each generation, we report the percentile change of $\widetilde{M}_{i,t}$ from the baseline economy. For the first generation, the average impact is relatively small since the 10 percent decline in productivity for all CBSAs does not alter the location choice of workers much, and it turns out to be a smaller effect. However, the generations of 2000 and 2010 show higher upward mobility on average. For generation 2000, individuals experience around 5 to 6 percentage increase in the upward mobility compared to the baseline economy. The logic for this result is the following. When the exogenous productivity is absent, the endogenous spillover in productivity becomes salient, and workers sort into the place with agglomeration. In addition, as we see in welfare results, a larger variation of real income growth creates workers' mobility both across locations and sectors. Together with these endogenous responses of workers, we see higher upward mobility on average, but with large variation in its gain. The variation in the change of intergenerational income mobility becomes large over time.

7.2 Amenities and Migration Barriers

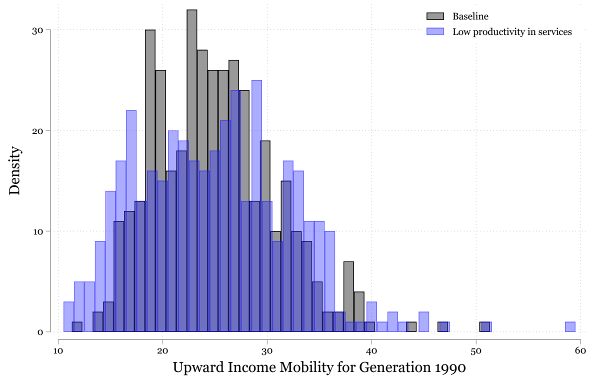
The second set of experiments undertakes the counterfactuals for the fundamental location characteristics in amenities. By construction, the variation of overall amenities across space include both the variation of fundamental benefit ($B^s_{i,t}$) and migration barrier ($M_{i,t}$). Then, we start with the lower migration barrier by 10 percent uniformly for every location. This benefits any workers in the economy, as it is isomorphic to an increase of benefit from residing and working in any particular location. Yet, workers' choices are not necessarily the same as the baseline since workers are ex-ante different in their origin, and the bilateral cost of migration defines the aggregate benefit of labor market access differently across workers. As another counterfactual about the migration barrier, we set a 10 percent lower migration barrier for top CBSAs. We define the top 50 CBSAs based on the total employment size in 1980, selecting them for the counterfactual experiments. Given that most migration occurs from small towns or cities to large cities, this counterfactual is of interest to consider whether such directed migration is important to explain the variation of structural change, welfare and upward mobility.

Panel B in Table 7.2 shows the results. In (iv), we can find TFP growth in the services sector. On average, service sector TFP exhibits 1.2 percent higher than the baseline economy in 1990, in-

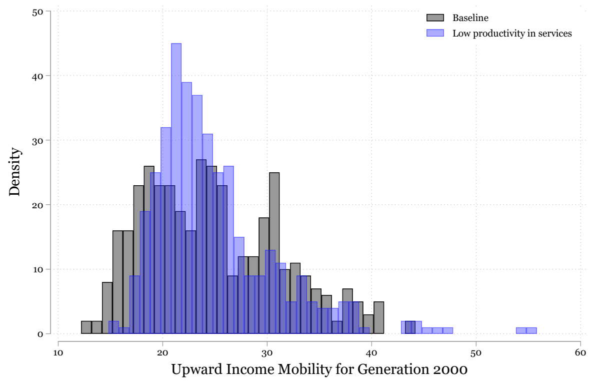
Figure 7.10: Welfare Effects and Intergenerational Income Mobility for the Productivity Change in Services



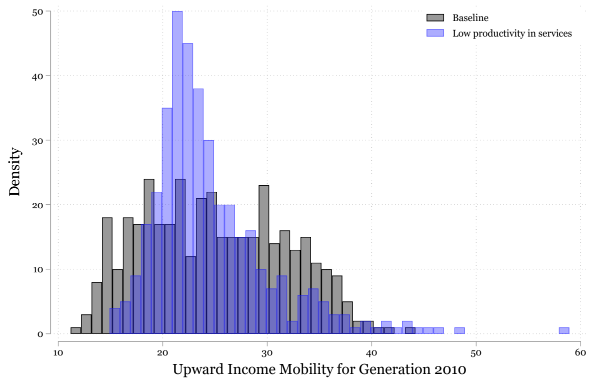
(a) Welfare effects



(b) Upward Income Mobility for Generation 1990



(c) Upward Income Mobility for Generation 2000



(d) Upward Income Mobility for Generation 2010

Note: These figures show the results for welfare and intergenerational income mobility for the counterfactual when fundamental productivity of all service industries (transport service, wholesale trade, retail, FIRE, health service, education and legal, communication service and other services) is dropped by 10 percent in 1990 and fixed over time. Panel (a) shows the welfare difference for the baseline and the counterfactual between two generations, $\mathrm{d}\ln V_{i,t}$. Blue dots (black triangles) show the welfare differences between generations 1990 and 2000 (2000 and 2010), respectively. In panels (b), (c) and (d), we report the distribution of upward income mobility for different three generations, generation 1990, 2000 and 2010. In each panel, gray bars show the distribution of the upward income mobility measure across CBSAs in the baseline, and the blue bars show that for the counterfactual economy.

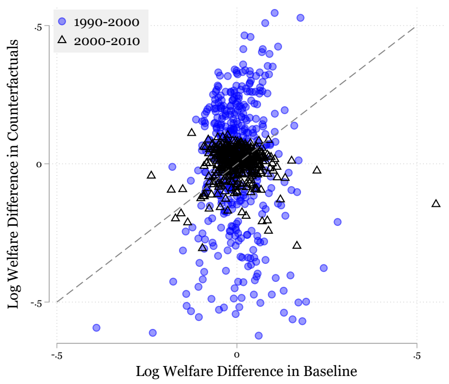
creasing to 4.5 percent in 2010. In contrast, the employment share of the services sector is dropped compared to the baseline economy and the change shows large variation. This is rationalized by the specialization of workers in CBSAs that exhibit relatively high amenity and productivity. Once the migration barrier is lower, workers are directed to such CBSAs and the clustering of workers counteracts the movement of workers across sectors due to the persistence in their job choice. More interestingly, comparing the results between counterfactuals (iv) and (v) finds that lowering the migration barriers for the top 50 CBSAs has a similar impact to lowering them for all CBSAs. This suggests that workers directed sorting to the large cities is essential to consider the role of the migration barriers in shaping the extent of structural change and TFP dynamics.

Turning to welfare, Panel B in Table 7.3 shows similar results as in the previous counterfactuals. The main driver of welfare dynamics is the change in the gains from consumption and their effects are persistent over generations. In Table 7.4, we find that upward mobility becomes high when we have a low migration barrier. This is consistent with the theory and quantitative analysis in the previous sections. Individuals are able to sort into the location with a higher return for any particular industry, and they have more opportunities to climb up the income ladder by the location choices. We also confirm that this mechanism is mainly at work for the migration to the large cities by comparing the similar magnitude in two counterfactuals (v) and (vi).

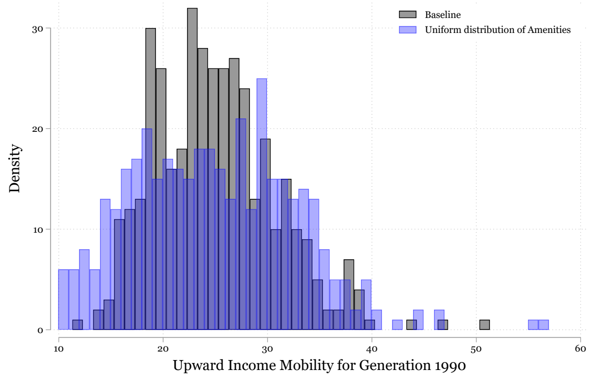
Spatial variation of amenities. Lastly, we investigate the role of differences in fundamental amenities across CBSAs. To this end, we perform the counterfactual in which overall amenities develop at the same rate across all CBSAs given any particular industry. As reported in (vi) in Table 7.2, this further benefits the TFP growth of the services sector. Once we turn off the difference in fundamental amenities among CBSAs, we predict a 7.1 percent increase in service sector TFP on average in 1990, and it becomes 14.16 percent in 2010. In Table 7.3, we find the largest welfare gains. In (vi), welfare dynamics are larger than the baseline economy by 2.8 percent for the generations 1990 and 2000, and it is 0.3 percent for generations 2000 and 2010. Seeing the decomposition of the effect, this substantial effect arises through the gains from a job opportunity. When equalizing the value of amenities, workers' industry choice and location choice are purely determined by the return of industry choice in the current location. Therefore, individuals achieve large gains from job opportunities. In Table 7.4, we also find substantial positive effects on intergenerational income mobility. This is also consistent with the benefit of job opportunities in the local labor market. Individuals are more likely to achieve a higher position of income rank compared to the previous generation. The measure of upward mobility becomes 9.2 percent higher for those in generation 2000 and 10.8 percent higher for generation 2010 on average. However, endogenous agglomeration of industries and ex-ante distribution of workers keep such gains substantially different across space.

Figure 7.11 presents the welfare effects and change in intergenerational income mobility for this counterfactual experiment. In Panel (a), the spatial variation in welfare differences between generations 1990 and 2000 is magnified in the counterfactual. Intuitively, equalizing amenities allows the first generation to change their location choices such that they move to productive and high real income places. This magnifies the differences of such gains among CBSAs, and, therefore, more spatial inequality in welfare gains. For the generations 2000 and 2010, the spatial variation of such gains becomes small since workers' location choices show the path dependency for each industry. Panel (b), (c) and (d) shows that the upward income mobility for generation 1990 exhibits a larger variation in the counterfactual than the baseline, while the negative impact on average. For other generations, the distribution becomes small in the counterfactual since the spatial variation in the labor mobility is less relative to the baseline once the geographical distribution of workers shows persistence after the change in the early period.

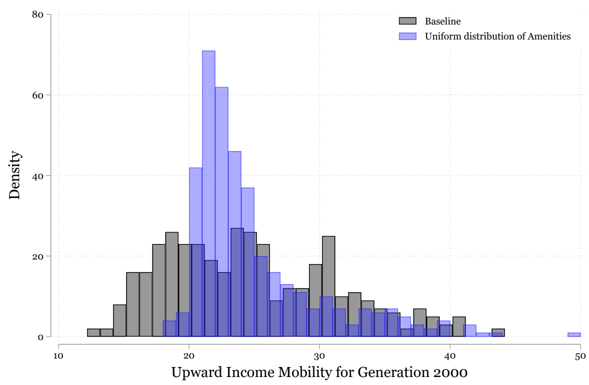
Figure 7.11: Welfare Effects and Intergenerational Income Mobility for the Uniform Distribution of Amenities



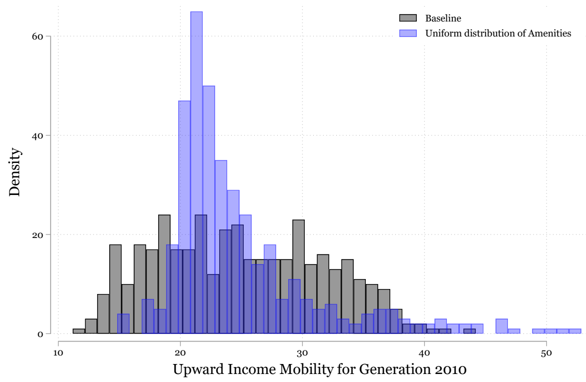
(a) Welfare effects



(b) Upward Income Mobility for Generation 1990



(c) Upward Income Mobility for Generation 2000



(d) Upward Income Mobility for Generation 2010

Note: These figures show the results for welfare and intergenerational income mobility for the counterfactual when overall amenities become uniform across CBSAs for given industry choices. Panel (a) shows the welfare difference for the baseline and the counterfactual between two generations, $\mathrm{d}\ln V_{i,t}$. Blue dots (black triangles) show the welfare differences between generations 1990 and 2000 (2000 and 2010), respectively. In panels (b), (c) and (d), we report the distribution of upward income mobility for different three generations, generation 1990, 2000 and 2010. In each panel, gray bars show the distribution of the upward income mobility measure across CBSAs in the baseline, and the blue bars show that for the counterfactual economy.

Table 7.2: Counterfactual Experiments – Impact on Service TFP and Change in Service Employment Share

Panel A			1990				2000				2010			
			Mean	SD	25prc	75prc	Mean	SD	25prc	75prc	Mean	SD	25prc	75prc
(i)	Low Productivity	Service TFP	−8.522	2.320	−10.035	−6.883	−7.515	4.255	−9.095	−5.485	−6.074	5.482	−8.809	−3.542
	in All Services	Service Emp. Share	−13.107	5.038	−16.470	−9.453	−17.467	5.473	−21.454	−13.602	−21.582	5.863	−26.231	−17.288
(ii)	Low Productivity in	Service TFP	−2.139	2.392	−3.629	−0.623	−1.141	2.681	−2.739	0.288	0.433	4.132	−1.587	2.448
	IT intensive services	Service Emp. Share	−13.843	5.115	−17.294	−10.126	−19.027	5.644	−23.355	−14.709	−23.272	6.075	−27.965	−18.357
(iii)	Low Productivity in	Service TFP	1.187	2.5834	−0.508	3.004	1.622	2.42	−0.034	3.3224	4.295	5.299	1.669	6.486
	Robot intensive manufacturing	Service Emp. Share	−14.347	5.143	−17.926	−10.598	−20.277	5.824	−24.6534	−16.077	−25.323	6.332	−30.241	−20.596
Panel B			1990				2000				2010			
			Mean	SD	25prc	75prc	Mean	SD	25prc	75prc	Mean	SD	25prc	75prc
(iv)	Low Migration Barrier	Service TFP	1.254	2.581	−0.439	3.065	1.761	2.42	0.104	3.47	4.506	5.339	1.85	6.632
	for All CBSAs	Service Emp. Share	−14.165	5.128	−17.667	−10.403	−19.812	5.704	−24.236	−15.567	−24.54	6.09	−29.26	−20.091
(v)	Low Migration Barrier	Service TFP	1.295	2.599	−0.417	3.044	1.827	2.455	0.16	3.469	4.588	5.369	1.904	6.758
	for Top 50 CBSAs	Service Emp. Share	−14.169	5.135	−17.656	−10.432	−19.823	5.712	−24.161	−15.492	−24.548	6.1	−29.166	−20.129
(vi)	Uniform Amenities	Service TFP	7.151	3.781	4.901	9.248	9.867	5.209	7.555	12.979	14.167	6.222	11.017	17.183
	Change across Space	Service Emp. Share	−10.867	8.401	−16.432	−5.82	−15.193	9.357	−21.35	−9.145	−15.017	10.303	−20.892	−9.246

Note: For each counterfactual scenario (i) to (vi), we report the percentage change of aggregate TFP in the services sector and the change of employment shares in the service sector from the baseline economy. For each year, 1990, 2000 and 2010, we show the mean, standard deviation, 25 percentile and 75 percentile values across 395 CBSAs in the U.S. economy. Units of all entries are percentages. Panel A report results of counterfactual exercises about productivity changes. Counterfactual (i) undertakes counterfactuals when fundamental productivity of all service industries (transport service, wholesale trade, retail. FIRE, health service, education and legal, communication service and other services) is dropped by 10 percent in 1990 and the fundamental productivity level is unchanged later. The counterfactual (ii) suppose that fundamental productivity of IT intensive services (FIRE and communication services) is lower than baseline by 10 percent in 1990 and fixed over time. In the counterfactual (iii), we do the analogous experiment for robot intensive industries (electric and computer industry and transport equipment). Panel B reports counterfactuals about amenities. Counterfactual (iv) undertakes the counterfactual in which migration barriers of all CBSAs are set to be 10 percent lower than the baseline in every period. In counterfactual (v), we set a lower migration barrier for 50 CBSAs defined based on population in 1980. The counterfactual (vi) assumes that overall amenities are uniform across locations. We compute the geometric mean of overall amenities across CBSAs for workers in each sector, and we implement the value for all locations in each period.

Table 7.3: Counterfactual Experiments – Impact on Welfare

Panel A				1990 – 2000				2000 – 2010			
				Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc
(i)	Low Productivity in All Services	Welfare		0.052	3.219	−1.687	2.112	0.05	3.17	−1.971	1.733
			Consumption	0.126	5.181	−3.033	2.231	0.125	4.962	−2.453	3.231
			Migration Gain	0.08	3.897	−1.433	2.433	0.085	4.18	−2.256	1.698
			Job Opportunity Gain	0.052	3.201	−1.933	2.129	0.037	2.715	−1.729	1.885
(ii)	Low Productivity in IT intensive services	Welfare		0.044	2.945	−1.637	1.888	0.051	3.173	−1.923	1.713
			Consumption	0.126	5.196	−3.23	2.383	0.126	4.98	−2.457	3.131
			Migration Gain	0.078	3.841	−1.275	2.404	0.086	4.217	−2.257	1.691
			Job Opportunity Gain	0.049	3.102	−1.816	2.129	0.036	2.666	−1.737	1.768
(iii)	Low Productivity in Robot Intensive manufacturing	Welfare		0.044	2.963	−1.643	1.963	0.054	3.284	−2.004	1.859
			Consumption	0.133	5.344	−3.457	2.461	0.142	5.282	−2.725	3.238
			Migration Gain	0.081	3.912	−1.51	2.326	0.1	4.554	−2.516	1.829
			Job Opportunity Gain	0.05	3.145	−1.772	2.046	0.039	2.766	−1.725	1.92
Panel B				1990 – 2000				2000 – 2010			
				Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc
(iv)	Low Migration Barrier for All CBSAs	Welfare		0.043	2.924	−1.584	1.887	0.052	3.202	−1.882	1.879
			Consumption	0.121	5.112	−3.364	2.374	0.128	5.000	−2.548	3.142
			Migration Gain	0.074	3.737	−1.391	2.467	0.087	4.255	−2.265	1.705
			Job Opportunity Gain	0.049	3.124	−1.819	2.05	0.037	2.686	−1.648	1.893
(v)	Low Migration Barrier for Top 50 CBSAs	Welfare		0.053	3.251	−1.917	1.895	0.053	3.247	−1.782	1.845
			Consumption	0.121	5.104	−3.399	2.68	0.128	5.004	−2.509	3.231
			Migration Gain	0.075	3.745	−1.519	2.419	0.087	4.26	−2.272	1.73
			Job Opportunity Gain	0.058	3.406	−1.945	2.165	0.038	2.727	−1.73	1.794
(vi)	Uniform Amenities Change across Space	Welfare		2.89	23.375	−13.812	21.407	0.344	7.944	−3.45	5.546
			Consumption	0.852	13.383	−9.98	9.665	0.116	4.737	−2.376	2.755
			Migration Gain	0.901	13.254	−7.365	9.403	0.675	11.274	−5.116	7.315
			Job Opportunity Gain	2.374	21.418	−14.399	18.771	0.244	6.851	−3.274	4.858

Note: For each counterfactual (i) to (vi), we report percentage change of the welfare differences over generations. Welfare differences are defined for between generation 2000 to 1990 and 2010 to 2000. We show the mean, standard deviation, 25 percentile and 75 percentile values of their changes across 395 CBSAs in the U.S. economy. Units of all entries are percentages. Panel A report results of counterfactual exercises about productivity changes. Panel B reports counterfactuals about amenities. See note of Table 7.2 for more details of counterfactuals.

Table 7.4: Counterfactual Experiments – Impact on Intergenerational Income Mobility

Panel A		1990				2000				2010			
		Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc
(i)	Low Productivity in All Services	−0.772	13.405	−7.408	8.752	5.612	30.591	−16.232	23.18	6.692	36.848	−19.431	27.432
(ii)	Low Productivity in IT intensive services	−0.799	13.596	−7.552	9.003	5.396	30.236	−17.408	23.461	5.971	35.18	−17.739	23.469
(iii)	Low Productivity in Robot intensive manufacturing	−0.82	13.792	−7.72	9.128	5.332	29.536	−16.303	23.62	5.802	35.517	−18.338	24.659
Panel B		1990				2000				2010			
		Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc
(iv)	Low Migration Barrier for All CBSAs	−0.821	13.754	−7.792	9.182	5.342	29.915	−16.806	23.297	5.621	34.982	−18.316	23.462
(v)	Low Migration Barrier for Top 50 CBSAs	−0.814	13.779	−7.6	9.264	5.322	29.886	−17.243	23.004	5.634	34.956	−18.656	23.911
(vi)	Uniform Amenities Change across Space	−1.202	14.948	−6.838	7.264	9.222	41.652	−24.22	34.772	10.829	51.562	−27.384	30.928

Note: For each counterfactual (i) to (vi), we report percentage change of the intergenerational income mobility from the baseline values. For each year, 1990, 2000 and 2010, we show the mean, standard deviation, 25 percentile and 75 percentile values across 395 CBSAs in the U.S. economy. Units of all entries are percentages. Panel A report results of counterfactual exercises about productivity changes. Panel B reports counterfactuals about amenities. See note of 7.2 for more details of counterfactuals.

8 Conclusion

The interplay between structural transformation in the aggregate and local economies is key to understanding spatial inequality and worker mobility. To look at this, we have developed a dynamic economic geography model with overlapping generations that accommodates the frictional adjustment of workers across locations and industries, non-homothetic preference and productivity spillovers in a tractable way. The theoretical framework provides insights into the cross-sectional disparity and intergenerational income inequality among workers that arise due to structural changes in the economy. We have calibrated the model with the U.S. economy and despite the high number of dimensions – on location, industry and time – the model structure allows us to back out productivity and amenities from the data. And this, in turn, enables us to quantitatively assess the importance of different mechanisms that drive spatial variation in total factor productivity (TFP), welfare dynamics, inequality and intergenerational income mobility. The dynamic nature of the spatial model therefore allows us to study phenomena that have received limited scrutiny but which are of fundamental interest in a country which is increasingly riven by growing inequality and barriers to upward mobility.

This paper allows us to understand how the structure of the spatial economy - through trade and migration, local labor market exposures and agglomeration - shapes individual outcomes. We begin to understand why in the same country, the citizens of San Jose are on entirely different trajectories than those in Cleveland. Why rising levels of inequality might constrain upward mobility as characterized by the Gatsby Curve. These are issues at the top of the policy agenda not just in the U.S. but in countries across the world. In effect, this paper is trying to open the black box of how the structure of economy not just across space but also across time can influence patterns of inequality and mobility in different locations. To do this, we perform counterfactual experiments using the parameterized model, which enables us to quantify the importance of technological progress and spatial variation in amenities in determining the pace of structural transformation across locations. Through such counterfactual analysis, we find that the productivity growth of industries that drive structural change and higher migration barriers limit upward mobility. In addition, the persistent variation in productivity and amenities across geographies is critical to explaining the regional disparity in TFP changes and workers' mobility. These results suggest they are critical to understanding how mobility can be encouraged and inequality in an economy that is increasingly dominated by services.

As seen in Figure 1.1, structural transformation shows uneven patterns across space in the U.S. While there has been sustained deindustrialization over the last half-century, manufacturing employment share remains high in most cities in the Rust Belt, including Buffalo, Cleveland, Detroit, Pittsburgh, and St. Louis. As late as the 1990s, the majority of Silicon Valley technology jobs were still in hardware, and the region was surrounded by fabrication plants building silicon semiconductors. Today, Silicon Valley has a very limited number of fabrication plants; however, it remains the dynamic global center of the communication service industries. Using the U.S. data, we show that the geographical variation of amenities and productivity spillovers are the main driver of such unevenness, and historical exposure to different industries creates persistence in the occupational structure. This paper demonstrates how different patterns of structural change across both space and time determine the geographical variation of welfare and the upward mobility of workers. Understanding this is critical to understanding how the U.S. as a whole and not just a few cities within it can regain the "land of opportunity" mantle.

The framework proposed is easily extended to quantify the effects of various shocks on local economies and workers within them in the long run. Amongst possible shocks, the interaction between locations and the rapidly changing international market is perhaps the most important to look at. Globalization and in particular the U.S. relationship with China is very much in the spotlight in terms of understanding why some cities in the U.S. have prospered whilst others have declined. The recent work with Italo Colantone and Gianmarco Ottaviano (Colantone et al. 2021) is looking at whether higher trade exposure to Chinese imports in the U.S. reduces social mobility, both in absolute and relative terms, conditioning for the initial level of inequality. The foundation of that paper is the framework developed in this paper, which allows us to quantify the redistributive impacts of the trade shock across space and time. Our model can also serve as a stepping-stone for analysis of the effects of other dynamic processes. Using the model, we can look at how environmental damages including climate change and air and water pollution might affect inequality and mobility across locations in the U.S. Another research avenue we are to pursue is applying my framework to locations within developing countries where the overall pace of structural change tends to be more rapid but where we understand little about distributional effects across space and time. The framework developed in this paper when combined with developing country data serves as an interesting laboratory for understanding variation in inequality and mobility. This understanding is fundamental to designing policies to equalize opportunities across locations within countries, something which is very much at the top of the global policy agenda as the world moves gradually out of the pandemic.

Appendices for Chapter 1

A Appendix: Model details

An index is omitted when it does not cause any confusion in some expressions. Subsection A.1 presents the details of the demand system featuring non-homothetic constant elasticity of substitution. Subsection A.2 provides micro foundations and derivation of the choice of industry job. Subsection A.2 explains the details of production side and Subsection A.4 provides problem of developers. Subsection A.5 discusses an extension of the baseline framework with introducing educational choices. Subsection A.6 discusses the comparison between the present model and the infinite horizon forward-looking model.

A.1 Demand system

The utility from consumption $\mathbb{C}$ is implicitly defined as a solution to

$$\sum_{s\in\mathcal{K}} u_s^0\left(\frac{c_s}{u_s^1(\mathbb{U})}\right) = 1 \tag{A.1}$$

$u_s^0(\cdot)$ is strictly decreasing and strictly convex function, and $u_s^1(\cdot)$ is monotonically increasing function. See Hanoch 1975 and Matsuyama 2019 for further discussion about the class of implicitly additive separable utility functions. The representation (A.1) nests (i) homothetic constant elasticity of substitution (CES) sectoral consumption across sectors, as well as (ii) non-homothetic CES structure (Comin et al. 2020). This paper adopts the latter by parametrizing two functions by

$$u_s^0(x) = \alpha_s^{(\sigma-1)/\sigma} x^{\frac{\sigma-1}{\sigma}}, \quad u_s^1(x) = x^{(\theta_s-\sigma)/(1-\sigma)}$$

where (α_s, σ, θ_s) are exogenous preference parameter. Throughout our analysis, these parameters satisfy:

$$\sigma \in (0,\ 1), \quad \theta_s \geq \sigma \tag{A.2}$$

where the sectoral consumptions are gross complement and utility function is concave. Alternatively, $\sigma \in (1, \infty)$ and $\sigma > \theta_s$ for the case of substitution between sectors in consumption.

Letting W be income of individuals and $\boldsymbol{p}$ be price vector, utility maximization problem of an individual is:

$$\boldsymbol{c} \in \arg\max \ \mathbb{U} \quad s.t. \quad \boldsymbol{p}\cdot\boldsymbol{c} \leq W \text{ and (A.1)} \tag{A.3}$$

Taking the first order condition associated with the dual problem of (A.3), the expenditure share becomes:

$$\psi_s = \alpha_s^{\sigma-1}(p_s/\mathcal{P})^{1-\sigma}(W/\mathcal{P})^{\theta_s-1} \tag{A.4}$$

with the aggregate price index for workers in local market (i,k) satisfying:

$$\begin{aligned} \mathcal{P}_i^k &= \left(\sum_{s\in\mathcal{K}} \alpha_s^{\sigma-1}(p_i^s)^{1-\sigma}(W_i^k/\mathcal{P}_i^k)^{\theta_s-1}\right)^{1/(1-\sigma)} \\ \iff \mathcal{P}_i^k &= \left(\sum_{s\in\mathcal{K}} \left(\alpha_s^{\sigma-1}(p_i^s)^{1-\sigma}(W_i^k)^{\theta_s-1}\right)^{\frac{1-\sigma}{\theta_s-\sigma}} (\psi_{s|ik})^{\frac{1-\theta_s}{\sigma-\theta_s}}\right)^{1/(1-\sigma)} \end{aligned} \tag{A.5}$$

The aggregate price index $\mathcal{P}$ is the unique solution to this fixed point equation given $(\boldsymbol{p}, \boldsymbol{W})$. The aggregate price index is the homogeneous of degree one in $(\boldsymbol{p}, \boldsymbol{W})$, that is the expenditure share is homogeneous of degree zero in $(\boldsymbol{p}, \boldsymbol{W})$. Expenditure on sector s, $E_s(\boldsymbol{p}, \boldsymbol{W})$ is increasing and concave in $\boldsymbol{p}$, and homogeneous of degree one in $(\boldsymbol{p}, \boldsymbol{W})$.

Define the average of θ_s weighted by expenditure shares:

$$\bar{\theta}(\boldsymbol{p}, \boldsymbol{W}) \equiv \sum_{s\in\mathcal{K}} \psi_s(\boldsymbol{p}, \boldsymbol{W})\theta_s \tag{A.6}$$

Using this, the elasticity of aggregate demand and sectoral expenditure with respect to sectoral price change are:

$$\frac{\partial \ln \mathbb{U}(\boldsymbol{p}, \boldsymbol{W})}{\partial \ln p_s} = -(1-\sigma)\frac{\psi_s(\boldsymbol{p}, \boldsymbol{W})}{\bar{\theta}(\boldsymbol{p}, \boldsymbol{W})}, \quad \frac{\partial \ln E_s(\boldsymbol{p}, \boldsymbol{W})}{\partial \ln p_s} = (1-\sigma)\left(1 - \frac{\theta_s}{\bar{\theta}(\boldsymbol{p}, \boldsymbol{W})}\psi_s(\boldsymbol{p}, \boldsymbol{W})\right) \tag{A.7}$$

The sector with higher expenditure share exhibits lower elasticity of expenditure to the change in price. Alternatively,

$$\frac{\partial \ln E_s(\boldsymbol{p}, \boldsymbol{W})}{\partial \ln W} = \theta_s - \frac{\theta_s-\sigma}{1-\sigma}\Big(\bar{\theta}(\boldsymbol{p}, \boldsymbol{W}) - 1\Big), \quad \mathcal{E}_W^\psi \equiv \frac{\partial \ln(\psi_s(\boldsymbol{p}, \boldsymbol{W})/\psi_{s'}(\boldsymbol{p}, \boldsymbol{W}))}{\partial \ln W} = \theta_s - \theta_{s'} \tag{A.8}$$

This describes how household expenditure on a particular good or service varies with household income. As (real) income changes, the expenditure pattern shifts between sectors differently. This feature of the Engel curve matters to explain the interplay between structural change and sectoral specialization along with growth. Lastly, we note that the elasticity of substitution conditional on the aggregate demand is constant:

$$\mathcal{E}_P^\psi \equiv \left.\frac{\partial \ln(\psi_s(\boldsymbol{p}, \boldsymbol{W})/\psi_{s'}(\boldsymbol{p}, \boldsymbol{W}))}{\partial \ln(p_s/p_{s'})}\right|_{W/\mathcal{P}} = 1-\sigma \tag{A.9}$$

The same proportional change in relative price affects consumption patterns in the same way as the homothetic CES demand system.

A.2 Geography of job opportunities

There are measure of $L^s_{i,t}$ workers in sector s in location i at period t. Without population growth, the total population of cohort $t+1$ (i.e., next generation) with origin i is $\bar{L}_{i,t+1} = \sum_{s\in\mathcal{K}} L^s_{i,t}$. The discussion is not altered when introducing the uniform birth rate across locations. The young generation is ex ante homogeneous and an individual has a unit time for job choice during the young period. Since the young generation does not obtain utility from consumption or leisure, there is no incentive to spend any time other than a job choice.

Consider young individuals in location i in period t. During the young period, an individual acquires information from existing workers in location i. Suppose one unit of time is divided into T spans and let $\Delta = 1/T$. In each span Δ, an individual spends the time to acquire information regarding a job in each sector. During time span Δ, an individual receives the valuable(positive) information about sector s with probability J^s_i, and the valueless(negative) information with probability $1 - J^s_i$. To achieve the probability J^s_i, an individual must spend time:

$$\mathcal{O}(J^s_i, L^s_{i,t}) = \Lambda_s \ln\left(\frac{1}{1-J^s_i}\right)\left(L^s_{i,t}\right)^{-\eta} \tag{A.10}$$

where Λ_s and η are strictly positive constant. Λ_s is a scalar for the time used for information acquisition, and η quantifies how much workers can save time when there are more existing workers in the local labor market.

Then, an individual decide time allocation across different sectors to maximize the logit of probabilities. The large value of logit corresponds to a large value of odds, and maximizing logit implies maximizing odds of acquiring positive information over acquiring negative information. Alternatively, people minimize the coefficient of variation for the number of valuable information they receive during the period. The coefficient of variation captures the relative variation of successful information acquisition over the average success rate given by $\sqrt{\frac{1-J^s_i}{J^s_i}}$. Thus, minimizing it is the same as maximizing the logit.

In summary, an individual solves:

$$J^s_{i,t} = \arg\max_{j^k_i\in(0,1)} \left\{\sum_k \ln\frac{j^k_i}{1-j^k_i} \quad \text{s.t.} \quad \sum_k \mathcal{O}(j^k_i, L^k_{i,t}) \le \Delta\right\} \tag{A.11}$$

Solution for this is:

$$J^s_{i,t} = \frac{(L^s_{i,t})^\eta}{\Lambda_s O_i} \tag{A.12}$$

where O_i is Lagrangian multiplier for individuals that solves:

$$\sum_{k \in \mathcal{K}} \Lambda_k \ln \left(\frac{\Lambda_k O_i}{\Lambda_k O_i - (L^k_{i,t})^\eta} \right) (L^k_{i,t})^{-\eta} = \Delta \tag{A.13}$$

Then, the probability that an individual of cohort $t+1$ successfully acquires valuable information during each time Δ is $J^s_{i,t}$. The probability of realization of T' successful information acquisition becomes ${}_T\mathrm{C}_{T'} \left(J^s_{it}\right)^{T'} \left(1 - J^s_{it}\right)^{T-T'}$ for $T' \leq T$. Taking its limit $\Delta \to 0$, for one unit of time, the ralization of the number of valuable information follows:

$$\mathcal{B}^s_{i,t}(m) = \frac{(J^s_{i,t})^m}{m!} \exp\bigl(-J^s_{i,t}\bigr) \tag{A.14}$$

The number of shocks that arrive to individuals of cohort $t+1$ exhibits the average $J^s_{i,t}$ and variance $J^s_{i,t}$. Intuitively, when $\eta > 0$, the average number is large as existing workers increase as marginal cost for acquiring positive information is low when $L^s_{i,t}$ is large. Further, many existing workers lead to a large variance of arrivals.

Next, the value of tastes follows Pareto distribution. The distribution is tractable and fitted to our context in which people have interests in the large value of tastes. Suppose that taste across employees has a distribution with density $g(\upsilon)$, where υ can be negative and positive values. Using constant $\phi > 1$, the taste value is transformed to $z = \upsilon^{-1/\phi}$ and its density is:

$$f(z) = g(\upsilon) \cdot \left| \frac{d\upsilon}{dz} \right| = \phi z^{-(\phi+1)} g(z^{-\phi})$$

People have interests in the large value of taste z. Therefore, the tail distribution of the large values z is approximated by density function and its associated distribution function:

$$f(z) \sim \phi g(0) z^{-(\phi+1)}, \quad F(z) \sim 1 - g(0) z^{-\phi}$$

Therefore, Pareto distribution for the taste values is a simple transformation of variables from general distribution.

Summarizing our discussion, individuals of generation $t+1$ receive taste shocks for each industry and the number of shocks $m^k_{i,t}$ is following Poisson distribution:

$$\mathcal{B}^s_{i,t}(m) = \mathbf{Pr}(m^s_{i,t} = m) = \frac{(J^s_{i,t})^m e^{-J^s_{i,t}}}{m!}, \quad J^s_{i,t} = \frac{(L^s_{i,t})^\eta}{\Lambda_s O_i} \tag{A.15}$$

Value of the each shock is supposed to be following Pareto distribution for every sector:

$$F(z) = 1 - (z/\underline{z})^{-\phi}, \quad \phi > 1 \tag{A.16}$$

The important assumption is that the number of arrival shocks is specific to pair of industry and location, while the size of shocks is independent to industry and location.

An individual picks up the largest value from the tastes. Its cumulative distribution function is:

$$\begin{aligned}\mathcal{F}^s_{i,t}(z) = \mathbf{Pr}\Big(z^s_{i,t} \leq z\Big) &= \sum_{m=1}^{\infty}\left(\prod_{m'=1}^{m}\mathbf{Pr}\Big(z^s_{i,t}(m') \leq z\Big)\right)\mathcal{B}^s_{i,t}(m) + \mathcal{B}^s_{i,t}(0) \\ &= \sum_{m=0}^{\infty}\big(1-(z/\underline{z})^{-\phi}\big)^m \frac{(J^s_{i,t})^m e^{-J^s_{i,t}}}{m!} \\ &= e^{-J^s_{i,t}(z/\underline{z})^{-\phi}}\end{aligned} \tag{A.17}$$

Define:

$$G^s_{i,t}(u) = \mathbf{Pr}\Big(\bar{U}^s_{i,t}z^s_{i,t} \leq u\Big) = e^{-\mathcal{V}^s_{i,t}u^{-\phi}} \quad \text{with} \quad \mathcal{V}^s_{i,t} = J^s_{i,t}(\underline{z}\bar{U}^s_{i,t+1})^{\phi} \tag{A.18}$$

The pattern of choosing industry s among cohort $t+1$ in location i becomes:

$$\begin{aligned}\mathbf{Pr}(\bar{U}^s_{i,t}z^s_{i,t} \geq \bar{U}^k_{i,t}z^k_{i,t},\ \forall k \neq s) &= \int_{\underline{u}}^{\infty} g^s_{i,t}(u)\prod_{k\neq s}G^k_{i,t}(u)\mathrm{d}u \\ &= \frac{\mathcal{V}^s_{i,t}}{\sum_{k\in\mathcal{K}}\mathcal{V}^k_{i,t}}\left[e^{-\sum_{k\in\mathcal{K}}\mathcal{V}^k_{i,t}u^{-\phi}}\right]_{\underline{u}}^{\infty} \\ &\to \frac{J^s_{i,t}(\bar{U}^s_{i,t+1})^{\phi}}{\sum_{k\in\mathcal{K}}J^k_{i,t}(\bar{U}^k_{i,t+1})^{\phi}} \quad (as\ \underline{z}\to 0)\end{aligned} \tag{A.19}$$

The last equation takes the minimum of Pareto distribution (i.e., lower bound of the Pareto distribution) to zero and expands its support to $(0,\infty)$. Therefore, the share of cohort $t+1$ in location i who choose industry s when they are young is:

$$\ln \varsigma^s_{i,t+1} = \ln\zeta_s + \eta\ln L^s_{i,t} + \phi\ln\bar{U}^s_{i,t+1} - \ln\left(\sum_{k\in\mathcal{K}}\zeta_k(L^k_{i,t})^{\eta}(\bar{U}^k_{i,t+1})^{\phi}\right) \tag{A.20}$$

where we denote $\zeta_s = 1/\Lambda_s$. The distribution of indirect utility satisfies:

$$1 - G_{i,t}(u) = 1 - \prod_{s\in\mathcal{K}}e^{-\mathcal{V}^s_{i,t}u^{-\phi}} = 1 - e^{-\mathcal{V}_{i,t}u^{-\phi}}, \quad \mathcal{V}_{i,t} = \sum_{s\in\mathcal{K}}\mathcal{V}^s_{i,t} \tag{A.21}$$

and the average welfare for the generation t born in i is:

$$\int_{\underline{u}}^{\infty} u\mathrm{d}G_{i,t}(u) = \int_0^{\mathcal{V}_{i,t}\underline{u}^{-\phi}}(y/\mathcal{V}_{i,t})^{-1/\phi}e^{-y}\mathrm{d}y \to \mathcal{V}_{i,t}^{1/\phi} \tag{A.22}$$

The average welfare among individuals of generation t who has an origin in location i is equalized ex ante because of the free mobility between sectors (i.e., self-selection) *ex ante*. Yet, there are idiosyncratic shocks in both location choice (i.e., idiosyncratic shocks in amenity) and idiosyncratic shocks in self-selection, so *ex post* utility of individuals is not equalized. Lastly, the measure of welfare is the *ex ante* expected utility of the cohort with adjustment of local labor market condition.

The adjustment is given by $O_{i,t}$ which is defined above. $O_{i,t}$ is Lagrangian multiplier for the cohort and it reflects the composition of workers among previous generation *before* sector choice. Define the ex-ante average utility of cohort $t+1$ with this adjustment by $V_{i,t+1}$:

$$\ln V_{i,t+1} = \frac{1}{|\mathcal{K}|}\sum_{s\in\mathcal{K}}\left(\ln\tilde{\zeta}_s + \frac{\eta}{\phi}\ln L^s_{i,t} - \frac{1}{\phi}\varsigma^s_{i,t+1} + \ln\bar{U}^s_{i,t+1}\right) \tag{A.23}$$

where $\tilde{\zeta}_s = \zeta_s^{1/\phi}$.

A.3 Production

This subsection summarizes the production side of the economy and regional trade pattern characterized by gravity structure with externalities. We start with the baseline specification. Then, we describe different models nested in the framework.

Baseline. Each final producers use intermediate goods available in location i. Their production is CES combination of intermediate goods with elasticity $\tilde{\kappa}$. The corresponding ideal price for final goods:

$$p^s_{i,t} = \left[\sum_{n\in\mathcal{N}}\int_{\mathbb{V}^s_{in,t}}\left(p^s_{in,t}(v)\right)^{1-\tilde{\kappa}}dv\right]^{1/(1-\tilde{\kappa})} \tag{A.24}$$

where $\mathbb{V}^s_{in,t}$ is set of available intermediate goods. The final goods are used for final consumption and materials in the production of intermediate goods. Intermediate goods producers have a constant return to scale technology with agglomeration economies. Share of labor and materials in inputs is given by β_s and $\beta_{ss'}$ respectively and $\beta_s + \sum_{s'\mathcal{K}\backslash 0}\beta_{ss'} = 1$ for any $s\in\mathcal{K}\backslash 0$. The cost minimization problem for firms leads to the unit cost of production:

$$c^s_{i,t}(\iota) = \frac{1}{a^s_{i,t}(\iota)}\frac{\Xi^s_{i,t}}{Z^s_{i,t}}, \quad \Xi_{is,t} = (w^s_{i,t})^{\beta_s}\prod_{s'\in\mathcal{K}\backslash 0}(p^{s'}_{i,t})^{\beta_{ss'}} \tag{A.25}$$

Consider regional trade of intermediate goods. Given iceberg trade cost $\tau^s_{int}\geq 1.$, prices of intermediate goods satisfy:

$$p^s_{i,t}(\iota) \in \min_{n\in\mathcal{N}}\left\{\tau^s_{in,t}c^s_{n,t}(\iota)\right\} \tag{A.26}$$

In the wake of Eaton and Kortum (2002), the productivity of firms $a^s_{it}(\iota)$ is following Fréchet distribution with unite location parameter and shape parameter $\kappa_s > 1$. The bilateral sector level trade pattern is given by:

$$\pi^s_{in,t} = \frac{(\tau^s_{in,t}\Xi^s_{n,t}/Z^s_{n,t})^{-\kappa_s}}{\sum_{\ell\in\mathcal{N}}(\tau^s_{i\ell,t}\Xi^s_{\ell,t}/Z^s_{\ell,t})^{-\kappa_s}}, \quad \forall i,n,\ell\in\mathcal{N} \quad \forall s\in\mathcal{K}\backslash 0 \tag{A.27}$$

Price of the final goods in location i and sector s is given by:

$$p_{i,t}^s = \Gamma_s^{-1} \cdot \left(\sum_{\ell \in \mathcal{N}} \left(\tau_{i\ell,t}^s \Xi_{\ell,t}^s / Z_{\ell,t}^s \right)^{-\kappa_s} \right)^{-1/\kappa_s}, \quad \Gamma_s \equiv \Gamma \left(\frac{1 - \tilde{\kappa} + \kappa_s}{\kappa_s} \right)^{1/(1-\tilde{\kappa})} \tag{A.28}$$

with assumption $1 - \tilde{\kappa} + \kappa_s > 0$. The market clearing condition is:

$$X_{i,t}^s = E_{i,t}^s + \sum_{s' \in \mathcal{K} \backslash 0} \beta_{s's} \sum_{n \in \mathcal{N}} \pi_{ni,t}^{s'} X_{n,t}^{s'}, \quad \forall i \in \mathcal{N}, \ \forall s \in \mathcal{K} \backslash 0 \tag{A.29}$$

NEG model for intermediate goods. The production technology of intermediate goods producers is increasing return to scale and they follow monopolistic competition. The final goods producers in location i and sector s has nested CES with two different sector specific elasticities between intermediate goods. Within the source location n, final good producers have constant elasticity of substitution between intermediate goods available in n, σ_s^F. While, final good producers have constant elasticity between source locations, σ_s^L.

The production cost of a intermediate producer ι in sector s in location i is:

$$c_{i,t}^s(\iota) = \left[F_s + \frac{y_{i,t}^s(\iota)}{Z_{i,t}^s} \right] \Xi_{i,t}^s \tag{A.30}$$

where F_s is fixed cost for production. Given the demand for intermediate variety, each intermediate firm produces different variety and monopolistic competition leads to price of variety in location i with constant markup. Letting $M_{i,t}^s$ be the measure of intermediate producers in location i and sector s,

$$\pi_{in,t}^s = \frac{(M_{n,t}^s)^{\frac{1-\sigma_s^L}{1-\sigma_s^F}} \left(\frac{\tau_{in,t}^s \Xi_{n,t}^s}{Z_{n,t}^s} \right)^{1-\sigma_s^L}}{\sum_{\ell \in \mathcal{N}} (M_{\ell,t}^s)^{\frac{1-\sigma_s^L}{1-\sigma_s^F}} \left(\frac{\tau_{i\ell,t}^s \Xi_{\ell,t}^s}{Z_{\ell,t}^s} \right)^{1-\sigma_s^L}}, \quad p_{i,t}^s = \frac{\sigma_s^F}{\sigma_s^F - 1} \left[\sum_{n \in \mathcal{N}} (M_{n,t}^s)^{\frac{1-\sigma_s^L}{1-\sigma_s^F}} \left(\frac{\tau_{in,t}^s \Xi_{n,t}^s}{Z_{n,t}^s} \right)^{1-\sigma_s^L} \right]^{\frac{1}{1-\sigma_s^L}} \tag{A.31}$$

Let $\bar{\ell}_{i,t}^s$ denote the average mass of labor of intermediate producers, and we obtain:

$$M_{i,t}^s = \frac{L_{i,t}^s}{\bar{\ell}_{i,t}^s F_s \sigma_s^F} \tag{A.32}$$

When $\bar{\ell}_{n,t}^s \to \bar{\ell}^s$ in the limit case, price becomes:

$$(p_{i,t}^s)^{1-\sigma_s^L} = \sum_{n \in \mathcal{N}} \left(\frac{1}{\bar{\ell}^s F_s \sigma_s^F} \right)^{\frac{1-\sigma_s^L}{1-\sigma_s^F}} \left[\frac{\sigma_s^F}{\sigma_s^F - 1} \frac{\tau_{in,t}^s \Xi_{nt}^s}{Z_{n,t}^s} (L_{nt}^s)^{\frac{1}{1-\sigma_s^F}} \right]^{1-\sigma_s^L} \tag{A.33}$$

NEG model with multi-sectors can be one representation of the baseline model in (A.27) and (A.28). Particularly, the trade elasticity with respect to trade cost is given by $1 - \sigma_s^L = -\kappa_s$, and larger amount of labor ($L^s_{i,t}$) leads to higher trade share. The market clearing condition is modified to:

$$X^s_{it} = E^s_{it} + \sum_{s' \in \mathcal{K} \setminus 0} \beta_{s's} \sum_{n \in \mathcal{N}} \frac{\sigma^F_{s'} - 1}{\sigma^F_{s'}} \pi^s_{nit} X^{s'}_{nt} \tag{A.34}$$

Heterogeneous firms with Pareto productivity. In Melitz (2003) model, there are mass of firms with different productivity that follows Pareto distribution. The Pareto distribution is σ^M_s and suppose that $\sigma^M_s + 1 > \sigma^F_s$. The final good producers is the same as in NEG framework. Under the monopolistic competition, firm set price with constant markup, and the price index for final good exported from n to i becomes:

$$(p^s_{in,t})^{1-\sigma^F_s} = M^s_{n,t} \left(\frac{\sigma^F_s}{\sigma^F_s - 1} \frac{\tau^s_{in,t} \Xi^s_{n,t}}{Z^s_{n,t}} \right)^{1-\sigma^F_s} \frac{\sigma^M_s (\underline{y}^s_{in,t})^{\sigma^F_s - \sigma^M_s - 1}}{1 + \sigma^M_s - \sigma^F_s} \tag{A.35}$$

Assume that firm's fixed cost for exporting is paid by composite goods of *destination* instead of origin. Therefore, the free entry condition pins down the productivity cutoff for export from n to i of sector s:

$$\underline{y}^s_{in,t} = \frac{\sigma^F_s}{\sigma^F_s - 1} \frac{\tau^s_{in,t} \Xi^s_{n,t}}{Z^s_{n,t}} \left(\frac{Y^s_{i,t}}{\sigma^F_s \Xi^s_{i,t} F_s} \right)^{\frac{1}{1-\sigma^F_s}} \frac{1}{p^s_{in,t}} \left(\frac{p^s_{in,t}}{p^s_{it}} \right)^{\frac{1-\sigma^L_s}{1-\sigma^F_s}} \tag{A.36}$$

Using this, $p^s_{in,t}$ satisfies:

$$\begin{aligned}(p^s_{in,t})^{(1-\sigma^L_s)\tilde{\sigma}_s} = &\left(\frac{\sigma^M_s}{1 + \sigma^M_s - \sigma^F_s} \right)^{\frac{(1-\sigma^F_s)(\sigma^L_s - 1)}{\sigma^M_s}} (M^s_{n,t})^{\frac{\sigma^L_s - 1}{\sigma^M_s}} \left(\frac{\sigma^F_s}{\sigma^F_s - 1} \frac{\tau^s_{in,t} \Xi^s_{n,t}}{Z^s_{n,t}} \right)^{1-\sigma^L_s} \\ &\times \left[\left(\frac{X^s_{i,t}}{\mu_s \Xi^s_{i,t} F_s} \right) (p^s_{i,t})^{1-\sigma^L_s} \right]^{\tilde{\sigma}_s - 1}\end{aligned}$$

where

$$\tilde{\sigma}_s \equiv 1 + \frac{(\sigma^L_s - 1)(1 + \sigma^M_s - \sigma^F_s)}{(\sigma^F_s - 1)\sigma^M_s} > 1$$

The free entry condition leads to the mass of firms proportional to labor. Namely,

$$M^s_{it} = \frac{\sigma^F_s - 1}{\sigma^F_s \sigma^M_s} \frac{L^s_{it}}{\underline{\ell}^s_{it} F_s}$$

and in the limit case with sufficiently small size of firms,

$$\pi^s_{in,t} = \frac{(L^s_{n,t})^{\frac{\sigma^L_s - 1}{\tilde{\sigma}_s \sigma^M_s}} \left(\frac{\tau^s_{in,t} \Xi^s_{n,t}}{Z^s_{n,t}} \right)^{\frac{1-\sigma^L_s}{\tilde{\sigma}_s}}}{\sum_{\ell \in \mathcal{N}} (L^s_{\ell,t})^{\frac{\sigma^L_s - 1}{\sigma^M_s \tilde{\sigma}_s}} \left(\frac{\tau^s_{i\ell t} \Xi^s_{\ell,t}}{Z^s_{\ell,t}} \right)^{\frac{1-\sigma^L_s}{\tilde{\sigma}_s}}} \tag{A.37}$$

A.4 Developers

Units of residential land endowment in location i are given by T_i, and it is fixed over time. Developers in each location produce structures that can be used for housing (i.e., residential use). The market is competitive, and there is a potentially large number of developers entering the market in each location. Developers require labor and stock of developed land. This implies that developers build a structure (i.e., housing) by adding structure to the previous stocks. Developers are homogeneous in production technology, and it is characterized by the neoclassical production function. Namely, the production function exhibits homogeneous of degree one and concave, and it satisfies Inada condition.

We let $f^H(\ell^0_{i,t},\ (1-\bar{h}_i)h_{i,t-1})$ be production function for developers, where $\bar{h}_i \in (0,1)$ is depreciation rate. Each developer in location i solves profit maximization problem for each unit of land development

$$\max_{\ell^0_{i,t}} \ p^0_{i,t} f^H(\ell^0_{i,t},\ (1-\bar{h}_i)h_{i,t-1}) - w^0_{i,t}\ell^0_{i,t} - r_{i,t} \tag{A.38}$$

where $r_{i,t}$ is competitive bidding price by developers to develop land. Each developer has to pay $r_{i,t}$ for land development to get permission. Developers are homogeneous in production technology and aggregate surplus extracted from developers is given by

$$R_{i,t} = r_{i,t}T_i = p^0_{i,t}\left(1 - \frac{\partial \ln f^H(L^0_{i,t},\ (1-\bar{h}_i)H_{i,t-1})}{\partial \ln L^0_{i,t}}\right) f^H(L^0_{i,t},\ (1-\bar{h}_i)H_{i,t-1}) \tag{A.39}$$

With the specification of Cobb-Douglas technology for f^H with constant share of labor χ, we have:

$$R_{i,t} = r_{i,t}T_i = (1-\chi)p^0_{i,t}H_{i,t}, \quad w^0_{i,t}L^0_{i,t} = \chi p^0_{i,t}H_{i,t} \tag{A.40}$$

Units of residential land endowment in location i is given by T_i and it is fixed over time. Then, revenue from land is distributed across local workers proportional to wage bill. This is Assumption 4. Then, land rent per worker is given by:

$$\widetilde{r}^s_{i,t} = R_{i,t} \times \frac{w^s_{i,t}}{\sum_j w^j_{i,t}L^j_{i,t}} = \frac{1-\chi}{\chi}\frac{w^0_{i,t}L^0_{i,t}}{\sum_j w^j_{i,t}L^j_{i,t}} w^s_{i,t} \tag{A.41}$$

A.5 Educational choices

The baseline model abstracts the endogenous choice of human capital. It can be extended to incorporate the additional dimension of endogenous heterogeneity. In particular, it is relevant to allow people to move to other places for the purpose of education. This subsection discusses a simple extension to this direction. The extension shares the spirit of the self-selection of education a lá Willis and Rosen (1979).

Consider generation t throughout the discussion here. People are ex-ante homogeneous. They choose the education level, graduate (G) or non-graduate (NG). If an individual chooses to be non-graduate, she stays in the city of birth. She can move to other cities for education if she chooses to be graduate. The timing is followings. An individual decides the sector. After the sector choice, she decides whether she goes to graduate or not, and the place of education if they choose to be graduate. If she decides to stay in the city, her skill in the future is at the non-graduate level. When she decides to leave the city for graduate education, her skill in the future is at the graduate level. She pays the cost of education and receives the education. Then, turning to the period t, she moves from the place of study to the other location to work and consume. To simplify the discussion, we suppose that amenities are different across locations but the same across workers in different sectors and education levels.

Solve the workers' decision backwards. A worker with education level e in sector s and location ℓ yields:

$$\widetilde{\mathcal{W}}_{\ell,t}^{s,e} = B_\ell \frac{W_{\ell,t}^{s,e}}{\mathcal{P}_{\ell,t}^{s,e}} \tag{A.42}$$

This is independent of the place of education. Workers with the same educational level ($e \in \{\mathrm{G}, \mathrm{NG}\}$) receive the same wage in a sector. Given the mobility costs and idiosyncratic location preference, the probability that an individual move to location ℓ from n is:

$$\lambda_{\ell n|s,t}^{e} = \frac{(\widetilde{\mathcal{W}}_{\ell,t}^{s,e}/D_{\ell n,t})^{\epsilon}}{\sum_{\ell' \in \mathcal{N}} (\widetilde{\mathcal{W}}_{\ell',t}^{s,e}/D_{\ell' n,t})^{\epsilon}} \tag{A.43}$$

and the conditional expected utility for an individual who chooses sector s and education level e in location n is given by:

$$\widetilde{U}_{n,t}^{s,e} = \left[\sum_{\ell \in \mathcal{N}} \left(\widetilde{\mathcal{W}}_{\ell,t}^{s,e}/D_{\ell n,t}\right)^{\epsilon}\right]^{1/\epsilon} \tag{A.44}$$

This captures the average return from education in location n for workers in sector s. Next, consider the educational choice of a worker. During period $t-1$, a worker has no cost if she chooses non-graduate, $e = \mathrm{NG}$. The expected utility for an individual who is born in location i and chooses $e = \mathrm{NG}$ is:

$$\widetilde{U}_{i,t}^{s,\mathrm{NG}} \times u^{s,\mathrm{NG}} \tag{A.45}$$

where $u^{s,\mathrm{NG}}$ is an idiosyncratic utility of non-graduate. When an individual chooses education level $e = \mathrm{G}$ in location n, the expected utility is:

$$\frac{\widetilde{U}_{n,t}^{s,\mathrm{G}} \times u_n^{s,\mathrm{G}}}{g_{ni,t-1}^{s}} \tag{A.46}$$

where $g_{ni,t-1}^{s}$ is common bilateral cost for education, and $u_n^{s,\mathrm{G}}$ is an idiosyncratic utility from graduate in location n. We let $\boldsymbol{u}^s = (u^{s,\mathrm{NG}}, u_1^{s,\mathrm{G}}, u_2^{s,\mathrm{G}}, \cdots, u_N^{s,\mathrm{G}})$ be the vector of idiosyncratic shocks in education choices. Assume that the idiosyncratic costs are drawn from the nested Gumbel distribution:

$$G^E(\boldsymbol{u};s) = \exp\left[-\left(u^{s,\mathrm{NG}}\right)^{-\alpha^E} - \left(\sum_{n\in\mathcal{N}}(u_n^{s,\mathrm{G}})^{-\alpha^G}\right)^{\alpha^E/\alpha^G}\right] \tag{A.47}$$

where α^E and α^G are exogenous parameters in the distribution. $\alpha^G > \alpha^E > 0$. With this distribution, the probability that an individual worker of sector s in location i chooses education choice $e = \mathrm{G}$ in location n:

$$\mathbb{G}_{i,t}^{s,n} = \frac{\left(\widetilde{U}_{n,t}^{s,\mathrm{G}}/g_{ni,t-1}^s\right)^{\alpha^G}}{\sum_{\ell\in\mathcal{N}}\left(\widetilde{U}_{\ell,t}^{s,\mathrm{G}}/g_{\ell i,t-1}^s\right)^{\alpha^G}} \times \frac{\left(\sum_{\ell\in\mathcal{N}}(\widetilde{U}_{\ell,t}^{s,\mathrm{G}}/g_{\ell i,t-1}^s)^{\alpha^G}\right)^{\alpha^E/\alpha^G}}{\left(\widetilde{U}_{i,t}^{s,\mathrm{NG}}\right)^{\alpha^E} + \left(\sum_{\ell\in\mathcal{N}}(\bar{U}_{\ell,t}^{s,\mathrm{G}}/g_{\ell i,t-1}^s)^{\alpha^G}\right)^{\alpha^E/\alpha^G}} \tag{A.48}$$

and the probability that an individual chooses $e = \mathrm{NG}$ is:

$$\mathbb{N}_{i,t}^s = \frac{\left(\widetilde{U}_{i,t}^{s,\mathrm{NG}}\right)^{\alpha^E}}{\left(\widetilde{U}_{i,t}^{s,\mathrm{NG}}\right)^{\alpha^E} + \left(\sum_{\ell\in\mathcal{N}}(\widetilde{U}_{\ell,t}^{s,\mathrm{G}}/g_{\ell i,t-1}^s)^{\alpha^G}\right)^{\alpha^E/\alpha^G}} \tag{A.49}$$

and the average utility:

$$\mathbb{W}_{i,t}^s = \bar{\gamma}\left(\left(\widetilde{U}_{i,t}^{s,\mathrm{NG}}\right)^{\alpha^E} + \left(\sum_{\ell\in\mathcal{N}}(\widetilde{U}_{\ell,t}^{s,\mathrm{G}}/g_{\ell i,t-1}^s)^{\alpha^G}\right)^{\alpha^E/\alpha^G}\right)^{1/\alpha^E} \tag{A.50}$$

Lastly, the choice of the sector depends on the average utility $\mathbb{W}_{i,t}^s$ and taste shocks as in the baseline model. Therefore, the probability of choosing sector s in location i is:

$$\varsigma_{i,t}^s = \zeta_s(L_{i,t-1}^s)^\eta\left(\frac{\mathbb{W}_{i,t}^s}{V_{i,t}}\right)^\phi \tag{A.51}$$

Together these probabilities, the mass of workers in location n in period t for sector s and graduate G is given by:

$$L_{n,t}^{s,\mathrm{G}} = \sum_{\ell\in\mathcal{N}}\sum_{i\in\mathcal{N}}\lambda_{n\ell|st}^{\mathrm{G}} \times \mathbb{G}_{i,t}^{s,\ell} \times \varsigma_{i,t}^s \times L_{i,t-1} \tag{A.52}$$

and that for non-graduate NG is given by:

$$L_{nt}^{s,\mathrm{NG}} = \sum_{\ell\in\mathcal{N}}\sum_{i\in\mathcal{N}}\lambda_{n\ell|st}^{\mathrm{NG}} \times \mathbb{N}_{i,t}^{s,\mathrm{NG}} \times \varsigma_{i,t}^s \times L_{i,t-1} \tag{A.53}$$

For the production side, consider the CES substitution between non-graduate and graduate. Specifically, the unit cost for production of tradable intermediates is:

$$\Xi_{n,t}^s = \left(\left(w_{n,t}^{s,\mathrm{NG}}\right)^{1-\beta} + \varpi_{s,t}\left(w_{n,t}^{s,\mathrm{G}}\right)^{1-\beta}\right)^{1/(1-\beta)} \tag{A.54}$$

where β controls the elasticity of substitution between non-graduate and graduate, and $\varpi_{s,t}$ is the sector specific premium for the graduates.

In this extension, the education choice leads to an additional channel creating inequality in income and welfare. Intuitively, if the birthplace has better access to graduate study (g^s_{ni}), young people are more likely to self-selection into graduate study. When they choose to start the graduate study, they show a high probability of moving to an educational place with better accessibility to the labor market ($D_{\ell n,t}$). Therefore, the birthplace affects the educational choice of workers.

A.6 Infinite horizon with perfect foresight

Recent papers (Dix-Carneiro 2014, McLaren 2017, Caliendo et al. 2019, Kleinman et al. 2021) propose the framework of individual agents in the infinite horizon with perfect foresight. This subsection discuss the comparison between our model and such framework by presenting a simple model of forward-looking agents and derive workers' choices. There is no exogenous population growth in the economy as in the baseline model.

For workers in location i and sector s, instantaneous utility is $\ln \mathcal{W}^s_{i,t}$ with amenity adjusted real income $\mathcal{W}^s_{i,t}$. An individual ω decides the location and sector for the next period together. An individual solves

$$v^s_{i,t}(\omega) = \ln \mathcal{W}^s_{i,t} + \max_{n\in\mathcal{N}} \max_{k\in\mathcal{K}} \left[\rho V^k_{n,t+1} - d_{ni,t+1} - \mu^k_{i,t+1} + \varphi^k_{ni,t+1}(\omega)\right] \tag{A.55}$$

where $d_{ni,t+1}$ is migration cost, $\mu^k_{i,t+1}$ is cost of choosing sector, and $\varphi^k_{ni,t+1}(\omega)$ is idiosyncratic shocks that are related to both location choice and sector choice, and $V^k_{n,t+1}$ is expected future utility:

$$V^k_{n,t+1} = \mathbb{E}\left[v^k_{n,t+1}(\omega)\right] \tag{A.56}$$

To derive the analytical formulae, assume that $\varphi^k_{ni,t+1}(\omega)$ is drawn from the independent Type I extreme distribution such that:

$$\Phi_{t+1}(\varphi) = \mathbf{Pr}(\varphi^k_{ni,t+1} \le \varphi) = e^{-(\varphi-\gamma\Omega)/\Omega}, \quad \phi_{t+1}(\varphi) = \frac{1}{\Omega} e^{-\varphi/\Omega} G(\varphi) \tag{A.57}$$

where γ is Euler-Mascheroni constant. Letting

$$\Upsilon^k_{ni,t+1} = \rho V^k_{n,t+1} - d_{ni,t+1} - \mu^k_{i,t+1}, \quad V^*_{i,t} = \max_{n\in\mathcal{N}} \max_{k\in\mathcal{K}} \Upsilon^k_{ni,t+1} + \varphi^k_{ni,t+1}(\omega),$$

we have

$$\mathbf{Pr}[\Upsilon^k_{ni,t+1} + \varphi^k_{ni,t+1}(\omega) \le v] = e^{-e^{\frac{1}{\Omega}\left(v - \Upsilon^k_{ni,t+1} - \gamma\Omega\right)}} \tag{A.58}$$

Therefore,

$$\Phi^*_{i,t}(v) = \mathbf{Pr}[V^*_{i,t} \leq v] = \prod_n \prod_k \Phi_{t+1}\Big(-\Upsilon^k_{ni,t+1} + v \Big) = e^{-e^{\left[-\frac{v-\gamma\Omega}{\Omega} + \ln \sum_n \sum_k e^{(\Upsilon^k_{ni,t+1}/\Omega)}\right]}}$$

so that

$$\mathbb{E}_{i,t}(v) = \Omega \ln \left[\sum_n \sum_k \exp(\Upsilon_{ni,t+1})^{1/\Omega} \right] \tag{A.59}$$

and

$$V^s_{i,t} = \ln \mathcal{W}^s_{i,t} + \Omega \ln \left[\sum_n \sum_k \exp\left(\rho V^k_{n,t+1} - d_{ni,t+1} - \mu^k_{i,t+1} \right)^{1/\Omega} \right] \tag{A.60}$$

In the followings, we transform variable such that $\varrho = \epsilon - \gamma\Omega$. The choice probability is:

$$\begin{aligned} \pi^s_{n|i,t+1} &= \int_{-\infty}^{\infty} \left(\prod_{\ell,k} e^{-e^{-\frac{1}{\Omega}\left(\varrho + \Upsilon^s_{ni,t+1} - \Upsilon^k_{\ell i,t+1}\right)}} \right) \frac{e^{-\varrho/\Omega}}{\Omega} \mathrm{d}\varrho \\ &= \int_{-\infty}^{\infty} \exp\left(-e^{-\frac{\varrho}{\Omega}} \sum_{\ell,k} e^{-\frac{1}{\Omega}\left(\Upsilon^s_{ni,t+1} - \Upsilon^k_{\ell i,t+1}\right)} \right) \frac{e^{-\varrho/\Omega}}{\Omega} \mathrm{d}\varrho \end{aligned} \tag{A.61}$$

Letting $\tilde{\varrho} = e^{-\varrho/\Omega}$,

$$\begin{aligned} \pi^s_{n|i,t+1} &= \int_0^{\infty} \exp\left(-\tilde{\varrho} \sum_\ell \sum_k e^{-\frac{1}{\Omega}\left(\Upsilon^s_{ni,t+1} - \Upsilon^k_{\ell i,t+1}\right)} \right) \mathrm{d}\tilde{\varrho} \\ &= \frac{\exp\big(\Upsilon^s_{ni,t+1}\big)^{1/\Omega}}{\sum_\ell \sum_k \exp\big(\Upsilon^k_{\ell i,t+1}\big)^{1/\Omega}} \\ &= \frac{\exp\big(\rho V^s_{n,t+1} - d_{ni,t+1} - \mu^s_{i,t+1}\big)^{1/\Omega}}{\sum_\ell \sum_k \exp\big(\rho V^k_{\ell,t+1} - d_{\ell i,t+1} - \mu^k_{i,t+1}\big)^{1/\Omega}} \end{aligned} \tag{A.62}$$

Workers in n and s in period $t+1$ is:

$$L^s_{n,t+1} = \sum_i \sum_k \pi^s_{n|i,t+1} L^k_{i,t} \tag{A.63}$$

The probability of choosing s among workers from i is:

$$\mathbb{S}^s_{i,t+1} = \left[\frac{\mathcal{W}^s_{i,t}}{\exp\big(V^s_{i,t} \mu^s_{i,t+1}\big)} \right]^{1/\Omega} \sum_n \exp\left(\rho V^s_{n,t+1} - d_{ni,t+1} \right)^{1/\Omega} \tag{A.64}$$

And among workers whose origin is i and who choose sector s, probability of moving to n is:

$$\mathbb{M}_{n|is,t+1} = \frac{\exp\left(\rho V^s_{n,t+1} - d_{ni,t+1}\right)^{1/\Omega}}{\sum_\ell \exp\left(\rho V^s_{\ell,t+1} - d_{\ell i,t+1}\right)^{1/\Omega}} \tag{A.65}$$

Since idiosyncratic shocks are independent across sector and location choices, they satisfy

$$\pi^s_{n|i,t+1} = \mathbb{S}^s_{i,t+1} \times \mathbb{M}_{n|is,t+1} \tag{A.66}$$

In the comparison between our baseline framework of overlapping generations and the forward-looking model, key probabilities are (A.64) and (A.65). First, the probability of migration among workers of sector s and origin i (A.65) takes a similar form to the probability ($\lambda_{ni|s,t+1}$) in the baseline model. The probability is determined by the expected value of location n and sector s in period $t+1$ and common friction in labor mobility. The main difference is the characterization of the expected utility in (A.60). In the baseline model, determinants of future value in workers' choice are fully characterized by the real income in the next period $t+1$, while the forward-looking model is, by its nature, the future value depends on the entire path of future in the economy. Second, the probability of choosing sector (A.64) depends on current local economy, common barrier of sector choice ($\mu^s_{i,t+1}$), and future return of sector s. The current value of real income and a common barrier correspond to the stickiness in workers' choice of sectors in the baseline model, and the future return of sector s corresponds to $\bar{U}^s_{i,t+1}$. Therefore, it also depends on the entire future path in the economy.

This difference plays an essential role in the transition process. Suppose that there is a transitory common shock in the economy. Since infinitely lived households (workers) choose the future path of mobility in a forward-looking way, taking into account future shocks, their choice of a given location is based on current real income and an option value associated with that location. Transitory effect changes the option value and adjustment of workers must take account for the updated path of shocks in the future, which leads to different speed of transition compared to the overlapping generation framework in which individuals that work only in the second period of their life. In addition, the baseline framework can isolate the relative importance of migration barrier, local labor market exposure, structural transformation, and agglomeration economies in workers' response to the common shocks since incorporating the externalities in the forward-looking model make it challenging to analyze the effect of change in option values along with the transition (e.g., Krugman 1991, Matsuyama 1991, Ottaviano 1999, Baldwin 2001).

At the same time, there is the equivalence between the two approaches when focusing on *backward* solution to back out the past fundamentals in the economy from *steady state*. By using the sequence of economic outcomes up to a certain period of the stationary steady state T, the model can be solved backward using the transition probabilities in both approaches. This process gives a path of shocks between period $T-1$ and T that rationalizes the change of states from period

$T-1$ to T. In both approaches, the change between two consecutive periods is fully characterized by shocks between these two periods so that we can map the past equilibrium as long as the path is unique. However, as being discussed above, *forward solution* shows the difference between the two approaches.

B Appendix: Equilibrium

This section describes the analytical characterization of the dynamic equilibrium in subsection B.1 and shows the system of equations for the transition of equilibrium and propose the solution method in subsection B.2. The last subsection B.3 discusses the steady state of the economy.

B.1 Solving for Dynamic Equilibrium

This subsection solves the equilibrium in period $t+1$ given the economy of period t. Given information of the time-varying fundamentals for both period t ($\boldsymbol{\mathcal{F}}_t$) and $t+1$ ($\boldsymbol{\mathcal{F}}_{t+1}$) and time-invariant fundamentals ($\bar{\boldsymbol{\mathcal{F}}}$) and conditional on the equilibrium in period t, the model is solved for the equilibrium in period $t+1$.

To economize notation, use the following notations for exogenous factors:

$$\mathbb{A}^s_{in,t} = (\tau^s_{in,t})^{-\kappa_s}(A^s_{i,t})^{\kappa_s}, \quad \mathbb{B}^s_{in,t} = (D_{in,t})^{-\varepsilon}(B^s_{i,t})^{\varepsilon}$$

The distribution of land rent in the local labor market is:

$$\mu_{i,t+1} = 1 + \frac{R_{i,t+1}}{\sum_s w^s_{i,t+1}L^s_{i,t+1}} = 1 + (1-\chi)\frac{p^0_{i,t+1}H_{i,t+1}}{\sum_s w^s_{i,t+1}L^s_{i,t+1}} \tag{B.1}$$

The income of workers is:

$$W^s_{i,t+1} = \mu_{i,t+1}w^s_{i,t+1} = \left[1 + (1-\chi)\frac{p^0_{i,t+1}H_{i,t+1}}{\sum_s w^s_{i,t+1}L^s_{i,t+1}}\right] w^s_{i,t+1} \tag{B.2}$$

Labor mobility implies that:

$$\begin{aligned} L^s_{in,t+1} &= \lambda^s_{in,t+1}\varsigma^s_{n,t+1}L_{n,t} \\ &= \mathbb{B}^s_{in,t+1}\left(\frac{W^s_{i,t+1}}{\mathcal{P}^s_{i,t+1}\bar{U}^s_{n,t+1}}\right)^{\epsilon} \zeta_s\left(L^s_{n,t}\right)^{\eta}\left(\frac{\bar{U}^s_{n,t+1}}{V_{n,t+1}}\right)^{\phi} L_{n,t} \end{aligned} \tag{B.3}$$

Employment size in location i and sector s in period $t+1$ is

$$L^s_{i,t+1} = \sum_{n\in\mathcal{N}} L^s_{in,t+1} \tag{B.4}$$

with

$$
\begin{aligned}
(\bar{U}^s_{i,t+1})^{\varepsilon} &= \sum_{\ell \in \mathcal{N}} \mathbb{B}^s_{\ell i,t+1} (W^s_{\ell,t+1})^{\varepsilon} (\mathcal{P}^s_{\ell,t+1})^{-\varepsilon} \\
(V_{i,t+1})^{\phi} &= \sum_{s \in \mathcal{K}} \zeta_s (L^s_{it})^{\eta} (\bar{U}^s_{i,t+1})^{\phi}
\end{aligned}
\tag{B.5}
$$

The expected utility conditional on sector choice ($\bar{U}^s_{i,t+1}$) and welfare measure ($V_{i,t+1}$) are determined by the real income, $\mathcal{W}^s_{i,t+1}$. (B.4) determines labor supply. In other words, equilibrium real income $\mathcal{W}^s_{i,t+1}$ and labor allocation $L^s_{i,t+1}$ solve (B.4) and (B.5) together. Combining them yields:

$$
L^s_{i,t+1} = (\mathcal{W}^s_{i,t+1})^{\varepsilon} \sum_{n \in \mathcal{N}} \mathbb{B}^s_{in,t+1} \frac{\zeta_s (L^s_{nt})^{\eta} (\bar{U}^s_{n,t+1})^{\phi}}{\sum_{k \in \mathcal{K}} \zeta_k (L^k_{nt})^{\eta} (\bar{U}^k_{n,t+1})^{\phi}} (\bar{U}^s_{n,t+1})^{-\varepsilon} L_{n,t} \tag{B.6}
$$

Manipulating this,

$$
\bar{U}^s_{i,t+1} = \left[\sum_{\ell \in \mathcal{N}} \left(\widetilde{B}^s_{\ell i,t+1} \frac{L^s_{\ell,t+1}}{\sum_{n \in \mathcal{N}} \left(\mathbb{B}^s_{\ell n,t+1} \frac{\zeta_s (L^s_{nt})^{\eta} (\bar{U}^s_{n,t+1})^{\phi}}{\sum_{k \in \mathcal{K}} \zeta_k (L^k_{nt})^{\eta} (\bar{U}^k_{n,t+1})^{\phi}} (\bar{U}^s_{n,t+1})^{-\varepsilon} L_{n,t} \right)} \right) \right]^{1/\varepsilon} \tag{B.7}
$$

when $\varepsilon < \phi$, there is unique mapping between the expected utility $\bar{U}^s_{i,t+1}$ and employment distribution $L^s_{i,t+1}$. $L^s_{i,t+1}$ is homogeneous of degree zero in $\bar{U}^s_{i,t+1}$. Labor supply $L^s_{i,t+1}$ is increasing in $\bar{U}^s_{i,t+1}$, and it is larger than marginal effect of other local labor market, $\bar{U}^k_{j,t+1}$ of $j \neq i$ and $k \neq s$. From (B.5), we can also uniquely map real income ($\mathcal{W}^s_{i,t+1}$) and welfare measure ($V_{i,t+1}$) to labor supply, $L^s_{i,t+1}$. In particular, real income solves:

$$
(\bar{U}^s_{i,t+1})^{\varepsilon} = \sum_{\ell \in \mathcal{N}} \mathbb{B}_{\ell i,t+1} (\mathcal{W}^s_{\ell,t+1})^{\varepsilon} \tag{B.8}
$$

$L^s_{i,t+1}$ is homogeneous of degree zero in $\mathcal{W}^s_{i,t+1}$. Inserting (B.3) into the productivity spillover yields

$$
Z^s_{i,t+1} = A^s_{i,t+1} \mathbb{Z}^s_{i,t+1} \tag{B.9}
$$

with

$$
\mathbb{Z}^s_{i,t+1} = \left[(\mathcal{W}^s_{i,t+1})^{\varepsilon} \sum_{n} \mathbb{B}^s_{in,t+1} \frac{\zeta_s (L^s_{n,t})^{\eta} (\bar{U}^s_{n,t+1})^{\phi}}{\sum_k \zeta_k (L^k_{n,t})^{\eta} (\bar{U}^k_{n,t+1})^{\phi}} (\bar{U}^s_{n,t+1})^{-\varepsilon} L_{n,t} Z^s_{n,t} \right]^{\rho} (L^s_{i,t+1})^{\gamma_s} \tag{B.10}
$$

(B.9) and (B.10) determine the productivity of period $t+1$ given previous labor allocation ($L^s_{i,t}$), previous productivity ($Z^s_{i,t}$) and employment of period $t+1$ ($L^s_{i,t+1}$) as the right-hand-side can be fully characterized by them.

And tradable goods price in location n and sector s becomes

$$(p^s_{n,t+1})^{-\kappa_s} = \Gamma_s^{-\kappa_s} \sum_i \mathbb{A}^s_{ni,t+1} \left[\mathbb{Z}^s_{i,t+1} (w^s_{i,t+1})^{-\beta_s} \left(\prod_{s' \in \mathcal{K} \backslash 0} (p^{s'}_{i,t+1})^{-\beta_{ss'}} \right) \right]^{\kappa_s} \tag{B.11}$$

Since

$$\left| \frac{\mathrm{d} \ln p^s_{i,t+1}}{\mathrm{d} \ln p^{s'}_{n,t+1}} \right| = \beta_{ss'} \pi^s_{in,t+1} < 1$$

for any combination of (i, s) and (n, s'), (B.11) is solved for unique prices $p^s_{i,t+1}$ conditional on wages $(w^s_{i,t+1})$ and productivity $(Z^s_{i,t+1})$. To simplify the discussion in the following, we set $\beta_{ss'} = 0$, that is $\beta_s = 1$ for any $s \in \mathcal{K} \backslash 0$. Therefore, (B.11) is reduced to

$$(p^s_{n,t+1})^{-\kappa_s} = \Gamma_s^{-\kappa_s} \sum_i \mathbb{A}^s_{ni,t+1} \left(\frac{w^s_{i,t+1}}{\mathbb{Z}^s_{i,t+1}} \right)^{-\kappa_s} \tag{B.12}$$

Price index satisfies:

$$\mathcal{P}^s_{i,t+1}{}^{1-\sigma} = \sum_{k \in \mathcal{K}} \alpha_k^{\sigma-1} (p^k_{i,t+1})^{1-\sigma} (W^s_{i,t+1}/\mathcal{P}^s_{i,t+1})^{\theta_k - 1} \tag{B.13}$$

Plugging (B.2) and (B.12) into this,

$$\left[1 + \frac{p^0_{i,t+1} \bar{H}_{i,t} (L^0_{i,t+1})^{\chi}}{\sum_s w^s_{i,t+1} L^s_{i,t+1}} \right]^{1-\sigma} (w^s_{i,t+1})^{1-\sigma} = \sum_k \widetilde{\alpha}_k \left[\sum_n \mathbb{A}^k_{in,t+1} \left(\frac{w^k_{n,t+1}}{\mathbb{Z}^k_{n,t+1}} \right)^{-\kappa_k} \right]^{-(1-\sigma)/\kappa_k} \mathcal{W}^s_{i,t+1}{}^{\theta_k - \sigma} \tag{B.14}$$

for all $s \in \mathcal{K}$, where $\widetilde{\alpha}_k = \alpha_k^{\sigma-1}$ and $\bar{H}_{i,t} = \nu_i (1-\chi)[(1-\bar{h}_i) H_{i,t}]^{1-\chi}$.

Inserting sectoral price and price index, the expenditure share becomes:

$$\psi^k_{s|n,t+1} = \widetilde{\alpha}_s \Gamma_s{}^{1-\sigma} \left[\sum_i \mathbb{A}^s_{ni,t+1} \left(\frac{w^s_{i,t+1}}{\mathbb{Z}^s_{i,t+1}} \right)^{-\kappa_s} \right]^{-(1-\sigma)/\kappa_s} W^k_{n,t+1}{}^{-(1-\sigma)} \mathcal{W}^k_{n,t+1}{}^{\theta_s - \sigma} \tag{B.15}$$

Labor market clearing condition for tradables is:

$$w^s_{i,t+1} L^s_{i,t+1} = \sum_n \pi^s_{ni,t+1} \left(\sum_k \psi^k_{s|n,t+1} W^k_{n,t+1} L^k_{n,t+1} \right) \tag{B.16}$$

where bilateral trade probabilities are

$$\pi^s_{ni,t+1} = \frac{\widetilde{A}^s_{ni,t+1} \left(w^s_{i,t+1} / \widetilde{Z}^s_{i,t+1} \right)^{-\kappa_s}}{\sum_{\ell \in \mathcal{N}} \widetilde{A}^s_{n\ell,t+1} \left(w^s_{\ell,t+1} / \widetilde{Z}^s_{\ell,t+1} \right)^{-\kappa_s}} \tag{B.17}$$

Together (B.15), (B.16) and (B.17) yields

$$\Gamma_s{}^{-(1-\sigma)} L^s_{i,t+1} = \frac{\widetilde{\alpha}_s}{w^s_{i,t+1}} \sum_n \left[\frac{\mathbb{A}^s_{ni,t+1}\left(w^s_{i,t+1}/\mathbb{Z}^s_{i,t+1}\right)^{-\kappa_s}}{\sum_{\ell\in\mathcal{N}} \mathbb{A}^s_{n\ell,t+1}\left(w^s_{\ell,t+1}/\mathbb{Z}^s_{\ell,t+1}\right)^{-\kappa_s}} \left(\sum_k \left(\sum_i \mathbb{A}^s_{ni,t+1}\left(\frac{w^s_{i,t+1}}{\mathbb{Z}^s_{i,t+1}}\right)^{-\kappa_s}\right)^{-(1-\sigma)/\kappa_s}\right. \right.$$
$$\left. \times \left(\left(1+\frac{p^0_{n,t+1}\bar{H}_{n,t}\left(L^0_{n,t+1}\right)^{\chi}}{\sum_j w^j_{n,t+1}L^j_{n,t+1}}\right) w^k_{n,t+1}\right)^{\sigma} \left(\mathcal{W}^k_{n,t+1}\right)^{\theta_s-\sigma} L^k_{n,t+1}\Bigg)\right] \tag{B.18}$$

for tradables. For sector 0, market clearing condition is

$$p^0_{i,t+1}{}^{\sigma} \bar{H}_{i,t} L^0_{i,t+1}{}^{\chi} = \widetilde{\alpha}_0 \sum_k \mathcal{W}^k_{i,t+1}{}^{\theta_0-\sigma} W^k_{i,t+1}{}^{\sigma} L^k_{i,t+1} \tag{B.19}$$

Manipulating this,

$$p^0_{i,t+1}{}^{\sigma} L^0_{i,t+1}{}^{\chi} \bar{H}_{i,t} \left[1+\frac{p^0_{i,t+1}\bar{H}_{i,t}(L^0_{i,t+1})^{\chi}}{\sum_j w^j_{i,t+1}L^j_{i,t+1}}\right]^{-\sigma} = \widetilde{\alpha}_0 \sum_k \mathcal{W}^k_{i,t+1}{}^{\theta_0-\sigma} w^k_{i,t+1}{}^{\sigma} L^k_{i,t+1} \tag{B.20}$$

The equilibrium in period $t+1$ is fully characterized by $(\bar{U}^s_{i,t+1}, \mathcal{W}^s_{i,t+1}, L^s_{i,t+1}, w^s_{i,t+1}, \widetilde{Z}^s_{i,t+1}, p^0_{i,t+1})$ that solve (B.7), (B.8), (B.10), (B.14), (B.18) and (B.20).

In order to discuss the uniqueness of equilibrium analytically, consider the conservative case: $\rho = 0$ and $\chi = 1$. $\rho = 0$ implies that productivity spillover happened locally, and $\chi = 1$ implies that supply of residential stocks is elastic. This implies:

$$W^s_{i,t+1} = w^s_{i,t+1}, \quad \mathbb{Z}^s_{i,t+1} = L^s_{i,t+1}{}^{\gamma_s} \tag{B.21}$$

and further, we can set $p^0_{i,t+1} = w^0_{i,t+1}$. (B.18) becomes

$$\Gamma_s{}^{-(1-\sigma)} L^s_{i,t+1} = \frac{\widetilde{\alpha}_s}{w^s_{i,t+1}} \sum_n \left[\frac{\mathbb{A}^s_{ni,t+1}(w^s_{i,t+1})^{-\kappa_s}(L^s_{i,t+1})^{\gamma_s\kappa_s}}{\sum_{\ell'} \mathbb{A}^s_{n\ell',t+1}(w^s_{\ell',t+1})^{-\kappa_s}(L^s_{\ell',t+1})^{\gamma_s\kappa_s}} \right.$$
$$\left. \times \left(\sum_\ell \mathbb{A}^s_{n\ell,t+1}(w^s_{\ell,t+1})^{-\kappa_s}(L^s_{\ell,t+1})^{\gamma_s\kappa_s}\right)^{-(1-\sigma)/\kappa_s} \left(\sum_k (w^k_{n,t+1})^{\sigma}(\mathcal{W}^k_{n,t+1})^{\theta_s-\sigma}L^k_{n,t+1}\right)\right] \tag{B.22}$$

Manipulating this yields

$$\Gamma_s^{-(1-\sigma)} {L_{i,t+1}^s}^{1-\gamma_s\kappa_s} = \frac{\widetilde{\alpha}_s}{{w_{i,t+1}^s}^{1+\kappa_s}} \sum_n \left[\mathbb{A}_{ni,t+1}^s \left(\sum_\ell \mathbb{A}_{n\ell,t+1}^s (w_{\ell,t+1}^s)^{-\kappa_s} (L_{\ell,t+1}^s)^{\gamma_s\kappa_s} \right)^{-1-(1-\sigma)/\kappa_s} \right.$$
$$\left. \times \left(\sum_k (w_{n,t+1}^k)^\sigma \left(\frac{L_{n,t+1}^k}{\sum_\ell \mathbb{B}_{n\ell,t+1}^k \varsigma_{\ell,t+1}^k (\bar{U}_{\ell,t+1}^k)^{-\varepsilon} L_{\ell,t}} \right)^{(\theta_s-\sigma)/\varepsilon} L_{n,t+1}^k \right) \right] \tag{B.23}$$

Suppose the following general form of equations:

$$F_i^j(\boldsymbol{x}) = (x_i^j)^a - \sum_n \left[z_{in}^j \left(\sum_\ell B_{n\ell}^j (x_\ell^j)^b \right)^{-c} \left(\sum_k C_n^k (D_n^k)^d (x_n^k)^e \right) \right] = 0 \tag{B.24}$$

where a, b, c, d and e are positive constant parameters. Let

$$Y_n^j = \sum_\ell B_{n\ell}^j (x_\ell^j)^b, \quad Z_n = \sum_k C_n^k (D_n^k)^d (x_n^k)^e$$

Then, we let:

$$f_i^j(\boldsymbol{x}) = a \ln x_i^j - \ln \sum_n z_{in}^j {Y_n^j}^{-c} Z_n = 0 \tag{B.25}$$

and

$$\frac{\partial f_i^j(\boldsymbol{x})}{\partial \ln x_p^j} = a\mathbb{I}_{[i=p]}$$
$$- \sum_n \frac{z_{in}^j {Y_n^j}^{-c} Z_n}{\sum_\ell z_{i\ell}^j {Y_\ell^j}^{-c} Z_\ell} \left[-bc \frac{B_{np}^j (x_p^j)^b}{Y_n^j} + \sum_k \frac{C_n^k (D_n^k)^d (x_n^k)^e}{Z_n} \left(d \frac{\partial \mathrm{d} \ln D_n^k}{\partial \ln x_p^j} + e \frac{\partial \ln x_n^k}{\partial \ln x_p^j} \right) \right]$$

Suppose that $a \leq 1$. Then, the gross substitute property holds for $F_i^j(\boldsymbol{x})$ when

$$bc - d - e < -1$$

Seeing (B.23), this condition corresponds to

$$\gamma_s \leq \frac{\theta_s - \sigma}{\kappa_s + (1-\sigma)} \left(1 + \frac{1}{\varepsilon} \right) \tag{B.26}$$

When gross substitute holds, unique solution exists when

$$a < -bc + d + e$$

In (B.23), this condition corresponds to

$$1-\gamma_s\kappa_s<\gamma_s\kappa_s\left(-1-\frac{1-\sigma}{\kappa_s}\right)+\frac{\theta_s-\sigma}{\varepsilon}+1 \quad \Longleftrightarrow \quad \gamma_s<\frac{1}{\varepsilon}\frac{\theta_s-\sigma}{1-\sigma} \tag{B.27}$$

Therefore, sufficient condition for unique solution of $L^s_{i,t+1}$ that solve (B.23) is

$$\gamma_s \leq \frac{\theta_s-\sigma}{\kappa_s+(1-\sigma)}\left(1+\frac{1}{\varepsilon}\right) \tag{B.28}$$

This condition is intuitive. When $\varepsilon \to \infty$, idiosyncratic shocks in migration is homogeneous and it leads to lower threshold for γ_s as weak agglomeration forces are required to avoid generating multiple equilibria. If θ_s becomes large, this condition becomes slack. This implies that large heterogeneity in consumption across workers of different income leads to more dispersion of workers to avoid multiple equilibria. Further discussion about (B.28) can be found later for quantification. Labor demand schedule solving (B.23) downward slope of wage. For labor supply, (B.7) argues that the labor supply schedule is upward slope of wage, therefore pinning down wage vector that clear the labor market.

B.2 Forward Solution for Transition in Dynamic Equilibrium

This subsection describes the solution method for the transition equilibrium. The aim is computing the transition process response to the changes in productivity ($\mathrm{d}\ln A^s_{i,t}$) and spatial frictions ($\mathrm{d}\ln \tau_{in}$ and $\mathrm{d}\ln D_{in}$). Fundamental amenities are kept unchanged, while including shock to amenity is straightforward. The outcome of our interests are the trajectory of endogenous variables in the economy conditional on some initial state. To this end, the process starts off with the guess of the trajectory of wage ($\mathrm{d}\ln w^s_{i,t+1}$), employment ($\mathrm{d}\ln L^s_{i,t+1}$) and housing price ($\mathrm{d}\ln p^0_{i,t+1}$).

Productivity. The initial guess for aggregate productivity change is

$$\mathrm{d}\ln Z^s_{i,t+1}=\mathrm{d}\ln A^s_{i,t+1}+\gamma_s \mathrm{d}\ln L^s_{i,t+1} \tag{B.29}$$

In the loop, the productivity change is updated using:

$$\mathrm{d}\ln Z^s_{i,t+1}=\mathrm{d}\ln A^s_{i,t+1}+\gamma_s \mathrm{d}\ln L^s_{i,t+1}+\rho\cdot \mathrm{d}\ln\left(\sum_n L^s_{in,t+1}Z^s_{n,t}\right) \tag{B.30}$$

where previous labor mobility ($L^s_{in,t}$), previous productivity ($A^s_{i,t}$, $Z^s_{i,t}$) are given, and change of labor mobility ($\mathrm{d}\ln L^s_{in,t+1}$) is updated in the loop.

Income. Zero profit condition for developing residential stock implies that

$$\mathrm{d}\ln p^0_{i,t+1}+\mathrm{d}\ln H_{i,t+1}=\mathrm{d}\ln w^0_{i,t+1}+\mathrm{d}\ln L^0_{i,t+1}=\mathrm{d}\ln R_{i,t+1} \tag{B.31}$$

and change in distribution of land rent is

$$\mu_{i,t} \exp(\mathrm{d}\ln \mu_{i,t+1}) = 1 + \frac{R_{i,t} \exp(\mathrm{d}\ln R_{i,t+1})}{\sum_s w^s_{i,t} L^s_{i,t} \exp\big(\mathrm{d}\ln w^s_{i,t+1} + \mathrm{d}\ln L^s_{i,t+1}\big)} \tag{B.32}$$

Together, change in income of workers is

$$\mathrm{d}\ln W^s_{i,t+1} = \mathrm{d}\ln \mu_{i,t+1} + \mathrm{d}\ln w^s_{i,t+1} \tag{B.33}$$

where change in total income is decomposed into change in land rent distribution and change in wage.

Trade. For tradable sectors, change in trade pattern shows

$$\mathrm{d}\ln \pi^s_{in,t+1} = -\kappa_s \mathrm{d}\ln \tau_{in,t+1} - \kappa_s \mathrm{d}\ln \Xi^s_{n,t+1} + \kappa_s \mathrm{d}\ln Z^s_{n,t+1} + \kappa_s \mathrm{d}\ln p^s_{i,t+1} \tag{B.34}$$

with

$$\mathrm{d}\ln \Xi^s_{n,t+1} = \beta_s \mathrm{d}\ln w^s_{n,t+1} + \sum_{s'} \beta_{ss'} \mathrm{d}\ln p^{s'}_{n,t+1} \tag{B.35}$$

and

$$\mathrm{d}\ln p^{s'}_{i,t+1} = \frac{1}{\kappa_s} \mathrm{d}\ln \pi^s_{ii,t+1} - \mathrm{d}\ln Z^s_{i,t+1} + \mathrm{d}\ln \Xi^s_{i,t+1} \tag{B.36}$$

Solutions to $N^2S + NS + NS$ equations of (B.34)-(B.36) are corresponding to the change of trade pattern and consumer price $(\mathrm{d}\ln \pi^s_{in,t+1}, \mathrm{d}\ln p^s_{i,t+1})$.

Demand shift. Turing to consumers' demand system, change in price index satisfies $N \times (S+1)$ equations:

$$\begin{aligned} &\exp\big[(1-\sigma)\mathrm{d}\ln \mathcal{P}^k_{i,t+1}\big] \\ &= \sum_s \psi^k_{si,t} \exp\big[(1-\sigma)\mathrm{d}\ln p^s_{i,t+1} + (1-\theta_s)\mathrm{d}\ln \mathcal{P}^k_{i,t+1} + (\theta_s - 1)\mathrm{d}\ln W^k_{i,t+1}\big] \end{aligned} \tag{B.37}$$

Given the previous expenditure patterns $(\psi^k_{si,t})$ and change in price and income $(\mathrm{d}\ln p^s_{i,t+1}, \mathrm{d}\ln W^s_{i,t+1})$, solving (B.37) gives change $\mathrm{d}\ln \mathcal{P}^k_{i,t+1}$ and $N(S+1)^2$ equations for change of expenditure patterns:

$$\mathrm{d}\ln \psi^k_{si,t+1} = (1-\sigma)\mathrm{d}\ln p^s_{i,t+1} - (\theta_s - \sigma)\mathrm{d}\ln \mathcal{P}^k_{i,t+1} + (\theta_s - 1)\mathrm{d}\ln W^k_{i,t+1} \tag{B.38}$$

In turn, change in real income and expenditure are

$$\mathrm{d}\ln \mathcal{W}^k_{i,t+1} = \mathrm{d}\ln W^k_{i,t+1} - \mathrm{d}\ln \mathcal{P}^k_{i,t+1} \tag{B.39}$$

and

$$\mathrm{d}\ln E^k_{si,t+1} = \mathrm{d}\ln \psi^k_{si,t+1} + \mathrm{d}\ln W^k_{i,t+1} \tag{B.40}$$

Labor mobility. Change in workers' mobility over space is described by

$$\mathrm{d}\ln\lambda_{in|s,t+1} = \mathrm{d}\ln\lambda_{nn|s,t+1} - \varepsilon \mathrm{d}\ln D_{in,t+1} + \varepsilon\Big(\mathrm{d}\ln\mathcal{W}^s_{i,t+1} - \mathrm{d}\ln\mathcal{W}^s_{n,t+1}\Big) \tag{B.41}$$

with the change in conditional average utility is:

$$\mathrm{d}\ln\bar{U}^s_{i,t+1} = -\frac{1}{\varepsilon}\mathrm{d}\ln\lambda_{ii|s,t+1} + \mathrm{d}\ln\mathcal{W}^s_{i,t+1} \tag{B.42}$$

For each sector, change in worker sorting between generation $t-1$ and generation t shows:

$$\mathrm{d}\ln\varsigma^s_{i,t+1} = \eta\cdot\mathrm{d}\ln L^s_{i,t} - \phi\cdot\mathrm{d}\ln V_{i,t+1} - \frac{\phi}{\varepsilon}\mathrm{d}\ln\lambda_{ii|s,t+1} + \phi\cdot\mathrm{d}\ln\mathcal{W}^s_{i,t+1} \tag{B.43}$$

By construction, they satisfy:

$$\sum_{n\in\mathcal{N}}\lambda_{ni|s,t}\cdot\exp\big(\mathrm{d}\ln\lambda_{ni|s,t+1}\big) = 1, \quad \sum_{s\in\mathcal{K}}\varsigma^s_{i,t}\cdot\exp\big(\mathrm{d}\ln\varsigma^s_{i,t+1}\big) = 1 \tag{B.44}$$

Together, change in worker distribution in the economy is specified by:

$$\mathrm{d}\ln L^s_{i,t+1} = \sum_{n\in\mathcal{N}}\frac{\lambda_{in|s,t}\varsigma^s_{n,t}}{L^s_{i,t}}\exp\big(\mathrm{d}\ln\lambda_{in|s,t+1} + \mathrm{d}\ln\varsigma^s_{n,t+1}\big)L_{n,t} \tag{B.45}$$

These $N^2(S+1)+N(S+1)+N(S+1)+N$ equations of (B.41)-(B.45) characterize the transition of the economy in terms of $(\mathrm{d}\ln\lambda_{ni|s,t+1},\ \mathrm{d}\ln\varsigma^s_{i,t+1},\ \mathrm{d}\ln L^s_{i,t+1},\ \mathrm{d}\ln V_{i,t+1})$.

Market clearing conditions. Change in production solves $N\times S$ market clearing conditions:

$$\begin{aligned}\mathrm{d}\ln X^s_{i,t+1} = &\sum_{s'\in\mathcal{K}\backslash 0}\beta_{s's}\sum_{n\in\mathcal{N}}\frac{\pi^{s'}_{ni,t}X^{s'}_{n,t}}{X^s_{i,t}}\exp\Big(\mathrm{d}\ln\pi^{s'}_{ni,t+1} + \mathrm{d}\ln X^{s'}_{n,t+1}\Big)\\ &+\sum_{k\in\mathcal{K}}\frac{E^k_{si,t}}{X^s_{i,t}}\exp\big(\mathrm{d}\ln E^k_{si,t+1}\big)\end{aligned} \tag{B.46}$$

Labor market clearing condition is given by $N\times S+N$ equations:

$$\mathrm{d}\ln w^s_{i,t+1} + \mathrm{d}\ln L^s_{i,t+1} = \sum_{n\in\mathcal{N}}\frac{\pi^s_{ni,t}X^s_{n,t}}{\sum_\ell\pi^s_{\ell i,t}X^s_{\ell,t}}\exp\big(\mathrm{d}\ln\pi^s_{ni,t+1} + \mathrm{d}\ln X^s_{n,t+1}\big) \tag{B.47}$$

$$\mathrm{d}\ln w^0_{i,t+1} + \mathrm{d}\ln L^0_{i,t+1} = \mathrm{d}\ln p^0_{i,t+1} + \mathrm{d}\ln H_{i,t+1} \tag{B.48}$$

And market clearing condition for residential stocks provides N equations:

$$\mathrm{d}\ln p^0_{i,t+1} + \mathrm{d}\ln H_{i,t+1} = \sum_{k\in\mathcal{K}}\frac{E^k_{0i,t}}{p^0_{i,t}H_{i,t}}\exp\big(\mathrm{d}\ln E^k_{0i,t+1}\big) \tag{B.49}$$

Together $N \times S + N + N$ equations are sufficient to determine the transition of $(\mathrm{d}\ln w^s_{i,t+1}, \mathrm{d}\ln p^0_{i,t+1})$.

Update vectors. We update $(\mathrm{d}\ln w^s_{i,t+1}, \mathrm{d}\ln L^s_{i,t+1}, \mathrm{d}\ln p^0_{i,t+1})$ until their convergence.

B.3 Steady State Equilibrium

In steady-state equilibrium, the exogenous time-variant factors are constant: $(\mathbf{D}, \boldsymbol{\tau}, \boldsymbol{A}, \boldsymbol{B})$ and the economy is characterized by the steady-state level of basic equilibrium vectors/matrix: $(\mathbf{p}, \mathbf{w}, \mathbf{L})$ and auxiliary equilibrium vectors and matrices $(\boldsymbol{\psi}, \boldsymbol{\pi}, \mathbf{X}, \boldsymbol{H}, \boldsymbol{V})$. Once economy reaches in the steady state, the aggregate variables are unchanged in the following periods but there is mobility of workers and goods.

Income and Price. Disposable income is uniquely determined by $(\mathbf{w}, \mathbf{L})$ such that $\mathbf{W} = \mathcal{F}_{\mathbf{W}}(\mathbf{w}, \mathbf{L})$. $\mathcal{F}_{\mathbf{W}}$ is continuous and it shows

$$\frac{\partial \ln W^s_i}{\partial \ln w^k_i} = \mathbb{I}_{[s=k]} - \frac{1-\chi}{\chi}\frac{\mu_i - 1}{\mu_i} y^s_i, \quad \text{with} \quad y^s_i = \frac{w^s_i L^s_i}{\sum_{k\in\mathcal{K}} w^s_i L^s_i} \tag{B.50}$$

Income is decreasing in other sector's wage. The small $1-\chi$ in development of land assures that the worker's income is increasing in wage.

Consider prices in steady state equilibrium. For sector 0, zero profit condition pins down $\boldsymbol{p}^0$ given $\boldsymbol{w}^0$. The price p^0_i is strictly increasing in w^0_i. For other sectors $s \in \mathcal{K}\backslash 0$, zero profit conditions and labor market clearing condition give the system of equations that characterize $\mathbf{p}$ conditional on $(\mathbf{w}, \mathbf{L})$. We have

$$(p^s_i)^{-\kappa_s} = \Gamma_s^{-\kappa_s} \sum_n (\tau^s_{in})^{-\kappa_s} (Z^s_n)^{\kappa_s} (w^s_i)^{-\beta_s \kappa_s} \prod_{s'\in\mathcal{K}\backslash 0} (p^{s'}_n)^{-\beta_{ss'}\kappa_s} \tag{B.51}$$

and

$$\left|\frac{\mathrm{d}\ln p^s_i}{\mathrm{d}\ln p^{s'}_n}\right| = \beta_{ss'}\pi^s_{in} < 1$$

for any combination of (i, s) and (n, s'). (B.51) pins down unique price matrix $\mathbf{p}$ conditional on $\mathbf{w}$ and productivity $\boldsymbol{Z}$. Productivity Z^s_n is a fixed point of:

$$Z^s_i = A^s_i \left(\sum_n L^s_{in} Z^s_n\right)^{\rho} \left(L^s_i\right)^{\gamma_s} \tag{B.52}$$

Taking its logarithm,

$$\ln Z^s_i = \ln A^s_i + \rho \ln \sum_n L^s_{in} Z_n + \gamma_s \ln L^s_i \tag{B.53}$$

For any $i, n \in \mathcal{N}$, we have:

$$\sum_n \left|\frac{\partial \ln Z^s_i}{\partial \ln Z^s_n}\right| = \sum_n \left|\rho \cdot \frac{L^s_{in} Z^s_n}{\sum_\ell L^s_{i\ell} Z^s_\ell}\right| < 1 \tag{B.54}$$

if $\rho < 1$. Therefore, price of tradables is uniquely determined by $(\mathbf{w}, \mathbf{L})$ such that $\mathbf{p} = \mathcal{F}_{\mathbf{p}}(\mathbf{w}, \mathbf{L})$ by continuous mapping $\mathcal{F}_{\mathbf{p}}$.

Consider the demand system. Given $\mathcal{F}_{\mathbf{p}}(\mathbf{w}, \mathbf{L})$ and $\boldsymbol{p}^0$, the aggregate price index $\{\mathcal{P}_i^k\}_{i\in\mathcal{N}}^{k\in\mathcal{K}}$ solves the $N \times (S+1)$ equations:

$$\sum_{s\in\mathcal{K}} \alpha_s^{\sigma-1} (p_i^s/\mathcal{P}_i^k)^{1-\sigma} (W_i^k/\mathcal{P}_i^k)^{\theta_s - 1} = 1 \quad \forall\, i \in \mathcal{N}, k \in \mathcal{K} \tag{B.55}$$

When the sign of $\theta_s - \sigma$ is same for any $s \in \mathcal{K}$, this allows us to characterize unique set of aggregate price $\{\mathcal{P}_i^k\}_{i\in\mathcal{N}}^{k\in\mathcal{K}}$ by solving this equation. Accordingly, we obtain the expenditure share matrix $\boldsymbol{\psi}$. The aggregate price index $\mathcal{P}_i^k$ is homogeneous of degree one in (**p**, **w**), and it is increasing in any sectoral price changes and income. Particularly, price index shows

$$\frac{\partial \ln \mathcal{P}_i^k}{\partial \ln p_i^s} > 0, \quad \frac{\partial \ln \mathcal{P}_i^k}{\partial \ln W_i^k} > 0, \quad \forall\, k, s \in \mathcal{K} \tag{B.56}$$

and expenditure share shows

$$\mathrm{d}\psi_s^p = \frac{\partial \ln \psi_{s|ik}}{\partial \ln p_i^s} = 1 - \sigma > 0, \quad \mathrm{d}\psi_s^w = \frac{\partial \ln \psi_{s|ik}}{\partial \ln W^{ik}} = \theta_s - 1 > 0, \quad \frac{\partial \ln \psi_{s|ik}}{\partial \ln \mathcal{P}_i^k} = \sigma - \theta_s < 0 \tag{B.57}$$

for any k and s. Real income is defined by $\mathcal{W}_i^s = W_i^s/\mathcal{P}_i^s$ and

$$\frac{\partial \ln \mathcal{P}_i^k}{\partial \ln \mathcal{W}_i^k} = \sum_{s\in\mathcal{K}} \frac{\psi_{s|ik}}{1-\sigma}(\theta_s - 1), \quad \frac{\partial \ln \psi_{s|ik}}{\partial \ln \mathcal{W}_i^k} = \theta_s - 1 > 0 \tag{B.58}$$

Combining the characterization of population movement, $\lambda_{ni|s}(\mathbf{w}, \mathbf{L})$, and industry choice $\varsigma_i^s(\mathbf{w}, \mathbf{L})$, define:

$$\mathcal{E}_{ni}^s(\mathbf{w}, \mathbf{L}) = \lambda_{ni|s}(\mathbf{w}, \mathbf{L}) \times \varsigma_i^s(\mathbf{w}, \mathbf{L}) \tag{B.59}$$

and labor supply is:

$$L_i^s = \sum_{n\in\mathcal{N}} \mathcal{E}_{in}^s(\mathbf{w}, \mathbf{L}) \left[\sum_k L_n^k\right] \equiv \mathcal{LF}_i^s(\mathbf{w}, \mathbf{L}), \quad \forall i \in \mathcal{N},\ \forall s \in \mathcal{K} \tag{B.60}$$

where $\mathcal{LF}_i^s(\cdot,\cdot)$ is continuous mapping. The equilibrium labor supply $\mathbf{L}$ solves this fixed point equation given wage $\mathbf{w}$. For any $0 < \mathbf{w} < \infty$ and $0 \le \mathbf{L} \le \bar{L}$, define:

$$\mathcal{E}_{\mathrm{M}}(\mathbf{w}) = \sup\ \mathcal{E}_{ni}^s(\mathbf{w}, \mathbf{L}) < 1, \quad \mathcal{LF}_{\mathrm{M}} = \sup\ \mathcal{E}_{\mathrm{M}}(\mathbf{w})\bar{L} \tag{B.61}$$

Using them, define compact and convex subset:

$$\boldsymbol{\Sigma}_L \equiv \left\{ l \in \mathbb{R}_+^{|\mathcal{N}|\times|\mathcal{K}|} \ :\ 0 \le l_i^s \le \mathcal{LF}_{\mathrm{M}}, \quad \forall\, i \in \mathcal{N}, s \in \mathcal{K} \right\} \subset \Delta^{|\mathcal{N}|\times|\mathcal{K}|}(\mathbb{R}_+) \tag{B.62}$$

Therefore, there exists solution $\mathbf{L}$ to (B.60) on $\mathbf{\Sigma}_L$ given $\mathbf{w}$. Then, refer $\mathcal{LS}(\mathbf{w})$ to a continuous function for labor supply given w, and the sufficient condition for unique $\mathbf{L}$ is:

$$\sup \left| \mathcal{E}^s_{ii}(\mathbf{w},\mathbf{L}) + \sum_{n\in\mathcal{N}} \mathrm{d}\mathcal{E}^s_{in}(\mathbf{w},\mathbf{L}) \cdot \vartheta_{in|s}(\mathbf{w},\mathbf{L}) \right| < 1, \quad \forall i \in \mathcal{N}, \quad s \in \mathcal{K} \tag{B.63}$$

where notations are

$$\mathrm{d}\mathcal{E}^s_{in}(\mathbf{w},\mathbf{L}) = \frac{\mathrm{d}\ln \mathcal{E}^s_{in}(\mathbf{w},\mathbf{L})}{\mathrm{d}\ln L^s_i}, \quad \vartheta_{in|s}(\mathbf{w},\mathbf{L}) = L^s_{in}/L^s_i$$

This condition argues that labor mobility across space is large enough not to generating the degenerated equilibrium, and the gains from agglomeration is small enough not to dominate the congestion forces that arise from immobile factor price or lower wage (i.e., excess supply of labor). When this is hold for a given $\mathbf{w}$, we can construct the unique labor supply function $\mathcal{LS}(\cdot) : \mathbb{R}^{N\times(S+1)}_{++} \rightarrow \mathbb{R}^{N\times(S+1)}_{++}$.

Next, consider the property of labor supply function. Taking any sequence $\mathbf{w}^m \rightarrow \mathbf{w}$, $\mathcal{LS}(\mathbf{w}^m) \rightarrow \mathcal{LS}(\mathbf{w})$. The supply function is bounded above by construction. Since the total population is fixed, $\bar{L}$, the supply function must satisfy $\mathcal{LS}(\mathbf{w}) \leq \bar{L}$ for any $\mathbf{w}$. In the steady state equilibrium,

$$\left(1 - \mathcal{E}^s_{ii}\right) \frac{\partial L^s_i}{\partial w^j_\ell} = \sum_{(n,k)\neq(i,s)} \mathcal{E}^s_{in} \left(\frac{L^k_n}{\mathcal{E}^s_{in}} \frac{\partial \mathcal{E}^s_{in}}{\partial w^j_\ell} + \frac{\partial L^k_n}{\partial w^j_\ell} \right) \tag{B.64}$$

for any combination of (i,s) and (ℓ,j). When

$$\left| \sum_{(n,k)\neq(i,s)} \mathcal{E}^s_{in} \frac{\partial L^k_n}{\partial w^s_i} \right| \geq \max_{(\ell,s')\neq(i,s)} \left| \sum_{(n,k)\neq(i,s)} \mathcal{E}^s_{in} \frac{\partial L^k_n}{\partial w^{s'}_\ell} \right| \tag{B.65}$$

for any i and s,

$$\left\{ \nabla_w \mathcal{LS}_{is}(\mathbf{w}) \right\}_{(i,s)} > \left\{ \nabla_w \mathcal{LS}_{is}(\mathbf{w}) \right\}_{(\ell,s')\neq(i,s)} \tag{B.66}$$

(B.65) is intuitive. The increase in wage w^s_i attracts more worker from other local labor market $(n,k) \neq (i,s)$ with higher probability $\mathcal{E}^s_{in}$, and this dominates the effect of other local market effect of increase in $w^{s'}_\ell$.

Market clearing condition. Now, consider labor demand function: $\mathcal{LD}(\cdot) : \mathbb{R}^{N\times(S+1)}_{++} \rightarrow \mathbb{R}^{N\times(S+1)}_{++}$. The market clearing conditions for goods lead to total value of production (X^s_i) and then labor market clearing condition leads to the labor demand function. Denote the matrix of variables:

$$\mathbf{X} = \left\{ X^s_i \right\}, \quad \boldsymbol{\pi} \equiv \left\{ \beta_{s's} \pi^{s'}_{ni}(\mathbf{w},\mathbf{L}) \right\}, \quad \mathbf{E} \equiv \left\{ \sum_k \psi^k_{s|i} W^k_i L^k_i \right\}$$

The goods market clearing condition is represented by

$$\widetilde{\boldsymbol{\pi}} \cdot \text{vec}(\mathbf{X}) = \text{vec}(\mathbf{E}) \quad \text{with} \quad \widetilde{\boldsymbol{\pi}} \equiv \mathbf{I} - \boldsymbol{\pi} \tag{B.67}$$

The elements in the diagonal of $\widetilde{\boldsymbol{\pi}}$ are strictly positive, and let $\widetilde{\boldsymbol{\pi}}_d$ denote matrix of the diagonal elements. Define:

$$\widetilde{\boldsymbol{\pi}}_d^{-1} > 0, \quad \mathbf{T} \equiv \widetilde{\boldsymbol{\pi}}\widetilde{\boldsymbol{\pi}}_d^{-1} - \mathbf{I}, \quad \boldsymbol{e} \equiv (\mathbf{I} - \mathbf{T}) \cdot \text{vec}\Big(\mathbf{E}\Big) > 0 \tag{B.68}$$

Then, the spectral radius of $|\mathbf{T}|$ is less than one when $\sum_{s'} \beta_{s's} < 1$, and followings are satisfied:

$$\begin{aligned}
&\sum_{n \neq i} |\mathbf{T}|_{in} \leq 1, \quad \rho(\mathbf{T}^2) = \rho(|\mathbf{T}|^2) < 1, \\
&\widetilde{\boldsymbol{\pi}}^{-1}\text{vec}(\boldsymbol{E}) = \widetilde{\boldsymbol{\pi}}_d^{-1}(\mathbf{I} + \mathbf{T})^{-1}\text{vec}(\mathbf{E}) = \widetilde{\boldsymbol{\pi}}_d^{-1}(\mathbf{I} + \mathbf{T})^{-1}(\mathbf{I} - \mathbf{T})^{-1}\boldsymbol{e}, \\
&(\mathbf{I} + \mathbf{T})^{-1}(\mathbf{I} - \mathbf{T})^{-1} = (\mathbf{I} - \mathbf{T}^2)^{-1} = \sum_{j=0}^{\infty} \mathbf{T}^{2j} \geq \mathbf{I},
\end{aligned} \tag{B.69}$$

where $\rho(\cdot)$ is spectral radius of matrix. Since there exists some positive number π^* such that $\boldsymbol{\pi}_d^{-1} \geq \pi^* \mathbf{I}$, we have

$$\widetilde{\boldsymbol{\pi}}^{-1} \cdot \text{vec}(\mathbf{E}) \geq \boldsymbol{\pi}_d^{-1} \cdot \boldsymbol{e} > 0. \tag{B.70}$$

Therefore, there is unique $\mathbf{X} > 0$ given $(\mathbf{w}, \mathbf{L})$: $\mathbf{X} = \mathcal{F}_{\mathbf{X}}(\mathbf{w}, \mathbf{L})$.

Letting

$$\mathbb{Z}_{in}^s = (\tau_{in}^s)^{-\kappa_s}(Z_n^s)^{\kappa_s} = (\tau_{in}^s)^{-\kappa_s}\left[A_i^s(L_i^s)^{\gamma_s}\left(\sum_n \mathcal{E}_{in}^s L_n Z_n^s\right)^{\rho}\right]^{\kappa_s} \tag{B.71}$$

For $s \in \mathcal{K} \backslash 0$ and $i \in \mathcal{N}$,

$$L_i^s = \frac{\beta_s}{w_i^s} \sum_{n \in \mathcal{N}} \frac{\mathbb{Z}_{ni}^s \left(\frac{(w_i^s)^{-\beta_s}}{\prod_{s' \in \mathcal{K}\backslash 0}(p_i^{s'}(\mathbf{w},\mathbf{L}))^{\beta_{ss'}}}\right)^{\kappa_s}}{\sum_{\ell \in \mathcal{N}} \mathbb{Z}_{n\ell}^s \left(\frac{(w_\ell^s)^{-\beta_s}}{\prod_{s' \in \mathcal{K}\backslash 0}(p_\ell^{s'}(\mathbf{w},\mathbf{L}))^{\beta_{ss'}}}\right)^{\kappa_s}} \Big\{\mathcal{F}_{\mathbf{X}}(\mathbf{w}, \mathbf{L})\Big\}_{ns} \tag{B.72}$$

and for development of residential stocks,

$$L_i^0 = \frac{\chi}{w_i^0} \sum_{s \in \mathcal{K}} \psi_{0|is} W_i^s L_i^s \tag{B.73}$$

Denote $\mathcal{J}(\cdot)$ be the mapping of $(\mathbf{w}, \mathbf{L})$ such that $\mathbf{L} = \mathcal{J}(\mathbf{w}, \mathbf{L})$ satisfying the above equations of the labor market clearing condition.

From above discussion, for any $\mathbf{w} > 0$ and $\mathbf{L} > 0$, $\mathbf{X} = \mathcal{F}_{\mathbf{X}}(\mathbf{w}, \mathbf{L}) > 0$ and there exists positive number $\underline{x} > 0$ such that $X_i^s \geq \underline{x}$ for any i and s. Further, for any $\mathbf{L} > 0$, define $\underline{z} > 0$ and $\overline{z} < \infty$ such that $\widetilde{Z}_{ni}^s \in [\underline{z}, \ \overline{z}]$ for any i, n and s. Solving for tradable prices can define $\underline{p} > 0$ and $\overline{p} < \infty$

such that tradable prices are within the range $[\underline{p}, \overline{p}]$. In addition, $\underline{\psi}$ can be constructed as in the same manner. Using them, define:

$$\underline{\mathcal{L}}^s \equiv \frac{\beta_s}{\overline{w}} \sum_n \frac{\underline{z}\overline{w}^{-\beta_s\kappa_s}\overline{p}^{-(1-\beta_s)\kappa_s}}{\overline{z}\underline{w}^{-\beta_s\kappa_s}\underline{p}^{-(1-\beta_s)\kappa_s}}\underline{x}, \quad \overline{\mathcal{L}}^s \equiv \max\left\{\bar{L}, \quad \frac{\beta_s}{\underline{w}} \sum_n \frac{\overline{z}\underline{w}^{-\beta_s\kappa_s}\underline{p}^{-(1-\beta_s)\kappa_s}}{\underline{z}\overline{w}^{-\beta_s\kappa_s}\overline{p}^{-(1-\beta_s)\kappa_s}}\overline{x}\right\}, \tag{B.74}$$

where

$$\underline{w} \equiv \min_{i\in\mathcal{N}}\left(\min_{s\in\mathcal{K}} w_{is}\right), \quad \overline{w} \equiv \max_{i\in\mathcal{N}}\left(\max_{s\in\mathcal{K}} w_{is}\right)$$

Further, define

$$\underline{\mathcal{L}}^0 \equiv \frac{\chi}{\overline{w}} \frac{\underline{w}}{1-\chi\underline{\psi}} \sum_{s\in\mathcal{K}\setminus 0} \underline{\psi}\underline{\mathcal{L}}^s, \quad \overline{\mathcal{L}}^0 \equiv \bar{L} \tag{B.75}$$

for sector 0. Together, define the simplex:

$$\mathbf{\Sigma}_D \equiv \left\{l \in \mathbb{R}_{++}^{N\times(S+1)} \ : \ \underline{\mathcal{L}}^s \leq l_i^s \leq \overline{\mathcal{L}}^s, \quad \forall i \in \mathcal{N}, \ s \in \mathcal{K}\right\} \subset \Delta^{|\mathcal{N}|\times|\mathcal{K}|}(\mathbb{R}_{++}) \tag{B.76}$$

This is compact and convex subset. Since $\mathcal{J}(\cdot)$ is continuous for any sequence of $\mathbf{L}$ given $\mathbf{w}$, $\mathbf{L}$ is characterized as a fixed point of the labor market clearing condition such that $\mathbf{L} = \mathcal{J}(\mathbf{L}, \mathbf{w})$. Given that $\mathcal{J}$ is continuous for $\mathbf{w}$ and parameters, and class C^1 for $\mathbf{L}$, the solution $\mathbf{L}$ is also continuous in $\mathbf{w}$ and parameters. The condition for unique labor demand function is given by:

$$\max_{s\in\mathcal{K}\setminus 0} \max_{(i,n)\in\mathcal{N}\times\mathcal{N}} \left|\frac{\partial \ln X_n^s}{\partial \ln L_i^s}\right| < 1 \tag{B.77}$$

The function $\mathcal{LD}(\cdot)$ exhibits downward slope for the local wage. First,

$$L_i^0 = \frac{1}{w_i^0} \frac{\chi}{1-\chi\psi_{0|i}^0\mu_i} \sum_{s\in\mathcal{K}\setminus 0} \psi_{0|i}^s \mu_i w_i^s L_i^s \tag{B.78}$$

that shows downward slope in local wage w_i^0 and non-decreasing in $\{w_i^s\}_{s\neq 0}$. Other sectors show:

$$\frac{\partial \ln L_i^s}{\partial \ln w_n^{s'}} = -\mathbb{I}_{\{n,s'\}=\{i,s\}} + \sum_{n\in\mathcal{N}} \frac{\mathbb{X}_{ni}^s}{\sum_{\ell\in\mathcal{N}} \widetilde{X}_{\ell i}^s} \frac{\partial \ln \mathbb{X}_{ni}^s}{\partial \ln w_n^{s'}} \tag{B.79}$$

where $\mathbb{X}_{ni}^s$ is value of export from location i to n in sector s. The last condition for the unique steady state is:

$$\Upsilon_{\{i,s\}}^{\{n,k\}} = \sum_{n\in\mathcal{N}} \frac{\mathbb{X}_{ni}^s}{\sum_{\ell\in\mathcal{N}} \mathbb{X}_{\ell i}^s} \frac{\partial \ln \mathbb{X}_{ni}^s}{\partial \ln w_n^k} \geq 0 \tag{B.80}$$

This implies the marginal wage change in local market $(n, k) \neq (i, s)$ induces the export of local market (i, s) to other markets on average, and it is because of the competitive advantage of (i, s). Since $\mathcal{LD}(\cdot)$ is homogeneous of degree zero in $\mathbf{w}$, the function exhibits gross substitutability under this condition.

Excess demand. Define the excess demand function for labor market such that $\mathcal{LZ}(\mathbf{w}) \equiv \mathcal{LD}(\mathbf{w}) - \mathcal{LS}(\mathbf{w})$. For this function, consider the following properties.

(i) $\mathcal{LZ}(\cdot)$ *is continuous and bounded below* : continuity follows by the augment that both $\mathcal{LD}(\cdot)$ and $\mathcal{LS}(\cdot)$ are continuous as in the above discussion. Since $\mathcal{LS}(\mathbf{w}) < \bar{L}$ for any $\mathbf{w}$ by construction, $\mathcal{LZ}(\mathbf{w})$ is bounded below for $\mathbf{w} \in \mathbb{R}_{++}^{|\mathcal{N}|\times|\mathcal{K}|}$.

(ii) $\mathcal{LZ}(\cdot)$ *satisfies Walras' law* : for any $\mathbf{w} \in \mathbb{R}_{++}^{|\mathcal{N}|\times|\mathcal{K}|}$, market clearing condition implies:

$$\begin{aligned} \boldsymbol{I}_{|\mathcal{N}|}^{\top}(\mathbf{w} \circ \boldsymbol{LZ}(\boldsymbol{w}))\boldsymbol{I}_{|\mathcal{K}|} = \sum_i \sum_{s\in\mathcal{K}\setminus 0} w_i^s L_i^s - \sum_i \sum_{s\in\mathcal{K}\setminus 0} \beta_s \sum_n \pi_{ni}^s \left(\sum_{k\in\mathcal{K}\setminus 0} \beta_{ks} X_n^k + E_n^s \right) \\ + \sum_i w_i^0 L_i^0 - \chi \sum_i E_i^0 = 0 \end{aligned} \tag{B.81}$$

(iii) $\mathcal{LZ}(\cdot)$ *satisfies the boundary condition such that:* $\max\ \{\mathcal{LZ}(\mathbf{w}^m)\} \to \infty$ *for any sequence* $\mathbf{w}^m \subset \mathbb{R}_{++}^{|\mathcal{N}|\times|\mathcal{K}|}$ *with* $\mathbf{w}^m \to \mathbf{w} \in \mathbb{R}_{+}^{|\mathcal{N}|\times|\mathcal{K}|}\setminus(\mathbb{R}_{++}^{|\mathcal{N}|\times|\mathcal{K}|} \cup \{\mathbf{0}\})$: for labor supply, $\mathcal{LS}(\mathbf{w}) \geq 0$ for every i and s, and $\{\mathcal{LS}(\mathbf{w})\}_{\{i,s\}} > 0$ for at least one element. Therefore, it is sufficient to show that $\{\mathcal{LD}(\mathbf{w}^m)\}_{\{i,s\}} \to \infty$ as $m \to \infty$ for such $\{i,s\}$. Suppose that $\{\mathcal{LD}(\mathbf{w}^m)\}_{\{i,s\}}^m$ is bounded above. Since $\{\mathcal{LD}(\mathbf{w}^m)\}_{\{i,s\}}^m \geq 0$ for all m, the sequence $\{\mathcal{LD}(\mathbf{w}^m)\}_{\{i,s\}}^m$ has convergent subsequence and we let $\{\mathcal{LD}(\mathbf{w}^m)\}_{\{i,s\}}^m$ be the subsequence and L_i^s be its limit.

(iv) $\mathcal{LZ}(\mathbf{w})$ *is homogeneous of degree zero in* $\mathbf{w}$: by construction, $\mathcal{LS}(\mathbf{w})$ is homogeneous of degree zero in $\mathbf{w}$. To verify this, $\mathbf{W}$ and $\mathbf{p}$ are homogeneous of degree one in $\mathbf{w}$, and it implies aggregate price index is homogeneous of degree zero in $\mathbf{w}$. Therefore, $\mathcal{E}(\cdot)$ is homogeous of degree zero in $\mathbf{w}$, and therefore labor supply is homogeneous of degree zero in $\mathbf{w}$. For demand of labor, $\mathcal{LD}(\mathbf{w})$ is also homogeneous of degree zero in $\mathbf{w}$. It follows immediately from labor market clearing conditions. They are homogeneous of degree zero in $\mathbf{w}$, and therefore solution for them $\mathcal{LD}(\mathbf{w})$ is also homogeneous of degree zero in $\mathbf{w}$.

(v) $\mathcal{LZ}(\cdot)$ *exhibits gross substitutability* : the gross substitute property follows under the assumptions discussed above.

In summary, steady state equilibrium uniquely exists under the *sufficient* conditions discussed in (B.65), (B.77) and (B.80).

C Appendix: Spatial dynamics of the economy

This section presents TFP measure and its change (subsection C.1), welfare (subsection C.2), wage changes (subsection C.3) and change in housing price (subsection C.4) in equilibrium. Given them, subsection C.5 considers measure of inequality in a local economy and subsection C.6 discusses measure of social mobility of workers and its changes.

C.1 Local Measured TFP

The measured total factor productivity (TFP) in location i and sector s at time t is given by:

$$\ln \delta_{i,t}^{s} = -\frac{1}{\kappa_s} \ln \pi_{ii,t}^{s} + \ln Z_{i,t}^{s} \tag{C.1}$$

and the overall productivity is

$$\ln Z_{i,t}^{s} = \ln A_{i,t}^{s} + \rho \ln \sum_n L_{in,t}^{s} Z_{n,t-1}^{s} + \gamma_s \ln L_{i,t}^{s} \tag{C.2}$$

The change of local TFP to exogenous productivity shock is:

$$\frac{\mathrm{d}\ln \delta_{i,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}} = 1 - \frac{1}{\kappa_s}\frac{\mathrm{d}\ln \pi_{ii,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}} + \rho \sum_n \frac{L_{in,t}^{s} Z_{n,t-1}^{s}}{\sum_\ell L_{i\ell,t}^{s} Z_{\ell,t-1}^{s}} \frac{\mathrm{d}\ln L_{in,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}} + \gamma_s \frac{\mathrm{d}\ln L_{i,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}} \tag{C.3}$$

Letting

$$\widetilde{z}_{in,t}^{s} = \frac{L_{in,t}^{s} Z_{n,t-1}^{s}}{\sum_\ell L_{i\ell,t}^{s} Z_{\ell,t-1}^{s}}, \quad \widetilde{l}_{in,t}^{s} = \frac{L_{in,t}^{s}}{\sum_\ell L_{i\ell,t}^{s}},$$

the change of local TFP is

$$\frac{\mathrm{d}\ln \delta_{i,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}} = 1 - \frac{1}{\kappa_s}\frac{\mathrm{d}\ln \pi_{ii,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}} + \sum_n \left(\rho \widetilde{z}_{in,t}^{s} + \gamma_s \widetilde{l}_{in,t}^{s}\right)\left(\frac{\mathrm{d}\ln \lambda_{in|s,t}}{\mathrm{d}\ln A_{i,t}^{s}} + \frac{\mathrm{d}\ln \varsigma_{n,t}^{s}}{\mathrm{d}\ln A_{i,t}^{s}}\right) \tag{C.4}$$

(C.4) gives the spatial variation of TFP growth along with the technological shock. The second term translates the comparative advantage in trade and its gain is different across locations. The third term captures the migration effects and persistency of workers' choice of industry. These effects change the TFP gains or losses through economies of scale and spillover through workers' mobility. Further, (C.2) can be expressed by

$$\ln \delta_{i,t}^{s} + \frac{1}{\kappa_s} \ln \pi_{ii,t}^{s} = \ln A_{i,t}^{s} + \rho \ln \sum_n L_{in,t}^{s} \left(\left(\pi_{nn,t-1}^{s}\right)^{1/\kappa_s} \delta_{n,t-1}^{s}\right) + \gamma_s \ln L_{i,t}^{s} \tag{C.5}$$

In the steady state, this implies:

$$\ln \delta_i^{s} + \frac{1}{\kappa_s} \ln \pi_{ii}^{s} = \ln A_i^{s} + (\gamma_s + \rho) \ln L_i^{s} + \rho \Delta_i^{s} + \rho \sum_n \frac{L_{in}^{s}}{L_i^{s}} \left(\ln \delta_n^{s} + \frac{1}{\kappa_s} \ln \pi_{nn}^{s}\right) \tag{C.6}$$

where $\Delta_i^s > 0$ is appropriate positive value. Letting

$$\widetilde{\boldsymbol{\delta}}_s = \left\{ \ln \delta_i^s + \frac{1}{\kappa_s} \ln \pi_{ii}^s \right\}, \quad \widetilde{\boldsymbol{A}}_s = \left\{ \ln A_i^s + (\gamma_s + \rho) \ln L_i^s + \rho \Delta_i^s \right\}, \quad \mathbf{L}_s = \left\{ \widetilde{l}_{in}^s \right\},$$

denote $N \times 1$ vectors and $N \times N$ matrix respectively, the equation leads to:

$$\widetilde{\boldsymbol{\delta}}_s = \left(\mathbf{I} - \rho \mathbf{L}_s\right)^{-1} \widetilde{\boldsymbol{A}}_s \tag{C.7}$$

If the spectral radius of $\rho \mathbf{L}_s$ is less than 1, the local TFP in the steady state is given by:

$$\ln \delta_i^s = -\frac{1}{\kappa_s} \ln \pi_{ii}^s + \sum_n \left\{ \sum_{m=0}^{\infty} (\rho \mathbf{L}_s)^m \right\}_{in} \left[\ln A_n^s + (\gamma_s + \rho) \ln L_n^s + \rho \Delta_n^s \right] \tag{C.8}$$

where $\{\sum_{m=0}^{\infty} (\rho \mathbf{L}_s)^m\}_{in}$ is $i - n$ th element of the matrix. The level of local TFP is decomposed into import penetration and spillover in productivity through labor mobility. The latter effect is governed by the matrix:

$$\mathbf{K} = \sum_{m=0}^{\infty} (\rho \mathbf{L}_s)^m = \sum_{m=0}^{\infty} \rho^m \left\{ \lambda_{in|s} \varsigma_n^s L_n^s \right\}^m \tag{C.9}$$

This is given in Proposition 2 in the main text.

C.2 Welfare Implications

Location choice probability satisfies:

$$\lambda_{ii|s,t} = \left(\frac{B_{i,t}^s \mathcal{W}_{i,t}^s}{D_{ii,t} \bar{U}_{i,t}^s} \right)^{\varepsilon} = \left[\frac{B_{i,t}^s \zeta_s^{1/\phi} \mathcal{W}_{i,t}^s \left(L_{i,t-1}^s \right)^{\eta/\phi}}{D_{ii,t} V_{i,t} \left(\varsigma_{i,t}^s \right)^{1/\phi}} \right]^{\varepsilon} \tag{C.10}$$

Using the expenditure share on tradable goods, this becomes:

$$\lambda_{ii|s,t} = \left[\frac{\zeta_s^{1/\phi} B_{i,t}^s \left(\alpha_s W_{i,t}^s \right)^{(1-\sigma)/(\theta_s - \sigma)} \left(L_{i,t-1}^s \right)^{\eta/\phi} \left(\psi_{s|i,t}^s \right)^{1/(\theta_s - \sigma)}}{\left(p_{i,t}^s \right)^{(1-\sigma)/(\theta_s - \sigma)} V_{i,t} \left(\varsigma_{i,t}^s \right)^{1/\phi}} \right]^{\varepsilon} \tag{C.11}$$

In the followings, $B_{i,t}^s$ is constant to focus on the endogenous mechanisms. Price of tradable final goods satisfy

$$\ln p_{i,t}^s = \ln \left[\Gamma_s (w_{i,t}^s)^{\beta_s} (\pi_{ii,t}^s)^{1/\kappa_s} \frac{1}{Z_{i,t}^s} \right] + \sum_j \beta_{sj} \ln p_{i,t}^j \tag{C.12}$$

Letting

$$\widetilde{\boldsymbol{p}}_{i,t} = \left\{ \ln p_{i,t}^s \right\}, \quad \widetilde{\mathbf{B}} = \left\{ \beta_{sk} \right\}, \quad \mathbf{C}_{i,t} = \left\{ \ln \Gamma_s (w_{i,t}^s)^{\beta_s} (\pi_{ii,t}^s)^{1/\kappa_s} \frac{1}{Z_{i,t}^s} \right\} \tag{C.13}$$

be corresponding vector and matrix. Then, price of final goods is:

$$\widetilde{\boldsymbol{p}}_{i,t} = \left(\mathbf{I} - \widetilde{\mathbf{B}} \right)^{-1} \boldsymbol{C}_{i,t} \tag{C.14}$$

Letting $\widetilde{\beta}_{sk}$ be elements of matrix $\left(\mathbf{I}-\widetilde{\mathbf{B}}\right)^{-1}$, price of tradable goods are:

$$p_{i,t}^{s}=\prod_{j}\left[\frac{\Gamma_j(w_{i,t}^{j})^{\beta_j}}{Z_{i,t}^{j}}(\pi_{ii,t}^{j})^{1/\kappa_j}\right]^{\widetilde{\beta}_{sj}} \tag{C.15}$$

Plugging this into above and manipulating it, we derive:

$$V_{i,t}=\bar{\gamma}_s\lambda_{ii|s,t}^{-1/\varepsilon}\left(\prod_j(\pi_{ii,t}^{j})^{-\widetilde{\beta}_{sj}/\kappa_j}\right)^{\widetilde{\theta}_s}\left(\prod_j\left(\frac{(w_{i,t}^{j})^{\beta_j}}{Z_{i,t}^{j}}\right)^{-\widetilde{\beta}_{sj}}\right)^{\widetilde{\theta}_s}{W_{i,t}^{s}}^{\widetilde{\theta}_s}{\psi_{s|i,t}^{s}}^{\widetilde{\theta}_s/(1-\sigma)}{\varsigma_{i,t}^{s}}^{-1/\phi}{L_{i,t-1}^{s}}^{\eta/\phi} \tag{C.16}$$

where $\widetilde{\theta}_s=(1-\sigma)/(\theta_s-\sigma)$ and $\bar{\gamma}_s$ is constant. Therefore, welfare of generation t born in i relative to that of generation $t-1$ born in i is:

$$\begin{aligned}\mathrm{d}\ln V_{i,t}=\frac{1}{S}\sum_{s\in\mathcal{K}\backslash 0}&\left[\widetilde{\theta}_s\left(\mathrm{d}\ln W_{i,t}^{s}-\sum_j\widetilde{\beta}_{sj}\left(\beta_s\mathrm{d}\ln\frac{w_{i,t}^{j}}{Z_{i,t}^{j}}+\frac{\mathrm{d}\ln\pi_{ii,t}^{j}}{\kappa_j}\right)\right)-\frac{\mathrm{d}\ln\lambda_{ii|s,t}}{\varepsilon}\right.\\&\left.+\widetilde{\theta}_s\frac{\mathrm{d}\ln\psi_{s|i,t}^{s}}{1-\sigma}-\frac{\mathrm{d}\ln\varsigma_{i,t}^{s}}{\phi}+\frac{\eta}{\phi}\mathrm{d}\ln L_{i,t-1}^{s}\right]\end{aligned} \tag{C.17}$$

Using local TFP change (C.4), it can be expressed by:

$$\begin{aligned}\mathrm{d}\ln V_{i,t}\propto\sum_{s\in\mathcal{K}\backslash 0}&\left[\widetilde{\theta}_s\beta_s\sum_j\widetilde{\beta}_{sj}\left(\mathrm{d}\ln\delta_{i,t}^{j}-\mathrm{d}\ln w_{i,t}^{j}\right)-\widetilde{\theta}_s\left((1-\beta_s)\sum_j\widetilde{\beta}_{sj}\frac{\mathrm{d}\ln\pi_{ii,t}^{j}}{\kappa_j}\right)\right.\\&\left.-\frac{\mathrm{d}\ln\lambda_{ii|s,t}}{\varepsilon}+\widetilde{\theta}_s\frac{\mathrm{d}\ln e_{s|i,t}^{s}}{1-\sigma}-\frac{\mathrm{d}\ln\varsigma_{i,t}^{s}}{\phi}+\frac{\eta}{\phi}\mathrm{d}\ln L_{i,t-1}^{s}\right]\end{aligned} \tag{C.18}$$

where $e_{s|i,t}^{s}=\psi_{s|i,t}^{s}w_{i,t}^{s}$. This is given in Proposition 3 in text. Alternatively, using expenditure on housing (sector 0), the welfare measure becomes

$$V_{i,t}=\bar{\gamma}_0\lambda_{ii|s,t}^{-1/\varepsilon}\left(\frac{W_{i,t}^{s}}{p_{i,t}^{0}}\right)^{\widetilde{\theta}_0}{\psi_{0|i,t}^{s}}^{\widetilde{\theta}_0/(1-\sigma)}{\varsigma_{i,t}^{s}}^{-1/\phi}{L_{i,t-1}^{s}}^{\eta/\phi} \tag{C.19}$$

for any $s\in\mathcal{K}$. Therefore,

$$\mathrm{d}\ln V_{i,t}=\widetilde{\theta}_0\left(\mathrm{d}\ln W_{i,t}^{s}-\mathrm{d}\ln p_{i,t}^{0}\right)-\frac{\mathrm{d}\ln\lambda_{ii|s,t}}{\varepsilon}+\widetilde{\theta}_0\frac{\mathrm{d}\ln\psi_{0|i,t}^{s}}{1-\sigma}-\frac{\mathrm{d}\ln\varsigma_{i,t}^{s}}{\phi}+\frac{\eta}{\phi}\mathrm{d}\ln L_{i,t-1}^{s} \tag{C.20}$$

Consider welfare loss of migration barrier. Letting $D_{in,t}=1$ for all i and n. Therefore, bilateral migration cost is negligible for any location pairs. Then, the average utility conditional on the sector choice is equalized across locations and location choice probabilities depend on only the destination:

$$\widetilde{\bar{U}}^s_{n,t} = \left[\sum_i \left(\widetilde{\mathcal{W}}^s_{i,t}\right)^\varepsilon\right]^{1/\varepsilon} = \widetilde{\bar{U}}^s_t, \quad \lambda_{in|s,t} = \left(\widetilde{\mathcal{W}}^s_{i,t}/\widetilde{\bar{U}}^s_t\right)^\varepsilon = \widetilde{\lambda}_{i|s,t} \tag{C.21}$$

and

$$\widetilde{\varsigma}^s_{n,t} = \frac{\zeta_s(L^s_{n,t-1})^\eta(\widetilde{\bar{U}}^s_t)^\phi}{\sum_j \zeta_j(L^j_{n,t-1})^\eta(\widetilde{\bar{U}}^j_t)^\phi} = \left[\frac{\zeta_s(L^s_{n,t-1})^{\eta/\phi}\widetilde{\mathcal{W}}^s_{n,t}(\widetilde{\lambda}^s_{n,t})^{-1/\varepsilon}}{\widetilde{V}_{n,t}}\right]^\phi \tag{C.22}$$

Comparing welfare for generation t when eliminating migration barrier,

$$\frac{\widetilde{V}_{i,t}}{V_{i,t}} = \left(\frac{\lambda_{ii|s,t}}{\lambda^s_{i,t}}\right)^{-1/\varepsilon}\left(\frac{\widetilde{\mathcal{W}}^s_{i,t}}{\mathcal{W}^s_{i,t}}\right)\left(\frac{\widetilde{\varsigma}^s_{i,t}}{\varsigma^s_{i,t}}\right)^{-1/\phi} \tag{C.23}$$

When migration cost is high enough, removing the cost leads to high expected utility, $\bar{U}^s_{i,t} < \widetilde{\bar{U}}^s_{i,t}$. Then, $\widetilde{\varsigma}^s_{i,t} > \varsigma^s_{i,t}$.

If taking $\phi \to \infty$, $\widetilde{V}_{i,t}/V_{i,t}$ is greater than the baseline where $\phi < \infty$. Therefore, gains from eliminating migration costs is large when $\phi \to \infty$. Intuitively, when $\phi \to \infty$, workers are homogeneous ex ante and it allows workers to choose the sector and location which returns highest return. Such less specificity of workers ex ante leads to larger welfare gains compared to $\phi < \infty$.

C.3 Local labor market

This subsection considers the income inequality in local labor market and wage changes given in Proposition 4. To simplify the discussion, suppose that $\mu_{it} = 1$ (i.e., absentee land ownership). The wage income share of each sector in local labor market is

$$\begin{aligned} y^s_{i,t} &= \frac{w^s_{i,t}L^s_{i,t}}{\sum_k w^k_{i,t}L^k_{i,t}} = f^s_{i,t}\frac{L_{i,t}}{\widetilde{W}_{i,t}}w^s_{i,t} = f^s_{i,t}\frac{w^s_{i,t}}{\bar{w}_{i,t}} \\ \mathrm{d}\ln y^s_{i,t} &= \mathrm{d}\ln Y^s_{i,t} - \sum_k y^k_{i,t-1}\mathrm{d}\ln Y^k_{i,t} - \frac{1}{2}Var_{\mathbf{y}}(\mathrm{d}\ln Y^k_{i,t}) \end{aligned} \tag{C.24}$$

The local labor supply shows changes over time:

$$\mathrm{d}\ln L^s_{i,t} = \varepsilon\mathrm{d}\ln\mathcal{W}^s_{i,t} + \mathrm{d}\ln\mathbb{H}^s_{i,t} \tag{C.25}$$

where $\mathbb{H}^s_{i,t}$ stands for labor market access:

$$\mathbb{H}^s_{i,t} = \sum_{n\in\mathcal{N}} D^{-\varepsilon}_{in,t}\left(\left(\frac{\zeta_s(L^s_{n,t-1})^\eta}{\varsigma^s_{n,t}}\right)^{-1/\phi}V_{n,t}\right)^{-\varepsilon}\varsigma^s_{n,t}L_{n,t-1} \tag{C.26}$$

The generalized CES price index implies:

$$(\theta_k - \sigma)\mathrm{d}\ln \mathcal{W}^s_{i,t} = \mathrm{d}\ln \psi^s_{k|i,t} - (1-\sigma)\mathrm{d}\ln p^k_{i,t} + (1-\sigma)\mathrm{d}\ln W^s_{i,t} \tag{C.27}$$

for any s and k. Letting

$$\bar{p}_{i,t} = \prod_{k\neq 0}(p^k_{i,t})^{\Psi^k_{i,t-1}}, \quad \bar{\theta}_{i,t-1} = \sum_{k\neq 0}\Psi^k_{i,t-1}\theta_k, \quad \bar{\psi}^s_{i,t} = \prod_{k\neq 0}\left(\psi^s_{k|i,t}\right)^{\Psi^k_{i,t-1}},$$

with

$$\Psi^s_{i,t} = \frac{\sum_k \psi^k_{s|it} W^k_{i,t} L^k_{i,t}}{\sum_{j\neq 0}\sum_k \psi^k_{j|i,t} W^k_{i,t} L^k_{i,t}}. \tag{C.28}$$

$\Psi^s_{i,t}$ is the aggregate expenditure share on sector s among total expenditure on tradable goods. Then, $\bar{p}_{i,t}$ is Törnqvist price index using pre-period expenditure share, and $\bar{\theta}_{i,t-1}$ is weighted average of Engel slope parameter. $\bar{\psi}^s_{i,t}$ is the information about how much workers of sector s's consumption pattern is distorted. To see this,

$$\bar{\psi}^s_{i,t} \leq \sum_j \Psi^j_{i,t-1}\psi^s_{j|i,t}$$

where the inequality holds with equality when $\psi^s_{j|i,t}$ is equalized across all $j \in \mathcal{K}\backslash 0$. By construction,

$$\begin{aligned}\mathrm{d}\ln \bar{p}_{i,t} &= \sum_k \Psi^k_{i,t-1}\mathrm{d}\ln p^k_{i,t}, \\ \mathrm{d}\ln \bar{\psi}^s_{i,t} &= \sum_k \Psi^k_{i,t-1}\mathrm{d}\ln \psi^s_{k|i,t}\end{aligned} \tag{C.29}$$

Then, using (C.25) and (C.27), change in employment is:

$$\mathrm{d}\ln L^s_{i,t} = \varepsilon\frac{1-\sigma}{\bar{\theta}_{i,t-1}-\sigma}\left(\mathrm{d}\ln W^s_{i,t} - \mathrm{d}\ln \bar{p}_{i,t}\right) + \frac{\varepsilon}{\bar{\theta}_{i,t-1}-\sigma}\mathrm{d}\ln \bar{\psi}^s_{i,t} + \mathrm{d}\ln \mathbb{H}^s_{i,t} \tag{C.30}$$

for all $s \in \mathcal{K}$. On the right-hand side, the first term is real income growth, the second term is change in consumption patterns and the last term is change in labor market access. Further, using (C.25), change in wage can be expressed by

$$\mathrm{d}\ln w^s_{i,t} = -\mathrm{d}\ln \mu_{i,t} + \mathrm{d}\ln \bar{p}_{i,t} - \frac{\mathrm{d}\ln \bar{\psi}^s_{i,t}}{1-\sigma} + \frac{\bar{\theta}_{i,t-1}-\sigma}{1-\sigma}\mathrm{d}\ln \mathcal{W}^s_{i,t} \tag{C.31}$$

The labor mobility implies that

$$\mathcal{W}^s_{i,t} = \left(\frac{\lambda_{ii|s,t}}{\lambda_{ni|s,t}}\right)^{1/\varepsilon}\mathcal{W}^s_{n,t}D^{-1}_{in,t}$$

Manipulating the probability of location choice, change in real income is approximated by:

$$\mathrm{d}\ln \mathcal{W}_{i,t}^{s} = \frac{1}{\varepsilon}\mathrm{d}\ln \lambda_{ii|s,t} + \sum_{n} \lambda_{ni|s,t-1}\Big(\mathrm{d}\ln \mathcal{W}_{n,t}^{s} - \mathrm{d}\ln D_{ni,t}\Big) \tag{C.32}$$

Let

$$\boldsymbol{\Lambda}_s = \Big\{\lambda_{ni|s,t-1}\Big\}_{n\in\mathcal{N}, i\in\mathcal{N}}, \quad \boldsymbol{W}_s = \Big\{\mathrm{d}\ln \mathcal{W}_{i,t}^{s}\Big\}_{i\in\mathcal{N}}$$
$$\boldsymbol{C}_s = \Big\{\frac{1}{\varepsilon}\mathrm{d}\ln \lambda_{ii|s,t} - \sum_{n} \lambda_{ni|s,t-1}\mathrm{d}\ln D_{ni,t}\Big\}_{i\in\mathcal{N}}$$

$\boldsymbol{\Lambda}_s$ is matrix of migration patterns of generation $t-1$, $\boldsymbol{W}_s$ is a vector of change in real income for workers in sector s, and $\boldsymbol{C}_s$ is a vector of change in other variables. Then, the change in real income is represented by:

$$\boldsymbol{W}_s = \Big(\mathbf{I} - \boldsymbol{\Lambda}_s^{\top}\Big)^{-1}\boldsymbol{C}_s \tag{C.33}$$

and Let $\widetilde{\lambda}_{in|s,t-1}$ refer to a element of the matrix $\Big(\mathbf{I} - \boldsymbol{\Lambda}_s^{\top}\Big)^{-1}$, change in real income is:

$$\mathrm{d}\ln \mathcal{W}_{i,t}^{s} = \sum_{n} \widetilde{\lambda}_{in|s,t-1}\left(\frac{1}{\varepsilon}\mathrm{d}\ln \lambda_{nn|s,t} - \sum_{\ell} \lambda_{\ell n|s,t-1}\mathrm{d}\ln D_{\ell n,t}\right) \tag{C.34}$$

If $\mathrm{d}\ln D_{ni,t} = 0$, change in real income is

$$\boldsymbol{W}_s = \frac{1}{\varepsilon}\Big(\mathbf{I} - \boldsymbol{\Lambda}_s^{\top}\Big)^{-1}\boldsymbol{\lambda}_s \tag{C.35}$$

where $\boldsymbol{\lambda}_s = \Big\{\mathrm{d}\ln \lambda_{ii|s,t}\Big\}$ is a vector of changes in non-migration probabilities. In addition, (C.15) implies that change of price is:

$$\begin{aligned}\mathrm{d}\ln p_{i,t}^{s} &= \sum_{j} \widetilde{\beta}_{sj}\left(\beta_j \mathrm{d}\ln w_{i,t}^{j} + \frac{1}{\kappa_j}\mathrm{d}\ln \pi_{ii,t}^{j} - \mathrm{d}\ln Z_{i,t}^{j}\right) \\ &= \sum_{j} \widetilde{\beta}_{sj}\left(\beta_j \mathrm{d}\ln w_{i,t}^{j} - \mathrm{d}\ln \delta_{i,t}^{j}\right)\end{aligned} \tag{C.36}$$

for $s \in \mathcal{K}\backslash 0$. Therefore, we have:

$$\mathrm{d}\ln \bar{p}_{i,t} = \sum_{s} \Psi_{i,t-1}^{s}\left[\sum_{j} \widetilde{\beta}_{sj}\Big(\beta_j \mathrm{d}\ln w_{i,t}^{j} - \mathrm{d}\ln \delta_{i,t}^{j}\Big)\right] \tag{C.37}$$

Combining (C.31), (C.35) and (C.37) yields

$$\begin{aligned}\mathrm{d}\ln w^s_{i,t} &= \sum_j \beta_j \widetilde{\Psi}^j_{i,t-1} \mathrm{d}\ln w^j_{i,t} - \sum_j \widetilde{\Psi}^j_{i,t-1} \mathrm{d}\ln \delta^j_{i,t} - \frac{\mathrm{d}\ln \bar{\psi}^s_{i,t}}{1-\sigma} \\ &+ \frac{1}{\varepsilon}\frac{\bar{\theta}_{i,t-1}-\sigma}{1-\sigma} \sum_n \widetilde{\lambda}_{in|s,t-1} \mathrm{d}\ln \lambda_{nn|s,t}\end{aligned} \tag{C.38}$$

where $\widetilde{\Psi}^j_{i,t-1} = \sum_s \Psi^s_{i,t-1}\widetilde{\beta}_{sj}$. Let

$$\boldsymbol{w}_i = \left\{\mathrm{d}\ln w^s_{i,t}\right\}_s, \quad \widetilde{\boldsymbol{\Psi}}_i = \left\{\beta_j \widetilde{\Psi}^j_{i,t-1}\right\}_j,$$

be $S \times 1$ vector of change in wages and $S \times S$ matrix. We also define

$$\begin{aligned}\widetilde{P}_i &\equiv -\sum_j \widetilde{\Psi}^j_{i,t-1}\mathrm{d}\ln \delta^j_{i,t}, \\ \boldsymbol{X}_i &\equiv \left\{-\frac{\mathrm{d}\ln \bar{\psi}^s_{i,t}}{1-\sigma} + \frac{1}{\varepsilon}\frac{\bar{\theta}_{i,t-1}-\sigma}{1-\sigma}\sum_n \widetilde{\lambda}_{in|s,t-1}\mathrm{d}\ln \lambda_{nn|s,t}\right\}_s\end{aligned}$$

$\widetilde{P}_i$ is scalar and $\boldsymbol{X}_i$ is $S \times 1$ vector. Using them, (C.38) is represented by:

$$\boldsymbol{w}_i = \widetilde{\boldsymbol{\Psi}}_i \boldsymbol{w}_i + \widetilde{P}_i + \boldsymbol{X}_i \tag{C.39}$$

Let ϱ^{sj}_i denote the element of inverse matrix $\left(\mathbf{I} - \widetilde{\boldsymbol{\Psi}}_i\right)^{-1}$. Then, change of wage is:

$$\mathrm{d}\ln w^s_{i,t} = \sum_j \varrho^{sj}_i \left(-\frac{\mathrm{d}\ln \bar{\psi}^j_{i,t}}{1-\sigma} + \frac{1}{\varepsilon}\frac{\bar{\theta}_{i,t-1}-\sigma}{1-\sigma}\sum_n \widetilde{\lambda}_{in|j,t-1}\mathrm{d}\ln \lambda_{nn|j,t} - \sum_k \widetilde{\Psi}^k_{i,t-1}\mathrm{d}\ln \delta^k_{i,t}\right) \tag{C.40}$$

This is hold for any tradables, and this expression is given in Proposition 4.

We see the special case to consider how employment share across different sectors in the local labor market is determined in equilibrium. Consider the economy without migration costs, $D_{in,t} = 1$ for all i and n. We also let $\zeta_s = \zeta$ for all s. Then, employment share in the local labor market is:

$$f^s_{i,t} \equiv \frac{L^s_{i,t}}{\sum_j L^j_{i,t}} = \frac{\lambda^s_{i,t}}{\sum_j \lambda^j_{i,t}} \times \frac{\sum_j \lambda^j_{i,t}}{\sum_j \lambda^j_{i,t} L^j_t} \times L^s_t \tag{C.41}$$

Using (C.22),

$$\frac{\lambda^s_{i,t}}{\sum_j \lambda^j_{i,t}} = \frac{(L^s_{i,t-1})^{\varepsilon\eta/\phi}(\mathcal{W}^s_{i,t})^{\varepsilon}(\varsigma^s_{i,t})^{-\varepsilon/\phi}}{\sum_j (L^j_{i,t-1})^{\varepsilon\eta/\phi}(\mathcal{W}^j_{i,t})^{\varepsilon}(\varsigma^j_{i,t})^{-\varepsilon/\phi}} \tag{C.42}$$

Therefore, relative labor share between sector s and j is:

$$\frac{f_{i,t}^s}{f_{i,t}^j} = \frac{L_t^s}{L_t^j}\left(\frac{\mathcal{W}_{i,t}^s}{\mathcal{W}_{i,t}^j}\right)^{\varepsilon}\left(\frac{\varsigma_{i,t}^s}{\varsigma_{i,t}^j}\right)^{-\varepsilon/\phi}\left(\frac{f_{i,t-1}^s}{f_{i,t-1}^j}\right)^{\varepsilon\eta/\phi} \tag{C.43}$$

When $\phi \to \infty$, workers are homogeneous ex-ante location choice and relative labor share is a combination of macro-level employment growth (L_t^s/L_t^j) and real income difference, $(\mathcal{W}_{i,t}^s/\mathcal{W}_{i,t}^j)$. The parameter ε controls the "curvature" of labor supply. When $\phi < \infty$, the relative propensity to choose the sector $(\varsigma_{i,t}^s/\varsigma_{i,t}^j)$ is negatively associated with the labor ratio. Intuitively, a higher propensity to sort into the sector implies a higher average utility of the sector. In turn, conditional on the real income difference, a higher average utility leads to a small probability of staying in the location. The last term is related to persistence in the labor market. Larger externalities in the labor market lead to the persistence of the local labor market.

C.4 Price Dynamics of Immobile Structure

This subsection derives first-order change of prices for the immobile structure. Assume that $\mu_{it} = 1$. Define the share of demand from workers of sector s in housing demand:

$$\xi_{i,t}^s \equiv \frac{\psi_{0|i,t}^s w_{i,t}^s L_{i,t}^s}{p_{i,t}^0 H_{i,t}} = \chi\psi_{0|i,t}^s \frac{y_{i,t}^s}{y_{i,t}^0} \tag{C.44}$$

The first-order change of demand for housing in location i:

$$\begin{aligned}\mathrm{d}\ln p_{i,t}^0 + \mathrm{d}\ln H_{i,t} &= \sum_{s\in\mathcal{K}}\xi_{i,t-1}^s\Big(\mathrm{d}\ln Y_{i,t} + \mathrm{d}\ln y_{i,t}^s + \mathrm{d}\ln\psi_{0|i,t}^s\Big)\\ &= \sum_{s\in\mathcal{K}}\xi_{i,t-1}^s\Big(\mathrm{d}\ln r_{i,t} - \mathrm{d}\ln y_{i,t}^0 + \mathrm{d}\ln y_{i,t}^s + \mathrm{d}\ln\psi_{0|i,t}^s\Big)\end{aligned} \tag{C.45}$$

where $\mathrm{d}\ln Y_{i,t}$ is given by the zero profit condition for developers. An increase in the stock of structure over time is given by:

$$\mathrm{d}\ln H_{i,t} = \chi\mathrm{d}\ln L_{i,t}^0 - (1-\chi)\sum_{\omega=1}^{t-1}(-\chi)^{\omega}\mathrm{d}\ln L_{i,t-\omega}^0 + (1-\chi)^t\ln(1-\nu_i) \tag{C.46}$$

Together, price change is

$$\mathrm{d}\ln p_{i,t}^0 = -\chi\mathrm{d}\ln L_{i,t}^0 + \mathrm{d}\ln r_{i,t} + \sum_{s\in\mathcal{K}}\xi_{i,t-1}^s\Big(\mathrm{d}\ln y_{i,t}^s - \mathrm{d}\ln y_{i,t}^0\Big) + \sum_{s\in\mathcal{K}}\xi_{i,t-1}^s\mathrm{d}\ln\psi_{0|i,t}^s \tag{C.47}$$

For the last term on the right-hand side,

$$\Big(\theta_0 - \bar{\theta}_{i,t-1}\Big)\mathrm{d}\ln\mathcal{W}_{i,t}^s = \Big(\mathrm{d}\ln\psi_{0|i,t}^s - \mathrm{d}\ln\bar{\psi}_{i,t}^s\Big) - (1-\sigma)\Big(\mathrm{d}\ln p_{i,t}^0 - \mathrm{d}\ln\bar{p}_{i,t}\Big) \tag{C.48}$$

Letting

$$\widetilde{\boldsymbol{p}}_t^0 = \left\{\frac{p_{i,t}^0}{\bar{p}_{i,t}}\right\}, \quad \widetilde{\boldsymbol{r}}_t = \left\{\frac{r_{i,t}}{\bar{p}_{i,t}}\right\},$$

be a vector of prices and land rents normalized by average prices, the first-order change of price for building structure is:

$$\begin{aligned} \sigma \mathrm{d} \ln \widetilde{\boldsymbol{p}}^0 &= -\chi \mathrm{d} \ln \boldsymbol{L}^0 + \mathrm{d} \ln \widetilde{\boldsymbol{r}} + \sum_{s \in \mathcal{K}} \boldsymbol{\xi}_{t-1} (\mathrm{d} \ln y_{i,t}^s - \mathrm{d} \ln y_{i,t}^0) \\ &+ \sum_{s \in \mathcal{K}} \boldsymbol{\xi}_{t-1} (\theta_0 - \bar{\boldsymbol{\theta}}_{t-1}) \mathrm{d} \ln \mathbf{W}_s + \sum_{s \in \mathcal{K}} \boldsymbol{\xi}_{t-1} \mathrm{d} \ln \bar{\boldsymbol{\psi}} \end{aligned} \tag{C.49}$$

On the right-hand side, the first term is an increase of supply controlled by the share of labor in the production of structure, the second term is a change in land rent, the third term is a change in income distribution with weighting demand share, and the last two terms together define the change in expenditure share.

C.5 Dynamics of Local Inequality

This part corresponds to Section 4.3 in main text.

The income distribution in location i is discrete and it is determined by the employment distribution across sectors. The average and variance of income in location i in period t is given by:

$$\bar{W}_{i,t} \equiv \sum_{s \in \mathcal{K}} f_{i,t}^s W_{i,t}^s = \mu_{i,t} \bar{w}_{i,t} \quad \text{and} \quad \mathrm{Var}_s(W_{i,t}^s) = \sum_{s \in \mathcal{K}} f_{i,t}^s (W_{i,t}^s)^2 - \bar{W}_{i,t}^2 \tag{C.50}$$

Using them, the coefficient of variation of income in location i is given by:

$$\mathbb{CV}_{i,t} = \frac{\sqrt{\mathrm{Var}_s(W_{i,t}^s)}}{\bar{W}_{i,t}} = \left(\frac{1}{\bar{W}_{i,t}^2} \sum_{s \in \mathcal{K}} \frac{L_{i,t}^s W_{i,t}^s}{L_{i,t}} W_{i,t}^s - 1\right)^{1/2} = \left[\sum_{s \in \mathcal{K}} y_{i,t}^s \left(\frac{y_{i,t}^s - f_{i,t}^s}{f_{i,t}^s}\right)\right]^{1/2}. \tag{C.51}$$

The coefficient of variation in location i is characterized by convex combination of wage difference using weight of income share. Further, manipulating this,

$$\mathbb{CV}_{i,t}^2 + 1 = \sum_{s \in \mathcal{K}} y_{i,t}^s \frac{w_{i,t}^s}{\bar{w}_{i,t}} = \sum_{s \in \mathcal{K}} f_{i,t}^s \left(\frac{w_{i,t}^s}{\bar{w}_{i,t}}\right)^2 \tag{C.52}$$

This is the monotonic transformation of the coefficient of variation, and therefore the right-hand-side can be used as a measure of intra-location income inequality. Let

$$\mathcal{I}_{i,t} = \sum_{s \in \mathcal{K}} y_{i,t}^s \frac{w_{i,t}^s}{\bar{w}_{i,t}} = \sum_{s \in \mathcal{K}} y_{i,t}^s \frac{y_{i,t}^s}{f_{i,t}^s} = \sum_{s \in \mathcal{K}} f_{i,t}^s \left(\frac{y_{i,t}^s}{f_{i,t}^s}\right)^2$$

Large value of this measure is corresponding to larger income inequality in the local labor market.

This is the convex combination of relative wage ($w^s_{i,t}/\bar{w}_{i,t}$) or ratio of income share and labor share ($y^s_{i,t}/f^s_{i,t}$). The change of this measure from period $t-1$ to period t is given by:

$$\begin{aligned} \mathrm{d}\ln \mathcal{I}_{i,t} &= \sum_{s\in\mathcal{K}} \underbrace{\frac{f^s_{i,t-1}(y^s_{i,t-1}/f^s_{i,t-1})^2}{\mathcal{I}_{i,t-1}}}_{\iota^s_{i,t-1}} \left(\mathrm{d}\ln y^s_{i,t} + (\mathrm{d}\ln y^s_{i,t} - \mathrm{d}\ln f^s_{i,t})\right) \\ &= \sum_{s\in\mathcal{K}} \iota^s_{i,t-1} \left(\mathrm{d}\ln y^s_{i,t} + \left(\mathrm{d}\ln w^s_{i,t} - \mathrm{d}\ln \bar{w}_{i,t}\right)\right) \\ &= \sum_{s\in\mathcal{K}} \iota^s_{i,t-1} \left[\mathrm{d}\ln y^s_{i,t} + \left(\mathrm{d}\ln Y^s_{i,t} - \sum_{k\in\mathcal{K}} y^k_{i,t-1}\mathrm{d}\ln Y^k_{i,t}\right) - \left(\mathrm{d}\ln L^s_{it} - \sum_{k\in\mathcal{K}} f^k_{it-1}\mathrm{d}\ln L^k_{it}\right)\right] \end{aligned} \tag{C.53}$$

where ι^s_{it-1} is the contribution of sector s in the income inequality in the previous generation. This captures the trend of income inequality in location i. In the parenthesis, the second term is relative income growth and the third term is relative employment growth.

The slope of Lorenz curve depends on the relative wage ($w^s_{it}/\bar{w}_{it}$) and it is equal to y^s_{it}/f^s_{it}, so that change of $(\boldsymbol{f}, \boldsymbol{y})$ are sufficient to characterize this. In particular, the difference between growth rate of wage and growth rate of labor across sectors changes the convexity of the Lorenz curve. The Gini coefficient for the local labor market for generation t becomes:

$$\begin{aligned} \mathrm{Gini}_{i,t} &\equiv \frac{1}{2\bar{w}_{i,t}} \sum_{s\in\mathcal{K}}\sum_{k\in\mathcal{K}} f^s_{i,t}f^k_{i,t}|w^s_{i,t} - w^k_{i,t}| \\ &\propto \sum_{s\in\mathcal{K}}\sum_{k\in\mathcal{K}} f^s_{i,t}f^k_{i,t}\left|\frac{y^s_{i,t}}{f^s_{i,t}} - \frac{y^k_{i,t}}{f^k_{i,t}}\right| = \sum_{s\in\mathcal{K}}\sum_{k\in\mathcal{K}} \left|y^s_{i,t}f^k_{i,t} - y^k_{i,t}f^s_{it}\right| \end{aligned} \tag{C.54}$$

Another index is Theil index. Theil index for local labor market i is defined as:

$$\mathrm{Theil}_{i,t} \equiv \sum_{s\in\mathcal{K}} f^s_{i,t}\frac{w^s_{it}}{\bar{w}_{i,t}} \ln\left(\frac{w^s_{i,t}}{\bar{w}_{i,t}}\right) = \sum_{s\in\mathcal{K}} f^s_{i,t}\frac{y^s_{i,t}}{f^s_{i,t}} \ln\left(\frac{y^s_{i,t}}{f^s_{i,t}}\right) \tag{C.55}$$

To sum, the model based inequality measure in the local labor market takes a form:

$$\mathcal{I}_{i,t} = \sum_{s\in\mathcal{K}} f^s_{i,t} \times G\left(\frac{y^s_{i,t}}{f^s_{i,t}}\right) \tag{C.56}$$

where $G(\cdot)$ is appropriate function for the slope of Lorenz curve. When $G(\cdot)$ is specified such that $G(x) = x^2$, it is based on the coefficient of variation. If $G(x) = x\ln x$, it is Theil index.

The following discussion focus on the change of measure (C.53) over time. Incorporating

$$\frac{1-\sigma}{\theta_0-\sigma}\sum_{s\in\mathcal{K}}\omega^s_{i,t-1}\mathrm{d}\ln W^s_{i,t} = \frac{1-\sigma}{\theta_0-\sigma}\mathrm{d}\ln \bar{W}_{i,t} - \frac{1-\sigma}{\theta_0-\sigma}\sum_{s\in\mathcal{K}}\omega^s_{i,t-1}\mathrm{d}\ln f^s_{it} \tag{C.57}$$

and

$$\frac{\mathrm{d}\ln\psi_{0|is,t}}{\theta_0-\sigma}=\frac{1-\sigma}{\theta_0-\sigma}\mathrm{d}\ln p_{i,t}^0+\frac{\sigma-1}{\theta_0-\sigma}\mathrm{d}\ln W_{i,t}^s+\mathrm{d}\ln\mathcal{W}_{i,t}^s \tag{C.58}$$

into (C.19) leads to:

$$\begin{aligned}&\sum_s\omega_{i,t-1}^s\left(\mathrm{d}\ln f_{i,t}^s+\left(\mathrm{d}\ln w_{i,t}^s-\mathrm{d}\ln\bar{w}_{i,t}\right)\right)\\&=\frac{\theta_0-\sigma}{1-\sigma}\left[-\mathrm{d}\ln V_{i,t}+\sum_s\omega_{i,t-1}^s\left(\frac{\eta}{\phi}\mathrm{d}\ln L_{i,t-1}^s-\frac{\mathrm{d}\ln\varsigma_{i,t}^s}{\phi}-\frac{1}{\varepsilon}\mathrm{d}\ln\lambda_{ii|s,t}+\mathrm{d}\ln\mathcal{W}_{i,t}^s\right)\right]\end{aligned} \tag{C.59}$$

where $\omega_{i,t-1}^s$ is appropriate weight such that $\sum_s\omega_{i,t-1}^s=1$. Manipulating this,

$$\begin{aligned}\frac{1-\sigma}{\theta_0-\sigma}\sum_s\omega_{i,t-1}^s\mathrm{d}\ln y_{i,t}^s=-\mathrm{d}\ln V_{i,t}+\sum_s\omega_{i,t-1}^s\Bigg[\frac{\eta}{\phi}\mathrm{d}\ln L_{i,t-1}^s-\frac{\mathrm{d}\ln\varsigma_{i,t}^s}{\phi}\\+\frac{1}{\varepsilon}\left(\sum_n\widetilde{\lambda}_{in|s,t-1}\mathrm{d}\ln\lambda_{nn|s,t}-\mathrm{d}\ln\lambda_{ii|s,t}\right)\Bigg]\end{aligned} \tag{C.60}$$

For variable $\boldsymbol{x}$, define the following notation for its normalization by weighted geometric mean:

$$\widetilde{x}_{i,t}^s\equiv\frac{x_{i,t}^s}{\prod_{s\in\mathcal{K}}(\boldsymbol{x}_{i,t}^s)^{\omega_{i,t-1}^s}} \tag{C.61}$$

Using this notation,

$$\frac{1-\sigma}{\theta_0-\sigma}\mathrm{d}\ln\widetilde{y}_{i,t}^s=-\frac{\eta}{\phi}\mathrm{d}\ln\widetilde{L}_{i,t-1}^s+\frac{1}{\phi}\mathrm{d}\ln\widetilde{\varsigma}_{i,t}^s+\frac{1}{\varepsilon}\widetilde{\lambda}_{is,t}^{\mathbf{N}} \tag{C.62}$$

where

$$\lambda_{is,t}^{\mathbf{N}}=\lambda_{ii|s,t}\prod_n\lambda_{nn|s,t}^{\widetilde{\lambda}_{in|s,t-1}} \tag{C.63}$$

Therefore,

$$\begin{aligned}\frac{1-\sigma}{\theta_0-\sigma}\left(\mathrm{d}\ln w_{it}^s-\mathrm{d}\ln\bar{w}_{it}\right)=-\frac{\eta}{\phi}\mathrm{d}\ln\widetilde{\boldsymbol{L}}_{t-1}+\frac{1}{\phi}\mathrm{d}\ln\widetilde{\boldsymbol{\varsigma}}_t+\frac{1}{\varepsilon}\mathrm{d}\ln\widetilde{\boldsymbol{\lambda}}_t^{\mathbf{N}}-\frac{1}{\theta_0-\sigma}\mathrm{d}\ln\widetilde{\boldsymbol{\psi}}_t^0\\-\frac{1-\sigma}{\theta_0-\sigma}\sum_{s\in\mathcal{K}}y_{it-1}^s\mathrm{d}\ln f_{it}^s\end{aligned} \tag{C.64}$$

Manipulating this,

$$\mathrm{d}\ln y_{it}^s=-\frac{\eta}{\phi}\frac{\theta_0-\sigma}{1-\sigma}\mathrm{d}\ln\widetilde{\boldsymbol{L}}_{t-1}+\frac{1}{\phi}\frac{\theta_0-\sigma}{1-\sigma}\mathrm{d}\ln\widetilde{\boldsymbol{\varsigma}}_t+\frac{1}{\varepsilon}\frac{\theta_0-\sigma}{1-\sigma}\mathrm{d}\ln\widetilde{\boldsymbol{\lambda}}_t^{\mathbf{N}}-\frac{1}{1-\sigma}\mathrm{d}\ln\widetilde{\boldsymbol{\psi}}_t^0+\mathrm{d}\ln\widetilde{\boldsymbol{f}}_t$$

Summarizing the discussion for the measure of income inequality based on the coefficient of variation, the change of income inequality at the local level over generations is:

$$\mathrm{d}\ln \mathcal{I}_{i,t} = \sum_{s\in\mathcal{K}} \iota^s_{i,t-1}\mathrm{d}\ln G^s_{i,t} - \underbrace{\sum_{s\in\mathcal{K}}\left(y^s_{i,t-1} - f^s_{i,t-1}\right)\mathrm{d}\ln L^s_{i,t}}_{\textit{Composition effect}} \tag{C.65}$$

with

$$\mathrm{d}\ln G^s_{i,t} = \underbrace{\mathrm{d}\ln \widetilde{Y}^s_{i,t}}_{\textit{Industry}} - \underbrace{\frac{\eta}{\phi}\frac{\theta_0-\sigma}{1-\sigma}\mathrm{d}\ln \widetilde{L}^s_{i,t-1}}_{\textit{Persistence}} + \underbrace{\frac{1}{\phi}\frac{\theta_0-\sigma}{1-\sigma}\mathrm{d}\ln \widetilde{\varsigma}^s_{i,t}}_{\textit{Sectoral choice}} + \underbrace{\frac{1}{\varepsilon}\frac{\theta_0-\sigma}{1-\sigma}\mathrm{d}\ln \widetilde{\lambda}^s_{ii,t}}_{\textit{Location choice}} - \underbrace{\frac{1}{1-\sigma}\mathrm{d}\ln \widetilde{\psi}^s_{0|i,t}}_{\textit{Expenditure on housing}} \tag{C.66}$$

where $\iota^s_{i,t-1}$ is the contribution of sector s in the income inequality among previous generation $t-1$, and let $\widetilde{x}^s_{i,t}$ refers the transformed variables using previous income share such that $\widetilde{x}^s_{i,t} \equiv \frac{x^s_{it}}{\prod_{s\in\mathcal{K}}(x^s_{it})^{y^s_{it-1}}}$.

This illustrates how income inequality in the local market is related to spatial structural change. The composition effect is a standard: employment shift from industry with relatively lower wage to higher wage suppress the inequality. Other than this, relative growth of industry ($\mathrm{d}\ln \widetilde{\boldsymbol{Y}}_t$), pre-trend of employment growth ($\mathrm{d}\ln \widetilde{\boldsymbol{L}}_{t-1}$), change of workers' sorting pattern in the sectoral choice ($\mathrm{d}\ln \widetilde{\boldsymbol{\varsigma}}_t$), difference in no-mobility workers ($\mathrm{d}\ln \widetilde{\boldsymbol{\lambda}}^{\mathbf{N}}_t$) and difference in expenditure share in housing ($\mathrm{d}\ln \widetilde{\boldsymbol{\psi}}^0_t$) shift the income inequality together. First, the industrial agglomeration creates the uneven labor adjustment process across industries in the first and second terms. The relative wage growth in the sector is positively associated with the sector's contribution to an expansion of income inequality. Without heterogeneity in sectoral choice ($\phi \to \infty$), these mechanisms of factor specificity are absent in the change of income inequality. A small probability of staying in the place is positively associated with a change in income inequality. When idiosyncratic shocks are more heterogeneous ($\varepsilon \to 1$), its contribution becomes large as workers must face a large gap in wage growth to stay conditional on industry choice. The last term says that the strong congestion force counteracts the positive composition effects.[20] Given that congestion forces are substantial when more substitutes, this countereffect is magnified by lower $1-\sigma$. So far, this gives the general equilibrium relationship between the different trends of income inequality in local labor markets along with spatial structural transformation, factor specificity, and labor mobility.

C.6 Intergenerational mobility

This subsection is appendix for the discussion about intergenerational income mobility in Section 4.3.

Notation. $W^{\mathbf{o}}_{i,t}(\omega)$ is income of individual worker ω of generation t working in location i. $W^{\mathbf{y}}_{i,t+1}(\omega)$ is income of individual of generation $t+1$ (i.e., children) who has origin in location i. In the model,

[20] We present the dynamics of price for immobile factors in Appendix C.4.

income distribution in the economy is the probability mass function, as the model derives the discrete finite number of possible income levels in the economy. Let $\mathbb{Y}_t$ denote the set of income levels in the economy and $\mathbb{Y}_{i,t}$ denote that in location i. $\mathcal{Q}_t(\cdot)$ is the probability distribution function for the income in period t in the whole economy in our model, and $\mathcal{Q}_{t+1}(\cdot)$ is that for period $t+1$. They are model based distribution, while we refer to $\mathcal{Q}^*_t(\cdot)$ and $\mathcal{Q}^*_{t+1}(\cdot)$ as those in data, so they are continuous distribution.

$\mathbb{P}_{i,t}(y)$ denotes the probability mass function for income y in location i at period t and $\mathbb{M}_{t+1}(y|i \to n)$ denotes the probability mass function of income level y among people of generation $t+1$ who move from location i to n. They satisfy:

$$\sum_{y\in\mathbb{Y}_{i,t}} \mathbb{P}_{i,t}(y) = 1, \quad \sum_{y\in\mathbb{Y}_t} \mathbb{M}_{t+1}(y|i \to n) = 1 \tag{C.67}$$

Measure of average upward mobility. Define:

$$\mathcal{R}^{\mathbf{o}}_{n,t} = \mathbb{E}[\mathcal{Q}^*_t(W^{\mathbf{o}}_{n,t}(\omega))] = \int_0^\infty \mathbb{P}_{n,t}(y)\mathcal{Q}^*_t(y)dy,$$
$$\mathcal{R}^{\mathbf{y}}_{n,t+1} = \mathbb{E}[\mathcal{Q}^*_{t+1}(W^{\mathbf{y}}_{n,t+1}(\omega))] = \sum_{\ell\in\mathcal{N}}\int_0^\infty \mathbb{M}_{t+1}(y|n \to \ell)\mathcal{Q}^*_{t+1}(y)dy$$

Then, a measure for the average upward mobility corresponding is the ratio of these two measures:

$$\mathcal{M}_{n,t+1} = \mathcal{R}^{\mathbf{y}}_{n,t+1}/\mathcal{R}^{\mathbf{o}}_{n,t} \tag{C.68}$$

This measure is related to the equilibrium variables in the model. First,

$$\mathcal{R}^{\mathbf{o}}_{i,t} = \sum_{s\in\mathcal{K}} f^s_{i,t} \times \mathcal{Q}_t(\mu_{i,t}w^s_{i,t}) \tag{C.69}$$

where $\mathcal{Q}_t(\mu_{i,t}w^s_{i,t})$ is the percentile of workers with income $\mu_{i,t}w^s_{i,t}$ in the entire economy. Given the probability mass function for the income in each location across sectors, $\mathbb{P}_{i,t}(y)$ is corresponding to the share of employment in different sector.

Next, income of generation $t+1$ from location i can yield $N \times (S+1)$ possible incomes in equilibrium. The proportion of each income level is identical to the choice probability of the industry and destination for work. Hence, the corresponding measure is:

$$\mathcal{R}^{\mathbf{y}}_{i,t+1} = \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{K}} \lambda_{ni|s,t+1}\varsigma^s_{i,t+1} \times \mathcal{Q}_{t+1}(\mu_{n,t+1}w^s_{n,t+1}) \tag{C.70}$$

This is the average rank of generation $t+1$ from the origin i. Combining them, the measure (C.68) becomes:

$$\mathcal{M}_{i,t+1} = \sum_{s\in\mathcal{K}} \varsigma^s_{i,t+1}\left[\sum_{n\in\mathcal{N}} \lambda_{ni|s,t+1}\frac{\mathcal{Q}_{t+1}(\mu_{n,t+1}w^s_{n,t+1})}{\sum_{k\in\mathcal{K}} f^k_{i,t}\mathcal{Q}_t(\mu_{i,t}w^k_{i,t})}\right] \tag{C.71}$$

Using the mass probability function, the rank of income is represented by:

$$\mathcal{Q}_{i,t}^{k} \equiv \mathcal{Q}_{t}(\mu_{i,t}w_{i,t}^{k}) = \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{K}} f_{n,t}^{s}\cdot \mathbf{1}_{Z}\Big(\mu_{n,t}w_{n,t}^{s} \leq \mu_{i,t}w_{i,t}^{k}\Big)\frac{L_{n,t}}{\bar{L}} \tag{C.72}$$

This is the percentile of workers with income $\mu_{i,t}w_{i,t}^{k}$ at the national level. Then, the measure (C.71) is rewritten by:

$$\mathcal{M}_{i,t+1} = \sum_{s\in\mathcal{K}} \varsigma_{i,t+1}^{s}\frac{\mathcal{Q}_{i,t}^{s}}{\sum_{k\in\mathcal{K}} f_{i,t}^{k}\mathcal{Q}_{i,t}^{k}}\frac{\mathcal{Q}_{i,t+1}^{s}}{\mathcal{Q}_{i,t}^{s}}\left[\sum_{n\in\mathcal{N}}\lambda_{ni|s,t+1}\frac{\mathcal{Q}_{n,t+1}^{s}}{\mathcal{Q}_{i,t+1}^{s}}\right] \tag{C.73}$$

This representation is very intuitive for understanding the upward mobility. The first term is the job opportunity in location i for generation $t+1$. The second term captures the local income inequality for generation t. The third term is growth of local labor market over generations represented by the rank-up for each sector. The last term in parenthesis is gains from labor mobility for generation $t+1$. This is given in the proposition in main text. High value of $\mathcal{M}_{i,t+1}$ implies that the next generation $(t+1)$ are *expected* to be climbing up the income ladder compared to the *average standard* of their parents (generation t). Its heterogeneity across space comes from the difference in each elements at work in (C.73).

To see the asymmetric effect between emigrants and stayers in the location i, consider the decomposition of (C.73) into different types of workers. First, let

$$\mathbb{Q}_{i,t+1}^{s} = \frac{\mathcal{Q}_{i,t}^{s}}{\sum_{k\in\mathcal{K}} f_{i,t}^{k}\mathcal{Q}_{i,t}^{k}}\frac{\mathcal{Q}_{i,t+1}^{s}}{\mathcal{Q}_{i,t}^{s}} \tag{C.74}$$

This part in (C.73) shows the relative wage growth of the sector in the local economy. Apart from workers' choice of industry and location, the industry growth of *ex-ante* high-wage sector is associated with an increase of upward mobility. Now, we straightforward obtain:

$$\mathcal{M}_{i,t+1} = \sum_{s\in\mathcal{K}}\mathbb{Q}_{i,t+1}^{s}\varsigma_{i,t+1}^{s} + \sum_{s\in\mathcal{K}}\mathbb{Q}_{i,t+1}^{s}\varsigma_{i,t+1}^{s}\sum_{n\in\mathcal{N}\setminus i}\lambda_{ni|s,t+1}\left(\frac{\mathcal{Q}_{n,t+1}^{s}}{\mathcal{Q}_{i,t+1}^{s}} - 1\right) \tag{C.75}$$

The first term is the sector specificity in the local labor market and income growth of natives in i (i.e., workers of generation $t+1$ who do not move to other locations). The second term is the location i's land of opportunity for emigrants. When location i has greater labor market access for the growing industries, this allows workers to climb up the income ladder by reallocation.

Absolute upward mobility. Another measure of upward mobility is the probability that children (workers of generation $t+1$) obtain higher income than their parents. To construct the measure, we introduce some additional notations. Given $\alpha \in (0,1)$ and for probability mass function of generation t, $\mathbb{G}_{i,t}(y)$, define

$$\begin{aligned}\inf\{y|\mathbb{G}_{i,t}(y)\geq\alpha\} &= \arg\min_{\nu}\sum_{\{s|\mu_{i,t}w^s_{i,t}\leq\nu\}} -f^s_{i,t}(1-\alpha)(\mu_{i,t}w^s_{i,t}-\nu)+\sum_{\{s|\mu_{i,t}w^s_{i,t}>\nu\}} f^s_{i,t}\alpha(\mu_{i,t}w^s_{i,t}-\nu)\\ &= \arg\min_{\nu}\sum_{s\in\mathcal{K}} f^s_{i,t}\Big(\alpha-\mathbf{1}(\mu_{i,t}w^s_{i,t}\leq\nu)\Big)\Big(\mu_{i,t}w^s_{i,t}-\nu\Big)\equiv\xi_{i,t}(\alpha)\end{aligned} \tag{C.76}$$

$\xi_{i,t}(\alpha)$ is the α-th quantile of income distribution of generation t in location i. For instance, $\xi_{i,t}(\alpha = 0.5)$ gives the median of income distribution. Using this, define the measure:

$$\mathcal{L}_{i,t+1}(\alpha)=\sum_{n\in\mathcal{N}}\sum_{y\in\mathcal{Y}_{t+1}}\lambda_{ni|s,t+1}\mathbb{M}_{t+1}(y|i\to n)\times\mathbf{1}_Y(y>\xi_{i,t}(\alpha)),\quad \alpha\in(0,1) \tag{C.77}$$

where $\mathbf{1}_Y(y>\xi_{i,t}(\alpha))$ is an indicator function that returns to one if income of generation $t+1$ is higher than the quantile level of parents, $\xi_{i,t}(\alpha)$. Further, using the probabilities in the model, this becomes

$$\mathcal{L}_{i,t+1}(\alpha)=\sum_{s\in\mathcal{K}}\sum_{n\in\mathcal{N}}\lambda_{ni|s,t+1}\mathbf{1}_W\Big(W^s_{n,t+1}>\xi_{i,t}(\alpha)\Big)\varsigma^s_{i,t+1},\quad \alpha\in(0,1) \tag{C.78}$$

This is the probability that workers of generation $t+1$ from i will get higher income relative to α-th quantile of income among previous generation. Compared to the measure (C.73), this measure does not rely on the income distribution at the national level. By construction, $\mathcal{L}_{i,t+1}(\alpha)$ takes the value between 0 and 1, and higher value corresponds to the higher *absolute* upward mobility. This measure is showing the similar pattern to the numerator of $\mathcal{M}_{i,t+1}$ when there are small variations in $\xi_{i,t}(\alpha)$ across locations given α.

Possibility of American Dream. The index (19) reflects the first-order moment of the income distribution. To see another measure of upward mobility from the bottom to the top, "American Dream", the next measure compares the people at the bottom of the quantile and the top of the quantile. Let $\mathcal{G}_{i,t}(\cdot):\mathbb{R}_+\to[0,1]$ refer the income distribution of among workers of generation t in i and $\mathcal{H}_{i,t+1}(\cdot):\mathbb{R}_+\to[0,1]$ denote the income distribution of generation $t+1$ whose *origin* is i. $\mathcal{Q}_t(\cdot)$ is probability distribution function for income in period t at the economy-wide. Using them, we define:

$$\varkappa_{i,t+1}=\frac{\mathcal{Q}_{t+1}\Big(\inf\{y|\mathcal{H}_{i,t+1}(y)\geq\overline{\alpha}\}\Big)}{\mathcal{Q}_t\Big(\inf\{y|\mathcal{G}_{i,t}(y)\geq\underline{\alpha}\}\Big)},\quad 0<\underline{\alpha}<<\overline{\alpha}<1 \tag{C.79}$$

The intuition for (C.79) is followings. The position in the whole economy for a worker of generation t in location i at $\underline{\alpha}$-th quantile is the denominator in (C.79). When location i is relatively wealthy in the economy, this returns a large value. The numerator of (C.79) identifies the position of the top income cohort originated from i. For workers of generation $t+1$ born in location i, their future income distribution is $\mathcal{H}_{i,t+1}(\cdot)$ and $\overline{\alpha}$-th quantile among the cohort is given by $\inf\{y|\mathcal{H}_{i,t+1}(y)\geq$

$\overline{\alpha}\}$. Therefore, a large value of (C.79) suggests the American Dream in location i: the top income people arise from the cohort of generation $t+1$ born in location i where the workers in the previous generation are relatively lower income group.

The value of quantile, $\underline{\alpha}$ and $\overline{\alpha}$ can be any value in $(0,1)$, to construct the robust measure of the upward mobility. For instance, $\underline{\alpha} = 0.20$ implies that a worker is at the bottom of income distribution in location i. Then, her position in the whole economy is given by the denominator in (C.79). Next, consider a worker of generation $t+1$ born in location i. For instance, $\overline{\alpha} = 0.80$ implies that a worker is at the top 80% among her cohort in the economy.

We let $\xi_{i,t}(\underline{\alpha})$ be the $\underline{\alpha}$-th quantile for income distribution in location i among workers of generation t. For generation $t+1$,

$$\begin{aligned}\inf\{y|\mathcal{H}_{i,t+1}(y) \geq \overline{\alpha}\} &= \arg\min_{\nu} \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{K}} \lambda_{ni|s,t+1}\varsigma^{s}_{i,t+1}\Big(\overline{\alpha} - \mathbf{1}(\mu_{n,t+1}w^{s}_{n,t+1} \leq \nu)\Big)\Big(\mu_{n,t+1}w^{s}_{n,t+1} - \nu\Big) \\ &\equiv \widetilde{\xi}_{i,t+1}(\overline{\alpha})\end{aligned} \tag{C.80}$$

Therefore, (C.79) is:

$$\varkappa_{i,t+1} = \frac{\mathcal{Q}_{t+1}\Big(\widetilde{\xi}_{i,t+1}(\overline{\alpha})\Big)}{\mathcal{Q}_{t}\Big(\xi_{i,t}(\underline{\alpha})\Big)} \tag{C.81}$$

for any $0 < \underline{\alpha} << \overline{\alpha} < 1$. where $\xi_{i,t}(\alpha)$ is α-th quantile of income distribution of workers among generation t in location i.

D Appendix: Simulated economy

This section provides a numerical illustration of an equilibrium. The goal is to understand the equilibrium implications discussed above more concretely. To this end, this section considers the simplest spatial economy. Imagine the hypothetical one-dimensional space (i.e., line economy) in which there are discrete locations over unit space. They locate with even geographical intervals. Specifically, there are 250 discrete locations over unit space $[0,1]$. The simulation requires specifications for some fundamental environment and parameters in the model.

D.1 Line economy

Trade Costs and Migration Costs. There are 250 locations on one-dimensional space. The geographical distance between the edge of the economy and the other edge is normalized to one. The geographical distances between any two neighboring locations on the line are exactly the same; therefore, cities are located with even geographical intervals.

There are four sectors: construction, manufacturing, non-tradable service and tradable service. For manufacturing, trade costs between locations i and n are given by $\tau^{\mathrm{M}}(i,n) = \exp\big(\tau^{\mathrm{M}}|i-n|\big)$

where $|i - n|$ is geographical distance between i and n. τ^{M} is trade efficiency of manufacturing and set to $\tau^{\mathrm{M}} = 0.15$. This is similar value to the findings in international trade. The functional form of trade costs for tradable services is the same as manufacturing sector. To eliminate exogenous factors that affect the difference between tradable services and the manufacturing sector, the parameter is set to the same number with manufacturing, $\tau^{\mathrm{S}} = 0.15$. For the nontradable goods, the parameter is set to $\tau^{\mathrm{R}} = 1.0$ that is sufficiently large value compared to other sectors.

Turning to the migration costs, suppose that they are the function of geographical distances and the functional form that is the same as trade costs. Namely, migration costs from i to n is $D_{in,t} = \exp(d|i - n|)$ with parameter value $d = -\ln 0.5$. This implies that the remaining utility for individuals after moving from edge to edge (i.e., unite distance) is 50 percent. The analysis for high migration costs assumes a higher value of d such that $d = -\ln 0.1$.

Productivity and Amenities. There is no variation in amenities across space: $B_{i,t} = 1$ for all i. In contrast, fundamental productivity for manufacturing exhibits differences across locations. The initial productivity of manufacturing in location i is: $A_0(i) = 1.20 - 0.40 \times i$, which implies that location at the left edge of the economy (location index 0) is highest and it is decreasing to the right edge of the economy, and productivity is minimum in the right edge of the economy (location index 1). The highest productivity is 50 percent higher than the lowest productivity. This rationalizes the concentration of manufacture in the early period in the economy and enables to eliminate the potential multiplicity of equilibria. If the initial productivity is a uniform distribution of fundamental productivity, potentially multiple equilibria arise where one sector is concentrated in either the central place or the edges.

In every city, the fundamental productivity of manufacturing grows at a constant rate, 5 percent in the baseline, whereas it becomes 10 percent when there are positive shocks. The fundamental productivity of both nontradable and tradable services is uniform across locations, and they are constant over time. Therefore, change in productivity for services arises through endogenous mechanisms – agglomerations.

Basic Parameters. Except for the parameters in the demand system, same values are assigned to other parameters across sectors to focus on the role of non-homothetic demand across sectors.

We set the followings for the sector specific parameter of Engel slope: $\theta_{\mathrm{M}} = 1.0$ for manufacturing, $\theta_{\mathrm{S}} = 1.5$ for tradable service, $\theta_{\mathrm{R}} = 0.8$ for nontradable service, and $\theta_{\mathrm{H}} = 0.8$ for housing. They are in line with numbers in Comin et al. (2020). Firms in the manufacturing and service sectors only use labor in production. For all sectors, trade elasticity is set to be 6.0. For the spillover in productivity, $\rho = 0$ and $\gamma_s = 0.2$. This implies that there is no productivity spillover across locations, and there is a pure scale economy for all sectors. The number of scale effects (γ_s) is relatively large compared to the agglomeration economies, but recent estimates by Bartelme et al. (2021) show similar values for some manufacturing industries.

The construction sector produces immobile structures (housing) by combining labor and the current stock of structures. The share of labor in production is $\chi = 0.35$ that is the same value we use in the calibration. The residential stock is depreciated in every period and we let $\bar{h}_i = 0.20$ based on Harding et al. (2007).

Parameters in Labor Supply. We use the same values for parameters in labor supply as in calibration. The elasticity of migration to real income is $\varepsilon = 1.5$ base on Fajgelbaum et al. (2019); the elasticity of sector choices to average utility $\phi = 2.5$; and local labor market exposure effect is $\eta = 0.8$.

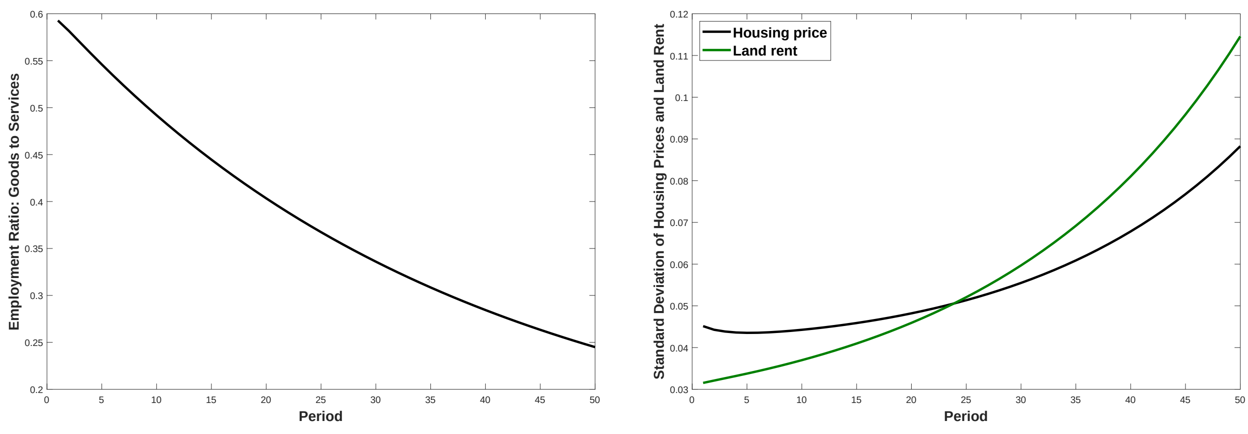
D.2 Solution methods and baseline results

We set out the characterization of the initial equilibrium. Given the initial distribution of productivity, we suppose that the initial equilibrium is in the steady state. Then, we change the fundamental productivity for the manufacturing sector and obtain the forward solution of the economy.

First, we solve the model for the steady state. We guess the steady state level of wage ($w^s_{i,0}$) and employment distribution ($L^s_{i,0}$). The law of motion for productivity ($Z^s_{i,0}$) gives the steady state productivity distribution. We use goods market clearing condition to obtain prices ($p^s_{i,0}$) and trade probabilities. Then, we update wages and employment by using labor market clearing conditions. Next, we follow the discussion in B.2 to obtain the dynamic equilibrium for the rest periods.

We first describe the baseline equilibrium and then proceed to see different scenarios. Figure D.1 displays the structural change and variation in housing prices in the aggregate economy. Panel D.1a confirms the shift of aggregate employment from goods (manufacturing sector and construction sector) to services (non-tradable service and tradable service). The employment ratio of goods to services declines over time, driven by the productivity growth of the manufacturing sector and nonhomothetic demand. Panel D.1b shows the standard deviation of housing prices and land rent. The spatial variation of housing prices increases as the agglomeration of services arises in the right edge of space along with the decline of the manufacturing sector.

Figure D.1: Structural Change and Variation of Housing Prices

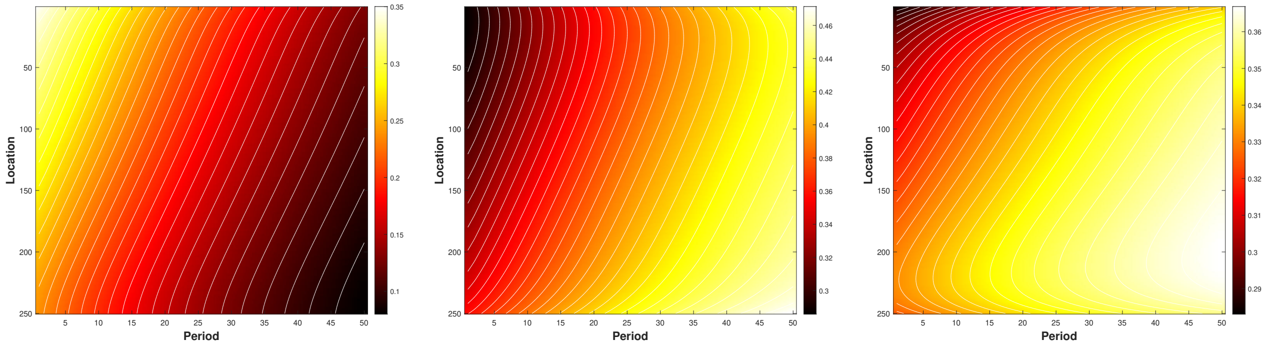


(a) Aggregate Employment Change

(b) Variation of Housing Prices and Land Rent

Figure D.2 shows the distribution of workers for three sectors. The gradation represents the share of employment in each sector for any particular location. The left edge locations keep the manufacturing sector's comparative advantages over time but move to service sectors. On the other hand, the right edge cities show specialization of nontradable services in early stages with shrinking of the manufacturing sector. This leads to demand shift to tradable service sectors due to non-homothetic preference. Therefore, the right cities give rise to tradable services. In the later period, the right-edge cities are specializing tradable services and the right-central cities see the agglomeration of nontradable services. The difference in the location of agglomeration between panel D.2b and D.2c is due to the difference in trade costs.

Figure D.2: Geography of Structural Change



(a) Manufacturing **(b)** Tradable Service **(c)** Nontradable Service

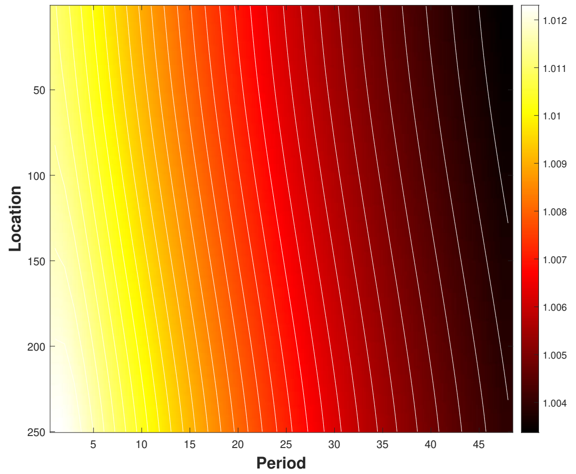
Figure D.3 graphics the welfare change for individuals between two generations. The first panel D.3a shows welfare difference between generations ($\mathrm{d}\ln V_{i,t}$). The overall welfare change for workers is high in the right edge places. The other three panels give the different margins to determine the welfare changes as we discussed in Proposition 3. Panel D.3b shows that the left places exhibit larger gains for workers from migration to the right locations where real income is high, and such gains decline over time as service sectors are dispersed in the later period.

Panel D.3c shows the spatial variation of gains from job opportunities. Individuals in the hot-colored locations benefit from the local labor markets in their choice of the sector. This margin takes an important role in the overall welfare changes in its magnitude. In contrast, individuals in the left-edge cities benefit less. The logic is clear; a higher probability to sort into the manufacturing sector leads to less opportunity to move to other labor markets. The last Panel D.3d shows the gains from trade. The right cities are net exporters of services to the left cities so that the measure of gains from trade exhibits large values in these areas. These three margins are compounded in the overall welfare changes in Panel D.3a.

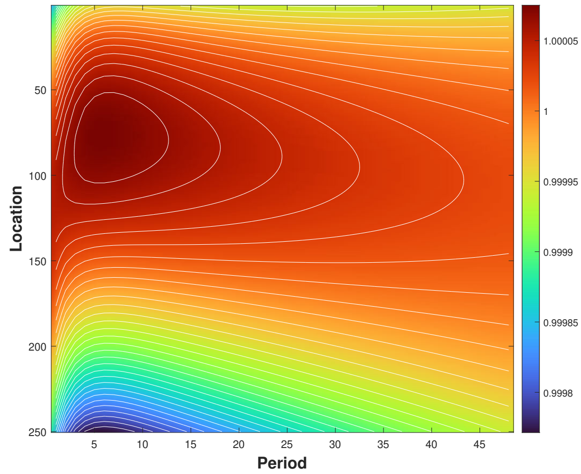
Next, we show inequality and upward mobility in Figure D.4. Panel D.4a confirms the intuition discussed in Appendix C.5. In the initial period, the left-edge cities show a concentration of the manufacturing sector, leading to lower inequality compared to the right places. Conditional on this fundamental pattern of sectoral contribution in income inequality, persistence of sorting and lower mobility keep the income inequality in the right cities high.

We investigate the implication of upward mobility in Panel D.4b where we show the measure

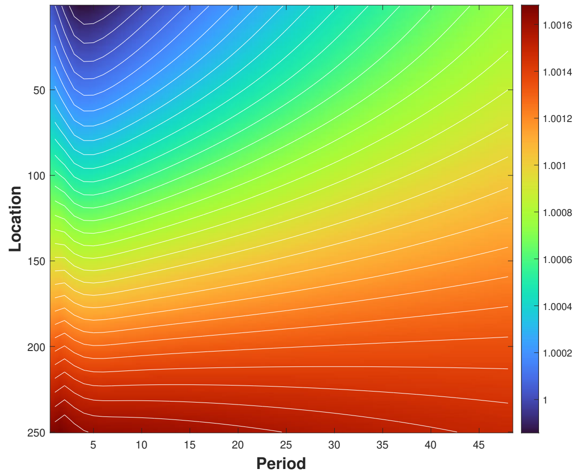
Figure D.3: Welfare Changes



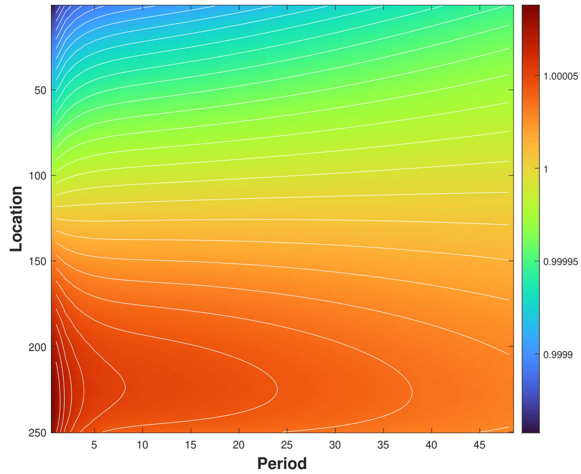
(a) Overall Welfare Change



(b) Migration Gains



(c) Job Opportunity Gains

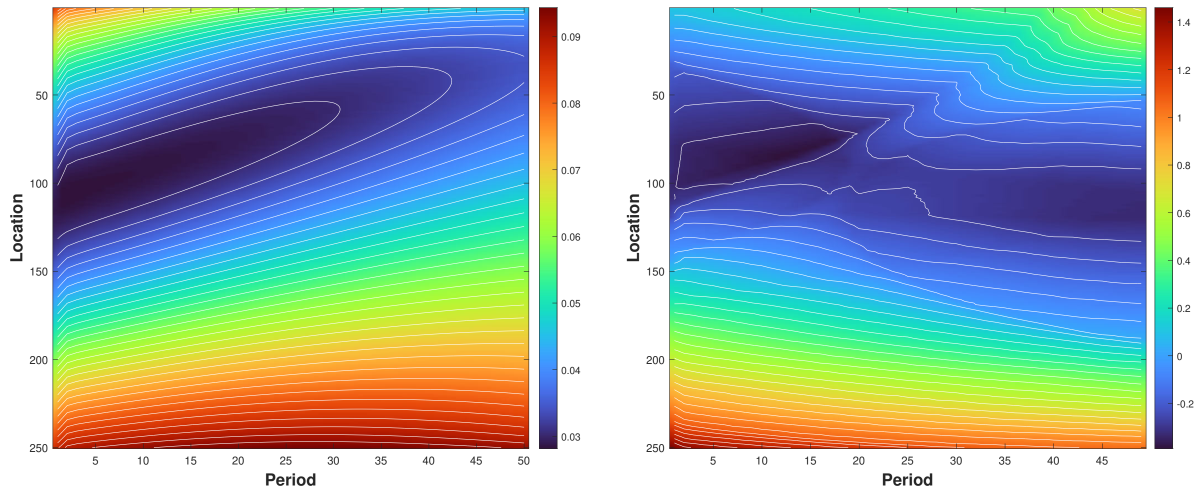


(d) Trade Gains

of the intergenerational mobility proposed in Proposition 5, $\ln \mathcal{M}_{i,t+1}$. We find that there is a huge difference in intergenerational mobility over space. In early periods, workers who originate from the service cities in the right area can climb up the position of income distribution compared to manufacturing cities. They are able to migrate to other cities and sort into the service sector with high likelihood. The central places exhibit the lowest upward mobility over time. The logic behind this is the low degree of dynamics among workers in the central places for both location choice and sector choice. Ultimately, the service cities exhibit lower upward mobility of workers. This is intuitive. In these cities, the service sector grows and more workers sort into both sectors of service. Then, conditional on the job opportunity, the change in the position of the income ladder becomes small. In contrast, the left-edge cities show higher upward mobility in the later period. This is because of the structural change from manufacturing to services and the spread of service sectors to the left cities where individuals from left-edge cities can move into.

Lastly, we show the additional results for the baseline economy. Figure D.5 display change in

Figure D.4: Inequality and Upward Income Mobility

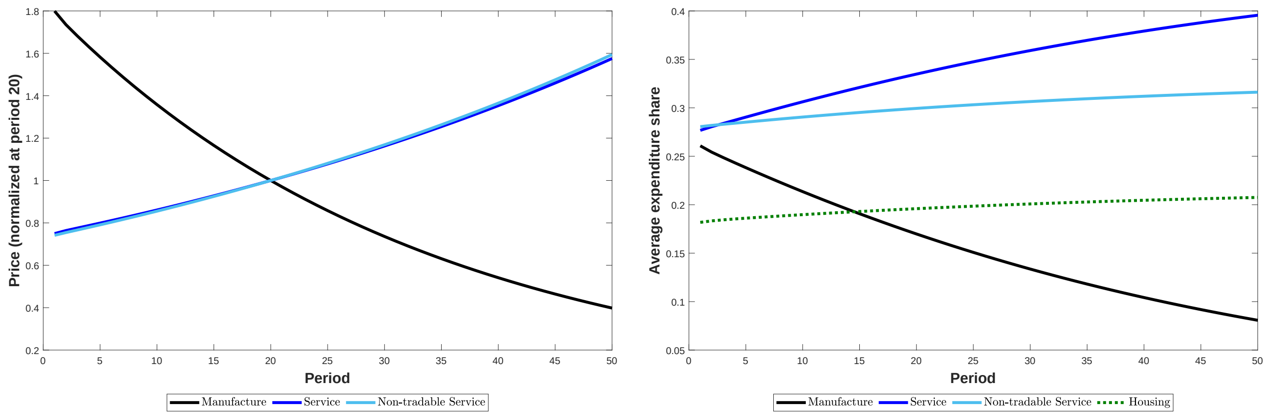


(a) Coefficient of Variation in Income

(b) Upward Mobility: $\ln \mathcal{M}_{i,t+1}$

average prices of manufacturing, tradable service and nontradable service sectors in Panel (a) and change in average expenditure share in the economy in Panel (b).

Figure D.5: Structural Change



(a) Prices

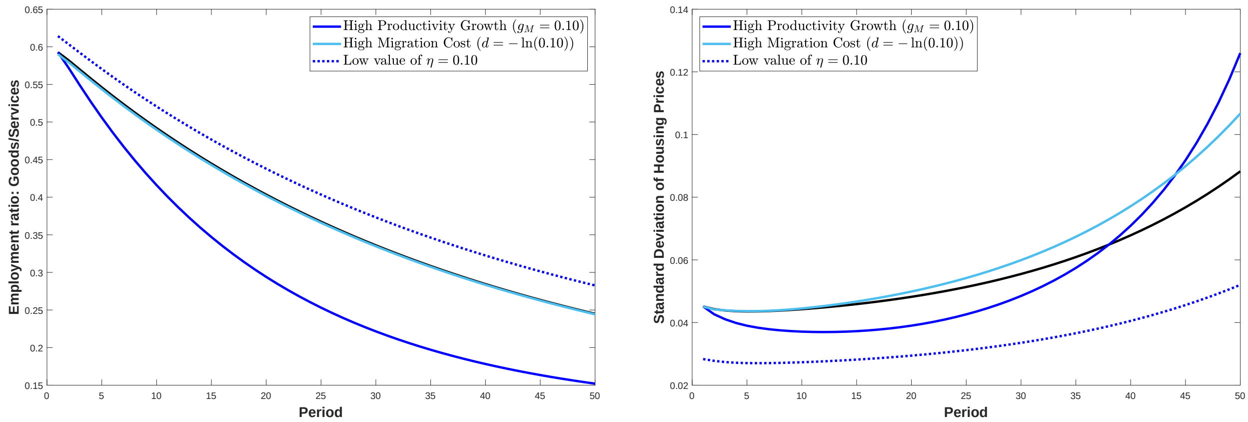
(b) Expenditure Share

These figures confirm the pattern of structural transformation in the aggregate economy. The relative price of service to the manufacturing sector increases over time. This results from the fundamental productivity growth of the manufacturing sector. The right-hand panel (b) shows the results of the non-homothetic demand system. The aggregate expenditure share in the economy shifts from the manufacturing sector to tradable services, while expenditure shares on the housing and non-tradable service sector are stable over time.

D.3 Productivity Shock and Migration Costs

Next, we consider the different two scenarios in the simple economy. First, suppose that there is technological progress in the manufacturing sector. In the baseline, we assume that the fundamental productivity of manufacturing grows at 5 percent each period. We now set 10 percent for the growth rate. Intuitively, this captures the continuous innovation in the manufacturing sector. Second, we consider high migration costs. In the baseline, we set $d = -\ln 0.5$. Now, we set $d = -\ln 0.1$, which implies that only 10 percent of utility remains when individuals migrate from the edge city to the other edge city. For these two scenarios, we consider how cross-sectional inequality and upward mobility are changed.

Figure D.6: Different Scenarios



(a) Aggregate Employment Change

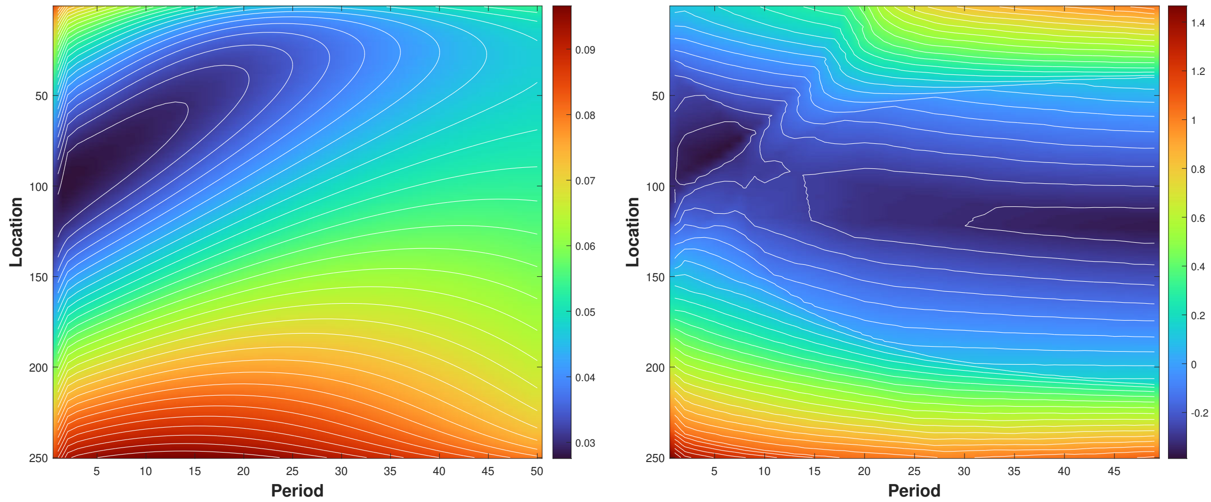
(b) Variation of Housing Prices

In Figure D.6, the left-hand panel D.6a shows the change of employment ratio corresponding to Figure D.1a, and the right-hand panel D.6b shows the variation of housing prices for different scenarios. When the productivity of manufacturing grows at a higher rate, the structural transformation from manufacturing to services proceeds and it generates more variation of housing prices due to the agglomeration of services. For the high migration costs, we can see a similar pattern of structural transformation in the macroeconomy. However, the spatial variation of housing prices becomes large compared to the baseline since high migration costs prevent individuals from adjusting their locations and agglomerations are reinforced.

We turn to inequality and upward mobility in Figure D.7.

The left-hand panel D.7a shows an increase in income inequality in the left cities due to the rapid structural transformation compared to the baseline. In contrast, the right cities show small income inequality due to the further specialization of tradable services than the baseline. The right-hand panel D.8a gives the spatial variation of upward mobility corresponding to Figure D.4b. Comparing these two figures, we find that central places exhibit lower upward mobility from the early period. The structural change due to the technological progress of manufacturing leads to specialization of workers in the edge cities and worse off individuals in the central cities in terms of mobility.

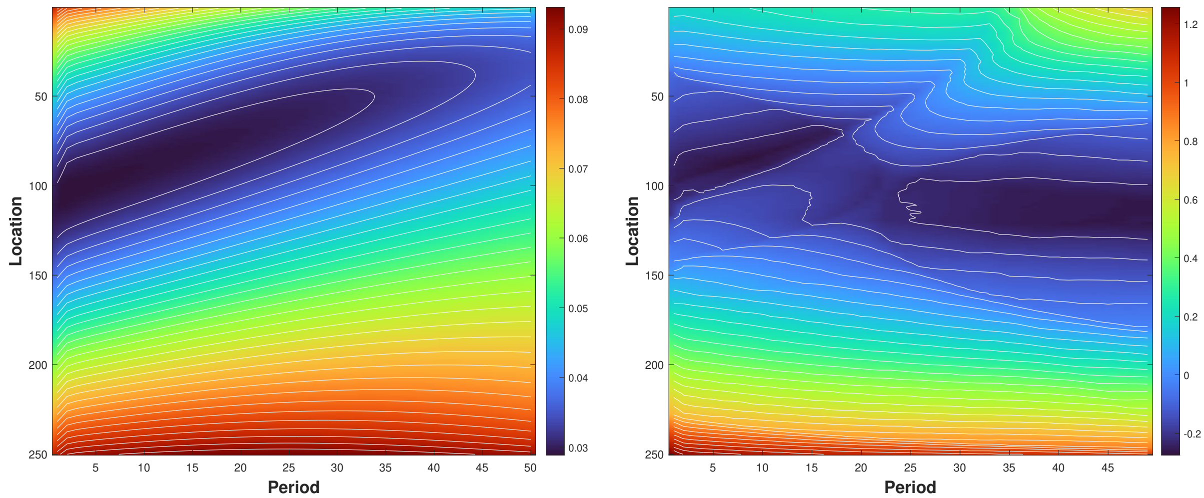
Figure D.7: Inequality and Upward Income Mobility When Productivity Growth of Manufacturing



(a) Coefficient of Variation in Income **(b)** Upward Mobility

Intuitively, this suggests the role of technology-driven structural transformation in the declining upward mobility of workers.

Figure D.8: Inequality and Upward Income Mobility When High Migration Cost



(a) Coefficient of Variation in Income **(b)** Upward Mobility

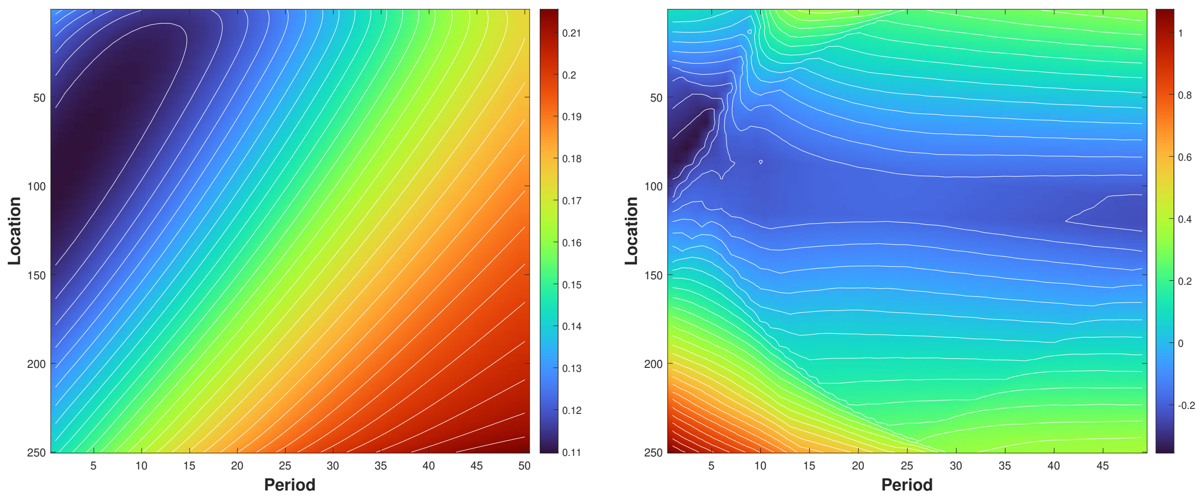
Next, Figure D.8 give these patterns for the case of high migration costs. The left-hand panel D.8a shows a similar pattern of inequality in the local labor market to the baseline results. This implies that the bilateral migration costs have a limited impact on inequality within the city. Nevertheless, the right-hand panel D.8b shows that upward mobility is small for most locations relative to baseline when migration is costly. With high migration frictions, workers are unable to leverage geographical mobility, and therefore, workers are less able to climb up the income ladder by moving across cities.

D.4 Role of Local Labor Market Exposure

Lastly, we exploit the simple economy to understand the role of local labor market exposure in the sector choice of workers by setting a lower value of η.

The parameter controls the limited exposure of individuals to the local labor market as we discussed in Assumption 2. As a contrast to the baseline value of η, we set $\eta = 0.10$. This implies that the effect of the previous generation in the local labor market has less impact on the choice of the sector. Figure D.9 give inequality and upward income mobility for the alternative parameter value.

Figure D.9: Inequality and Upward Income Mobility When $\eta = 0.10$



(a) Coefficient of Variation in Income

(b) Upward Mobility

First, we see significant income inequality compared to the baseline. All locations show an increase in income inequality over time. When the persistent effect of local labor market exposure is weak, less specificity to sector and location fosters the mobility of workers and it leads to less specialization of workers in equilibrium. Therefore, we see a rise in income inequality within a city and less variation of housing prices (Panel D.6b).

In the right-hand panel D.9b, we see the relatively low intergenerational income mobility. The less specialization and structural transformation lead to a small variation of real income across locations and a small gap in expected return from sector choice between the tradable service sector and non-tradable service sector. Together, the intergenerational income mobility becomes low and shows small geographical variation. Overall, the local labor market exposure impacts inequality and upward mobility by the direct effect in the sector choice of workers and an indirect effect through the specialization of workers in local labor markets.

E Appendix: Quantification of the model

Here, we explain the procedure of calibration for the quantitative analysis. We start with the description of data in subsection D where we define the economy. We explain the calibration of the parameters in subsection E.2. Using the data and parameters, we discuss how to use the model structure to obtain fundamentals in our model in subsection E.3. We first explain how to invert the model for steady state and then extend it to the inversion of dynamics in fundamentals: fundamental productivity and amenities over time.

E.1 Data

Cities. We focus on U.S. Core Based Statistical Areas (CBSAs). We use the definition of CBSAs based on Census 2010. Each CBSA consists of a unique county or multiple counties anchored by an urban center of at least 10,000 population and adjacent counties. We use 395 CBSAs in the calibration, where we are able to compute wages and employment throughout time. The list of 395 CBSAs includes all metropolitan areas in the U.S., excluding Alaska and Hawaii and some large micropolitan areas. Since the definition of CBSAs is based on the social and economic linkages between counties and commuting, we take them as our units of cities in the U.S. economy. Throughout time, we fix the definition of CBSAs to exclude the potential problem arising from the change of geographical size that is outside our model.

Industries. We consider the fixed set of industries throughout time. The economy consists of three different groups of industries. We let $\mathcal{K}^{\mathrm{M}}$ refer to the set of manufacturing industries, $\mathcal{K}^{\mathrm{S}}$ refer to the set of service industries, and $\mathcal{K}^{0}$ refer to the single sector related to the immobile structure in the model. We use a crosswalk between industry codes to define each industry for different years based on 4 digit SIC 87. We assign industries to each group as follows. The group of manufacturing sector $\mathcal{K}^{\mathrm{M}}$ consists of: Food, beverage, and tobacco product (4 digit: 2000 to 2141); Textile, textile product mills, apparel, leather, and allied product (4 digit: 2200 to 2399); Wood product, paper, printing, and related support activities (4 digit: 2400 to 2796); Chemical, petroleum, rubber and coal products, and nonmetallic mineral product (4 digit: 2800 to 3299); Metal and fabricated metal product (4 digit: 3300 to 3499); Machinery (4 digit: 3500 to 3599); Computer and electronic product, and Electrical equipment and appliance (4 digit: 3600 to 3699); Transportation equipment (4 digit: 3700 to 3799); Furniture and related product, and Miscellaneous manufacturing (4 digit: 3800 to 3999).

The group of industries $\mathcal{K}^{\mathrm{S}}$ consists of: Transport services and storage (4 digit: 4000 to 4789); Wholesale trade (4 digit: 5000 to 5199); Retail (4 digit: 5200 to 5999); Finance, insurance and real estate (4 digit: 6000 to 6799); Health service and social services (4 digit: 8000 to 8099, 8300 to 8399); Legal service and education service (4 digit: 8100 to 8299); Communication service (4 digit: 4800 to 4971); Other local services (4 digit: 7000 to 7999, 8400 to 8811).

The construction sector includes 4 digit: 1500 to 1799. We do not include agriculture, forestry, fishing (4 digit: 0100 to 0971), mining (4 digit: 1000 to 1499) and the rest (4 digit over 9000) in our analysis since these sectors show a small share of employment in the period we analyze.

Wages and employment. Wages and employment are essential to calibrate the model. We construct wages and employment by industry and CBSA. Our data source for employment is the County

Business Pattern (CBP) in 1980, 1990, 2000 and 2010. Following procedures to impute employment counts by county and 4-digit SIC 87 industry in David et al. (2013) and using the methodology in Acemoglu et al. (2016), we imputed employment for each county. After the imputation, we aggregate them at the CBSA level to define industry employment.

For wages, we use the American Community Survey (ACS) and decennial censuses. The datasets are downloaded from IPUMS using standardized variables. For the years 1980, 1990 and 2000, we exploit a 5 sample of the respective censuses. For the year 2010, we are based on ACS data. Within each CBSA, we compute the log of average wages across counties for each industry. Wages are inflated to the year 2010 using the Personal Consumption Expenditure Index. Through this process, we can obtain 395 CBSAs that we focus on.

E.2 Parameters

We first set parameters in the nonhomothetic demand system based on the literature. Then, we calibrate the set of parameters in technology by using the information on input-output linkages. Then, we exploit the gravity equation of trade to estimate trade elasticity for manufacturing sectors, while for the service sector, we use the values from the literature. Then, we use the parameters in labor mobility across locations and sectors by exploiting the equilibrium relationships.

Demand parameters

We set the elasticity of substitution $\sigma = 0.4$ as a baseline value. This is consistent with the traditional values in the macroeconomic literature on structural transformation. We have non-homothetic demand system, and it has two sets of parameters across different sectors. Namely, we have:

$$\sum_j \alpha_j^{(\sigma-1)/\sigma} \left(\frac{c_j}{\mathbb{U}^{(\theta_j - \sigma)/(1-\sigma)}} \right)^{(\sigma-1)/\sigma} = 1 \tag{E.1}$$

The parameter θ_j defines the sector-specific Engel slope and α_j is the shift of expenditure. Following Comin et al. (2020), we normalize $\theta_j = 1$ if sector j is manufacturing sector. For the service sectors and residenti, we set: $(\theta_j - \sigma)/(1 - \sigma) = 1.75$ which is the middle in the range of estimates in Table I in Comin et al. (2020). This implies that $\theta_j = 1.375$ for service sectors. For the consumption of residential stocks, we set $\theta_0 = 1$. This implies that we captures the demand driven structural transformation at the aggregated level between manufacturing, service, and housing by setting three different parameters.

Turning to the scale parameters (α_j), we also consider three different parameters for each aggregation. We set $\alpha_j = 3.0$ for manufacturing. For the other two aggregated sectors, we match the parameter α_j such that the average expenditure share is matched to the aggregate expenditure share. These parameters of the shifter are constant over time.

Technology parameters

The BEA table allows us to specify the following identity for input-output in each sector at period t (year):

$$Y_t^s = \sum_j Q_t^{sj} + IM_t^s + LW_t^s + VS_t^s + TX_t^s \tag{E.2}$$

where Y_t^s is value of output of sector s, Q_t^{sj} is value of input purchases from other sectors j, IM_t^s is value of import in intermediate goods for sector s, LW_t^s is value of labor compensation of sector s, VS_t^s is value of gross operating surplus in sector s, and TX_t^s is value of taxes on production in sector s. We use this identity to set the parameters in production technology. Before adjustment, we map our definition of industries to NAICS codes.

First, we adjust the identity by subtracting the import of intermediate goods since the baseline model does not consider the international trade of intermediate goods. We subtract IM_t^s from both input purchases and output. Next, we consider the term, VS_t^s. Following Caliendo et al. (2018), we decompose this term into two parts. We deduct 17 percent of value added of sector s from the operating surplus, adding this to the material purchase by splitting across different sectors proportional to its input share. Then, the rest of the surplus is assumed to be exploited by the landlords. The value of the remaining surplus is added to the real estate sector and other local services proportional to their ratio for each sector. When the 17 percent of value added is greater than the operating surplus, we deduct the total surplus from the right-hand side of (E.2) and put it in the real estate sector and other local services proportional to their ratio. Given this, we compute the labor share and other input shares including input-output linkages in the straightforward way:

$$\beta_s = \frac{LW_t^s}{Y_t^s - TX_t^s}, \quad \beta_{sj} = \frac{Q_t^{sj}}{Y_t^s - TX_t^s} \tag{E.3}$$

We average them over five years, 2011-15. For the production technology of residential stocks, we set $\chi = 0.35$ based on labor compensation in the construction sector.

Gravity of trade

The regional trade in the model takes a form of gravity equation. We assume that the regional trade costs are given by:

$$\ln \tau_{in,t}^s = \bar{\delta} + \delta_s \ln \text{dist}_{in} + \varepsilon_{in,t}^s(\tau) \tag{E.4}$$

where dist_{in} is geographical distance between n and i in kilometers and $\varepsilon_{in,t}^s(\tau)$ is other factors that are orthogonal to other characteristics. Given this, trade pattern (A.27) leads to:

$$\ln X_{in,t}^s = \mathbb{D}_{i,t}^s + \mathbb{O}_{n,t}^s - \kappa_s \delta_s \ln dist_{in} + \varepsilon_{in,t}^s \tag{E.5}$$

with

$$\mathbb{D}_{i,t}^s = \kappa_s \ln p_{i,t}^s + \kappa_s \ln \Gamma_s \tag{E.6}$$

and

$$\mathbb{O}^s_{n,t} = -\kappa_s \ln \Xi^s_{n,t} + \kappa_s \ln Z^s_{n,t} \tag{E.7}$$

where $\mathbb{D}^s_{i,t}$ factors destination characteristics and $\mathbb{O}^s_{n,t}$ factors origin characteristics, respectively. Therefore, the coefficient estimated in the restricted gravity (E.5) gives an information about $\kappa_s\delta_s$ that is a composite of Fréchet shape parameter ($\kappa_s > 1$) and industry specific parameter for in trade costs, δ_s.

For commodities, we estimate (E.5) for $\kappa_s\delta_s$ using U.S. Commodity Flow Survey in 2012. As we cannot back to the past, we only use the cross sectional data. Once we obtain the estimate $\widehat{\kappa_s\delta_s}$, we compute Fréchet shape parameter κ_s given δ_s. We assign value of δ_s for commodities based on literature. Ramondo et al. (2016) proposed the value of trade cost elasticity with respect to distance, 0.27, for international trade. This value is close to the estimates in Hummels (2001). Further, Eaton and Kortum (2002) use the relationship between international trade and prices to estimate the Fréchet shape parameter 8.28 and coefficient of gravity equation such that 1.10. This implies the trade cost elasticity is around 0.13. Our analysis is the domestic trade, therefore we use the lower value of the cost elasticity $\delta_s = 0.125 = 1/8$ for all sectors in both manufacturing and services.

For the service sectors, we do not direct observation of bilateral trade values. Therefore, we rely on estimates by Anderson et al. (2014). For non-tradables (i.e., retail), we set ∞ for the trade cost elasticity and Fréchet parameter is set to be 5.0 that is in around the middle of estimates for trade elasticities. Table E.1 summarizes numbers. The values of trade elasticity are in the range of estimates from the trade literature (Head and Mayer 2014, Simonovska and Waugh 2014). In addition, manufacturing shows a larger value of elasticity relative to services except for health and education services, which are consistent with findings in Gervais and Jensen (2019).

Labor Mobility

In the model, the conditional probability that people migrate from location n to i conditional on the industry choice s is given by:

$$\lambda_{in|s,t} = (D_{in,t})^{-\varepsilon}(D_{nn,t})^{\varepsilon}(\mathcal{W}_{is,t})^{\varepsilon}(\mathcal{W}_{ns,t})^{-\varepsilon}\lambda_{nn|s,t} \tag{E.8}$$

Therefore, mass of workers who move from n to i is:

$$L^s_{in,t} = (D_{in,t})^{-\varepsilon}(\mathcal{W}_{is,t})^{\varepsilon}(\mathcal{W}_{ns,t})^{-\varepsilon}\mathcal{E}^s_{nn,t}L_{n,t-1} \tag{E.9}$$

Then, we obtain:

$$\ln L^s_{in,t} - \ln L^s_{nn,t} = -\varepsilon \ln D_{in,t} + \varepsilon(\ln B_{i,t} - \ln B_{n,t}) + \varepsilon\left(\ln \frac{W^s_{i,t}}{\mathcal{P}^s_{i,t}} - \ln \frac{W^s_{n,t}}{\mathcal{P}^s_{n,t}}\right) \tag{E.10}$$

Table E.1: Gravity coefficients estimated for commodities

(1) Industry	(2) $\kappa_s\delta_s$ (GC)	(3) $\kappa_s\delta_s$ (Route)	(4) κ_s	(5) Source
1. Food/Beverage/Tobacco	.990	.996	7.92	CFS 2012
2. Textile/Apparel	.824	.834	6.59	CFS 2012
3. Wood/Paper/Printing	1.13	1.134	9.04	CFS 2012
4. Chemical/Petro/Coal/ Nonmetalic	1.035	1.04	8.28	CFS 2012
5. Metal	1.029	1.036	8.23	CFS 2012
6. Machinery	.803	.812	6.42	CFS 2012
7. Electric/Computer	.626	.638	5.00	CFS 2012
8. Transport Equipment	.961	.966	7.68	CFS 2012
9. Miscellaneous Manufacture	.816	.828	6.53	CFS 2012
10. Transportation Service	.617	.617	4.94	Anderson et al. (2014)
11. Wholesale Trade	1.379	1.391	11.03	CFS 2012
12. Retail	∞	∞	5.0	–
13. FIRE	.678	.678	5.42	Anderson et al. (2014)
14. Health Service	1.42	1.42	11.36	Anderson et al. (2014)
15. Education and Legal	1.01	1.01	8.08	Anderson et al. (2014)
16. Communication Service	.297	.297	2.38	Anderson et al. (2014)
17. Other Services	.724	.724	5.79	Anderson et al. (2014)

Note: This table reports the estimated gravity coefficients and inferred trade elasticities for relevant industries. Column (2) uses the great circle distance for distance, and column (3) uses the route distance. In column (4), we compute trade elasticities based on estimates in column (2).

Therefore, for the small difference of real income, difference of labor mobility becomes:

$$\varepsilon = \frac{l^s_{in,t} - l^s_{nn,t}}{(\tilde{w}^s_{i,t} - \tilde{p}^s_{i,t}) - (\tilde{w}^s_{n,t} - \tilde{p}^s_{n,t})} \tag{E.11}$$

where $l^s_{in,t}$, $\tilde{w}^s_{i,t}$ and $\tilde{p}^s_{i,t}$ are log of corresponding variables. Hence, ε reflects the elasticity of local labor supply across different locations to the real income, and we set $\varepsilon = 1.5$ that lies in the middle of the estimates in Fajgelbaum et al. (2019) for the U.S. economy.

Now, we consider the friction in labor mobility. We cannot identify the spatial friction of labor mobility, $D_{in,t}$, directly in the data. Therefore, we derive the spatial friction in the process of model inversion, which we discuss later. To this end, we assume that the friction takes the following form:

$$D_{in,t} = (\text{dist}_{in})^{\widetilde{\delta}} M_{i,t}, \quad \forall i \neq n \tag{E.12}$$

where $\widetilde{\delta}$ is positive constant and $M_{i,t}$ is positive value that explains the migration barrier for workers who choose i. (E.9) implies that mass of workers moving from n to i is:

$$L^s_{in,t} = (\text{dist}_{in})^{-\varepsilon\widetilde{\delta}} (M_{i,t})^{-\varepsilon} \left(\frac{\varsigma^s_{n,t}}{\bar{U}^s_{n,t}} \right)^{\varepsilon} (\mathcal{W}^s_{i,t})^{\varepsilon} L_{n,t-1} \tag{E.13}$$

Taking logs, (E.13) can be written as:

$$\ln L^s_{in,t} = \mathbb{W}^s_{i,t} - \varepsilon\widetilde{\delta} \ln \text{dist}_{in} + \mathbb{H}^s_{n,t} \tag{E.14}$$

where

$$\mathbb{W}^s_{i,t} \equiv \varepsilon \ln \mathcal{W}^s_{i,t} - \varepsilon \ln M_{i,t} \tag{E.15}$$

and

$$\mathbb{H}^s_{n,t} \equiv \varepsilon \ln \varsigma^s_{n,t} - \varepsilon \ln \bar{U}^s_{n,t} + \ln L_{n,t-1} \tag{E.16}$$

contain source location and industry characteristics, and destination and industry characteristics, respectively.

We use (E.14) to obtain estimated value of $\varepsilon\widetilde{\delta}$. To this end, we need information on labor mobility between locations for workers in different sectors. We use ACS 5 year sample data between 2006-2010 and 2011-2015. The ACS data allows us to know the current location (county), previous location (county), and industry of workers in the sample. We extract workers in our sectors and map their locations to the CBSA level. Then, we focus on workers who moved between different CBSAs during the sample periods and compute average distances at the aggregation of state level. Therefore, the final data contains the number of workers in each sector who move from state to state and the average distance of the mobility pattern. We construct the data for 5-year period 2006-10, 2011-15 and 10-year period, 2006-2015. Using the data, we estimate (E.14) by ordinary least squares (OLS). We replace $\mathbb{W}^s_{i,t}$ and $\mathbb{H}^s_{n,t}$ by origin-sector indicators and destination-sector indicators, respectively.

Table E.2 shows the estimates of $\varepsilon\widetilde{\delta}$. The estimates are similar to the findings for intra-national migration elasticity to distance in Bryan and Morten (2019) for Indonesia. Compared to Allen and Arkolakis (2018) for migration cost in U.S. history, estimates are small. This difference arises from the different periods of our data. For the old period, it would be large because of the higher moving cost per unit of distance. Based on the results, our preferable values for migration elasticity is 0.75, and therefore we set $\widetilde{\delta} = 0.50$ in our analysis.

Table E.2: Coefficients estimated for workers mobility

	(1) Year 2006-10	(2) Year 2011-15	(3) Year 2006-15
ln dist	-.743	-.728	-.806
	(0.0296)	(0.0317)	(0.0282)
Observations	11,292	11,374	14,852

Note: Robust standard errors in parenthesis.

We consider the parameters in labor supply, η and ϕ. We let $\mathbb{U}^s_{n,t+1} \equiv \zeta_s^{1/\phi} \bar{U}^s_{n,t+1}$. Then, we use the structural equations in our model. As we discussed in Appendix B.1 and using migration costs

(E.12), our model derives:

$$\mathbb{U}^s_{i,t+1} = \left(\sum_{\ell \in \mathcal{N}} \left((\text{dist}_{\ell i})^{-\varepsilon\widetilde{\delta}} \frac{L^s_{\ell,t+1}}{\sum_{n\in\mathcal{N}} \left(\text{dist}_{\ell n}^{-\varepsilon\widetilde{\delta}} \frac{\left(L^s_{nt}\right)^{\eta}\left(\mathbb{U}^s_{n,t+1}\right)^{\phi}}{\sum_{k\in\mathcal{K}}\left(L^k_{nt}\right)^{\eta}\left(\mathbb{U}^k_{n,t+1}\right)^{\phi}} \left(\mathbb{U}^s_{n,t+1}\right)^{-\varepsilon} L_{n,t} \right)} \right) \right)^{1/\varepsilon} \tag{E.17}$$

for every s and i. This is labor mobility clearing condition. (E.17) gives the relationship between employment $\{L^s_{i,t+1}\}$, $\{L^s_{i,t}\}$, geographical distance and $\{\mathbb{U}^s_{i,t+1}\}$. Then, we solve $N \times (S+1)$ system of equations (E.17) for $\{\mathbb{U}^s_{i,t+1}\}$.

First, we use the above parameters (ε, $\widetilde{\delta}$) and guess two key parameters (ϕ, η). Implementing the observation of employment $\left(\{L^s_{i,t+1}\}, \{L^s_{i,t}\}\right)$, we solve $N \times (S+1)$ equations for $\{\mathbb{U}^s_{i,t+1}\}$ as a fixed point of (E.17). We denote this as $\widehat{\mathbb{U}}^s_{i,t+1}$. Then, we compute the inferred probability in industry choice,

$$\widehat{\varsigma}^s_{n,t+1} = \frac{\left(L^s_{n,t}\right)^{\eta}\left(\widehat{\mathbb{U}}^s_{n,t+1}\right)^{\phi}}{\sum_k \left(L^k_{n,t}\right)^{\eta}\left(\widehat{\mathbb{U}}^k_{n,t+1}\right)^{\phi}} \tag{E.18}$$

where we use employment in observation, $L^s_{n,t}$. Further, by construction, we have:

$$\bar{U}^s_{i,t+1} = \left(\sum_{\ell} \text{dist}_{\ell i}^{-\varepsilon\tilde{\delta}} \left(\mathcal{W}^s_{\ell,t+1}/M_\ell\right)^{\varepsilon} \right)^{1/\varepsilon} \tag{E.19}$$

Therefore, we can write:

$$\mathbb{U}^s_{i,t+1} = \left(\sum_{\ell} \text{dist}_{\ell i}^{-\varepsilon\tilde{\delta}} \left(\widetilde{T}^s_{\ell,t}\right)^{\varepsilon} \right)^{1/\varepsilon} \tag{E.20}$$

where we let

$$\widetilde{T}^s_{\ell,t} \equiv \frac{\mathcal{W}^s_{\ell,t}}{\zeta_s^{1/\phi} M_\ell}$$

We define:

$$\widetilde{\mathbb{U}}_{t+1} \equiv \left\{ \left(\mathbb{U}^s_{i,t+1}\right)^{\varepsilon} \right\}, \quad \mathbb{D} \equiv \left\{ \text{dist}_{\ell i}^{-\varepsilon\widetilde{\delta}} \right\}, \quad \mathbb{T}_{t+1} \equiv \left\{ \left(\widetilde{T}^s_{\ell,t}\right)^{\varepsilon} \right\} \tag{E.21}$$

$\widetilde{\mathbb{U}}_{t+1}$ is $N \times (S+1)$ matrix of average utility conditional on industry choice, $\mathbb{D}$ is $N \times N$ matrix of distances, and $\mathbb{T}_{t+1}$ is $N \times (S+1)$ matrix of adjusted real income by migration cost. (E.20) implies that we can compute $\widehat{\widetilde{T}}^s_{\ell,t}$ by using these matrix:

$$\widehat{\mathbb{T}}_{t+1} = \left(\mathbb{D}^{\top}\right)^{-1} \widehat{\widetilde{\mathbb{U}}}_{t+1} \tag{E.22}$$

where $\widehat{\widetilde{\mathbb{U}}}_{t+1}$ is matrix $\widetilde{\mathbb{U}}_{t+1}$ by substituting $\{\widehat{\mathbb{U}}^s_{i,t+1}\}$. After we compute this, we derive the model inferred location choice probabilities:

$$\widehat{\lambda}_{in|s,t+1} = \left(\text{dist}_{in}^{-\widetilde{\delta}} \times \frac{\widehat{\widetilde{T}}^{s}_{i,t+1}}{\widehat{\mathbb{U}}^{s}_{n,t+1}}\right)^{\varepsilon} \tag{E.23}$$

Combining (E.18) and (E.23), we obtain the mobility of workers between two locations during period t and $t+1$:

$$\widehat{L}_{in,t+1} = \sum_{s} \widehat{\lambda}_{in|s,t+1} \widehat{\varsigma}^{s}_{n,t+1} L_{n,t} \tag{E.24}$$

This is the labor mobility from n to i for any particular generation $t+1$ predicted in the model. Using (E.24), we compute $\widehat{\vartheta}_{in,t+1} = \widehat{L}_{in,t+1} / \sum_{\ell \neq i} \widehat{L}_{\ell n,t+1}$ for $i \neq n$, which is the predicted pattern of migration from location n in the model.

In turn, we exploit IRS county-to-county migration data and aggregate the flow of people to the CBSA pairs which we use. We process this for two time periods, 1990-2000 and 2000-2010. The period 1990-2000 corresponds to the movement of workers in generation 2000, while the period 2000-2010 corresponds to the movement of workers in generation 2010. We compute $\vartheta_{in,t+1} = L_{in,t+1} / \sum_{\ell \neq n} L_{\ell n,t+1}$ where $L_{in,t+1}$ is the migration from source n to destination i. This is the pattern of labor mobility given the source location.

Then, we argue that the pattern of emigration in the data is equal to the pattern predicted in the model. Namely, we consider the following moment condition:

$$\mathbb{E}\left[\left(\vartheta_{in,t+1} - \widehat{\vartheta}_{in,t+1}\right) \times \mathbf{I}_{in}\right] = 0$$

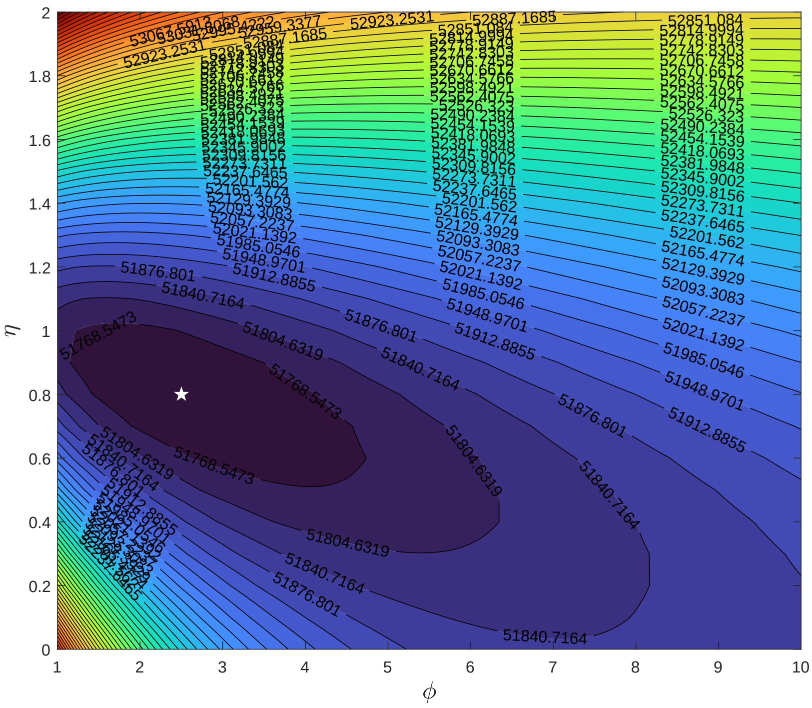
The underlying assumption for this is that any errors between the observed pattern of migration and the migration pattern predicted in the model are unrelated to the level of distances within the same range of distances. In particular, we define 6 ranges of distances between two locations, and we use the moment condition. The figure E.1 shows the moment conditions based on the sum of the squared difference in the migration patterns across the distance ranges. We construct contours for the 1800 combinations of values of (ϕ, η) in the parameter space (shown on the horizontal and vertical axes). The figure shows a unique global minimum in the parameter space. We use the parameter $\phi = 2.50$ and $\eta = 0.80$ in our analysis.

Economies of scale

The parameters for economies of scale, $\boldsymbol{\gamma} = \{\gamma_s\}$, are fixed over time. The estimation of these parameters is not straightforward as it is impossible to decompose the overall productivity $\{Z^{s}_{i,t}\}$ into the contribution of the fundamentals $\{A^{s}_{i,t}\}$ and the endogenous factors. Instead of using the empirical approach, we use the model restriction to pin down the parameters, $\{\gamma_s\}$.

As we discussed in Appendix B.1, the sufficient condition for parameters to characterize unique

Figure E.1: Moment Condition for $\|\widetilde{\vartheta}_{in,t+1} - \widehat{\vartheta}_{in,t+1}\|$



dynamic equilibrium when $\rho = 0$ and $\chi = 1$ is given by (B.28). We state the condition again here:

$$\gamma_s \leq \frac{\theta_s - \sigma}{\kappa_s + (1-\sigma)} \left(1 + \frac{1}{\varepsilon}\right) \tag{E.25}$$

This condition gives the upper bound of the parameters, $\boldsymbol{\gamma}$. In our model, small parameter ε leads to dispersion of workers in their location choice, which turns out to be the dispersion force of workers. For the trade elasticity, a small parameter of κ_s implies that production is more dispersed across locations as idiosyncratic productivity shocks have a large variation. On the demand side, a large value of θ_s means that expenditure share expands more as real income grows. Then, we have more heterogeneity in consumption patterns across workers with different incomes. Therefore, this leads to the dispersion of workers. If we introduce the partially inelastic supply of structure, $\chi < 1$, it also leads to additional dispersion force.

To understand the condition (E.25) more concretely, we think about the extreme cases. When $\kappa_s \to \infty$, the right-hand-side of (E.25) becomes zero, and therefore we need no economies of scale, $\gamma_s = 0$. In words, when there is no variation in idiosyncratic shocks in productivity, all workers are located in the fundamentally productive places, and economies of scale lead to multiplicity of equilibrium. When $\varepsilon \to \infty$, the right-hand-side of (E.25) becomes strictly small. If idiosyncratic taste shocks are homogeneous, weak agglomeration forces are required to avoid the multiple equilibria. Instead, suppose that $\theta_s = 1$, $\gamma_s = 1/\kappa_s$ and $\varepsilon \to \infty$. As we described in Appendix A.2, the parameter $\{\gamma_s\}$ is the inverse of the trade elasticity for the conventional models of trade: Armington model, NEG model and Melitz model with Pareto. $\varepsilon \to \infty$ implies that workers' taste shocks are homogeneous as in the conventional models of Ricardian trade and NEG. Then, the condition (E.25) is reduced to be:

$$\kappa_s \leq \frac{1-\sigma}{\sigma}$$

Therefore, trade elasticity must be sufficiently small, and the elasticity of substitution is relatively small to avoid the black-hole equilibria.

In our framework, we allow $\rho \neq 0$. If we introduce the spillover in productivity ($\rho > 0$), it leads to agglomeration forces in our model since favorable locations attract workers while the remoted places lose as they do not benefit from the inflow of workers. Therefore, we take the conservative values that satisfy the condition (E.25) when $\varepsilon \to \infty$. This implies that the equilibrium is unique when idiosyncratic shocks for location choice are homogeneous. Namely, we set γ_s such that:

$$\gamma_s = \frac{\theta_s - \sigma}{\kappa_s + (1-\sigma)} \tag{E.26}$$

Table E.3 gives parameters across different sectors (E.26).

Table E.3: Parameters $\{\gamma_s\}$

(1) Industry	(2) γ_s	(3) Combes et al. (2012)	(4) Bartelme et al. (2021)
1. Food/Beverage/Tobacco	.070	.064	.22
2. Textile/Apparel	.083	.040	.12
3. Wood/Paper/Printing	.062	.084	.13
4. Chemical/Petro/Coal/ Nonmetalic	.068	.073	.15
5. Metal	.068	.069	.09
6. Machinery	.085	.083	.24
7. Electric/Computer	.107	.076	.08
8. Transport Equipment	.072	.086	.18
9. Miscellaneous Manufacture	.084	–	–
10. Transportation Service	.176	–	–
11. Wholesale Trade	.084	–	–
12. Retail	.174	–	–
13. FIRE	.162	–	–
14. Health Service	.082	–	–
15. Education and Legal	.112	–	–
16. Communication Service	.328	–	–
17. Other Services	.153	–	–

In the table, column (2) reports the parameter values of $\{\gamma_s\}$ that are given as a upper bound of the condition (E.26). Column (3) reports the values of $\{\gamma_s\}$ in Table I from Combes et al. (2012). Column (4) reports estimated values in Bartelme et al. (2021) for manufacturing sector as a reference. The values in column (2) are close to those in column (3) while lower than those in column (4). This reflects the spatial scale of the economy. As a baseline, we use the parameters in column (2).

Lastly, we need parameter ρ, which governs the spillover in productivity over time. We do not have a direct estimation for the value, so we calibrate the value during the inversion of fundamental productivity in our model. We discuss that in the next section.

E.3 Calibration of fundamentals

Inversion of fundamentals in the steady state equilibrium

Our goal in this subsection is to solve the model for the time-variant environment of the economy conditional on our information about the local labor markets. To this end, we compute the baseline level of the environment as we need to obtain the endogenous variables for the baseline economy that we cannot directly observe in the data at the disaggregated level.

We drop the subscript t for the steady state equilibrium. Suppose that we have data for wage $\{w_i^s\}$, workers $\{L_i^s\}$ and price of housing $\{p_i^0\}$ in the steady state. Then, we obtain values of fundamentals in space: (i) migration cost adjusted amenity; (ii) fundamental productivity; (iii) fundamental features in the development of residential stocks. We explain the procedures and relevant results step by step. We suppose that economy is in the steady state in 2010.

Step 1: Development and income. Developers' problem implies that revenue from land and surplus distributed across workers are:

$$R_i = r_i T_i = \frac{1-\chi}{\chi} w_i^0 L_i^0, \quad \mu_i = 1 + \frac{R_i}{\sum_j w_i^j L_i^j} \tag{E.27}$$

Thus, income for workers in location i and sector s is:

$$W_i^s = \mu_i w_i^s \tag{E.28}$$

Further, in the steady state, we obtain:

$$H_i = \nu_i^{1/\chi} (1 - \bar{h}_i)^{(1-\chi)/\chi} L_i^0 \tag{E.29}$$

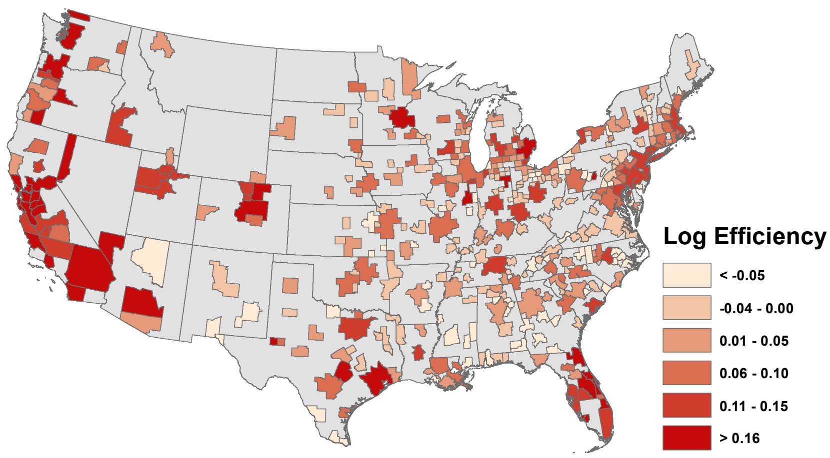
and we also derive

$$\widetilde{\nu}_i \equiv \nu_i (1 - \bar{h}_i)^{1-\chi} = \exp\Big(\chi\Big(- \ln\chi + \ln w_i^0 - \ln p_i^0\Big)\Big) \tag{E.30}$$

As we discussed previously in the subsection E.2, we use $\chi = 0.35$. We also implement the price of housing in 2010 into $\{p_i^0\}$. Our data for the housing price comes from Federal Housing Finance Agency (FHFA). We exploit the Hosing Price Index (HPI) of the all-transactions index across CBSAs.

We also implement wage of sector 0 (i.e., construction sector) in 2010 for $\{w_i^0\}$. Then, (E.30) gives value of fundamental efficiency, $\{\widetilde{\nu}_i\}$, for CBSAs. In Figure E.2, We can see the variation of $\{\widetilde{\nu}_i\}$ for our sample CBSAs. There is a large variation in $\{\widetilde{\nu}_i\}$. Intuitively, locations with large value $\widetilde{\nu}_i$ implies persistency of residential stocks (i.e., small $\bar{h}_i$) and high efficiency of construction (i.e., large ν_i). In equilibrium, they are related to the variation of housing supply and price change.

Figure E.2: Distribution of Development Efficiency



Note: This map shows variation of log scale exogenous efficiency in production of residential stock ($\ln \widetilde{\nu}_i$) across CBSAs.

<u>**Step 2: Inversion of overall endogenous productivity.**</u> We solve for overall productivity in location i and sector s for tradables, Z_i^s. Guess the vector of overall productivity, $\{Z_i^s\}$. Letting $\tau_{in}^s \equiv \text{dist}_{in}^{\bar{\delta}}$ based on the discussion in the subsection E.2, we compute price vector of tradables:

$$\Gamma_s^{\kappa_s}\left(p_i^s\right)^{-\kappa_s} = \sum_n \left(\left(\tau_{in}^s\right)^{-\kappa_s}\left(Z_n^s\right)^{\kappa_s}\left(\left(w_n^s\right)^{\beta_s}\prod_j\left(p_n^j\right)^{\beta_{sj}}\right)^{-\kappa_s}\right) \tag{E.31}$$

Solution for this is $\{p_i^s\}$. Further, sector level trade probability is:

$$\pi_{in}^s = \left(\tau_{in}^s\right)^{-\kappa_s}\left(\frac{Z_n^s}{p_i^s}\right)^{\kappa_s}\left(\left(w_n^s\right)^{\beta_s}\prod_j\left(p_n^j\right)^{\beta_{sj}}\right)^{-\kappa_s} \tag{E.32}$$

Once we have $\left(\{p_i^s\}, p_i^0, \{W_i^s\}\right)$, we solve the following $N \times N \times (S+1)$ dimensional fixed point system of equations:

$$\left(\mathcal{P}_i^s\right)^{1-\sigma} = \sum_j \alpha_j^{\sigma-1}\left(p_i^j\right)^{1-\sigma}\left(\frac{W_i^s}{\mathcal{P}_i^s}\right)^{\theta_j-1} \tag{E.33}$$

for non-homothetic price index, $\{\mathcal{P}_i^s\}$. Using this, we obtain the implied expenditure for sector s in location i such that

$$E_i^s = \sum_j \alpha_s^{\sigma-1}\left(p_i^s\right)^{1-\sigma}\left(\mathcal{P}_i^j\right)^{\sigma-\theta_s}\left(W_i^j\right)^{\theta_s}L_i^j = \sum_j \alpha_s^{\sigma-1}\left(p_i^s\right)^{1-\sigma}\left(\mathcal{P}_i^j\right)^{\sigma}\left(\mathcal{W}_i^j\right)^{\theta_s}L_i^j \tag{E.34}$$

Now, we use the market clearing condition. The market clearing condition states:

$$X_i^s = \sum_j \beta_{js}\sum_n \pi_{ni}^j X_n^j + E_i^s \tag{E.35}$$

for final goods production $\{X_i^s\}$. Labor market clearing condition implies that

$$X_i^s = \sum_j \beta_{js} \frac{w_i^j L_i^j}{\beta_j} + E_i^s \tag{E.36}$$

Combining (E.32), (E.34), (E.36), we obtain:

$$\begin{aligned} \frac{w_i^s L_i^s}{\beta_s} = \sum_n \left(\tau_{ni}^s\right)^{-\kappa_s} \left(\frac{Z_i^s}{p_n^s}\right)^{\kappa_s} \left(\left(w_i^s\right)^{\beta_s} \prod_j \left(p_i^j\right)^{\beta_{sj}}\right)^{-\kappa_s} \\ \times \left(\sum_j \beta_{js} \frac{w_n^j L_n^j}{\beta_j} + \sum_j \alpha_s^{\sigma-1} \left(p_n^s\right)^{1-\sigma} \left(\mathcal{P}_n^j\right)^{\sigma-\theta_s} \left(W_n^j\right)^{\theta_s} L_n^j\right) \end{aligned} \tag{E.37}$$

Manipulating this,

$$\begin{aligned} Z_i^s = \left(\frac{w_i^s L_i^s}{\beta_s}\right)^{\frac{1}{\kappa_s}} \left(\left(w_i^s\right)^{\beta_s} \prod_j \left(p_i^j\right)^{\beta_{sj}}\right) \\ \times \left(\sum_n \left(\tau_{ni}^s p_n^s\right)^{-\kappa_s} \sum_j \left(\beta_{js} \frac{w_n^j L_n^j}{\beta_j} + \alpha_s^{\sigma-1} \left(p_n^s\right)^{1-\sigma} \left(\mathcal{P}_n^j\right)^{\sigma} \left(\frac{W_n^j}{\mathcal{P}_n^j}\right)^{\theta_s} L_n^j\right)\right)^{-1/\kappa_s} \end{aligned} \tag{E.38}$$

We solve (E.38) for $\{Z_i^s\}$. The inner loop calculation for this procedure gives inferred overall productivity, $\{Z_i^s\}$, and other endogenous variables used in the inner loop, $\left(\{p_i^s\}, \{\mathcal{P}_i^s\}\right)$.

Step3: Inversion of amenities and labor mobility. Now, we consider labor mobility. Once the economy reaches the steady state, the mass of workers in the local labor market becomes constant. Yet, we have the move of workers due to idiosyncratic shocks.

We have three fundamentals here. First, we have fundamental amenity, $\{B_i^s\}$. Second, as we have introduced migration barrier in each location in (E.12), $\{M_i\}$. These two fundamentals decide the exogenous gains for workers who choose the destination, and it is impossible to isolate them. Further, we have another fundamental in the industry choice, $\{\zeta_s\}$.

Therefore, we consider the construction of inferred fundamental for location choice,

$$\Omega_i^s = B_i^s \times \frac{1}{M_i} \times \zeta_s^{1/\phi} \tag{E.39}$$

that conflates them. We begin with the guess of $\{\Omega_i^s\}$. Letting $\widetilde{D}_{in} \equiv \text{dist}_{in}^{-\widetilde{\delta}}$ based on estimation in the subsection E.2, we compute the adjusted average real income:

$$\bar{\mathbb{U}}_n^s = \left(\sum_i \left[\Omega_i^s \widetilde{D}_{in} \frac{W_i^s}{\mathcal{P}_i^s}\right]^{\varepsilon}\right)^{1/\varepsilon} = \zeta_s^{1/\phi} \bar{U}_n^s \tag{E.40}$$

We also compute the probability of labor mobility conditional on the sector choice and the probability of sectoral choice:

$$\lambda_{in|s} = \left(\frac{\Omega_i^s \widetilde{D}_{in}}{\bar{\mathbb{U}}_n^s} \left[\frac{W_i^s}{\mathcal{P}_i^s} \right] \right)^{\varepsilon}, \quad \varsigma_n^s = \frac{\left(L_n^s\right)^{\eta} \left(\bar{\mathbb{U}}_n^s\right)^{\phi}}{\sum_j \left(L_n^j\right)^{\eta} \left(\bar{\mathbb{U}}_n^j\right)^{\phi}} \tag{E.41}$$

Then, we use labor mobility condition:

$$L_i^s = \sum_n \lambda_{in|s} \varsigma_n^s L_n \tag{E.42}$$

Plugging the above equations into this yields:

$$\Omega_i^s = \left(\frac{1}{L_i^s} \sum_n \left(\frac{\widetilde{D}_{in}}{\bar{\mathbb{U}}_n^s} \left[\frac{W_i^s}{\mathcal{P}_i^s} \right] \right)^{\varepsilon} \frac{\left(L_n^s\right)^{\eta} \left(\bar{\mathbb{U}}_n^s\right)^{\phi}}{\sum_j \left(L_n^j\right)^{\eta} \left(\bar{\mathbb{U}}_n^j\right)^{\phi}} L_n \right)^{-1/\varepsilon} \tag{E.43}$$

In the right-hand-side, we use (E.40) inside the loop. Inner loop for this step gives inferred fundamental amenity, $\{\Omega_i\}$ and other endogenous variables. In particular, we derive:

$$L_{in}^s = \lambda_{in|s} \varsigma_n^s L_n \tag{E.44}$$

$\{L_{in}^s\}$ is mass of workers in sector s who move from n to i.

Step4: Inversion of productivity. As a last step, we consider the inversion of fundamental productivity and calibration of parameter ρ. In the steady state, overall productivity satisfies:

$$Z_i^s = A_i^s \left(\sum_n L_{in}^s Z_n^s \right)^{\rho} \left(L_i^s \right)^{\gamma_s} \tag{E.45}$$

Therefore, the exogenous fundamental productivity is:

$$\ln A_i^s = \ln Z_i^s - \rho \ln \left(\sum_n L_{in}^s Z_n^s \right) - \gamma_s \ln L_i^s \tag{E.46}$$

We implement overall productivity $\{Z_i^s\}$ in Step 2, inferred labor mobility $\{L_{in}^s\}$ in Step 3 and employment in data $\{L_i^s\}$ and parameters $\{\gamma_s\}$ we discussed in the subsection E.2 into this. To estimate ρ, we use the following moment conditions:

$$\mathbb{E}\left[\left(\ln A_i^s - \frac{1}{N} \sum_n \ln A_n^s - \frac{1}{S} \sum_k \ln A_i^k \right) \times \mathbb{I}_g \right] = 0, \quad g \in \mathbb{G}_0, \mathbb{G}_1, \cdots, \mathbb{G}_P \tag{E.47}$$

where $\mathbb{I}_g$ is an indicator that the labor market potential of location i is in the group of g. The group of locations is defined by the accessibility of the location i. We ordered locations by the sum of population in other places with an inverse of bilateral migration cost as weights. Namely, for location i and sector s, we compute the measure

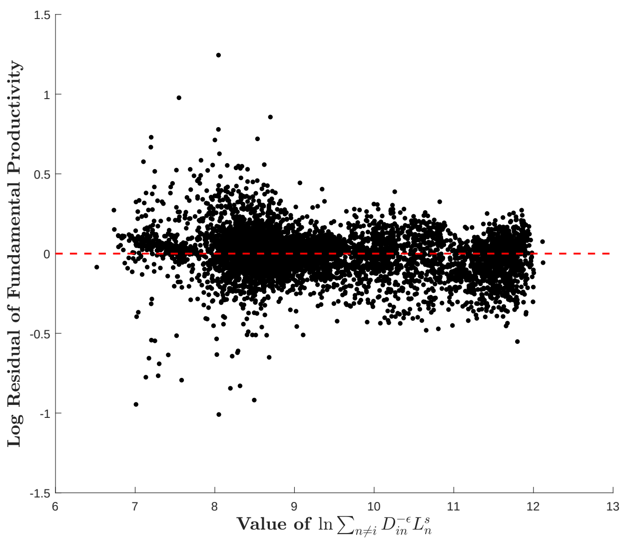
$$\sum_{n \neq i} \widetilde{D}_{in}^{\varepsilon} L_n^s \tag{E.48}$$

Then, we define the group of location and sector pairs based on this measure. As a baseline, we use $P = 20$ groups defined by 5 percentile of the measure. The moment conditions assume that the location and sector specific exogenous part after eliminating averages is not systematically related to the labor market access of the location as the spatial dependence of productivity is captured by the endogenous terms in (E.46) through labor mobility. We use (E.47) and search parameter ρ that minimize the distances of the moment conditions.

Figure E.3 shows the log of residuals for inverted productivity in (E.47) in the vertical axis and value of (E.48) in the horizontal axis for different pairs of CBSAs and sectors.

Figure E.4 displays value of the objective function of (E.47). The estimated ρ that minimizes the objective function is given by the dashed line, $\widehat{\rho} = 0.0284$. In the computation, we define the grid of ρ over $[0, 0.5]$ by 0.0001. In Figure E.3, we only report the value of objective functions in the range $[0, 0.05]$ as it is increasing in other region.

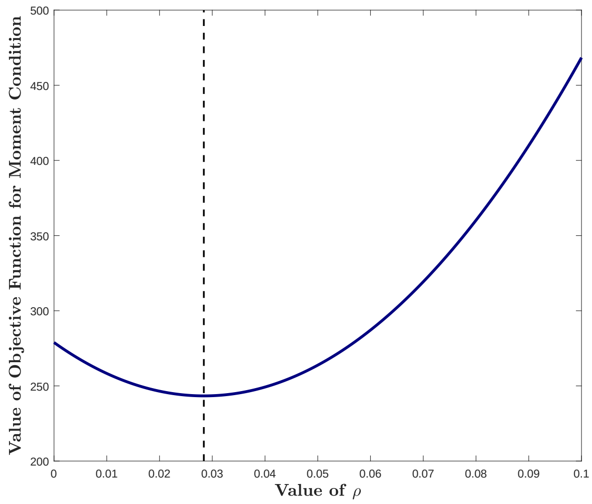
Figure E.3: Moment Conditions (E.47): Value of residuals



Computing past fundamentals

Our aim is solving the model for time-variant environment of the economy conditional on our information about the local labor markets. To this end, we compute the path of $\{A_{i,t}^s\}$. In period $t = 2010$, we assume $A_{i,t}^s = A_i^s$ that is derived in (E.46) and it is unchanged after then. So, we compute the path of $\{A_{i,T}^s\}_{T=t-1,t-2,\cdots}$ back to the previous periods. For fundamental amenities, our interests are dynamics of residential amenity $\{B_{i,t}^s\}$ and migration barrier $\{M_{i,t}\}$. Our obser-

Figure E.4: Moment Conditions (E.47): Estimate of $\widehat{\rho}$



vation over periods is the equilibrium wage and employment in 2000, 1990 and 1980, $\{w^s_{i,t}\}$ and $\{L^s_{i,t}\}$. Given other exogenous environments and parameters discussed above, they are sufficient to compute the pattern of fundamental productivity and amenities in the past, starting from the steady state equilibrium.

Step1: Housing and land market clearing conditions. Given $\left(\{w^0_{i,t-1}\}, \{L^0_{i,t-1}\}\right)$, zero profit condition and distribution of land rent implies:

$$R_{i,t-1} = \frac{1-\chi}{\chi} w^0_{i,t-1} L^0_{i,t-1}, \quad W^s_{i,t-1} = \left(1 + \frac{R_{i,t-1}}{\sum_j w^j_{i,t-1} L^j_{i,t-1}}\right) w^s_{i,t-1} \tag{E.49}$$

Using zero profit condition, in the steady state equilibrium (i.e., $t = 2010$) or any period t, we compute the stock of residential stocks such that:

$$H_{i,t} = \frac{1}{\chi} \frac{w^0_{i,t} L^0_{i,t}}{p^0_{i,t}} \tag{E.50}$$

Once we obtain this, we compute the residential stock in the previous period that solves:

$$\ln H_{i,t-1} = \frac{1}{1-\chi} \left(\ln H_{i,t} - \ln \widetilde{\nu}_i - \chi \ln L_{i,t}\right) \tag{E.51}$$

where we use $\{\widetilde{\nu}_i\}$ in the subsection E.3. Then, market clearing condition leads to price in period $t-1$:

$$p^0_{i,t-1} = \frac{1}{\chi} \frac{w^0_{i,t-1} L^0_{i,t-1}}{H_{i,t-1}} \tag{E.52}$$

This procedure obtain the path of $\left(\{p^0_{i,t-1}\}, \{H_{i,t-1}\}, \{R_{i,t-1}\}, \{W^s_{i,t-1}\}\right)$ in equilibrium that are not directly observable.

Step2: Overall productivity path. To derive the overall productivity in the past, we guess the path of productivity, $\{\mathrm{d}\ln Z^s_{i,t}\}$. Therefore, we guess $\{Z^s_{i,t-1}\}$, given pre-determined $\{Z^s_{i,t}\}$. Then, we compute price $\{p^s_{i,t-1}\}$ that solve:

$$\mathrm{d}\ln p^s_{i,t} \equiv \ln p^s_{i,t} - \ln p^s_{i,t-1} = -\frac{1}{\kappa_s}\ln\left(\frac{\sum_n\left(\tau^s_{in}\left(w^s_{n,t}\right)^{\beta_s}\prod_j\left(p^j_{n,t}\right)^{\beta_{sj}}\right)^{-\kappa_s}\left(Z^s_{n,t}\right)^{\kappa_s}}{\sum_n\left(\tau^s_{in}\left(w^s_{n,t-1}\right)^{\beta_s}\prod_j\left(p^j_{n,t-1}\right)^{\beta_{sj}}\right)^{-\kappa_s}\left(Z^s_{n,t-1}\right)^{\kappa_s}}\right) \tag{E.53}$$

and we compute the trade pattern $\{\pi^s_{in,t-1}\}$ such that:

$$\mathrm{d}\ln\pi^s_{in,t} \equiv \ln\pi^s_{in,t} - \ln\pi^s_{in,t-1} = \kappa_s\left(\mathrm{d}\ln Z^s_{n,t} - \mathrm{d}\ln p^s_{i,t} - \mathrm{d}\ln w^s_{n,t} + \sum_j\beta_{sj}\mathrm{d}\ln p^j_{n,t}\right) \tag{E.54}$$

Given income and price in period $t-1$, $\left(\{p^s_{i,t-1}\},\{W^s_{i,t-1}\}\right)$, we solve for aggregate price index $\{\mathcal{P}^s_{i,t-1}\}$ as in (E.33). Then, we use the static equation of market clearing conditions, as in (E.38), to solve for the overall productivity $(\{\widehat{Z}^s_{i,t-1}\})$ that rationalize observed wage and number of workers as an equilibrium.

Step3: Casting the workers' move and path of amenities. The procedures Step 1 and 2 allow us to compute the spatial distribution of prices, real income and overall productivity, starting from the steady state level set to the year 2010. Next, we use the model structure forward from the initial period. This allows us to derive the path of location attractiveness. We start from the guess of overall attractiveness of location and sector, $\Omega^j_{i,t}$ as in the subsection E.3: $\Omega^j_{i,t} \equiv B^j_{i,t}\zeta_j^{1/\phi}/M_i$. This overall amenity becomes large when the value of utility benefit from residential amenity for workers in location i ($B^j_{i,t}$) is high, migration barrier of location i (M_i) is small and sector level taste parameter (ζ_j) is large for workers in the sector.

Guess $\Omega^s_{i,t}$. Given income ($W^s_{i,t}$) derived in Step1 and aggregate price index ($\mathcal{P}^s_{i,t}$) derived in Step2, we compute the average real income

$$\widehat{\mathbb{U}}^s_{n,t} = \left(\sum_i\left(\widetilde{D}_{in}\Omega^s_{i,t}\frac{W^s_{i,t}}{\mathcal{P}^s_{i,t}}\right)^{\varepsilon}\right)^{1/\varepsilon} \tag{E.55}$$

Then, we compute

$$\widehat{\Omega}^s_{i,t} = \left(\frac{1}{L^s_{i,t}}\sum_n\left(\frac{\widetilde{D}_{in}}{\widehat{\mathbb{U}}^s_{n,t}}\left(\frac{W^s_{i,t}}{\mathcal{P}^s_{i,t}}\right)\right)^{\varepsilon}\frac{\left(L^s_{n,t-1}\right)^{\eta}\left(\widehat{\mathbb{U}}^s_{n,t}\right)^{\phi}}{\sum_j\left(L^j_{n,t-1}\right)^{\eta}\left(\widehat{\mathbb{U}}^j_{n,t}\right)^{\phi}}L_{n,t-1}\right)^{-1/\varepsilon} \tag{E.56}$$

We update $\Omega^j_{i,t}$ until $\|\widehat{\Omega}^j_{i,t} - \Omega^j_{i,t}\| < \varepsilon$ for sufficient small number ε and appropriate norm $\|\cdot\|$.

This procedure allows us to cast the workers' choice across locations and sectors predicted in the model. This is essential as an overidentification test to assess the performance of our model for workers' choice. In particular, we compute two probabilities:

$$\widehat{\lambda}_{in|s,t} = \left(\frac{\widehat{\Omega}^s_{i,t} \widetilde{D}_{in}}{\widetilde{\mathbb{U}}^s_{n,t}} \left(\frac{W^s_{i,t}}{\mathcal{P}^s_{i,t}} \right) \right)^{\varepsilon}, \quad \text{with} \quad \widehat{\mathbb{U}}^s_{n,t} = \left(\sum_i \left(\widetilde{D}_{ni} \widehat{\Omega}^s_{i,t} \frac{W^s_{i,t}}{\mathcal{P}^s_{i,t}} \right)^{\varepsilon} \right)^{1/\varepsilon} \tag{E.57}$$

and

$$\widehat{\varsigma}^s_{n,t} = \frac{\left(L^s_{n,t-1}\right)^{\eta} \left(\widehat{\mathbb{U}}^s_{n,t}\right)^{\phi}}{\sum_j \left(L^j_{n,t-1}\right)^{\eta} \left(\widehat{\mathbb{U}}^j_{n,t}\right)^{\phi}} \tag{E.58}$$

where we use pre-period population in data ($L^s_{i,t-1}$). We lastly compute

$$\widehat{L}^s_{in,t} = \sum_n \widehat{\lambda}_{in|s,t} \widehat{\varsigma}^s_{n,t} L_{n,t-1} \tag{E.59}$$

where $L_{n,t-1}$ is total number of workers in data for previous generation. (E.59) is predicted move of workers for generation t.

<u>**Step 4: Fundamental productivity.**</u> For the initial period, we set $A^s_{i,t} = Z^s_{i,t}$. In our setting, it is applied for 1980. For other period, we compute

$$\ln \widehat{A}^s_{i,t} = \ln \widehat{Z}^s_{i,t} - \widehat{\rho} \ln \left(\sum_n \widehat{L}^s_{in,t} \widehat{Z}^s_{n,t-1} \right) - \gamma_s \ln L^s_{i,t} \tag{E.60}$$

where $\widehat{\rho}$ is obtained in the subsection E.3 and $\{\widehat{Z}^s_{i,t}\}$ are computed in Step2.

====================================

F Appendix: Calibration Results

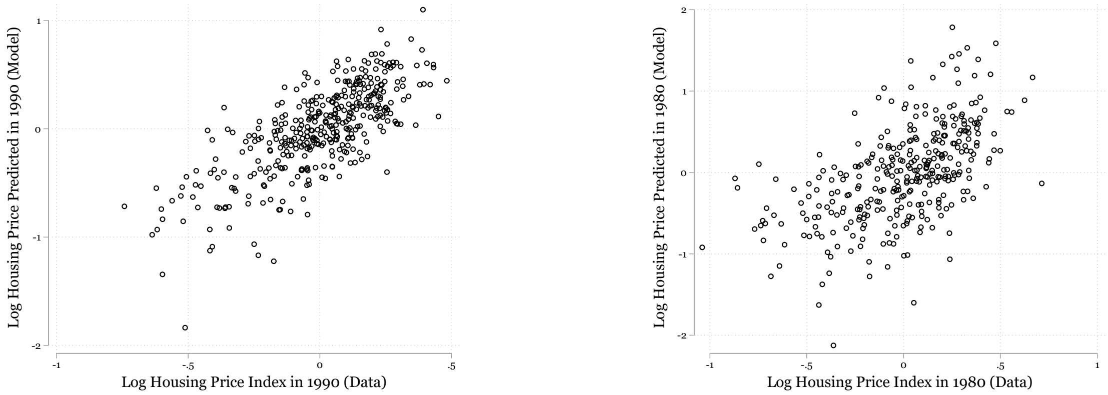
This section presents the results of calibration that we explained in the previous section. The subsection F.1 shows the results for an inversion of development, amenities and productivity. We also compute TFP. The subsection F.2 shows the welfare and intergenerational mobility in the baseline.

F.1 Inverted Environment

Housing Prices and Development Efficiency

We can gauge our model specification in (E.51) and (E.52) by comparing the predicted value of $\{\widehat{p}^0_{i,t}\}$ in the past. Among 395 CBSAs in our calibration, FHFA data for housing prices are limited for the past years. Therefore, we compare housing prices predicted in the model for past years, 1980 and 1990, and those in data for limited number of CBSAs.

Figure F.1: Housing Prices



(a) 1990

(b) 1980

Figures F.1 show such comparison. Panel F.1a shows the comparison between predicted price and limited data for 1990, where values are demeaned log price. The number of CBSAs is 386. The dashed line shows a 45-degree line. This confirms that the model prediction for housing prices is closely related to the observation. The correlation between model inferred prices and limited data for 1990 is 0.746. Panel F.1b displays those for 1980. The correlation becomes 0.568 for 331 CBSAs.

Figure F.2: Efficiency of Development and Housing Prices

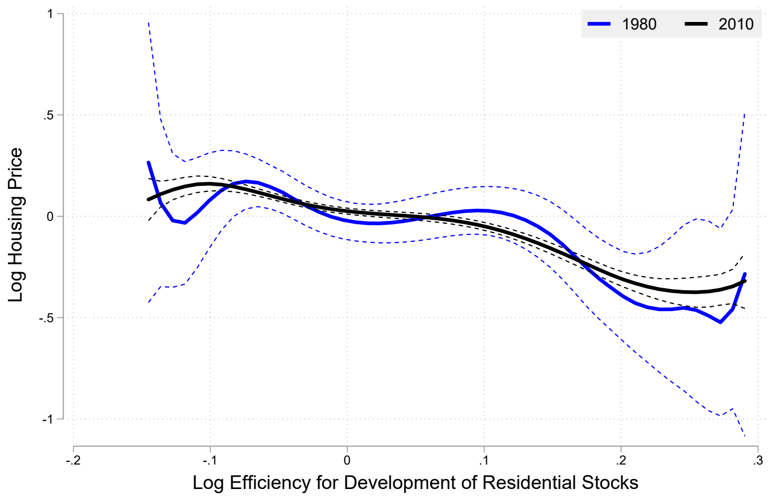


Figure F.2 shows the relationship between efficiency of development for housing ($\widetilde{\nu}_i$) inverted in E.3 and housing prices for 1980 and 2010. Housing prices are demeaned log prices, and the line shows a fitted polynomial fitted line. Intuitively, a large value of efficiency for development leads to lower housing prices *ceteris paribus*.

Amenities

Using the local data, we obtain the local amenities for workers, $\Omega^s_{i,t}$. For each year, 2010, 2000 and 1990, Table F.1 reports the mean and standard deviation of the logarithm of the amenity vector across different sectors. We find the difference in their variations across industries, and the standard deviation becomes large in the last period compared to the previous periods.

Table F.1: Summary of Local Amenities $\{\Omega^s_{i,t}\}$

	2010		2000		1990	
Industry	Mean	S.D.	Mean	S.D.	Mean	S.D.
0. Construction	-0.004	1.054	0.028	0.840	-0.004	0.905
1. Food/Beverage/Tobacco	0.006	0.944	0.025	0.770	0.022	0.881
2. Textile/Apparel	-0.047	1.082	-0.026	0.883	0.018	0.938
3. Wood/Paper/Printing	0.014	1.002	0.007	0.813	0.027	0.858
4. Chemical/Petro/Coal/ Nonmetalic	0.007	1.007	0.026	0.786	-0.002	0.831
5. Metal	0.036	0.989	0.058	0.805	0.028	0.864
6. Machinery	0.003	1.010	-0.033	0.823	0.014	0.879
7. Electric/Computer	0.013	0.998	-0.036	0.867	-0.032	0.928
8. Transport Equipment	0.026	0.916	-0.007	0.788	0.003	0.834
9. Miscellaneous Manufacture	-0.006	1.029	0.004	0.827	0.035	0.841
10. Transportation Service	-0.015	1.016	-0.043	0.885	-0.002	0.893
11. Wholesale Trade	0.023	0.928	-0.001	0.784	0.000	0.885
12. Retail	-0.016	1.047	0.022	0.788	-0.008	0.913
13. FIRE	0.013	0.962	0.012	0.752	-0.024	0.895
14. Health Service	-0.023	0.998	0.004	0.741	-0.001	0.843
15. Education and Legal	0.008	1.008	-0.019	0.852	-0.013	0.920
16. Communication Service	-0.035	1.073	-0.008	0.856	-0.049	0.998
17. Other Services	-0.005	1.006	-0.013	0.863	-0.012	0.934

Large value of amenity in location i and sector s ($\Omega^s_{i,t}$) is associated with large number of workers of sector s. Therefore, the average of $\Omega^s_{i,t}$ at the CBSA level is related to the total size of CBSA. To see this, we compute geometric mean of local amenities, $\widetilde{\Omega}_{i,t} = \left(\prod_s \Omega^s_{i,t}\right)^{1/|\mathcal{K}|}$ for each CBSA. Figure F.3 shows the positive relationship between the average of amenities and size of CBSA for 1990 and 2010. Table F.2 list the top 15 CBSAs showing highest value of average amenities for each period.

Table F.3 reports relationship between inverted average amenities and characteristics of CBSAs. In column (1), we include the average temperature of CBSAs in January and July, which are constructed in Rappaport (2007). Column (2) includes annual precipitation from Rappaport (2007) and the air quality index from U.S. Environmental Protection Agency. Column (3) adds a log of the crime rate per 100,000 inhabitants in 2010 from Uniform Crime Reporting (UCR). Column (4) also includes the share of commuters who take more than 30 minutes to the workplace. We confirm that the amenities are related to weather characteristics, but most inverted amenities are less related to other characteristics. This is similar to findings in Glaeser et al. (2016).

Figure F.3: Amenities and Total Employment Size

Note: The employment is normalized by the total employment in the economy for each period.

Table F.2: CBSAs with Highest Average Amenities

Rank	1990	2000	2010
1	Pittsfield, MA	Oxnard-Thousand Oaks-Ventura, CA	Pittsfield, MA
2.	Oxnard-Thousand Oaks-Ventura, CA	Pittsfield, MA	Oxnard-Thousand Oaks-Ventura, CA
3.	Plymouth, IN	Paducah, KY-IL	Pittsburg, KS
4.	Pittsburg, KS	Plymouth, IN	Plymouth, IN
5.	Salinas, CA	Pittsburg, KS	Paducah, KY-IL
6.	Paducah, KY-IL	Pine Bluff, AR	Pine Bluff, AR
7.	Santa Maria-Santa Barbara, CA	Portland-South Portland, ME	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
8.	Portland-South Portland, ME	Pittsburgh, PA	Portland-South Portland, ME
9.	Santa Rosa, CA	Portland-Vancouver-Hillsboro, OR-WA	Pittsburgh, PA
10.	Portland-Vancouver-Hillsboro, OR-WA	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	Portland-Vancouver-Hillsboro, OR-WA
11.	Pittsburgh, PA	Phoenix-Mesa-Scottsdale, AZ	Phoenix-Mesa-Scottsdale, AZ
12.	San Luis Obispo-Paso Robles-Arroyo Grande, CA	Oshkosh-Neenah, WI	Oshkosh-Neenah, WI
13.	Pine Bluff, AR	Pensacola-Ferry Pass-Brent, FL	Pensacola-Ferry Pass-Brent, FL
14.	Searcy, AR	Port St. Lucie, FL	Port St. Lucie, FL
15.	Santa Cruz-Watsonville, CA	Ottawa-Peru, IL	Ottawa-Peru, IL

Table F.3: Average Amenities and Characteristics

	(1)	(2)	(3)	(4)
		$\ln \widetilde{\Omega}_{i,2010}$		
Jan. Temperature	0.206**	0.244**	0.365**	0.479*
	(0.0996)	(0.106)	(0.148)	(0.275)
Jul. Temperature	-1.073*	-1.236**	-1.281**	-1.499
	(0.546)	(0.565)	(0.646)	(0.977)
Precipitation		0.00988	0.00582	-0.0228
		(0.0817)	(0.0882)	(0.110)
Air Quality		0.185	0.230	0.297
		(0.153)	(0.164)	(0.226)
Violent Crime			-0.0006	0.0272
			(0.103)	(0.140)
Property Crime			-0.0411	-0.0606
			(0.183)	(0.280)
Long Commuting				0.276
				(0.852)
Observations	278	256	229	128
R^2	0.015	0.027	0.037	0.080

Productivity

Step 2 in the subsection E.3 yields overall productivity ($Z^s_{i,t}$) for past years: 1980, 1990, 2000 and 2010. In addition, we obtain trade patterns ($\pi^s_{in,t}$). Using them, we can compute TFP for each sector and location as we discussed in the subsection C.1, $\ln \delta^s_{i,t}$.

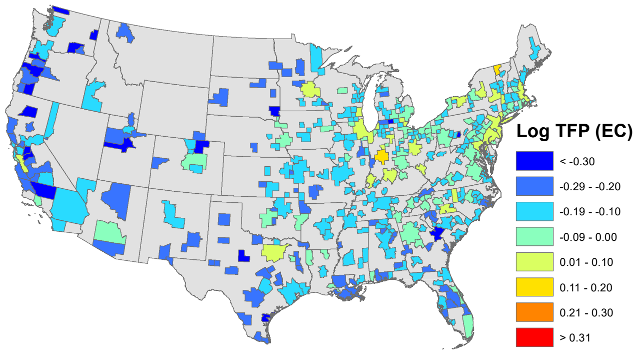
Figure F.4 and F.5 show the spatial distribution of log of TFP for two different industries: electric and computer industry and finance, insurance and real estate (FIRE). In both maps, blue colored CBSAs show lower TFP, while red colored CBSAs show higher TFP. We see large variation of TFP across CBSAs for each industry, and it has shown the transition over periods.

From these maps, we can identify an increase in the spatial variation of TFP for the electric and computer industry, along with the development of clusters in California and large metropolitan areas on the East coast. We also see the different evolution of TFP for FIRE. Over time, there has been a remarkable increase in its level and variation. The industry has seen a significant development on the East coast (New York metropolitan area) and in large cities that are the hub of the financial market in each region (Chicago, Dallas, Atlanta and Nashville) from 1980 to 1990. Then, these clusters show persistent development over time, while some other inland cities also have seen a rise in FIRE.

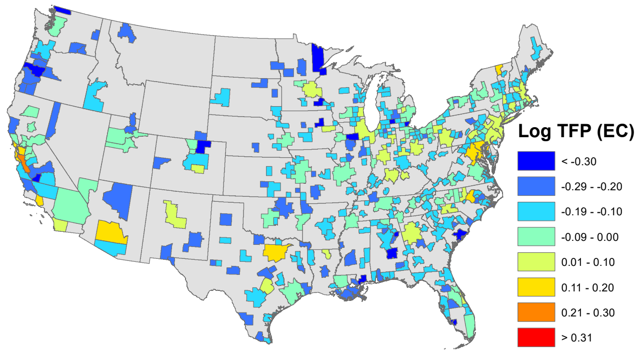
In Table F.4, we report the standard deviation of the measured TFP and inverted fundamental productivity across CBSAs. The fundamental productivity shows the large spatial variation relative to TFP. This implies that the covariance of the import penetration ($\pi^s_{ii,t}$) and fundamental productivity is significant. Intuitively, firms in a city demand for own products more when the location exhibits high fundamental productivity.

Table F.5 reports CBSAs that have shown largest TFP changes between different years in Table F.5. For each sector, we compute $\mathrm{d}\ln\delta^s_{i,t}$ and we identify CBSAs that exhibited largest value of $\mathrm{d}\ln\delta^s_{i,t}$ during each period 1980-1990, 1990-2000 and 2000-2010.

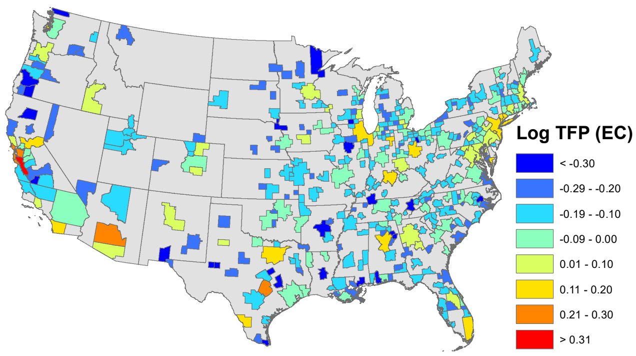
Figure F.4: Geography of TFP: Electric and Computer Industry



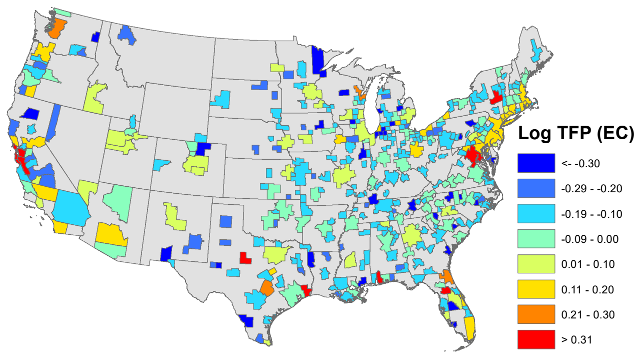
(a) 1980



(b) 1990



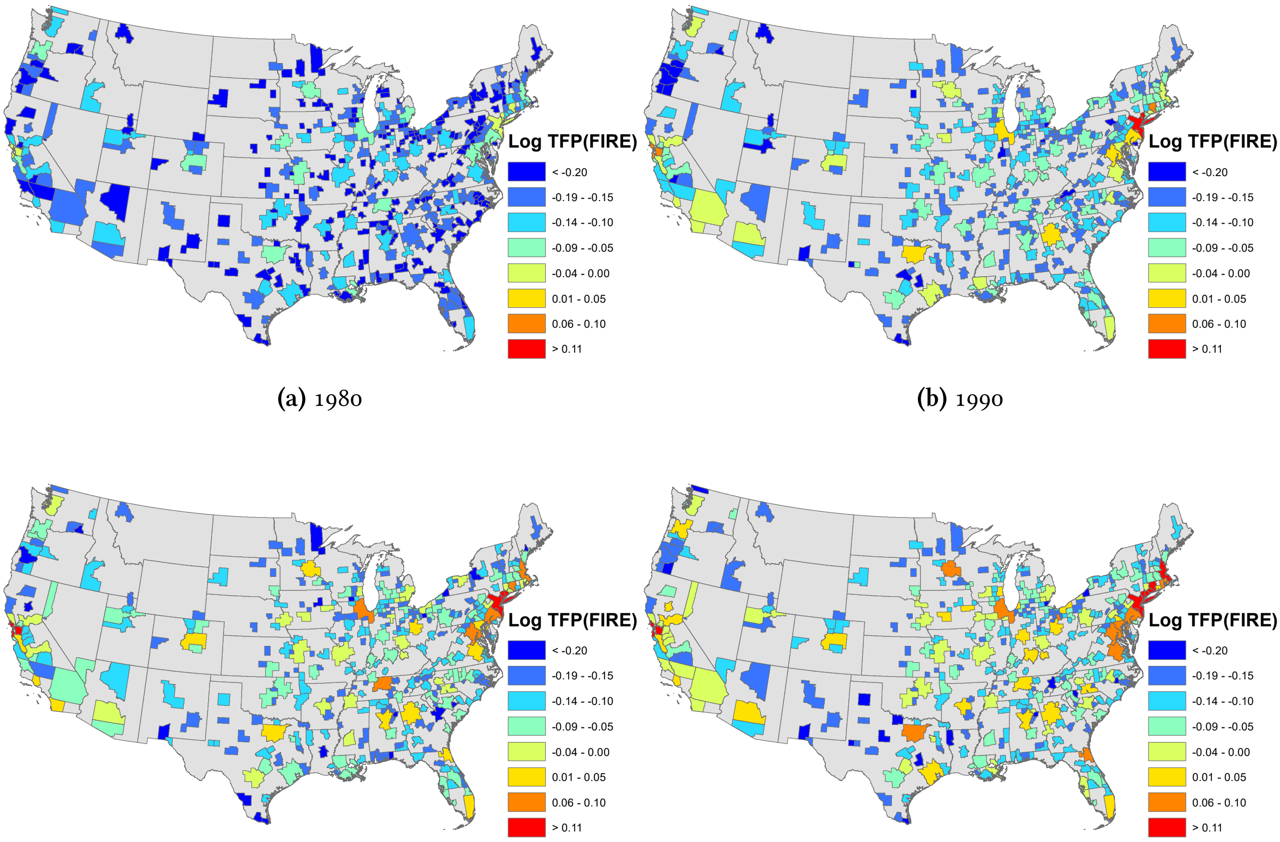
(c) 2000



(d) 2010

Note: These maps show the measured TFP for the electric and computer industry in our calibration of the baseline economy.

Figure F.5: Geography of TFP: Finance, Insurance and Real Estate



(a) 1980 **(b)** 1990

(c) 2000 **(d)** 2010

Note: These maps show the measured TFP for the finance, insurance and real estate (FIRE) industry in our calibration of the baseline economy.

Table F.4: Spatial Variation of TFP and Fundamental Productivity

	1980		1990		2000		2010	
Industry	S.D.(δ_i^s)	S.D.(A_i^s)	S.D.(δ_i^s)	S.D.(A_i^s)	S.D.(δ_i^s)	S.D.(A_i^s)	S.D.(δ_i^s)	S.D.(A_i^s)
1. Food/Beverage/Tobacco	.037	.050	.042	.102	.051	.098	.055	.103
2. Textile/Apparel	.077	.115	.096	.102	.100	.136	.164	.225
3. Wood/Paper/Printing	.040	.044	.044	.115	.053	.108	.059	.111
4. Chemical/Petro/Coal/Nonmetalic	.035	.044	.038	.120	.052	.114	.057	.108
5. Metal	.045	.053	.042	.114	.050	.110	.093	.144
6. Machinery	.060	.107	.073	.105	.087	.108	.082	.111
7. Electric/Computer	.087	.209	.112	.112	.159	.143	.178	.179
8. Transport Equipment	.044	.060	.047	.111	.054	.094	.061	.097
9. Miscellaneous Manufacture	.090	.131	.078	.111	.094	.121	.105	.124
10. Transportation Service	.067	.207	.063	.159	.067	.150	.070	.155
11. Wholesale Trade	.038	.038	.047	.153	.058	.148	.072	.156
12. Retail	.045	.045	.059	.269	.068	.261	.060	.270
13. FIRE	.054	.158	.082	.164	.102	.158	.103	.145
14. Health Service	.055	.056	.066	.150	.060	.142	.068	.148
15. Education and Legal	.081	.093	.097	.160	.100	.160	.110	.142
16. Communication Service	.064	.615	.073	.120	.092	.157	.103	.191
17. Other Services	.102	.183	.119	.128	.171	.156	.142	.141

Note: This table reports the standard deviation of measured TFP ($\delta_{i,t}^s$) and fundamental productivity ($A_{i,t}^s$) for any particular year and industry.

Table F.5: CBSAs with largest TFP growth for each industry

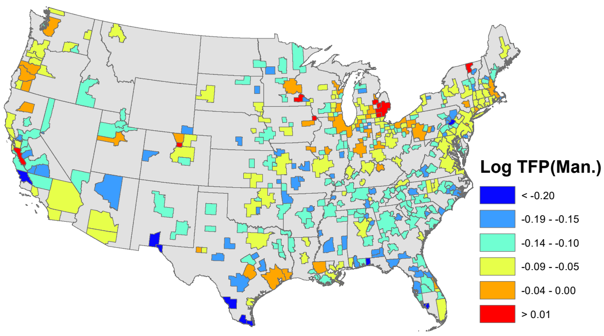
Industry	1980-1990	1990-2000	2000-2010
1. Food/Beverage/Tobacco	Crestview-Fort Walton Beach-Destin, FL	Bridgeport-Stamford-Norwalk, CT	Roseburg, OR
2. Textile/Apparel	Wausau, WI	Gainesville, FL	Kansas City, MO-KS
3. Wood/Paper/Printing	Las Vegas-Henderson-Paradise, NV	Bridgeport-Stamford-Norwalk, CT	Flagstaff, AZ
4. Chemical/Petro/Coal/ Nonmetalic	Cape Coral-Fort Myers, FL	Bridgeport-Stamford-Norwalk, CT	Bremerton-Silverdale, WA
5. Metal	San Luis Obispo-Paso Robles-Arroyo Grande, CA	Macon, GA	Worcester, MA-CT
6. Machinery	Port St. Lucie, FL	San Luis Obispo-Paso Robles-Arroyo Grande, CA	Goldsboro, NC
7. Electric/Computer	Abilene, TX	Waterloo-Cedar Falls, IA	Daphne-Fairhope-Foley, AL
8. Transport Equipment	Gadsden, AL	Midland, TX	Waterloo-Cedar Falls, IA
9. Miscellaneous Manufacture	Tallahassee, FL	Abilene, TX	Las Cruces, NM
10. Transportation Service	Memphis, TN-MS-AR	Jacksonville, FL	Roseburg, OR
11. Wholesale Trade	Rochester, MN	Bridgeport-Stamford-Norwalk, CT	Gainesville, FL
12. Retail	Concord, NH	Fayetteville-Springdale-Rogers, AR-MO	Abilene, TX
13. FIRE	Anniston-Oxford-Jacksonville, AL	Bridgeport-Stamford-Norwalk, CT	Chico, CA
14. Health Service	Atlanta-Sandy Springs-Roswell, GA	Salisbury, MD-DE	Champaign-Urbana, IL
15. Education and Legal	Midland, TX	Gainesville, GA	Rochester, MN
16. Communication Service	Atlanta-Sandy Springs-Roswell, GA	Bridgeport-Stamford-Norwalk, CT	Niles-Benton Harbor, MI
17. Other Services	Augusta-Richmond County, GA-SC	Bridgeport-Stamford-Norwalk, CT	New Orleans-Metairie, LA

Having TFP of each industry, we compute the TFP aggregated to large sector level: manufacturing and services. We compute TFP of aggregated level:

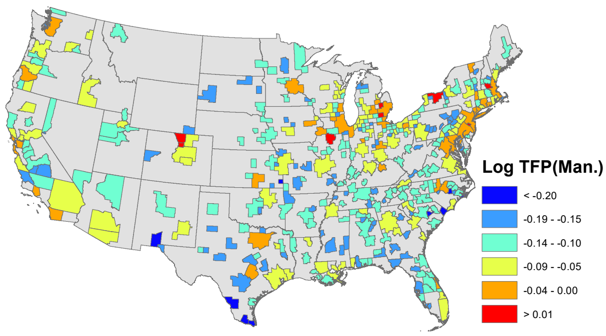
$$\delta_{i,t}^{K} = \sum_{j \in K} \frac{X_{i,t}^{j}}{\sum_{k \in K} X_{i,t}^{k}} \delta_{i,t}^{j} \tag{F.1}$$

where K is aggregate level of sector and $X_{i,t}^{j}$ is value of production of sector j in location i. We compute two aggregate sectors, the manufacturing sector and the services sector, for K. Figure F.6 show log of TFP of manufacturing sector in different period: 1980, 1990, 2000 and 2010. Red colored areas show high TFP for the manufacturing sector, while blue colored CBSAs exhibit low TFP for the manufacturing sector. As we can see in the maps, cities around the Rust Belt show persistence in their relatively high productivity in manufacturing, while the South and East coast areas show growth in productivity in manufacturing. These differences across space reflect the set of industries in the manufacturing sector across different cities.

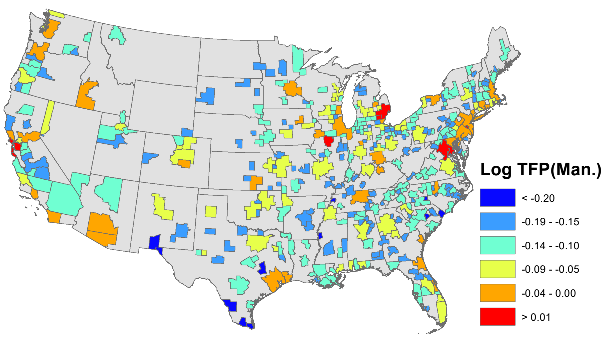
Figure F.6: TFP of Manufacturing Sector



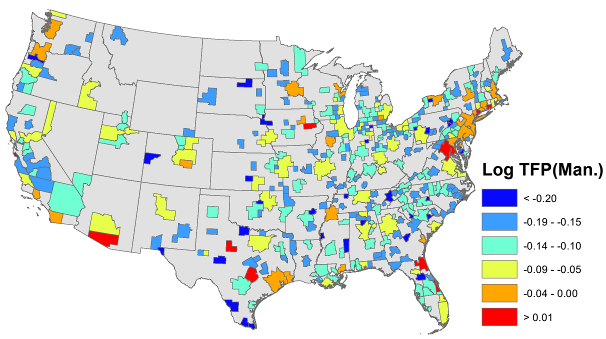
(a) 1980



(b) 1990



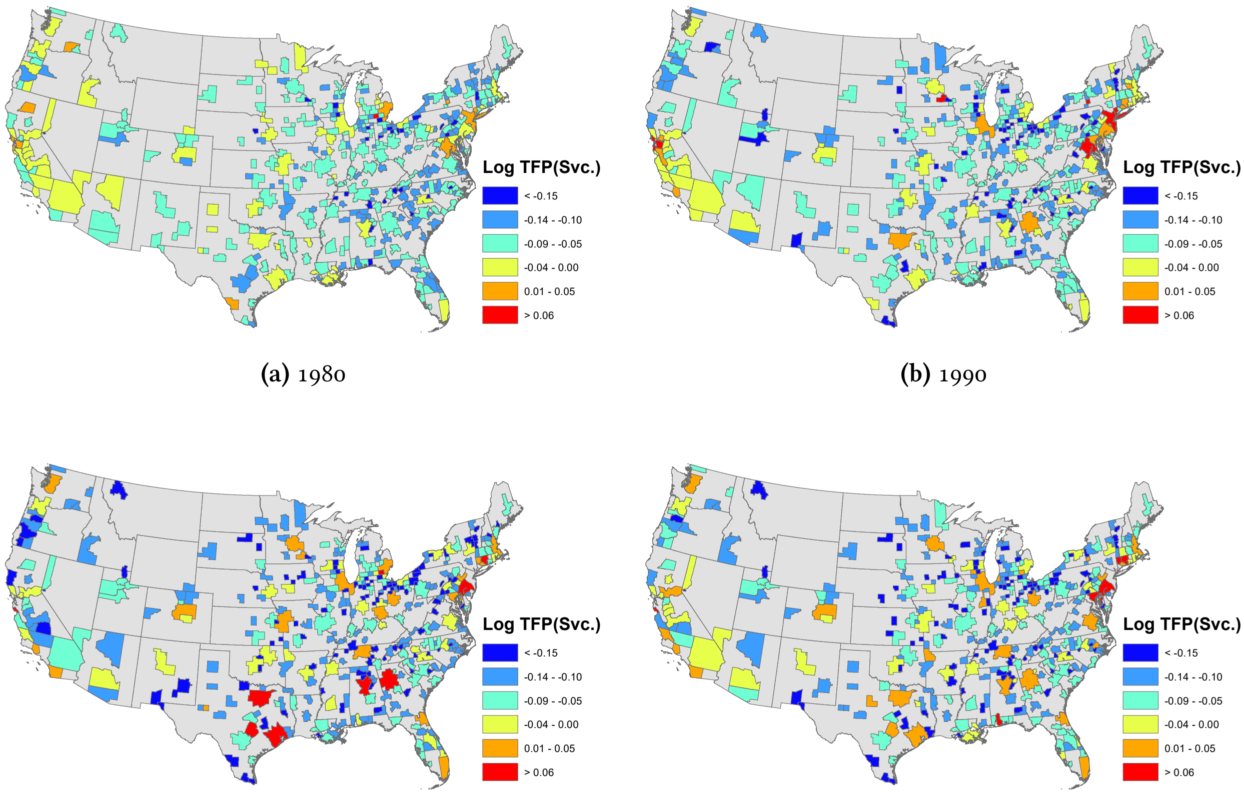
(c) 2000



(d) 2010

In turn, figure F.7 show the log of TFP of the service sector in different periods. Red colored areas show high TFP for the service sector, while blue colored CBSAs exhibit low TFP. We can see the TFP growth over time in the U.S. economy with clustering. Throughout time, the TFP of services grows in large cities, while there are variations across regions. From 1980 to 1990, the services grow in cities on the East coast and the West coast. The period 1990 to 2000 exhibits growth of TFP in the South. In the last period, 2000-10, the persistent growth in these areas led to the country's service growth.

Figure F.7: TFP of Services Sector



(a) 1980 **(b)** 1990

(c) 2000 **(d)** 2010

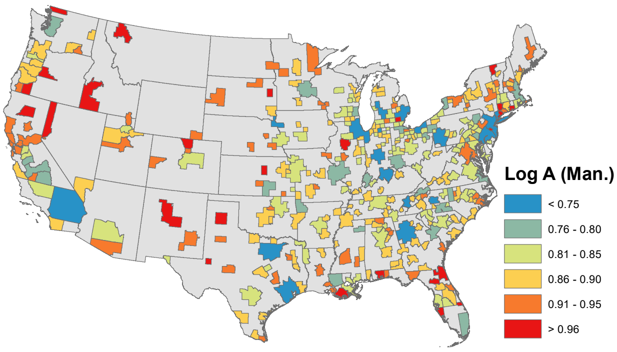
Fundamental productivity. The last step of inversion for fundamental environment in the economy is inversion of fundamental productivity ($A^s_{i,t}$) for each industry. Following the definition of aggregated level TFP, we can also define fundamental productivity aggregated at manufacturing and service:

$$A^K_{i,t} = \sum_{j \in K} \frac{X^j_{i,t}}{\sum_{k \in K} X^k_{i,t}} A^j_{i,t} \tag{F.2}$$

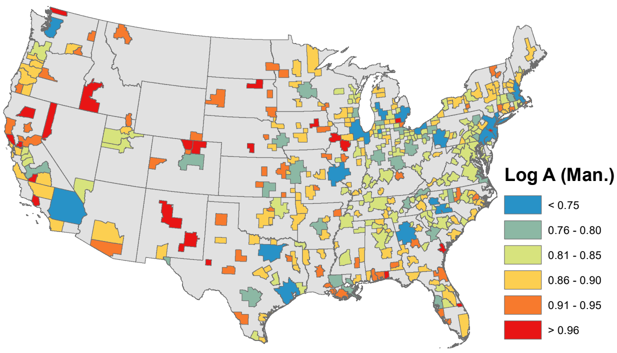
Figure F.8 show fundamental productivity of manufacturing sector for 1990, 2000 and 2010. The hot (Red) colored CBSAs exhibit relatively high exogenous productivity, while cool (Blue) colored CBSAs show relatively low exogenous productivity for the manufacturing sector. We can confirm the strong agglomeration forces when we compare these maps and previous maps of TFP. In the Rust belt, we see lower fundamental productivity, which is offset by the agglomeration of workers in the manufacturing sector. This is important in our model since more concentration of workers instead of exogenous advantage leads to persistent over generations in their job opportunity. Figure

F.9 display the distribution of fundamental productivity for the services sector.

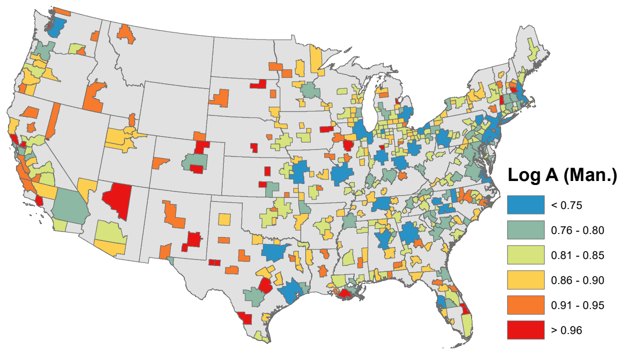
Figure F.8: Fundamental Productivity of Manufacturing Sector



(a) 2010

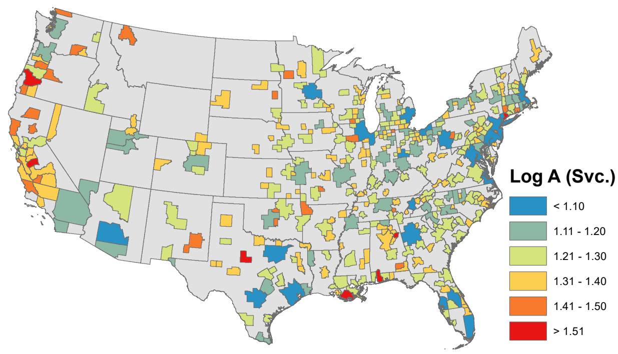


(b) 2000

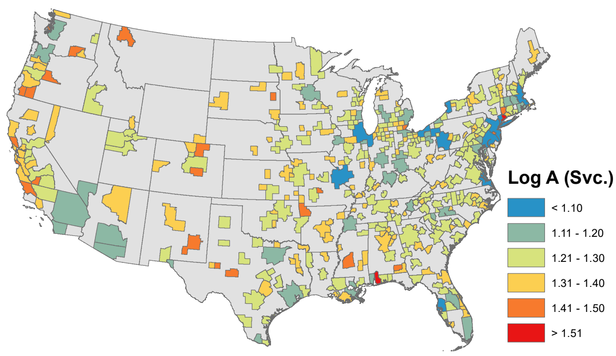


(c) 1990

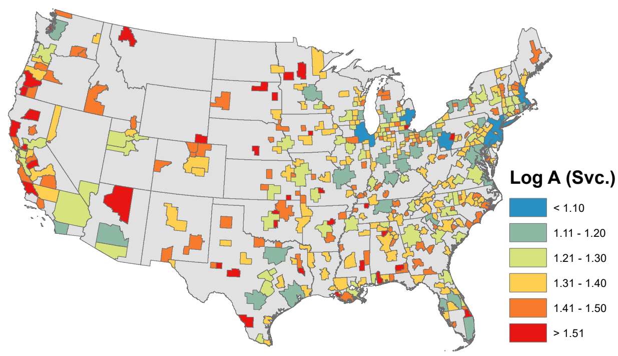
Figure F.9: Fundamental Productivity of Service Sector



(a) 2010



(b) 2000



(c) 1990

F.2 Welfare and upward mobility

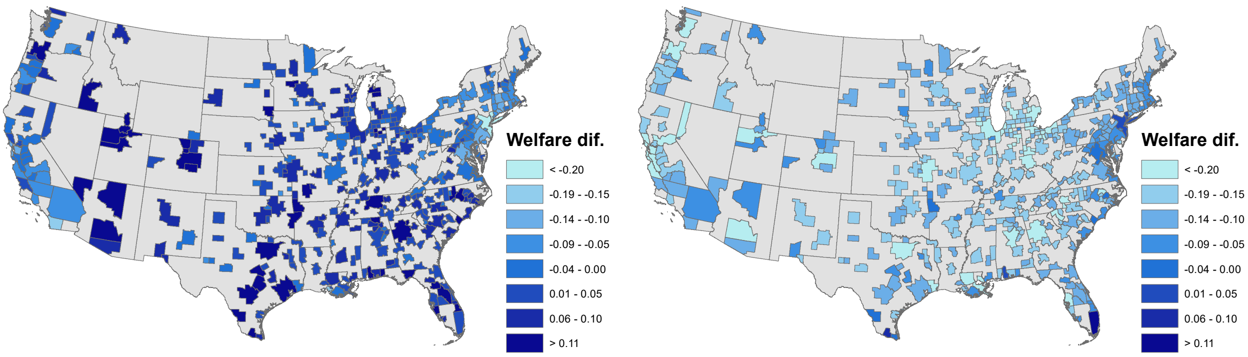
Figure F.10 displays the distribution of welfare differences between individuals who have the same origin of CBSA. Panel (a) shows the welfare difference between generation 2000 and 1990, and panel (b) is for generation 2010 and 2000. Figures F.11 show the distribution of welfare changes. Most locations exhibit a welfare decline from 2000 to 2010. This reflects the lower wage growth and higher increases in housing prices during the period, while the effects show large variation across locations.

The income inequality in each local labor market is evaluated by coefficient of variation (CV) in income. For each year, Figure F.12 shows the distribution of CV in income for 395 CBSAs. The horizontal axis is log of CV in income. In our theory, we assume that income of workers in location i and sector s is proportional to their wage and distribution of surplus μ_i is same across workers in the same location. Therefore, CV in income within CBSA is equal to that of wages.

Next, we see the upward mobility for individuals of each generation. Figure F.13 shows, for each CBSA, the income percentile for generation t working in the CBSA and the income percentile for generation $t+1$ who have origin in the CBSA.

For each CBSA, the horizontal axis shows the average income percentile of workers (i.e., old generations) in the country; the vertical axis shows the average income percentile of individuals in the next generation. The black colored ones show the relationship for generations 1980 and 1990,

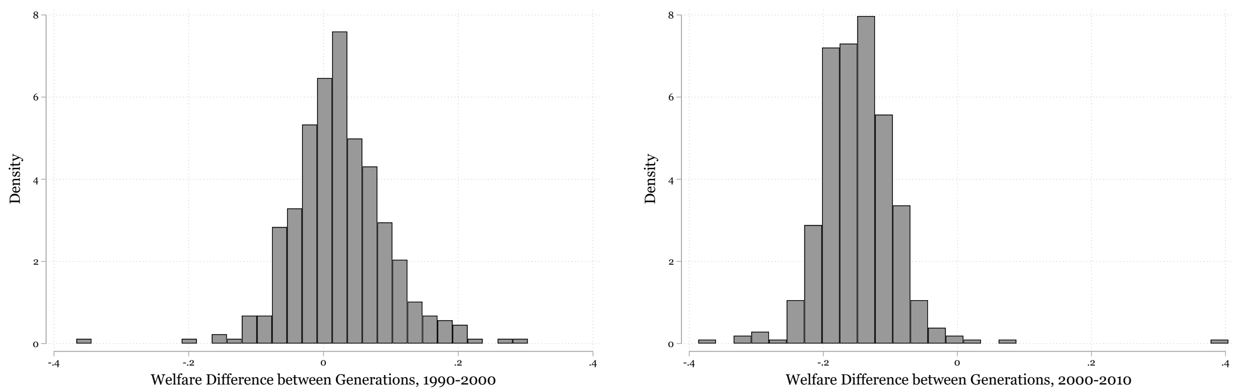
Figure F.10: Welfare Differences



(a) $\mathrm{d}\ln V_{i,2000}$ **(b)** $\mathrm{d}\ln V_{i,2010}$

Note: These figures show the spatial pattern of welfare differences between generations.

Figure F.11: Distribution of Welfare Dynamics

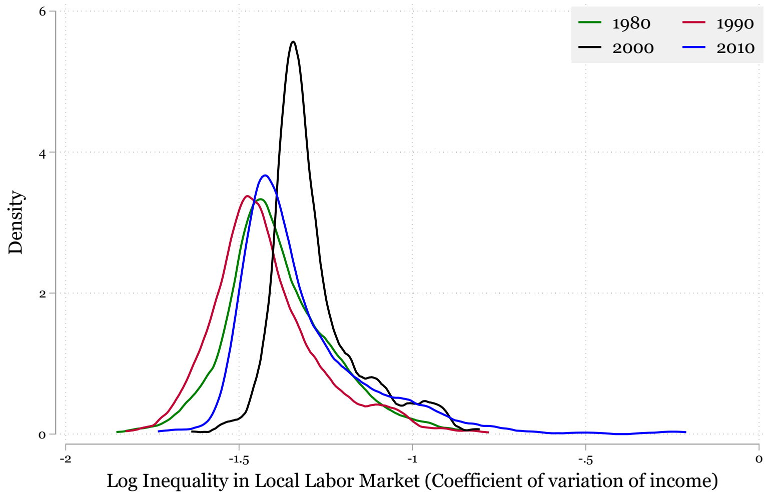


(a) 1990-2000 **(b)** 2000-2010

while red colored ones show that for 2000 and 2010. Each circle represents the size of generation 1980 and 2000 respectively, and the dashed line is the 45-degree line. Therefore, locations above the reference line show upward mobility of the generation compared to their previous generation, while those below the line are the places with relatively low upward mobility. From this figure, we find that large CBSAs exhibit lower upward mobility in 2000-10 compared to 1980-90, which leads to lower upward mobility on average. As we discuss in the main text, we compute $\widetilde{\mathcal{M}}_{i,t} = (\mathcal{M}_{i,t}/\bar{\mathcal{M}}_t) \times 25$. Figure F.14 show the distribution of the measure for two generations, 1990 and 2010. We find large variation in the measure across locations for both generations.

We also see the relationship between local inequality and upward mobility. Figures F.15 show such relationship for CBSAs in 1990, 2000 and 2010. In Figure F.15a displays the upward mobility for generation 1990 on the vertical axis and CV of income in 1980 on the horizontal axis. Figure F.15b and F.15c display the same relationships for generation 2000 and generation 2010 respectively. They are related to the *Great Gatsby curve* in U.S., which shows the negative relationships between local inequality and upward mobility for individuals who have the origin in the CBSA.

Figure F.12: Distribution of Inequality in CBSA



Lastly, we compare our measure and the measure by Chetty et al. (2014) that is computed by exploiting the microdata of the U.S. samples. Figure F.16 shows the comparison between the model predicted average income percentile for workers of generation 1990 and the absolute upward mobility measure in Chetty et al. (2014) across locations. Their absolute upward mobility measures the expected income rank for people born in 1980-82, which is based on income in 2011-12 relative to that of their parents in 1996-2000 and defines the expected income rank for children from families with below-median parents' income in the national distribution. We use their measures at the MSA level. In the figure, we only use the CBSAs that correspond to their metropolitan areas. The size of each circle shows the size of the location.

As we see in this figure, the average income percentile for workers of generation 1990 is related to their measure of upward mobility. This implies that there is a correlation between the aggregate measures of the possibility of upward mobility for workers and the micro evidence across cities in the U.S, with relatively large opportunities in large cities.

Figure F.17 displays comparison of our measures of upward mobility ($\widetilde{M_{i,t}}$) and the absolute upward measure from Chetty et al. (2014). The correlations show that our measure of upward income mobility based on the aggregate data and model structure is related to the results based on the individual level data in the sample of Chetty et al. (2014) at the city level.

Figure F.13: Average Income Percentile of Individuals

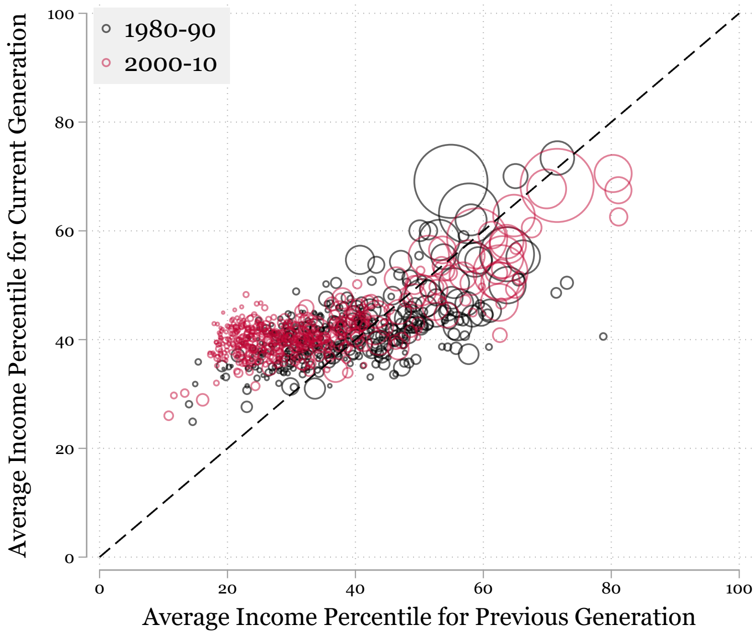
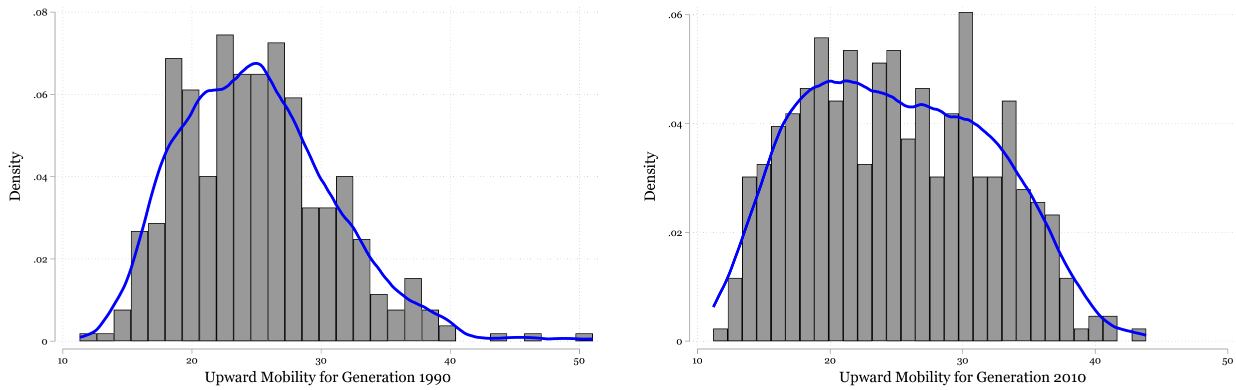


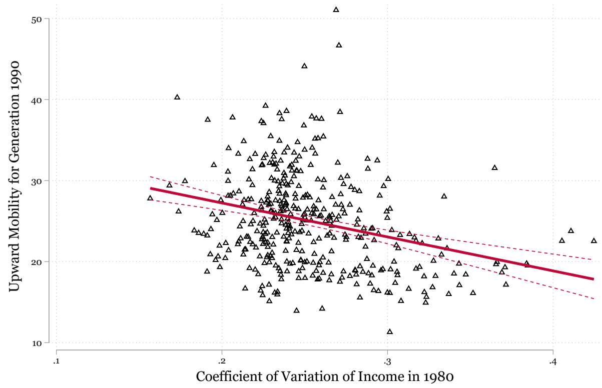
Figure F.14: Distribution of Upward Mobility



(a) Generation 1990

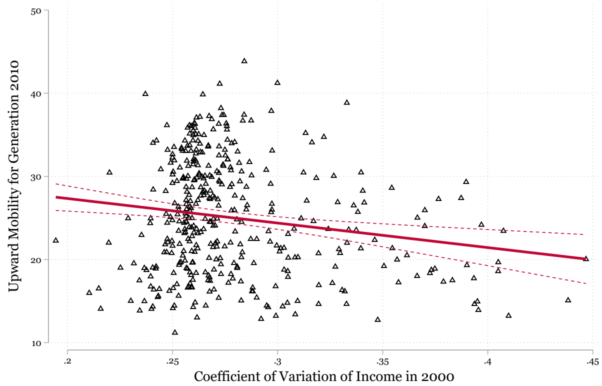
(b) Generation 2010

Figure F.15: Inequality and Upward Mobility



(a) 1990

(b) 2000



(c) 2010

Figure F.16: Average Income Percentile of Children and Measure of Absolute Upward Mobility by Chetty et al. (2014)

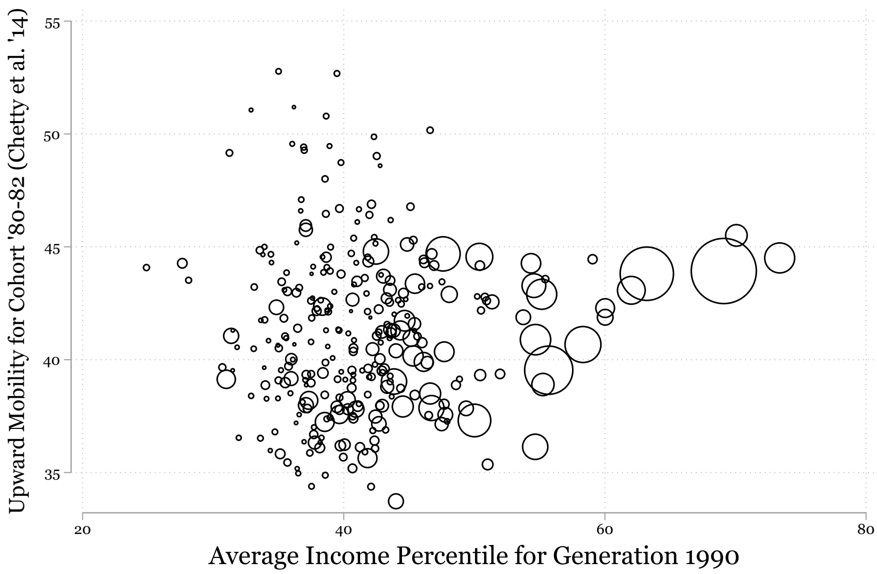
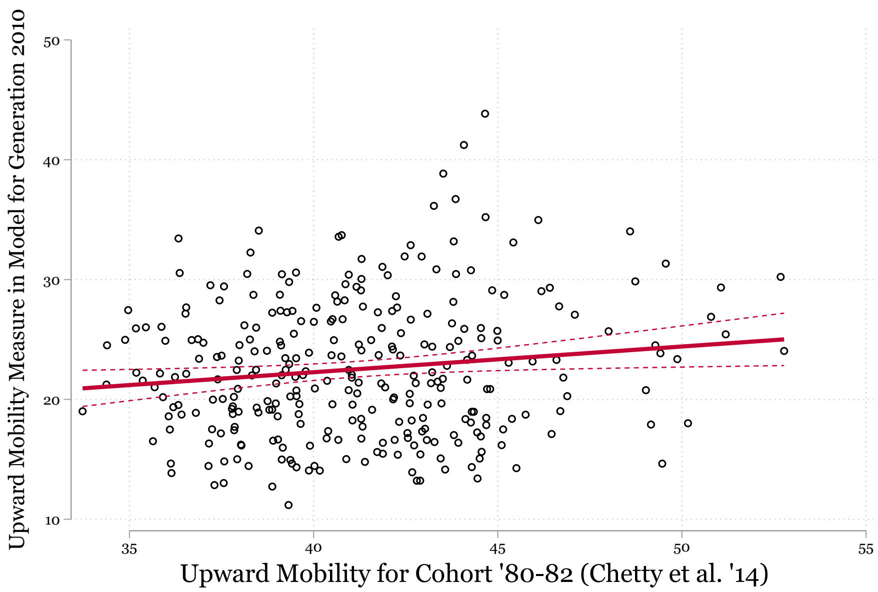


Figure F.17: Aggregate Average Upward Mobility and Measure of Absolute Upward Mobility by Chetty et al. (2014)



G Appendix: Counterfactuals

Spatial variation of fundamental productivity. We can consider another counterfactual for fundamental productivity: eliminating the variation of fundamental productivity differences on geography. We set the fundamental productivity of industry j in period t such that $\widetilde{A}_t^j = \left(\prod_i A_{i,t}^j \right)^{1/N}$. The average productivity grows over time, but the evolution is same across CBSAs.

Table G.6 reports the percentage change of aggregate TFP in the services sector and the change of employment shares in the service sector from the baseline economy. For each year, 1990, 2000 and 2010, we show the mean, standard deviation, 25 percentile and 75 percentile values across 395 CBSAs in the U.S. economy. Units of all entries are percentages. The counterfactual undertakes when productivity growth of all industries is uniform across CBSAs. We take the geometric mean of fundamental productivities across locations for each sector in each period, and we assume that all CBSAs experience the same rate of fundamental productivity growth. Compared to the baseline economy, service sector TFP shows 7.6 percent higher in 1990 and reaches 15.1 percent in 2010 on average, and most CBSAs exhibit TFP growth in the service sector. This is consistent with the implication of the agglomeration economies. Once we abstract the exogenous variation, industrial locations are subject to strong agglomeration forces due to the spillover in productivity. Since the strength of local agglomeration economies is strong for service sectors (γ_j), we can see an increase of service sectors TFP and an increase of the standard deviation. The decline of employment share is small relative to other counterfactuals about productivity shocks in main text, implying that the benefit of agglomerations counteracts the slow structural change.

In Table G.7, welfare differences are defined for between generation 2000 to 1990 and 2010 to 2000. We show the mean, standard deviation, 25 percentile and 75 percentile values of their changes across 395 CBSAs in the U.S. economy. Units of all entries are percentages. The endogenous spillover works for the welfare changes. Comparing the numbers to counterfactuals (i) to (iii) in the main text, both average welfare change and change in standard deviation show significant increases.

In Table G.8, we report percentage change of the intergenerational income mobility from the baseline values. We find the positive effects on the upward mobility of workers. These findings conclude that the spatial variation of productivity mitigates the polarization of locations in terms of welfare dynamics and upward income mobility. Without such exogenous productivity differences, individual consequences in terms of intergenerational income mobility are crucially shaped by the place they have an origin, while they may benefit on average.

Table G.6: Counterfactual Experiments – Impact on Service TFP and Change in Service Employment Share

	1990				2000				2010			
	Mean	SD	25prc	75prc	Mean	SD	25prc	75prc	Mean	SD	25prc	75prc
Service TFP	7.639	7.473	3.041	11.855	7.891	7.702	3.559	12.236	15.167	11.354	7.888	22.172
Service Emp. Share	−10.195	5.274	−13.712	−6.698	−13.972	6.295	−18.615	−9.7854	−18.194	6.687	−22.77	−13.72

Table G.7: Counterfactual Experiments – Impact on Welfare

	1990 − 2000				2000 − 2010			
	Mean	SD	25 prc	75 prc	Mean	SD	25 prc	75 prc
Welfare	0.09	4.232	−2.277	2.507	0.067	3.664	−2.179	2.234
Consumption	0.228	6.858	−3.65	3.735	0.229	6.772	−3.324	3.781
Migration Gain	0.136	5.176	−2.165	2.94	0.164	5.796	−2.898	2.829
Job Opportunity Gain	0.068	3.686	−1.818	2.349	0.06	3.454	−2.214	2.498

Table G.8: Counterfactual Experiments – Impact on Intergenerational Income Mobility

1990				2000				2010			
Mean	SD	25prc	75prc	Mean	SD	25prc	75prc	Mean	SD	25prc	75prc
−1.089	15.73	−9.982	10.029	6.399	44.263	−22.205	24.941	7.511	52.466	−26.36	29.747

H Appendix: Additional figures

This section provides additional figures that are related to the elements in our model.

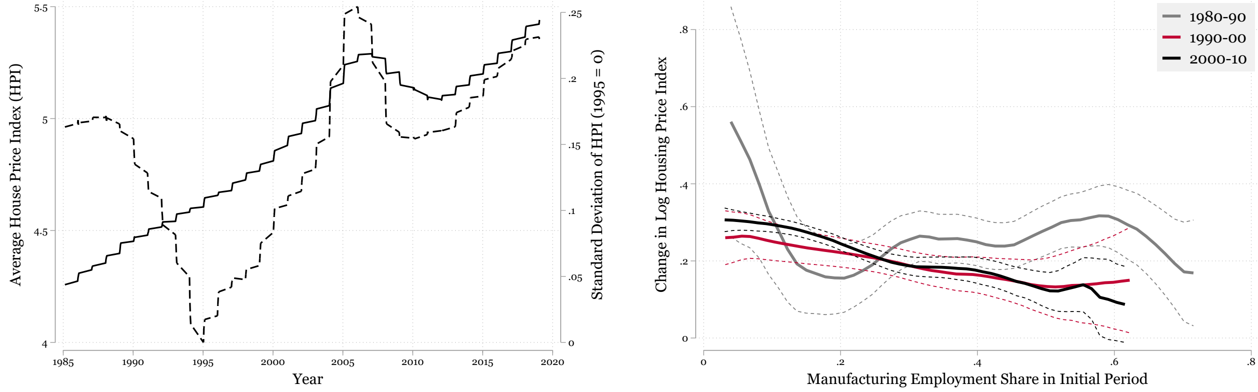
Housing prices. The population distribution is uneven across cities and agglomerations change the value of local amenities, which is reflected in land and housing prices. Figure F.1 shows the changes in the U.S. economy in the average and standard deviation of the house price index and its relation to the employment structure.

The left-hand panel F.1a shows the change in the average house price index (HPI) and the standard deviation of HPI across MSAs. The black line (corresponding to the left axis) is the average HPI, and the dashed line (corresponding to the right axis) is the standard deviation of HPI across MSAs. The standard deviation decreased before 1995, while it increased over time until around 2007. This dropped after the financial crisis, but it again increased in the last decade.

The right-hand panel F.1b confirms that changes in housing prices are related to the initial employment patterns. We focus on the employment share of the manufacturing sector in the initial period. In the 1980-90 period, housing prices were high in the place where services were concentrated and where structural transformation proceeded from manufacturing to services. During this period, the structural transformation in a large number of cities drives the decline in the variation of housing prices. In the later periods, 1990-2000 and 2010-2010, the negative relationship confirms that housing prices have grown in the service cities, where workers are concentrating. This creates an increase in the variation of housing prices over time.

These variations in the housing prices and the underlying local amenities are essential margins that account for the welfare disparity by place occurring in the structural transformation phase. In the model, we introduce the different values of amenities for workers by location and sector and developers that supply residential stocks. This allows us to see the variation of the price of residential stocks across space and time and how it is related to the pattern of structural transformation.

Figure F.1: Spatial Variation of Housing Prices in the U.S.



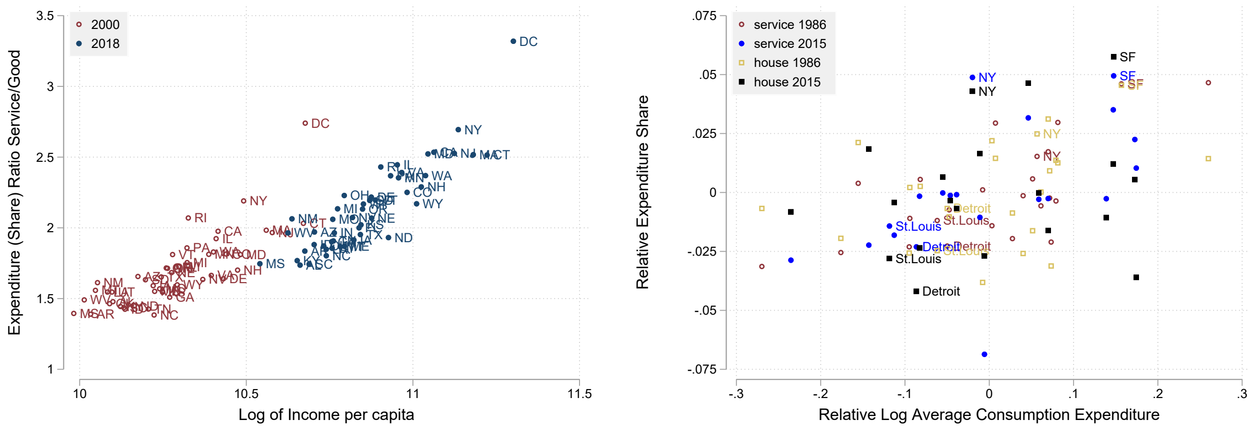
(a) Variation of House Price Index **(b)** Housing Price and Employment Change

Note: Panel (a) shows the change in the average house price index (HPI) and the standard deviation of HPI across MSAs. The black line (corresponding to the left axis) is the average HPI, and the dashed line (corresponding to the right axis) is the standard deviation of HPI across MSAs. HPI is normalized to the first quarter of 1995; therefore, the standard deviation of HPI in the quarter is zero. Panel (b) shows the relationship between the change in the log of HPI and the the manufacturing employment share. Different lines show the polynomial fitted line across MSAs, and the dotted lines are 95% confidence intervals. The data source for HPI is Federal Housing Finance Agency (FHFA).

Expenditure. We consider the different slopes of the Engel curve of consumers by geography and sector as a driver of consumption-led growth in our model.

The idea behind this can be seen in Figure F.2 that shows (i) the relation between employment pattern and income per capita across large U.S. states and (ii) the expenditure composition for large MSAs after eliminating the trend and MSA specific factor.

Figure F.2: Spatial Variation of Expenditure



(a) Across States **(b)** Across Top MSAs

Panel (a) shows the consumption expenditure ratio between services and goods and the log of income per capita across the U.S. states (we drop two states, Hawaii and Alaska), and its comparison between 2000 (red points) and 2018 (blue points). The data source is BEA personal consumption expenditure. Panel (b) shows the variation of expenditure share and size of consumption expenditure across top MSAs in the U.S. The x-axis indicates de-meaned the expenditure share of service (including housing) and housing in 1986 and 2015, respectively. The y-axis indicates the de-meaned log of average consumption expenditure in 1986 and 2015. The data source is Consumer Expenditure Survey in the U.S. We classify the expenditure into different categories: goods, housing, transport, and non-housing services. Goods include alcoholics, food, apparel and tabacco. Non-housing service includes the rest of the expenditure after deducting other categories. It includes healthcare, entertainment, personal care, reading, education and insurance. Service is defined as a sum of non-housing service and housing. From these figures, we can find the heterogeneity in expenditure composition and its change over time. In our model, beyond the conventional channel of structural transformation, the expenditure shift leads to the dynamics of agglomeration and dispersion. Therefore, the response of expenditure patterns to the policy shocks is an essential margin that interacts with the dynamic gains from agglomeration.

Chapter 2

How Useful are Quantitative Urban Models for Cities in Developing Countries? Evidence from Dhaka

1 Introduction

A new class of quantitative urban models allows us to capture the rich heterogeneity of real world cities and evaluate the impact of different policy interventions. With few exceptions these models have been used to analyze cities in developed countries, where rich data at a fine geographical scale is available. However, the majority of the world's urban population lives in cities in developing countries for which we often only observe very limited data from censuses and other traditional data sources. An obvious question is therefore how useful quantitative urban models are when it comes to guiding policy analysis in cities in developing countries.

In this paper, we calibrate a quantitative urban model that builds closely on the literature following Ahlfeldt et al. (2015) on a typical developing country city, the city of Dhaka in Bangladesh. Building on recent work by Kreindler and Miyauchi (2021), who use increasingly available mobile phone data together with Google travel times to estimate commuting costs in Dhaka, we show how their results can be combined with newly available satellite data on building heights to estimate the key structural parameters of an urban model. With these estimates we show how the model can be used to estimate the price of floor space and land in each ward of Dhaka, which are prices that are difficult to reliably measure for many developing country cities. To illustrate how the model can be used for policy analysis, we examine two model counterfactuals: (i) an increase in density of the city modeled as an increase in the floor space supply elasticity and (ii) a reduction in commuting times in Dhaka due to a new radial road.

The main idea behind our approach of using satellite data on heights is that information on the variation in quantities can be used to infer variation in prices. While it is often impossible to obtain reliable data on land or floor space prices for cities in developing countries, newly available data on building heights can be interpreted as the equilibrium outcome of the interplay between the demand and supply of floor space. Demand for floor space is created by both firms and residents in each ward of the city. We combine these two sources of demand for floor space with the help of our model to estimate the demand curve for floor space in each location, which allows us to infer both the floor space supply elasticity and the price of floor space and land.

The paper presents six main results. First, we estimate a floor space supply elasticity for Dhaka of 1.45 which is not far below the average elasticity for U.S. cities estimated by Saiz (2010). Second, we estimate that land prices in Dhaka vary by nearly an order of magnitude across different locations in the city. Third, we estimate productivity and amenities in each ward of Dakha, which vary in plausible ways, using information on land prices and model-derived wages. Fourth, we use a simple cross-sectional moment condition to estimate an elasticity of productivity with respect to the employment density of 0.045, which is in the middle of estimates for this parameter for cities in developed countries. Fifth, we show that changes in density modeled as an increase in the floor space supply elasticity have a number of surprising implications, including that average commuting times in the city remain the same despite a much larger agglomeration of workers and residents in the center of the city. Finally, we show that a new road would attract both residents and employment to wards in the immediate proximity of the new road at the expense of other parts of the city.

Our paper contributes to several strands of literature. First, we build on the recent urban literature following Ahlfeldt et al. (2015), which has developed quantitative models of the internal organisation of cities. Recent contributions to this literature include Allen et al. (2015), Redding and Rossi-Hansberg (2017), Monte et al. (2018), Owens III et al. (2020), Heblich et al. (2020), Tsivanidis (2020). We use a benchmark model from the literature to demonstrate how the key parameters of this model can be estimated in a data-scarce developing country context exploiting mobile phone data and satellite data on building heights.

Second, our paper is closely related to the literature that estimates housing supply elasticities. A seminal contribution to this literature is Saiz (2010) who estimates housing supply elasticities for a large number of U.S. cities and shows how the substantial variation in housing supply elasticities depends on features of the natural geography. Ahlfeldt and McMillen (2018) is a recent contribution that examines the economics of extremely tall buildings and Ahlfeldt and Barr (2022) provides a synthesis of this literature. Baum-Snow and Han (2021) is a recent contribution that estimates housing supply elasticities across thousands of US census tracks using variation in labor demand shocks to commuting destinations as a source of exogenous variation. Our main contribution relative to this literature is to develop a methodology that uses satellite data on building heights in combination with a quantitative urban model to estimate a housing supply elasticity in settings where housing prices are not observed. Due to the widespread lack of reliable house price data, we know very little about housing supply elasticities in developing country cities where the majority of the world's urban population resides.

Third, our paper contributes to the large literature that uses satellite data to understand economic phenomena. See Donaldson and Storeygard (2016) for a recent survey of this literature.

Henderson et al. (2011) is a prominent example of the large number of papers that uses night-lights as a proxy for GDP. Burchfield et al. (2006) is an early example of the literature that has used satellite data to determine different forms of land-use. There are only few economic applications of satellite data on building heights with Henderson et al. (2021) being an important example. Our contribution relative to this literature is to show how satellite data on building heights can be combined with a quantitative urban model to estimate key parameters, such as the floor space supply elasticity, and infer floor space and land prices, for which there are few reliable conventional data sources in many developing countries.

Fourth, our paper is related to the growing literature using mobile phone data. Kreindler and Miyauchi (2021) use data on the flow of mobile phones between phone towers in Dhaka together with Google driving times to estimate a gravity equation for commuting flows. They use the destination fixed effects of the gravity equation as an estimate of wages in different parts of Dhaka. Miyauchi et al. (2021) use fine-grained GPS location data from mobile phone apps to estimate commuting and consumption trips in Tokyo and quantify the importance of consumption trips and trip chaining for quantitative urban models. Allen et al. (2020) combine mobile phone data with credit card transaction data for Barcelona to estimate the impact of tourism on different parts of the city. Our contribution to this literature is to demonstrate how mobile phone flow data can be combined with satellite height data to estimate key parameters of an urban model in data sparse settings.

The rest of the paper proceeds as follows. Section 2 presents the model. Section 3 describes the data for Dhaka, and Section 4 discusses the calibration of the model and our estimates of key structural parameters. Section 5 presents the results of the two policy counterfactuals. Section 6 concludes. The Appendix provides additional derivations and a detailed description of the data and additional information on the calibration of the model.

2 Theoretical Framework

In this section, we outline a simple quantitative urban model, which builds closely on the literature on quantitative urban models that has developed in the wake of Ahlfeldt et al. (2015). We consider a city that consists of a discrete number of locations denoted by $n, i \in \mathcal{S} = \{1, 2, \cdots, N\}$.[1] Each location is endowed with a fixed amount of land, K_i. Locations differ in terms of their productivity and residential amenities, which can both be exogenously given or a function of externalities. A transport network connects all locations and determines the commuting times between all locations in the city.

The city is populated by a continuum of ex ante homogeneous workers with a total city population of L. Workers supply one unit of labor and consume both housing, which we refer to as floor space, and a freely tradable final good. The city is embedded in a wider economy that offers a constant level of utility $\mathbb{U}$ to workers. Conditional on moving to the city, workers receive an in-

[1]We typically use the subscript n to refer to the place of residence of a worker and subscript i to refer to her place of work. Where necessary, we also use the subscript k to denote residence locations and j for workplaces.

dependent idiosyncratic taste shock for each combination of workplace and residence in the city. Taking into account this taste shock, wages, house prices and the cost of commuting, workers make optimal location choices in the city. Production of the freely tradable final good is carried out by perfectly competitive firms who use labor, floor space and final goods as material. Their productivity depends on employment density. Free entry implies that all firms make zero profits. Floor space is supplied by a competitive construction sector using both land and freely tradable, which is available at a fixed price that we normalize to one. Firms and residents compete over floor space in each location of the city. Land in the city is owned by land owners, who, for simplicity, only consume the freely tradable final good. The Appendix A presents the details of each element in the model.

2.1 Workers

Workers in the city consume both the freely tradable final good and floor space. We assume that their preferences take the Cobb-Douglas form, so that the indirect utility for a worker ω residing in location n and working in location i is:

$$V_{ni}(\omega) = \frac{b_{ni}(\omega) B_n}{d_{ni}} \frac{w_i}{P_n^{\alpha} Q_n^{1-\alpha}} \tag{1}$$

where w_i is the wage the worker earns at location i, P_n denotes the price of the final good in location n, and Q_n denotes the price of residential floor space in location n. The worker faces a commuting cost between locations n and i of $d_{ni} \geq 1$. We assume that commuting costs are a constant elasticity function of travel time, $\ln d_{ni} = \delta \ln t_{ni}$, where t_{ni} is the time it takes to travel from location n to i in minutes and δ is the elasticity of commuting costs with respect to travel time. Workers enjoy a common residential amenity B_n in location n and receive an idiosyncratic shock to their utility for each combination of workplace and residence in the city $b_{ni}(\omega)$. Intuitively, the indirect utility of a worker depends positively on the real wage that the worker receives, which is the ratio of wages to house price and the price of the final good; negatively on the commuting costs between her residence and work location; and positively on the level of common and idiosyncratic amenities.

We assume that the idiosyncratic utility shocks $b_{ni}(\omega)$ are independent draws from a Fréchet distribution with a cumulative probability distribution:

$$\text{Prob}\Big[b_{ni}(\omega) \leq b\Big] = e^{-b^{-\varepsilon}}, \tag{2}$$

where $\varepsilon > 1$ is the shape parameter of the distribution which regulates the variance of the idiosyncratic utility shock, with larger values of the shape parameter corresponding to a smaller variance of the shocks. After observing the realizations for their idiosyncratic utility draw for each residence and workplace combination in the city, workers choose the residence and workplace combination that maximizes their indirect utility (1). As shown in the Appendix A this results in bilateral commuting probabilities:

$$\lambda_{ni} = \frac{(B_n w_i)^{\varepsilon} \left(d_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\varepsilon}}{\sum_{k\in\mathcal{S}} \sum_{j\in\mathcal{S}} (B_k w_j)^{\varepsilon} \left(d_{kj} P_k^{\alpha} Q_k^{1-\alpha}\right)^{-\varepsilon}}. \tag{3}$$

The numerator of this expression shows that the probability that a worker chooses a particular combination of workplace and residence depends positively on the residential amenities and wages of this combination and negatively on commuting costs, the price of residential floor space and the price of the final good. The attractiveness of any bilateral residence and workplace combination is compared to all other possible such pairs ("multilateral resistance" in the trade literature) in the denominator.

The model also yields a simple expression for the conditional probability of commuting from n to i conditional on living in location n:

$$\lambda_{ni|n} = \frac{(w_i/d_{ni})^{\varepsilon}}{\sum_{j\in\mathcal{S}} (w_j/d_{kj})^{\varepsilon}}. \tag{4}$$

The conditional commuting probability depends only on the wage earned at workplace i and the cost of commuting to this workplace d_{ni} and is independent of the characteristics of a residence location, such as its amenity level B_n or residential floor space price Q_n. The total number of residents R_n in location n is equal the sum of the unconditional commuting probabilities (3) involving this residence location times the total population of the city:

$$R_n = \sum_{i\in\mathcal{S}} \lambda_{ni} L. \tag{5}$$

Similarly the total number of workers L_i in location i is equal to the sum of the unconditional commuting probabilities (3) involving this workplace times the total population of the city:

$$L_i = \sum_{n\in\mathcal{S}} \lambda_{ni} L. \tag{6}$$

The total population of the city (L) depends on the attractiveness of the city relative to the wider economy, which offers reservation level of utility $\mathbb{U}$. We assume that labor supply to the city is a simple constant elasticity function:

$$L = \left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma} \bar{L}, \quad \sigma > 1 \tag{7}$$

where $\bar{L}$ is a constant, σ is the elasticity of labor supply to the city and $\bar{U}$ is the expected utility of moving to the city. This expected utility takes the form:

$$\bar{U} = \bar{\gamma} \left(\sum_{n\in\mathcal{S}} \sum_{i\in\mathcal{S}} (B_n w_i)^{\varepsilon} \left(d_{ni} P_n^{\alpha} Q_n^{1-\alpha}\right)^{-\varepsilon} \right)^{1/\varepsilon}, \tag{8}$$

where $\bar{\gamma} = \Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right)$ is the value of the Gamma function and the expectation is taken over the idiosyncratic utility shocks to each combination of residence and workplace, which are only revealed once a worker has decided to move to the city.[2] For large values of σ, even small increases in the expected utility of moving to the city relative to the utility offered by the wider economy would result in large inflows of workers. In contrast, as σ approaches zero the population of the city is constant.

2.2 Production

In each location of the city, perfectly competitive firms produce a composite final good using labor, floor space and intermediate inputs. Their production function takes the Cobb-Douglas form:

$$Y_i = A_i(L_i)^{\beta}(H_i^B)^{\gamma}(X_i)^{1-\beta-\gamma}, \quad \beta,\ \gamma \in (0,1), \tag{9}$$

where L_i is labor input in location i, H_i^B is floor space used for production in location i, and X_i denotes final goods used as an intermediate input. The expenditure share A_i is a Hicks-neutral productivity shifter that determines the productivity of location i. We assume that the productivity of a location can depend both on externalities and fundamental factors, such as access to a river. In particular, we assume that:

$$A_i = a_i\left(\frac{L_i}{K_i}\right)^{\chi}, \tag{10}$$

where the fundamental advantage in the overall productivity a_i is exogenous. In this specification overall productivity is assumed to increases with the employment density in a location, L_i/K_i, and χ is the elasticity of productivity with respect to employment density and regulated the strength of agglomeration forces in production. As the spatial units in our analysis are relatively large, we abstract from externalities that might spillover from other locations in the city.

Cost minimization of firms combined with free entry and zero profits implies the following relationship between input prices and the output price of firms in each location i:

$$P_i = \frac{\psi}{A_i} w_i^{\beta} q_i^{\gamma}, \tag{11}$$

where q_i is the price of commercial floor space and ψ collects a number of constants.[3] We assume that the final good is freely tradable in the city and hence its price is the same everywhere, and we use it as the numeraire.

[2] This timing assumption implies that workers move to the city based on the expected value of living in the city and hence might prefer to leave the city once their utility draws are revealed. Heblich et al. (2020) relax this assumption and consider a world where workers choose from all locations in the economy simultaneously.

[3] Specifically, the constant is: $\psi = \beta^{\beta}\gamma^{\gamma}(1-\beta-\gamma)^{1-\beta-\gamma}$

2.3 Developers

There are a large number of perfectly competitive developers who combine land and freely tradable capital to produce floor space. Location n in the city is endowed with K_n units of land. In each location, we assume that T_n units of land are available for construction and treat T_n as a policy parameter, which depends on the number of parks, for example, and other open spaces in the city.[4] The floor space produced by developers in each location can be used by residents or firms. We assume that there are no distortions in the allocation of floor space between these two uses and therefore arbitrage requires that the price of residential floor space (Q_n) and commercial floor space (q_n) is the same in each location, $Q_n = q_n$.[5]

The construction costs of a building of height h_n on one unit of land is $\xi(h_n) = \kappa_n {h_n}^{\nu}$, where we assume that $\nu > 1$ and κ_n is a cost shifter that we allow to vary across locations.[6] When developers maximize profits $\pi_n = Q_n h_n - \xi(h_n)$ the height of buildings in location n is:

$$h_n = \left(\frac{Q_n}{\nu \kappa_n} \right)^{\frac{1}{\nu - 1}}. \tag{12}$$

This floor space supply function implies that variation in the height of buildings in the city reflects both the variation of floor space prices and variation in the cost shifters κ_n across locations. As we discuss in detail in Section 4.3, we treat these unobserved cost shifters as structural residuals in our empirical analysis. Note that (12) implies that the elasticity of building height with respect to the common price of floor space (Q_n) is $1/(\nu - 1)$. The total amount of floor space H_n in a location is the height of buildings multiplied with the area available for construction, i.e. $H_n = h_n T_n$.[7] With a large number of developers competing for land in each location, the price of a unit of land r_n will be equal to the equilibrium profits π_n^* that developers make by developing a unit of land:

$$r_n = \pi_n^* = \kappa_n (\nu - 1) \left(\frac{Q_n}{\nu \kappa_n} \right)^{\frac{\nu}{\nu - 1}}, \tag{13}$$

which implies that total land rents received by land owners in location n are therefore $r_n T_n$.

[4] It is natural to assume that the available land area for development is smaller than the total land endowment: $T_n \leq K_n$. In our data, the size of the developed land area is smaller than the total area size in every ward.

[5] The assumption that the price of residential and commercial floor space is equalized relies on there being no locations in which floor space is either entirely used by residents or firms. For applications in developing countries, where we are typically dealing with larger spatial units, such specialization is empirically unlikely and we see both employment and residents in all locations of Dhaka in our empirical application.

[6] As shown in the Appendix A.3, this cost function is derived from a Cobb-Douglas production function using developed land and capital. Combes et al. (2021) have provided evidence for the Cobb-Douglas production function of housing.

[7] For ease of exposition we refer to H_n as the amount of floor space in a location even though it is strictly the volume of housing. We measure the volume of housing in cubic meters in the calibration below. If we assume that floors are typically 3 m high including the floor space, then dividing the estimates of the volume of housing by a factor of three re-scales this volume measure to a measure of floor space.

2.4 Commuter and Floor Space Clearing

To close the model, we introduce two market clearing conditions. Commuter market clearing requires that the number of workers that are working in location i is equal to the number of workers commuting to this location in equilibrium:

$$L_i = \sum_{n \in \mathcal{S}} \lambda_{ni|n} R_n \tag{14}$$

The right-hand side of this equation is the product of the residential population in each location n of the city with the conditional commuting probability (4) from residence n to workplace i.

The second market clearing condition is a floor space clearing condition:

$$\frac{(1-\alpha)\bar{w}_n R_n}{Q_n} + \frac{\gamma}{\beta}\frac{w_n L_n}{Q_n} = H_n \tag{15}$$

The first term on the left-hand side is demand for floor space from residents, while the second term is demand for floor space from firms. The sum of these two sources of demand has to be equal to the supply of floor space H_n. Commercial demand for floor space is a function of the wage bill paid by firms in this location. Demand for floor space from residents depends on the average income of residents $\bar{w}_n$, which is a function of their commuting choices and wages in the entire city. In particular, average labor income of the residents in location n is $\bar{w}_n = \sum_{i \in \mathcal{S}} \lambda_{ni|n} w_i$ where $\lambda_{ni|n}$ is given in (4).

2.5 City Equilibrium and Model Inversion

We are now in a position to define a competitive equilibrium of the city economy:

DEFINITION 1 (City Equilibrium) *Given the set of structural parameters of the city (α, β, γ, ε, δ, ν, χ, σ), a level of utility in the wider economy ($\mathbb{U}$), the constant ($\bar{L}$) in the labor supply function (7) and the constants (κ_n) in the floor space supply equation (12), and fundamentals in each location of the city ($\boldsymbol{a}$, $\boldsymbol{B}$, $\boldsymbol{T}$), and a matrix of travel times between each location in the city ($\boldsymbol{t}$), the general equilibrium of a city is characterized by the total city population L and vectors of wages ($\boldsymbol{w}$), land rents ($\boldsymbol{r}$), floor space prices ($\boldsymbol{Q}$) residential populations ($\boldsymbol{R}$), employment at each workplace ($\boldsymbol{L}$), total floor spaces ($\boldsymbol{H}$) that are determined by a system of seven equations: utility maximization and optimal residential choice of by workers (5); commuter market clearing (14); zero profit condition in production (11); supply of floor space (12); zero profits of developers (13); floor space market clearing (15) and labor supply to the city (7).*

We will use this definition of the competitive equilibrium of the city when we compute model counterfactuals to evaluate the impact of a particular policy. When we do such counterfactual analysis, the multiplicity of equilibria becomes a problem. Intuitively, workers may be concentrated in the location just because their clustering is self-fulfilling. Then, an alternative equilibrium could exist where workers are concentrated in another location. As we extensively discuss in the Appendix B, we establish the existence and uniqueness of the competitive equilibrium in our model.

PROPOSITION 1 (Existence and Uniqueness) *(i) The competitive equilibrium exists. (ii) When $\gamma = 0$, the sufficient condition for the unique competitive equilibrium is $\chi \leq \frac{\beta}{2\varepsilon+1}$. (iii) In other cases, we may have a multiplicity of equilibria.*

The first part states that the existence of the competitive equilibrium is guaranteed for any parameters value and positive values of fundamentals in the model. The second statement implies that, when there is no land input in production, we have a unique equilibrium if either the agglomeration force (χ) is sufficiently small or the value of the parameter of Fréchet distribution (ε) is sufficiently small, which implies that there is a large variation in the idiosyncratic shocks to location choices. In general, however, we may have a multiplicity of equilibria. The sufficient condition for the uniqueness in the general case is the gross substitute property of the excess demand function for land, which are more likely to hold when congestion forces dominate the agglomeration forces – the local spillover is small, the parameter of Fréchet distribution is large, expenditure share on floor space ($1 - \alpha$) is large, and cost share on floor space (γ) is large.

To calibrate the model to data, we use a procedure of model inversion. The basic idea is to use the observed values of endogenous variables together with the structure of the model to infer other endogenous variables and the fundamental productivity (A_i) and residential amenities (B_n) of each location in the city. This approach is formalized in the following proposition, which builds closely on similar results in Ahlfeldt et al. (2015):

PROPOSITION 2 (Inversion) *Conditional on the city-wide structural parameters (α, β, γ, ε, δ, ν), data on employment ($\boldsymbol{L}$), residents ($\boldsymbol{R}$) and built-up area ($\boldsymbol{T}$) in each location of the city, and travel times ($\boldsymbol{t}$) between all locations in the city, we can solve for unique values of all other endogenous variables and the unobserved productivity ($\boldsymbol{A}$) and residential amenities ($\boldsymbol{B}$) in each location. Furthermore, given an elasticity of productivity with respect to density (χ) and the land area of each location ($\boldsymbol{K}$), we can uniquely decompose productivity ($\boldsymbol{A}$) into the contribution of spillovers and unobserved location fundamentals ($\boldsymbol{a}$).*

A proof of this proposition is contained in the Appendix C. This proposition implies that we can solve for all other endogenous variables and obtain unique values for the unobserved vector of productivity ($\boldsymbol{A}$) and residential amenities in each location ($\boldsymbol{B}$) if we observe level of employment, the number of residents in each location and travel times between each of these locations. Note that this result holds also if we are in the parameter range where the model has multiple equilibria.

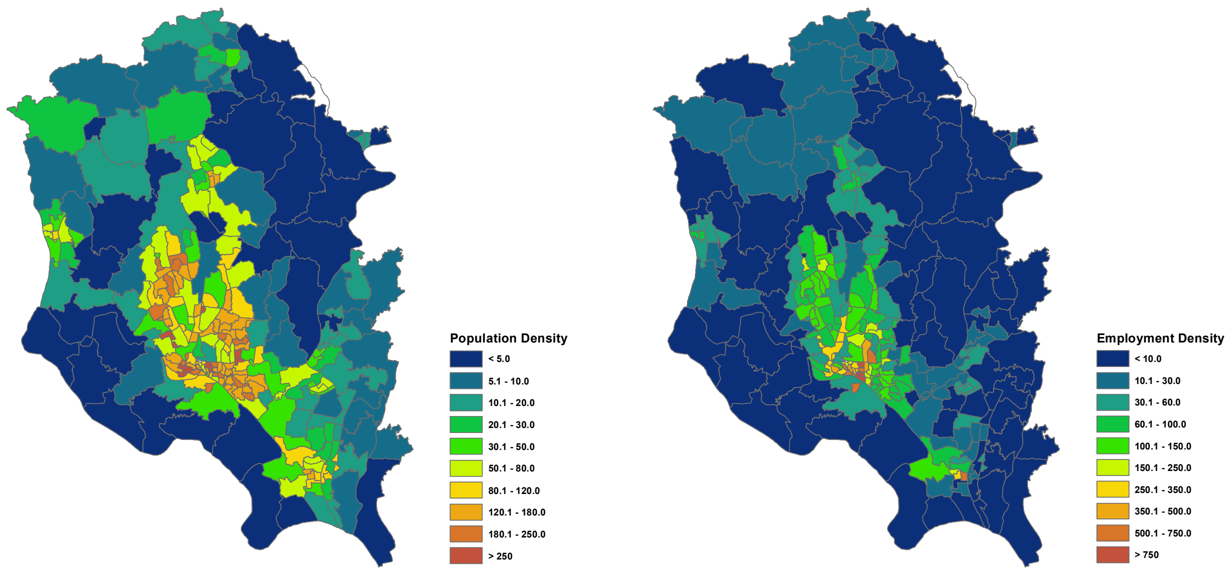
3 Data for Dhaka

In this section, we describe the cross-section of data required for the calibration of the model. The basic population and employment data have previously been used in Bird et al. (2018) and Bird and

Venables (2019). In addition to this data we use a number of additional types of data to inform key parameters of the model. In particular, the cell phone data exploited in Kreindler and Miyauchi (2021) informs our choice of the commuting parameters of the model, and the new satellite data on buildings enables us to estimate the floor space supply elasticity. Additional information about the data that we are using is contained in Appendix D.

Area and spatial units. The first component of our data is a definition of the metropolitan area of Dhaka and a partition of this area into a set of locations. We follow Bird and Venables (2019) in our definition of the metropolitan area of Dhaka. They partition the metropolitan area into 266 wards ("unions") which are based on the Bangladesh Population and Housing Census 2011. The total area of the city is approximately 1465 km^2, and the size of the wards varies considerably. The central part of Dhaka is divided into many smaller wards, while wards on the outskirts of Dhaka are typically much larger. The smallest ward is just 0.023 km^2, while the largest ward is 48.37 km^2, and the average size is 5.51 km^2. The area covered by this definition of Dhaka should contain most of the functional city of Dhaka, as many of the peripheral wards have very low densities of employment and residents.[8]

Figure 3.1: Population and Employment Density in Dhaka



(a) Population Density (person per hectare) **(b)** Employment Density (person per hectare)

Note: This figure shows the population and employment density across 264 wards of Dhaka. Density is expressed as persons per hectare ($ha = 0.01\text{km}^2$). The population data comes from the 2010 Population Census for Dhaka while the employment data comes from the 2013 Employment Census for Dhaka.

[8]The spatial disaggregation of Dhaka that we are using should be sufficiently detailed for most interesting policy questions. A disaggregation of the city into smaller spatial units quickly increases the computational cost of the model and may also increase the measurement error in the data for a developing country city.

Population and employment. Dhaka has experienced rapid population growth since the partition of India in 1947, and it continues to attract new population at a rapid pace. The second component of our data is counts of employment at workplace, which we often refer to as employment for simplicity, and employment at the place of residence, which we often refer to as residents for simplicity, for each location in the city. We use data on the population from the 2010 Population Census while the employment data comes from the 2013 Employment Census and both of these datasets have previously been used in Bird and Venables (2019). To convert population counts to employment at the place of residence, the population data is scaled down with a constant labor force participation rate. In this scaling, we implicitly assume that everyone working in the city also has a place of residence in the city and does not commute to work across the outer boundary of our definition of Dhaka. Given the low density of the districts on the periphery of our definition of Dhaka this should be a plausible approximation.[9]

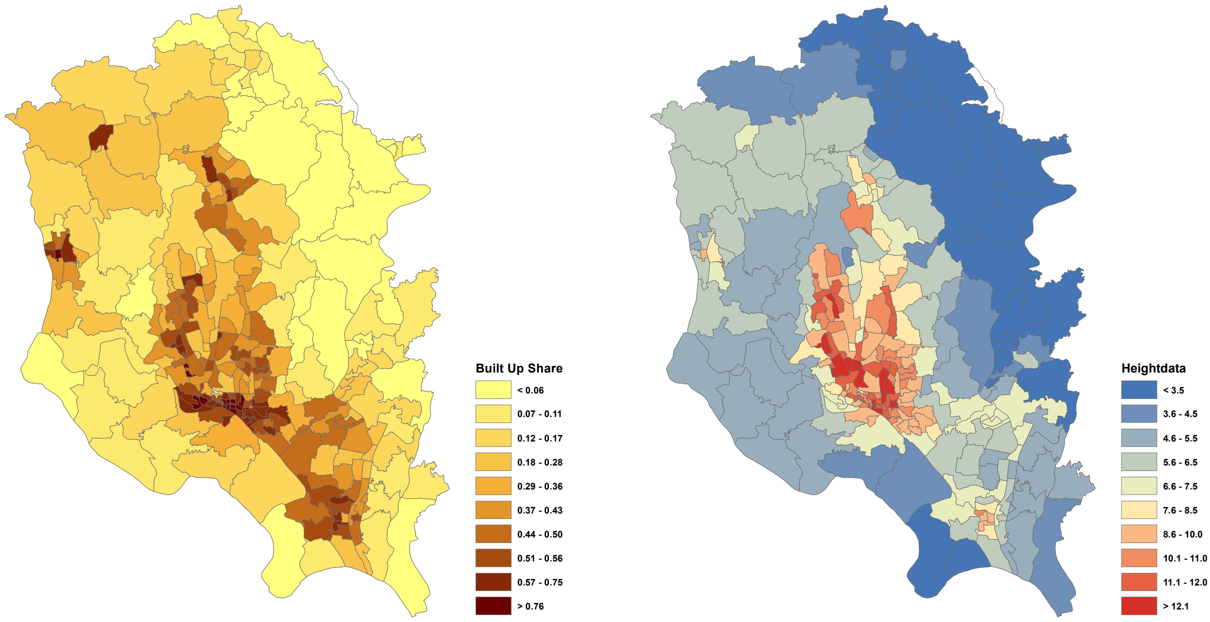
Figure 3.1 shows the density of population and employment across the wards of Dhaka.[10] The figure shows that both population and employment density vary sharply across locations within the city. In line with data for other cities, the variation in employment density is substantially larger than the variation in population density. Employment density is highest in the old center of Dhaka and there is also an employment spike in the south of the city close to the harbor area. Population density is highest in areas that are surrounding the highest employment density wards. Figure D.2 in the Appendix shows the variation in population density and employment density as a function of distance from the center of the city. These figures show that close to the city center employment density exceeds population density.

Satellite data on heights and built-up area. The third component of our data is satellite data on the built-up area and the average height of buildings in all wards of Dhaka from the German Aerospace Center (DLR). The data is based on their World Settlement Footprint 3D product (WSF-3D). The original satellite data, the 12m WSF-3D layer, provides information about the height of all buildings in 12 meter $\times$ 12 meter pixels. Using the height data of all 12 meter WSD-3D pixels that belong to a building structure in a ward, the average height of buildings at the ward level is computed. We also obtain the total built-up area within the ward. Based on these two measures, the total volume of all buildings within a ward is the product of the total built-up area and the average building height in the ward. Figure 3.2 shows both the built-up area and average height estimated by the DLR for each ward of Dhaka. The figure shows substantial variation both in the height of buildings and the share of the built-up land area. In large parts of the periphery of the metropolitan

[9]We use 264 of the 266 wards in our analysis as one very small ward in the center has neither employment nor population and one ward on the periphery of Dhaka has inconsistent data on built-up area, employment and population.

[10]The population, employment and built-up area data for a small ward on the north-east fringe of the metropolitan area of Dhaka has a ratio of employment and population to built-up area that is two orders of magnitude higher than in any other ward. We therefore exclude this ward, which has low levels of population and employment, from the analysis and show this ward as a white area on all maps. There is also a very small ward in the center of Dhaka which has neither employment nor population and this area is also shown as a white area on the maps. This leaves us with 264 wards for our analysis.

Figure 3.2: Satellite Data on Build up Area and Height of Buildings



(a) Share of Developed Land in the DLR data

(b) Average Height in the DLR data (meters)

Note: This figure shows the share land in each ward that is built-up and the average height of building as estimated by the DLR from satellite images.

area of Dhaka, less than 6 percent of the land area of wards is covered by buildings, while in several central wards over three-quarters of the land area is built-up. The variation in average heights is similarly striking, with average heights of less than 5.5 meters in much of the metropolitan area with much larger heights of buildings in the very center of Dhaka. Building heights at the 90th percentile of the height distribution are 11.8 meters, while those at the 10th percentile are just 3.6 meters.

Commuting time The fourth component of our data is information on travel times between each ward in our data. To estimate these travel times, we use the data of Kreindler and Miyauchi (2021), who compute the car travel times between 1,859 cell phone towers in Dhaka using the Google Maps API.[11] We have merged the locations of these mobile phone towers into our wards and computed the travel time between wards n and i as the average travel time across all pairs of mobile phone towers in these two wards. There are a small number of wards that do not contain any cell phone towers. For these wards, we have selected the three closest cell phone towers to the ward and use these to compute travel times to and from these wards. Finally, we need to make an assumption about the own travel time in each ward and assume that it is 5 minutes. As discussed further in subsection 4.2 below, this assumption of travel times generates an average own commuting probability of 0.366 in the baseline calibration, which is in line with the own commuting probability in the 2001 census for

[11] We are very grateful to Gabriel Kreindler and Yuhei Miyauchi for sharing this data along with the coordinates of the cell phone towers with us.

the Greater London Area using local authorities as to the spatial units. If data for own commuting flows in each ward in Dhaka was available, this information could be used to calibrate a different own travel time for each ward in Dhaka.[12]

Cell phone data. Kreindler and Miyauchi (2021) use cell phone data for Dhaka to infer commuting flows between different mobile phone masts in Dhaka. They use this data together with the data on car travel times discussed above to estimate the spatial decay of commuting flows in Dhaka. They also report a number of checks on the representativeness of the cell phone commuting flows. We use the results of their analysis to specify the structural parameters governing commuting behavior as discussed in detail in subsection 4.2 below.

4 Calibrating the Model on Dhaka

In this section, we discuss how we calibrate the model using our data for Dhaka. In Section 4.1 we summarize the key structural parameters of the model and provide an intuitive discussion of what data we are using to inform the value of each of these parameters. The following subsections go step-by-step through the calibration of the model. In Section 4.2 we discuss how the estimates of Kreindler and Miyauchi (2021) for commuting costs in Dhaka using cell phone data can be used to estimate wages in every ward of Dhaka in our model. These wage estimates are combined in Section 4.3 with our satellite data to estimate the elasticity of floor space supply in Dhaka. The estimated floor space supply elasticity is used in Section 4.4 to estimate floor space and land prices in each ward of Dhaka. The estimated wage and floor space prices are used in Section 4.5 to invert the model and back out the level of productivity and residential amenities in each ward of Dhaka. Finally, in Section 4.6 we use a simple moment condition to estimate the elasticity of productivity with respect to employment density. The Appendix C contains additional information on each step of the calibration process.

4.1 Overview of the Calibration

The urban model developed in Section 2 has a total of 8 structural parameters, which are summarized in Table 4.1 below. These parameters can be grouped into four groups of parameters. The first group contains the elasticity of commuting costs with respect to travel time (δ), the shape parameter of the Fréchet shocks to location choices (ε) and the elasticity of the labor supply to the city (σ). The product of the first two parameters determines how quickly commuting flows in the city fall with travel time. Kreindler and Miyauchi (2021) estimate how quickly commuting flows decline with travel time in Dhaka using commuting flows observed in mobile phone movements between cell phone towers. They combine this with evidence on the spatial variation in wages in Dhaka to decompose the overall effect of travel time on commuting flows into the contribution of δ and ε.

[12] In settings where no reliable census data on employment and residents population is available the cell phone flow data used by Kreindler and Miyauchi (2021) and others is an alternative source to estimate such data.

Table 4.1: Structural Parameters of the Model

Parameter	Definition
δ	Elasticity of commuting cost with respect to travel time
ε	Fréchet shape parameter of the idiosyncratic taste shocks to location choices
σ	Elasticity of labor supply to the city with respect to average utility in the city
ν	Elasticity of construction costs with respect to building height
χ	Elasticity of productivity with respect to employment density
α	Consumer spend share α on the tradeable good and $1-\alpha$ on floor space
β	Firm expenditure share on labor
γ	Firm expenditure share on floor space

Note: This table summarizes the structural parameters of the model presented in Section 2.

The elasticity of labor supply to the city captures in a simple reduced form way how open the city is to migration from the wider economy. We set this parameter to central estimates of the spatial labor supply elasticity in the literature.

The second key structural parameter of the model is the elasticity of building costs with respect to height ν, which implies a floor space supply elasticity of $1/(\nu-1)$. We show how this elasticity can be identified using the information on the volume of buildings in each ward of the city from our satellite data. Intuitively, information on quantities, such as the volume of buildings, can be used to infer prices given the structure of demand for floor space by residents and firms in the model. The floor space supply elasticity is the key model parameter that determines whether an increase in demand for floor space leads to higher buildings or whether a city remains low rise and sprawls further out. While there are a number of estimates of the floor space supply elasticity of cities in developed countries, we know little about floor space supply elasticities in the developing world due to the limited availability of reliable price data in these settings.

The third key structural parameter of the model is the elasticity of productivity with respect to employment density (χ), which governs the strength of agglomeration forces in production in the city. If these agglomeration forces are strong, employment in the city will cluster more strongly in the center of the city. We estimate this elasticity using the distribution of productivity in Dhaka implied by the model together with a simple moment condition. We discuss the advantages and disadvantages of this approach relative to the existing literature in Section 4.6.

The final group of structural parameters is the expenditure share of consumers on the tradable good (α) and floor space ($1-\alpha$) and the expenditure share of firms on labor (β), floor space (γ) and final good used as an intermediate input ($1-\beta-\gamma$). To set the expenditure share of consumers on housing, we use information from a household survey for Dhaka. There is no direct evidence of firm expenditure shares for Dhaka and we set the parameters of the production function to central parameters in the literature.

In the following subsections, we exploit the recursive structure of the model to estimate the structural parameters of the model in several steps. Having estimated subsets of the structural parameters, we can use these estimates to infer the values of key endogenous variables such as wages or floor space prices. Finally, we use estimated wages and floor space prices together with the structural parameters of the model to infer productivity and residential amenities in each location of the city.

4.2 Commuting Flows and Implied Wages

The first step of the calibration uses the commuting market clearing condition (14) to solve for wages that are consistent with a competitive equilibrium in the model. This approach was first used in Ahlfeldt et al. (2015). Combining the commuting market clearing condition with the conditional commuting probabilities in (4) and the assumed relationship between commuting costs and travel time, $d_{ni} = (t_{ni})^{\delta}$, results in:

$$L_i = \sum_{n \in \mathcal{S}} \frac{w_i^{\varepsilon} t_{ni}^{-\varepsilon\delta}}{\sum_{j \in \mathcal{S}} w_j^{\varepsilon} t_{nj}^{-\varepsilon\delta}} R_n \tag{16}$$

In our data, we observe both employment in each location (L_i), residents in each location (R_n) and travel times between each location of the city (t_{ni}). We can therefore solve this system of N equations for a unique vector of wages (w_i) in each employment location i if we have values for the two structural parameters governing commuting flows: the shape parameter of Fréchet distribution of location tastes (ε) and the elasticity of commuting costs with respect to time (δ). Intuitively, there is a unique wage vector (up to a multiplicative scalar) that generates commuting flows from each place of residence so that the number of commuters arriving at each work location equals the number of workers that we observe working at this destination in the data.

Kreindler and Miyauchi (2021) use commuting flows observed in mobile phone data for Dhaka to estimate an elasticity of commuting flows with respect to the travel time of approximately -2.5. They decompose this overall effect of travel time on commuting flows using the information on the dispersion of wages in Dhaka into a Fréchet shape parameter of approximately 8.0 and the elasticity of commuting costs with respect to the travel time of 0.31. Using these estimates, figure D.4 in the Appendix shows our estimated wage distribution in Dhaka. In the estimated wages, the ratio between the 95th percentile and the 5th percentile is around 1.53 and the interquartile range is [1.08, 0.92]. Using the estimated wages, we can also compute income per capita for each residential place ($\bar{w}_n$). Figure D.6 in the Appendix shows that locations close to concentrations of employment have relatively high income per capita.[13]

[13] Kreindler and Miyauchi (2021) interpret the destination fixed effects of their commuting gravity estimates as the wage in each location in Dhaka. As their spatial units are not identical to ours, we cannot directly compare their wage estimates to ours. While closely related, both approaches do not necessarily result in the same wage estimates even with identical spatial units. Intuitively, their approach uses actual commuting flows, including bilaterals with zero flows, to estimate a gravity regression. Our approach, which is based on Ahlfeldt et al. (2015), imposes a gravity structure on commuting flows and uses this assumption to solve for wages that are consistent with the distribution of residents and employment.

4.3 Estimating the Floor Space Supply Elasticity

The key innovation of our calibration approach is to use the satellite data on the volume of buildings in each location to estimate the floor space supply elasticity in Dhaka. In a second step, we use the estimated floor space supply elasticity to solve for floor space and land prices in each location of the city. Our starting point to estimate the floor space supply elasticity is to combine the floor space clearing condition (15) with the housing supply function (12), taking into account that $H_n = h_n T_n$, which yields:

$$\frac{(1-\alpha)\bar{w}_n R_n}{Q_n} + \frac{\gamma}{\beta}\frac{w_n L_n}{Q_n} = \left(\frac{Q_n}{\nu \kappa_n}\right)^{\frac{1}{\nu-1}} T_n. \tag{17}$$

The first term on the left-hand side of (17) captures demand for floor space from residents, which depends on the number of residents (R_n), the average residential income in this location ($\bar{w}_n$), the price of floor space (Q_n) and the share of income that residents spend on floor space ($1-\alpha$). Using data from the 2016 Household Income and Expenditure Survey (HIES) for Bangladesh, we set $1-\alpha$ to be 0.25, which is approximately the sum of average household expenditure on housing and rent (17.25%), household effects (3.03%) and lighting and fuel (5.02%) reported in this survey.[14]

The second term on the left-hand side of (17) is demand for floor space by firms, which is a function of the wage bill in a location ($w_n L_n$) and the parameters β and γ, which are the expenditure share of firms on labor and floor space respectively, and the price of floor space (Q_n). We assume that the share of labor in firm expenditure is 0.6 and the share of expenditure on floor space is 0.2. This implies that the share of expenditure on the final good as an intermediate input is 0.2, which is broadly in line with values reported in Table 5 of Valentinyi and Berthold (2008) for the United States. The right-hand side of (17) is the supply of floor space in the model, which is a function of the build up areas (T_n), the floor space supply elasticity ($1/(\nu-1)$), the price of floor space (Q_n) and the constant of the housing supply function (κ_n), which we allow for varying across locations. We treat the built-up area in each location T_n as an exogenous parameter that we take from the satellite data, but in a richer model, this parameter could be endogenously determined.

To estimate the floor space supply elasticity, we use a moment condition on the location specific intercept of the housing supply function κ_n, which we treat as structural residuals of the model. Treating κ_n as a structural residual implies that the model can exactly match the observed distribution of building heights across wards of Dhaka. In particular, by substituting the inverse of the housing supply function (12) into the market clearing condition for floor spaces (17) and replacing the predicted height (h_n) with the observed height (h_n^*) in the data, we can express the structural

[14] Heblich et al. (2020) use a related approach in which they multiply (15) with Q_n and combine this equation with data on the value of land and buildings in each borough of London, i.e., the product of the price of floors space (Q_n) and the volume of buildings (H_n), and use this to estimate wages in each borough of London. While they do not observe the price and volume of buildings separately, we directly observe the volume of buildings in each location from satellite data and use this information to estimate the floor space supply elasticity ($1/(\nu-1)$) and the price of floor space in each location (Q_n).

residual in each location κ_n as:

$$\ln \kappa_n = -\nu \ln h_n^* - \ln \nu - \ln T_n + \ln \left[(1-\alpha)\overline{w}_n R_n + \frac{\gamma}{\beta} w_n L_n \right]. \tag{18}$$

The key idea behind our moment condition to estimate the floor space supply elasticity is to minimize the contribution that variation in the structural residuals (κ_n) makes to explaining the observed height distribution of Dhaka. In particular, we require that across 10 density deciles the contribution of the structural residuals towards explaining the observed height distribution is minimized. We compute these density deciles by adding up the total employment and residents in a ward and divide this sum by the area of the ward. These density deciles are a flexible way of capturing the distance of a ward to the center of Dhaka without having to impose an arbitrary location as the center of the city. Formally, we require that:

$$\mathbb{E}\Big[\mathbb{I}_n(k) \times (\ln \kappa_n - \overline{\ln \kappa})\Big] = 0, \tag{19}$$

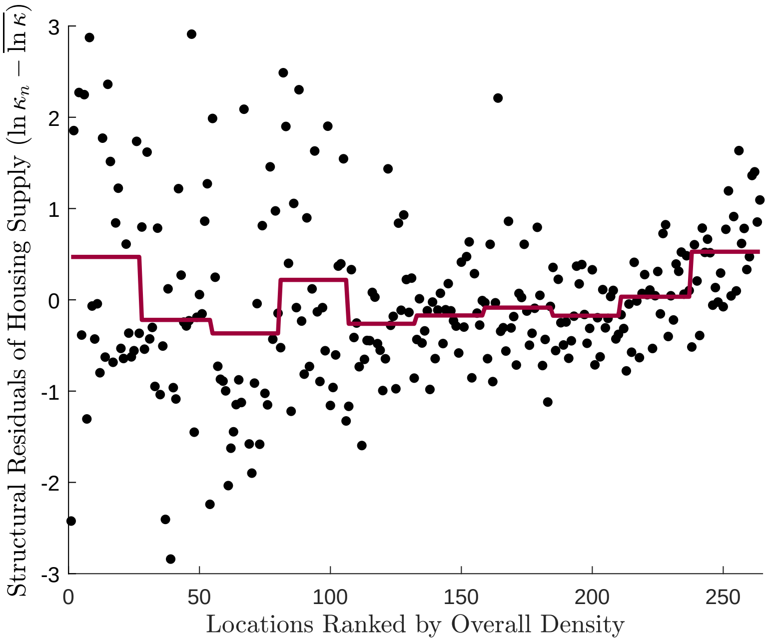
where $\mathbb{I}_n(k)$ are indicator variables for each of the k grid cells and $\overline{\ln \kappa}$ is the average of the logarithm of the structural residuals. Intuitively, this moment condition requires that variation in building heights across our grid cells is explained by variation in the demand for floor space rather than variation in the average level of the structural residuals κ_n. In other words, we assume that variation in average heights across these 10 grid cells is due to shifts in the demand for floor space rather than shifts in the supply curve for floor space due to variation in the average value of κ_n across these grid cells.

We think that this identifying assumption has at least three attractive features. First, the identifying variation in demand for floor space implied by our model is very large. Using the left-hand side of (17), the demand for floor space at the 90th percentile of wards is more than 10 times larger than the demand at the 10th percentile of wards. The vast majority of this substantial variation is due to the highly uneven distribution of employment and population across wards of Dhaka rather than spatial variation in our estimated wages. Second, essentially using the height gradient of the city, which has developed over many decades of construction activity, to identify the floor space supply elasticity ensures that we capture a long-run supply elasticity, and do not rely on a small source of variation that is also often only observed for short time periods. Third, Dhaka is not known for possessing well-enforced urban planning that is able to enforce planning restrictions that vary substantially across locations in Dhaka. This suggests that the large differences in the height of buildings are primarily driven by differences in demand rather than shifts in the supply curve due to regulatory differences that are correlated with our grid cells.

Using this moment condition, we estimate an elasticity of building costs with respect to building height (ν) of 1.69, which implies a floor space supply elasticity ($1/(\nu-1)$) of 1.45. This elasticity is somewhat smaller than the average housing supply elasticity that Saiz (2010) estimates across US cities but larger than his estimates for supply constrained large cities, such as San Francisco and

New York.[15] It is important to bear in mind that our estimated floor space supply elasticity captures both technological and regulatory factors. The floor space supply elasticity in Dhaka could be lower than the US average due to higher costs of building tall structures in developing countries such as Bangladesh or due to restrictions on the supply of tall structures in Dhaka. We return to this question when we consider model counterfactuals in Section 5 below.

Figure 4.3: Examining the Housing Supply Moment Condition



Note: Each observation is the logarithm of the estimated cost shifter of the floor space supply function in a ward (κ_n) minus the overall average for this parameter across all wards. The red line is the average value of this difference in each of the 10 cells used by the moment condition. Total density is the sum of employment and residents divided by the overall area of a ward.

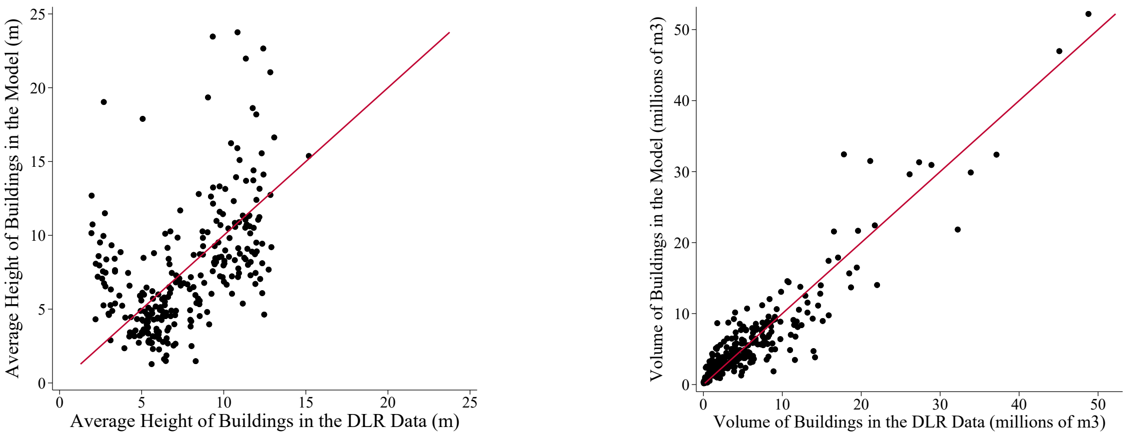
Figure 4.3 examines the fit of our housing supply model, i.e. the ability of the model to explain the variation of heights in Dhaka with a constant floor space supply elasticity. Each point on the graph is the logarithm of the estimated cost shifter of the floor space supply function (κ_n) in a ward minus the overall average for this parameter across all wards. The red line is the average value of this difference in each of our 10 grid cells. If the red line was a straight line at zero, this would imply that our model can perfectly explain the height variation in Dhaka across grid cells with our constant elasticity floor space supply function, without any variation in the average value of the structural residuals across grid cells. In addition, if all points were to lie on top of this red line, this would imply that the constant floor space supply elasticity can also perfectly explain all the variation in building heights within each grid cell, which is highly unlikely given the many idiosyncratic factors not captured by our model that could influence the average height of buildings in a particular ward.

Figure 4.3 has at least two striking features. First, the variance of the structural residuals (κ_n) is

[15]Saiz (2010) estimates a population weighted average housing supply elasticities of 1.75 for U.S. MSAs, while the estimated housing supply elasticity of New York is just 0.76 and 0.66 for San Francisco.

much higher in the lower density wards of Dhaka. One explanation for this finding is that measurement error in the satellite data is plausibly additive, suggesting that the percentage measurement error in heights is much more pronounced at lower heights than for taller buildings. Another possible explanation for this finding is that our population and employment come from earlier years than the satellite data. Rapid growth in employment or population particularly at the periphery of the city could have added further measurement error to our estimates at lower densities. Second, while the red line is not a perfectly straight line, it is close to zero for much of the upper part of the density distribution. However, in the highest grid cell, the average structural residual is somewhat more positive, which implies that the model needs a leftward shift in the supply curve to explain the lower heights observed in the data relative to what a constant elasticity supply curve would suggest.[16]

Figure 4.4: Average Height and Volume of Buildings: Data versus Model



(a) **(b)**

Note: The left-hand panel compares the average height of buildings in meters in the satellite data to the average height predicted by the model when we set the cost shifter of the floor space supply function (κ_n) to its average value. The right-hand side panel compares the volume of the housing stock in millions of cubic meters in the satellite data to the volume predicted by the model again setting (κ_n) to its average value. The red lines are 45-degree lines.

Figure 4.4 is a complementary approach to examining the fit of our housing supply function. The left-hand side panel compares the height of buildings across wards in the satellite data to the model predicted heights when we use the best fit value of the floor space supply elasticity and set the shifters of the floor space supply function κ_n to their average value in Dhaka. While the correlation between the satellite height data and the model-predicted heights is not perfect, the two are clearly strongly correlated and locations lie close to the 45-degree line. The correlation coefficient between the height predicted by the model and the height in the data is 0.496. Similar to Figure, 4.3, the figure shows that the correlation between the model heights and heights observed in the data is lower at

[16]The finding that the floor space supply elasticity in the very center of Dhaka is plausibly smaller than in other parts of the city is consistent with the finding of Baum-Snow and Han (2021) who find that the floor space supply elasticity of US cities increases with distance to the center.

lower densities. Furthermore, the model predicts higher buildings in the very center of Dhaka if we do not allow variation in the supply shifter κ_n. The right-hand panel of Figure 4.4 repeats this exercise for the volume of buildings. Here the correlation between the model and the data is much higher. Note that we use the built-up area in the data also in the model. The same built-up area therefore enters the volume calculations on either axis of the graph.

To explore the robustness of our estimate of the floor space supply elasticity, we examine how sensitive this estimate is to the value of ε, which controls the variance of the idiosyncratic shocks to combinations of the workplace and residential locations. When we set ε to be 3.0, which is in the middle of estimates by Tsivanidis (2020), we estimate a floor space supply elasticity of 1.37. If we instead set ε to 6.0, which is close to the value estimated by Heblich et al. (2020), we obtain a floor space supply elasticity of 1.43. While variation in the value of ε changes the estimated distribution of wages in the city considerably, this only changes the estimated floor space supply elasticity marginally because much of the variation in demand for floor space is due to variation in the distribution of employment and residents and not variation in wages across wards.

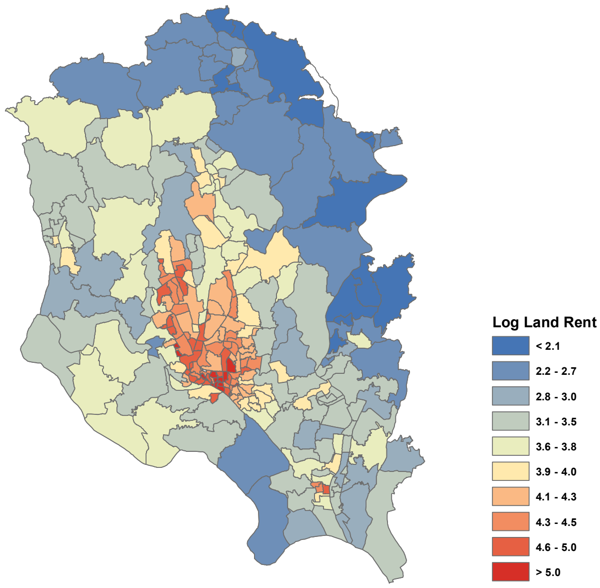
4.4 Estimating Floor Space and Land Prices

We now turn to estimating the floor space and land prices consistent with market clearing in the housing market. In principle, we can estimate the price of floor space in each ward of Dhaka without knowing the floor space supply elasticity, which we have estimated in the previous subsection. The left-hand side of the housing market clearing condition (15) is a function of the distribution of population and employment across wards and the expenditure shares in the utility and production function. If we substitute the volume of housing in each ward observed in the satellite data on the right-hand side of (15), we can solve this system of equations for floor space prices in each ward. With these floor space prices, we can, in turn, also solve for the implied land prices using the zero profit condition of developers (13).

However, using this approach results in often very noisy estimates of floor space and land prices, particularly in the lower density periphery of Dhaka. The reason is the likely larger measurement error in the satellite data in lower density areas and the mismatch in timing between our key data components that we have already discussed in the context of Figures 4.3 and 4.4 above. Suppose, for example, that the satellite data underestimates the height of buildings in a location substantially. The model rationalizes the low heights with a large value of the shifter of the floor space supply function κ_n) resulting in a leftward shift of the housing supply function. This implies that floor space prices will be very high in this ward to depress the floor space demand of both residents and workers that reside in this ward according to the census so that their demand matches the mistakenly low estimate of floor space supply.

To overcome this problem and essentially smooth our estimated floor space and land price estimates, we use our estimated floor space supply elasticity. In particular, we assume that the housing supply function in each ward has our estimated floor space supply elasticity and the average value

Figure 4.5: Estimated Land Rents



Note: This map shows the logarithm of the estimated land rents that result from imposing our best fit floor space supply elasticity on the model.

of κ in the grid cell in which the ward is located. Our estimate of the floor space price in a ward is the floor space price at which this housing supply function supplies buildings of the height that we observe in the satellite data. Using the zero profit condition of developers (13), we can then also estimate the implied land prices. This approach is similar to using the predicted values of a regression line rather than the actual observations to separate the underlying relationship from random measurement error.[17]

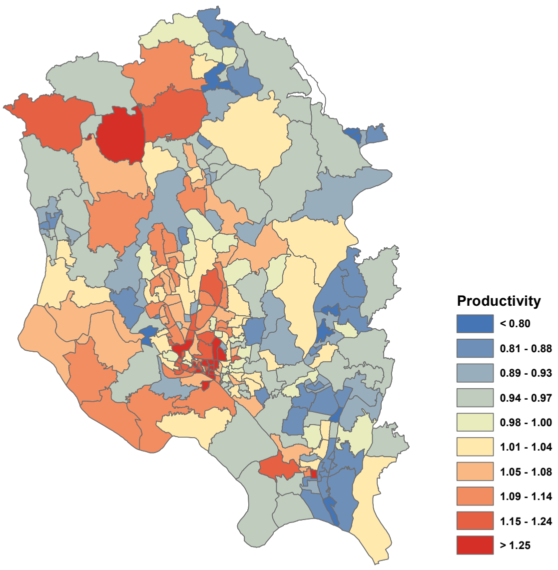
Figure 4.5 shows the land prices that we estimate for each ward in Dhaka using this procedure. The figure shows substantial variation in land prices that is closely correlated with employment and population density in the city. The estimated land prices vary by a factor of 8.5 between the 10th and 90th percentile of the distribution. In contrast, the estimated variation in floor space prices only varies by a factor of 2.8 between the 10th and 90th percentile. The much smaller variation in floor space prices than in land prices highlights the role of building height as a margin of adjustment to respond to increases in the demand for floor spaces. In the Appendix, Figure D.11 displays the land rents as a function of distance to the center, which suggests that Dhaka is largely a mono-centric city.

[17]The floor space prices estimated in this way combined with the number of residents and firms imply a level of demand for floor space from (15). This level of demand will now no longer exactly match the supply of floor space in a location as we have constrained the variation in the supply shifter κ. To make the calibration internally consistent for model counterfactuals, we adjust the built-up area in each ward until the floor space clearing condition (15) holds again.

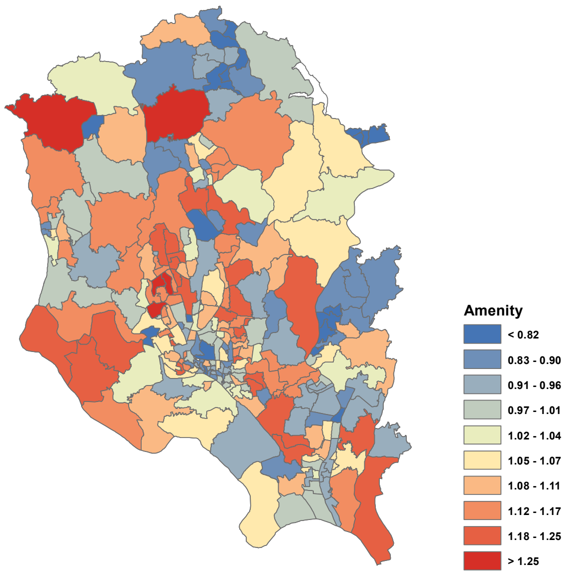
4.5 Estimating Productivity and Amenities

Having estimated wages and floor space prices in each ward of Dhaka, we are in a position to use the results of Proposition 2 to invert the model to determine the productivity and residential amenity level of each ward that are consistent with the observed equilibrium and the structural parameters of the model. Intuitively, in a zero profit equilibrium, firm productivity must be high in locations where firms pay high wages and floor space prices. Similarly, given spatial arbitrage residents in locations with high floor space prices and low levels of commuting market access to work locations must be enjoying high residential amenities.

Figure 4.6: Productivity and Residential Amenities in the City



(a) Workplace Productivity



(b) Residential Amenities

Note: The left-hand panel shows the estimated level of productivity (A_i) in each ward and the right-hand panel shows the level of the estimated residential amenities (B_n).

The left-hand panel of Figure 4.6 displays the level of productivity in each ward of Dhaka implied by our model.[18] We find that productivity is high in central locations, such as in the central business district area (Motijheel) and in areas with a dense cluster of businesses (Gulshan and Tejgaon). Productivity declines dramatically as we move away from the city center. There are also several notable clusters with higher productivity on the outskirts of Dhaka. Among them is the city of Naratanganj in the south of Dhaka, which is the oldest port in Bangladesh and the hub of the country's textile industry. Productivity is also high in the north of the city where high technology industries have located, such as in Gazipur and Kaliakair. The variation in productivity across wards is substantial. The ward at the 90th percentile of the productivity distribution has around 36 percent

[18]Our estimates of overall productivity in a ward differ from the existing literature in that we have estimated both wages and the price of floor space while the existing literature typically observes proxies for floor space prices. Our productivity estimates therefore leverage the structure of the model even more than the existing literature.

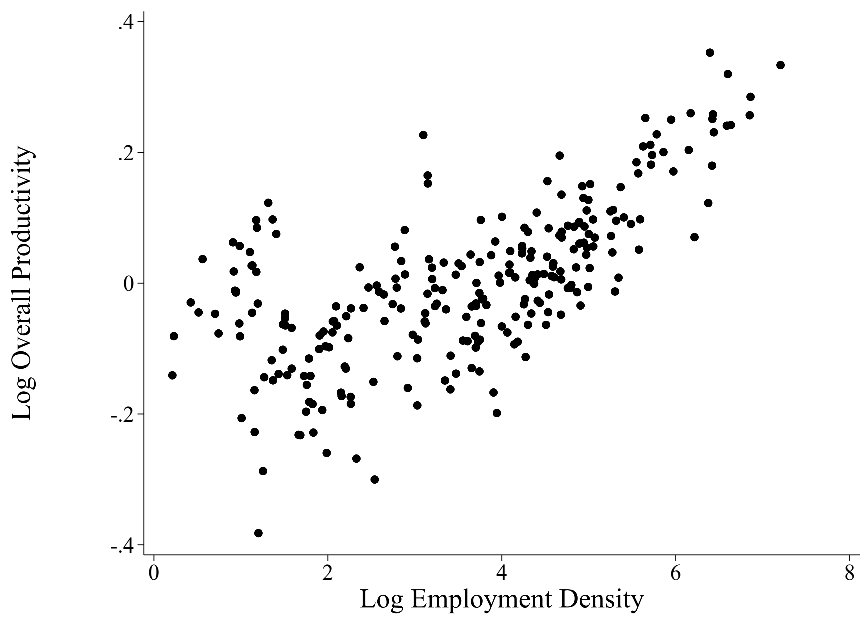
higher productivity than the ward at the 10th percentile.

The right-hand panel of Figure 4.6 displays the level of residential amenities in each ward of Dhaka implied by our model. We find that residential amenities are not highest in the locations where the productivity distribution peaks. The highest residential amenities are often adjacent to areas of high productivity. The variation in residential amenities is even larger than the variation in productivity, with the ward at the 90th percentile having 40 percent higher amenities than the ward at the 10th percentile.

4.6 Estimating the Elasticity of Productivity with Respect to Density

The overall level of productivity in each ward that we have estimated in the previous subsection is a combination of the fundamental productivity advantages of a location and spillovers as captured by (10). How strong productivity spillovers depend on the value of χ. Figure 4.7 shows the correlation between the logarithm of the overall level of productivity A_i and the logarithm of employment density in each ward. The figure reveals a suggestive log linear relationship between the overall level of productivity in a ward and the employment density in the ward.

Figure 4.7: Productivity and Employment Density



Note: Graph shows the correlation between the logarithm of the model estimated productivity in each ward and the employment density in this ward.

To separate how much of this variation in overall productivity is driven by spillovers and variation in fundamental productivity, the existing literature has used changes in employment density triggered by exogenous shocks such as the construction of new transport infrastructure in Tsivanidis (2020). With only one cross section of data and no obvious exogenous shock, we will use a simpler moment condition to separate the contribution of fundamentals and spillovers. In particular we assume that the fundamental level of productivity a_i is uncorrelated with the same 10 grid cells that we have used to estimate the floor space supply elasticity. In other words, this moment condition imposes that changes in overall productivity across these grid cells are driven by spillovers rather than variation in fundamentals.

While this is a strong assumption to identify the role of spillovers, we think it has at least three attractive features. First, this condition can be used in data sparse environments where we do not have access to panel data. Second, Dhaka lies on a largely featureless plane and it is difficult to point to natural features of the geography that could make the central parts of Dhaka inherently more productive than more peripheral areas. Third, the condition uses cross-sectional variation, which plausibly reflects a steady state of the city. Using changes over a relatively short period of time as often used in this literature, makes it much harder to argue that we are observing changes between two steady states.

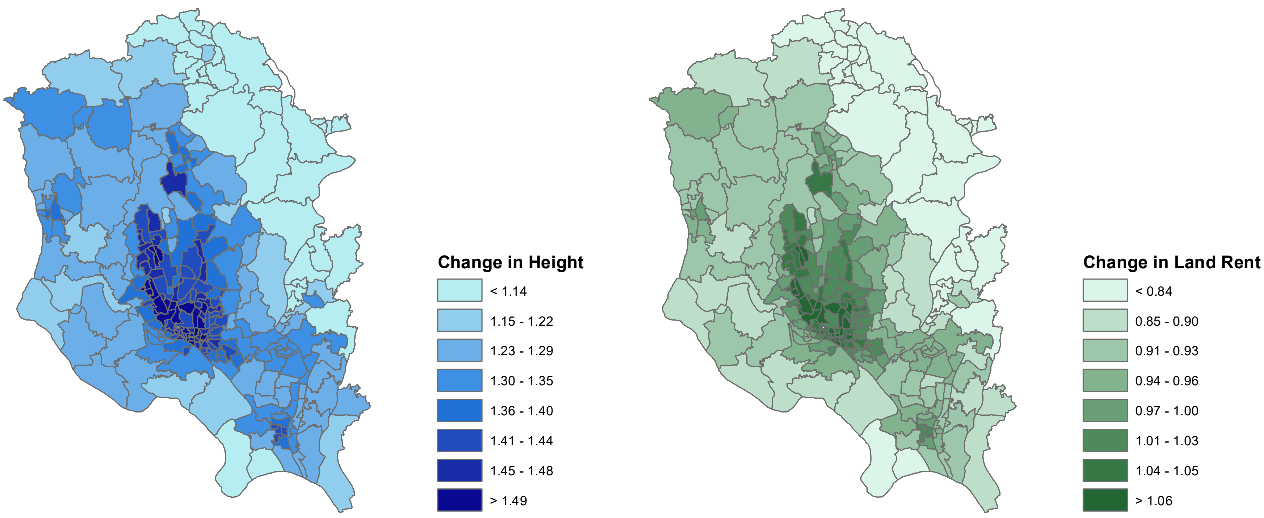
Using this moment condition, we estimate an elasticity of productivity with respect to employment density (χ) of 0.045. This estimate lies in the (0, 0.08) range for the elasticity of labor productivity with respect to density reported in the recent meta analysis in Ahlfeldt and Pietrostefani (2019). This estimate is substantially smaller than the elasticity of productivity with respect to the employment estimated by Tsivanidis (2020) for Bogota. If agglomeration forces in Dhaka were as large as these estimates suggest, fundamental productivity a_i would have to decline rapidly as one approaches the center of Dhaka, which seems counterintuitive.

5 Counterfactuals

Having estimated the key structural parameters of the model using data from Dhaka we now illustrate the usefulness of the model for policy analysis with two model counterfactuals. In particular, we consider the effects of two changes: (i) an increase in the floor space supply elasticity and (ii) the building of a north-south road through the city. The floor space supply elasticity reflects both the state of the local building technology and local geological conditions, as well as regulatory constraints on the construction of new buildings. Our first counterfactual explores the effects of a uniform increase in the floor space supply elasticity by 25 percent. The second counterfactual investigates the effects of a traditional infrastructure policy, the construction of a new north-south radial road through Dhaka, which reduces travel times along this corridor.

For both counterfactuals we solve for a new steady-state equilibrium of the model holding all other parameters of the model at their baseline level. If productivity and residential amenities in the city are exogenous or agglomeration forces in production are below the sufficient condition characterized by Proposition 1, the model has a unique equilibrium. Outside this parameter range, the model can exhibit multiple equilibria. For the estimated parameter values we have not been able to find evidence of multiple equilibria using different starting values. In these counterfactuals, we can either hold the total population of the city constant or allow the size of the city to adjust endogenously. We concentrate on the case of an open city where the total population of Dhaka changes according to (7) with an assumed labor supply elasticity to the city of $\sigma = 2$. We also mainly concentrate on the case where productivity is endogenous and set the elasticity of productivity to employment density to the value of 0.045 that we have estimated in Section 4.6.

Figure 5.8: Increasing the Building Supply Elasticity in the City: Height and Land Rent



(a) Change in Average Height

(b) Change in Land Rent

Note: The left-hand panel shows the counterfactual changes in average height of buildings relative to the DLR data and the right- hand panel shows those in estimated land rent relative to the baseline.

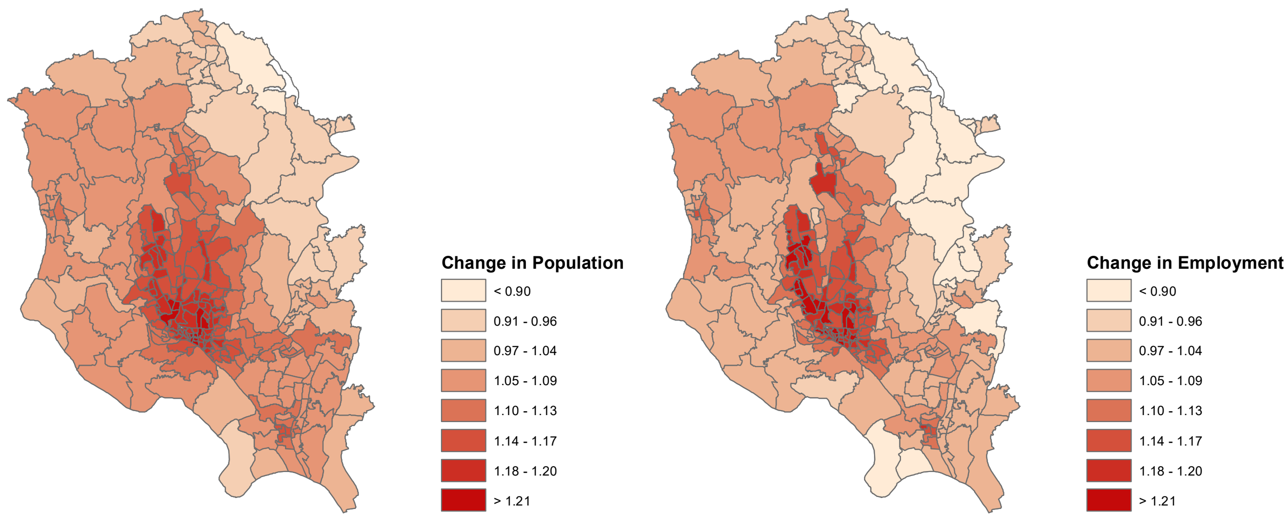
5.1 Increasing the Floor Space Supply Elasticity

In our first counterfactual we increase the floor space supply elasticity by 25 percent relative to the value of 1.45 that we have estimated in Section 4. The floor space supply elasticity depends both on the available building technology, geological conditions in the city and planning constraints. The counterfactual illustrates the outcomes of an increase in the floor space supply elasticity without taking a stand on whether this change would be achieved through an improvement in building technology or regulatory changes. An increase in the floor space supply elasticity is a simple way to capture the idea of "densification" policies, which are popular among many urban commentators and decision makers. The impact of this change is more complex than one would expect and highlights the value of using a general equilibrium framework to evaluate the impact of policy changes in cities.

Figures 5.8 display the counterfactual impact of a 25 percent increase in the floor space supply elasticity on building heights and land rents. Panel (a) shows that building heights increase in all wards of Dhaka in response to this change, but the impact is much more pronounced in the most central wards where average building heights increase by more than 50 percent. Intuitively, even though the floor space supply elasticity increases by the same amount in all locations of the city, this change is much more valuable in locations where buildings are tall already before this change.

In contrast, in low-rise areas of the city making it easier to build taller buildings has little impact. Panel (b) shows that land rents in the periphery of Dhaka fall in response to this change, while land rents in the center increase. To assess the average impact of this change on land owners we compute the change in total land revenue ($\sum_n r_n T_n$) in the city, which falls by 3.0 percent. Intuitively, an increase in the floor space supply elasticity reduces the importance of land in the housing production function and leads, on average, to a fall in the demand for land. This occurs despite the fact that the total city population increases significantly in this counterfactual.[19]

Figure 5.9: Increasing the Building Supply Elasticity in the City: Population and Employment



(a) Change in Residents **(b)** Change in Employment

Note: The left-hand panel shows the counterfactual changes in population, and the right-hand panel shows those in employment relative to the baseline data.

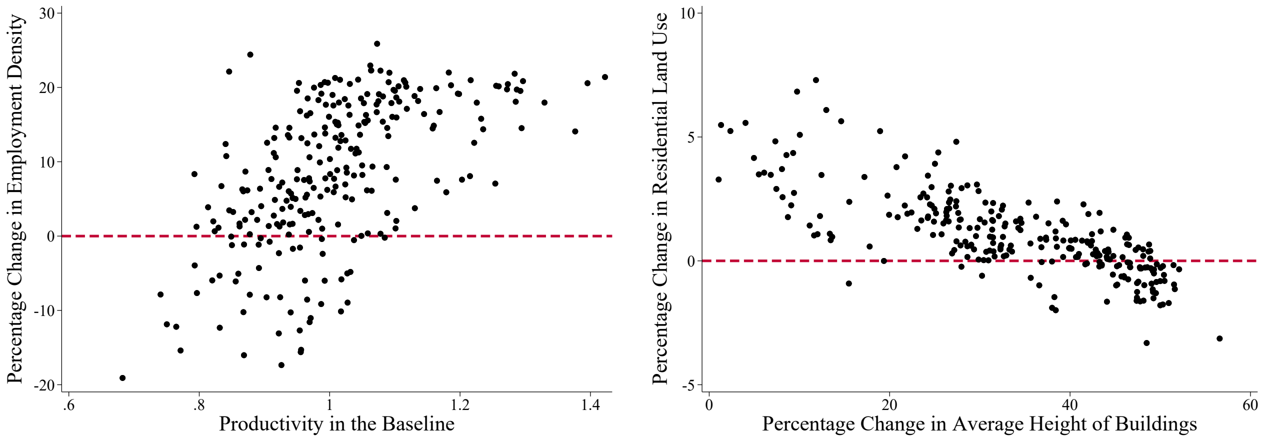
Figure 5.9 displays the impact of a 25 percent increase in the floor space supply elasticity on the distribution of residents and employment. Panels (a) and (b) of the figure show that residents and employment both increase by more than 20 percent in the central wards of Dhaka and decrease in the peripheral wards. The decrease in residents and employment in the periphery of the city occurs even though the total population of Dhaka increases by 10.1 percent in this counterfactual.

Figure 5.10 and 5.11 unpack the general equilibrium impact of an increase in the floor space supply elasticity further. Panel (a) of Figure 5.10 shows that increases in employment density are strongly correlated with the productivity of wards in the baseline, suggesting that higher density allows more workers to work in high productivity locations of the city. Panel (b) shows that in locations that experience greater increases in the average height of buildings (i.e. more central wards of Dhaka) the share of land used for residential purposes declines. Despite the large increases in the number of residents in central wards of the city in the counterfactual, there is an even larger increase in employment in these locations. As a result the center specializes further in being a place

[19]Consistent with this we find that total land revenue falls by 13.2 percent if we compute the same counterfactual in the closed city version of the model.

of employment while more peripheral wards become more residential. Intuitively, central locations are very productive but do not have equally high residential amenities. As density in the center increases due to higher average building heights, more workers are within commuting distance from the most productive locations of the city making these locations even more attractive for firms to locate in. With endogenous productivity, this increased employment density further increases productivity, in the process enhancing the attractiveness of these locations for employment further.

Figure 5.10: Examining the Employment and Residential Land Use Changes



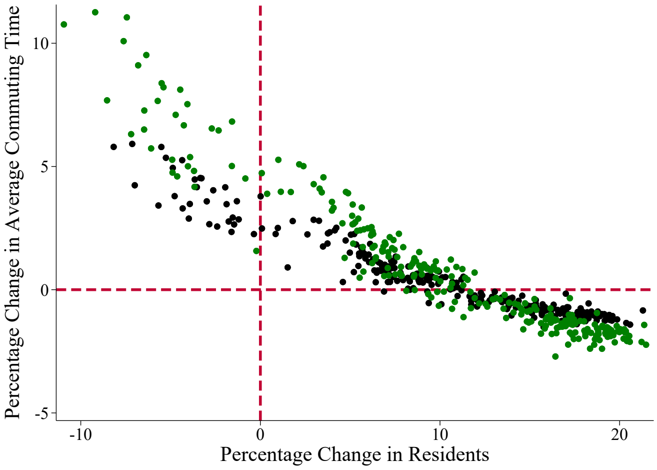
(a) Changes in Employment

(b) Changes in Residential Land

Note: Each point shows the change in a ward in response to a 25 percent increase in the floor space supply elasticity. The left-hand panel shows the relationship between the percentage change in employment density in a ward and its productivity (A_i) in the baseline. The right-hand panel shows the relationship between the percentage change in the share of land used for residential purposes and the percentage change in the average height of buildings.

Figure 5.11 shows the implications of the pattern visible in Figure 5.10 for the change in average commuting in the counterfactual. Figure 5.11 considers both the case where productivity is exogenous (black dots) and the case where productivity is a function of the employment density (green dots). The figure shows that while average commuting times decline in the counterfactual for residents of wards that gain residents (which are wards close to the center of Dhaka), the opposite is true in more peripheral wards. If the peripheral wards become more residential while the central wards become more specialized in employment, some residents in the peripheral wards will have to commute longer distances to work. This pattern holds for both the case of exogenous productivity and endogenous productivity. Due to this asymmetric change in average commuting times across wards, average commuting times in the city remain largely unchanged between the baseline and the counterfactual with the average commuting time weighted by the number of commuters increasing by 0.84 percent. These results highlight that policies based on "densification" of cities to reduce the need for commuting may not be successful once the full general equilibrium impact of higher density is taken into account.

Figure 5.11: Changes in Average Commuting Times of Residents



Note: This figure shows the relationship between the percentage change in average commuting times of residents in each ward and the percentage change in population. The black dots show the results of the model counterfactual with exogenous productivity (i.e., productivity A_i is fixed), and the green dots show the results of the model counterfactual with the estimated agglomeration forces in production ($\chi = 0.045$). For the latter, the 95 percentile of the average commuting time change is 7.10 percent, the average is 0.84 percent, and the 5 percentile is -1.99 percent.

We now turn to the welfare effects of the economic changes triggered by an increase in the floor space supply elasticity. Our welfare measure for workers is their expected utility of residing in the city $\bar{U}$ defined in (8). Combining the unconditional commuting probabilities (3) with the expected utility of residing in the city (8) we can express the expected utility of workers in the city as:

$$\bar{U} = \bar{\gamma} \frac{B_n}{d_{nn}} \frac{w_n}{P_n^{\alpha} Q_n^{1-\alpha}} \lambda_{nn}^{-\frac{1}{\varepsilon}} \tag{20}$$

Therefore, for any location n in the city, welfare of workers is a combination of exogenous location characteristics (B_n/d_{nn}), the real income of workers ($w_n P_n^{-\alpha} Q_n^{-(1-\alpha)}$) and the non-commuting probability (λ_{nn}). Note that this equation implies that while the different components of (20) can vary across locations in the city, expected utility is the same for all locations. Higher values of, for example, residential amenities (B_n) in a location relative to another location have to be exactly compensated by another component on the right-hand side of (20).

To compute welfare changes between the baseline equilibrium and the counterfactual equilibrium we can write (20) in relative changes as:

$$\hat{\bar{U}} = \frac{\hat{B}_n}{\hat{d}_{nn}} \times \frac{\hat{w}_n}{\hat{P}^{\alpha} \hat{Q}_n^{1-\alpha}} \times \hat{\lambda}_{nn}^{-\frac{1}{\varepsilon}} \tag{21}$$

where the hat notation denotes the ratio between the counterfactual value of a variable and its baseline value. We use (E.10) to decompose the change in expected utility into three components: (i) changes in residential amenities and commuting costs, which will not change in this counterfactual (and therefore $\hat{B}_n$ and $\hat{d}_{nn}$ are equal to one) (ii) the change in real income, which is the ratio of wages and the price index consisting of goods prices and the price of floor space and (iii) the change in the own commuting probability. Taking the geometric mean of both sides of (E.10) over all N locations in the city results in:

$$\hat{\bar{U}} = \widehat{\tilde{w}}\hat{\tilde{P}}^{-\alpha}\hat{\tilde{Q}}^{-(1-\alpha)}\hat{\tilde{\lambda}}_{\mathrm{N}}^{-\frac{1}{\varepsilon}}, \tag{22}$$

where $\tilde{x}$ refers to geometric mean of variable x_i. This shows that the overall change in worker welfare consists of the change in average real income and the change in the average non-commuting probability.

Table 5.2: Welfare Effects of a Change in the Floor Space Supply Elasticity

	(1)	(2)	(3)	(4)
	Closed City		Open City	
Counterfactuals (relative to baseline)	Exogenous Productivity	Productivity Spillovers	Exogenous Productivity	Productivity Spillovers
Average Welfare of Workers ($\bar{U}$)	1.0627	1.0719	1.0520	1.0533
– Average Non-Commuting Probability ($\widetilde{\lambda_{\mathrm{N}}^{-1/\varepsilon}}$)	1.0010	1.0087	1.0020	1.0033
– Average Real Income	1.0616	1.0627	1.0499	1.0499
Total Land Revenue ($\sum_n r_n T_n$)	0.8605	0.8671	0.9660	0.9703
Total Population ($\bar{L}$)			1.1067	1.1095

Note: These counterfactuals explore an increase in the floor space supply elasticity by 25 percent. For each counterfactual, we report the change relative to the baseline equilibrium (with 1 meaning no change). In counterfactuals (1) and (2), we assume that city population is fixed. Counterfactual (1) supposes that overall productivity (A_i) is fixed at the baseline level. In counterfactual (2), we assume that productivity is a function of employment density in each ward. In counterfactuals (3) and (4), we consider the open city where workers can move into and out of the city. Counterfactual (3) again assumes that overall productivity is fixed while counterfactual (4) allows for endogenous productivity.

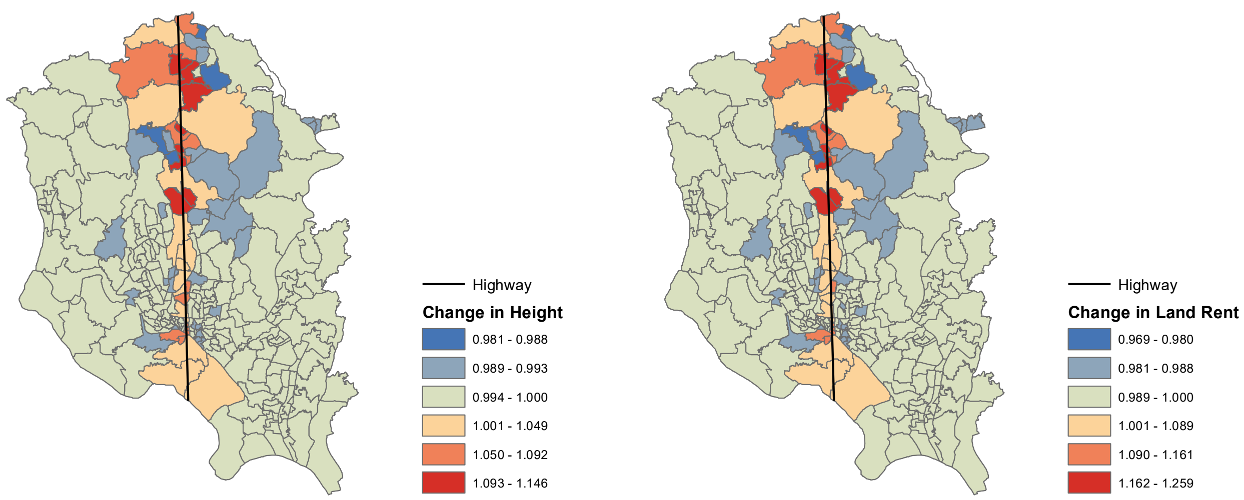
Table 5.2 reports the welfare changes relative to the baseline for our counterfactual 25 percent increase in the floor space supply elasticity. Columns (1) and (2) assume that the city is closed while Columns (3) and (4) assume that the total population of Dhaka changes according to (7) with an assumed labor supply elasticity to the city of $\sigma = 2$. Columns (1) and (3) assume that productivity (A_i) is exogenous while columns (2) and (4) allow productivity to change in response to changes in employment density. The results in this table have three main features. First, worker expected utility increases by between 5.3 and 7.2 percent across all four scenarios that we have considered. Second, most of this welfare gain is driven by changes in real income rather than changes in the non-commuting probabilities. Finally, total land revenue is substantially higher when the total population of the city is allowed to increase in response to this change, although it is below its baseline also in this case. The reason is that the increase in city population by just over 10 percent increases demand for floor space and hence land rents.

5.2 Changes in Travel Time

Our second counterfactual examines the impact of the construction of a hypothetical new north-south road through Dhaka. We assume that travel times will be reduced by 25 percent for any pair of wards intersected by this north-south road. We also adjust the travel times for wards that are not intersected by the new road if a faster indirect connection between two wards that involves traveling to a ward intersected by the road and then along the new road to a final destination is now available. The key purpose of this counterfactual is to demonstrate that infrastructure improvements will have important general equilibrium effects on the city, which are difficult to predict in the absence of a quantitative model.

Figure 5.12 shows the hypothetical new road as a straight black line and displays its effects on building heights and land prices. Panel (a) shows that average building heights increase by more than 10 percent in some of the wards intersected by the new road. Panel (b) shows that while land prices fall very slightly in large parts of the city, they increase in most but not all of the wards intersected by the new road with the largest increases being around 25 percent. Figure 5.13 displays the changes in residents and employment as a result of the new road. The figure has a number of striking features. First, the road acts as a magnet for both employment and population while areas further away from this new piece of infrastructure marginally decline. First, population declines in some of the very central wards intersected by the road and increases in more peripheral wards consistent with the idea that faster transportation speeds allow residents to relocate to higher amenity locations while still commuting to their previous places of work. Finally, the increases in employment are clustered in a small number of locations that are close to concentrations of employment in the baseline.

Figure 5.12: Impact of a New North-South Road: Height and Land Rent

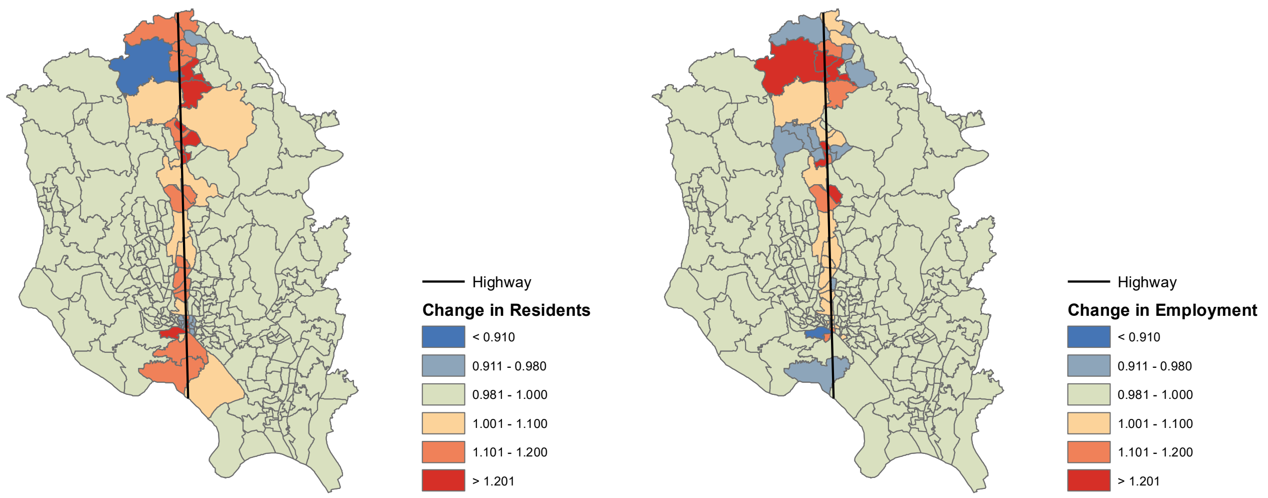


(a) Change in Average Height

(b) Change in Land Rent

Note: The left-hand panel shows the counterfactual changes in the average height of buildings relative to the DLR data, and the right-hand panel shows those in estimated land rent relative to the baseline.

Figure 5.13: Impact of a New North-South Road: Population and Employment



(a) Change in Residents **(b)** Change in Employment

Note: The left-hand panel shows the counterfactual changes in population, and the right-hand panel shows those in employment relative to the baseline data. The black line is the hypothetical highway.

Table 5.3 presents the impact of the new road on worker welfare and the income of land owners. The highway raises aggregate welfare of workers by just under 0.5 percent in the closed city version of the model. A change in the average non-commuting probability accounts for the majority of this increase in welfare, while average real income effects are modestly positive. Total land revenue marginally increases. When we allow the population of Dhaka to adjust in response to this shock, the population of the city grows by approximately 0.8 percent. This population growth causes land revenue to increase by nearly 1 percent if productivity is endogenous. Worker welfare still increases substantially also in this case, but the gains are smaller relative to the gains in the case of a closed city. The distribution of the benefits of the new road between workers and land owners is determined by the elasticity of labor supply to the city. A larger labor supply elasticity to the city implies that a larger share of the benefits of the new road are capitalized into land rents rather than worker welfare.

Table 5.3: Welfare Effects of a New North-South Road

	(1)	(2)	(3)	(4)
	Closed City		Open City	
Counterfactuals (relative to baseline)	Fix Productivity	With Productivity Spillovers	Fix Productivity	With Productivity Spillovers
Average Welfare of Workers ($\bar{U}$)	1.0054	1.0053	1.0037	1.0040
– Average Non-Commuting Probability ($\widetilde{\lambda_{\mathrm{N}}^{-1/\varepsilon}}$)	1.0049	1.0054	1.0049	1.0054
– Average Real Income	1.0005	1.0000	0.9988	0.9987
Total Land Revenue ($\sum_n r_n T_n$)	1.0007	1.0017	1.0080	1.0097
Total Population ($\bar{L}$)			1.0073	1.0080

Note: These counterfactual exercises assume the construction of the new highway from the north to the south of Dhaka, reducing travel times along the network by 25 percent relative to the current travel times. In counterfactuals (1) and (2), we assume the closed city. Counterfactual (1) supposes that overall productivity (A_i) is fixed at the baseline level. In counterfactual (2), we also allow spillovers in productivity so that overall productivity changes with employment density. In counterfactual (3) and (4), we consider the open city where workers can move into and out of the city. Counterfactual (3) fixes overall productivity as in (1), and counterfactual (4) allows productivity spillovers as in (2).

This counterfactual analysis of the hypothetical highway through Dhaka illustrates that improvements in commuting speeds will benefit not only commuters on the road network but will also be capitalized into land prices in general equilibrium. These effects are highly unequal, with land rents close to the new road increasing while other parts of Dhaka experience small declines. Our counterfactual analysis suggests that the welfare effects of local improvements in the speed of commuting are well approximated by the change in non-commuting probabilities.

6 Conclusion

The urban populations of developing countries are increasing rapidly, making it imperative to develop better tools to model the effects of policy interventions in developing country cities. While recent advances in quantitative spatial modeling can capture the rich heterogeneity of real world cities, many developing-country cities have very limited data from traditional data sources to inform such models. To overcome this challenge, we propose a method for calibrating a quantitative urban model of a typical developing-country city using mobile phone data and satellite data on building heights using Dhaka in Bangladesh as our application. Having estimated the key structural parameters of the model we show how the estimated model can be used for policy analysis. We use the calibrated models to quantify the impacts of two policies in Dhaka, an increase in the floor space supply elasticity and the construction of a new radial road that reduces travel times.

We believe that the approach we suggested in this paper can be a stepping stone for the future analysis of a wide range of urban policies, particularly in developing countries. Amongst possible policies, clearance or creation of slums is perhaps the most important to look at in developing countries. Given the recent finer satellite data for buildings and their quality, our framework can be easily extended to include informal housing. A more challenging way forward is introducing the dynamic aspects into the framework where people's expectation for future growth of the city drives the dynamic adjustment of economic activities across the space and building process of the city. The dynamic model, when combined with developing country data, serves as an interesting laboratory for understanding the long run effects of urban policy. Calibration of such a dynamic urban model is ambitious and would be crucial for designing sustainable cities in developing countries as the world moves out of the pandemic.

Appendices for Chapter 2

A Appendix: The baseline framework

We start with a relatively simple model that provides the idea of spatial equilibrium in a city. Particularly, we consider the urban structure for the economy with homogeneous agents, single sector in production, and competitive developers of land. The number of locations within a city is finite, and each location is indexed by $i, j, k, n \in \mathcal{S} = \{1, 2, \cdots, N\}$. Each location is endowed with exogenous value of amenities $\{B_n\} \in \mathbb{R}^N_{++}0$ and productivity $\{A_i\} \in \mathbb{R}^N_{++}$. There is a continuum of homogeneous agents in a city, L.

Subsection A.1 describes the utility of workers and their decision in location choices. Subsection A.2 explains the production side. There is a single sector that produces consumption goods on the production side, and we assume free trade. The production technology is given by Cobb-Douglas technology combining labor, land and capital. Subsection A.3 provides details of developers' problem and characterize the supply of floor spaces. Subsection A.4 discusses the details of competitive equilibrium. The last subsection A.5 explains how to consider the labor supply function of the city in a large economy.

A.1 Workers

Individuals choose a residential place n and workplace i in a city and receive utility from residential amenities and disutility from commuting. They also receive an idiosyncratic shock from location decisions. Preference is represented by Cobb-Douglas utility function:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega) B_n}{d_{ni}} \left(\frac{C_{ni}(\omega)}{\alpha} \right)^{\alpha} \left(\frac{H_{ni}(\omega)}{1-\alpha} \right)^{1-\alpha} \tag{A.1}$$

where $b_{ni}(\omega)$ is an idiosyncratic shock, d_{ni} is disutility from commuting which takes the value greater than one, $C_{ni}(\omega)$ is consumption of goods, and $H_{ni}(\omega)$ is consumption of housing (floor spaces for residence).

Utility maximization leads to consumption expenditure and housing expenditure of individuals:

$$P_n C_{ni}(\omega) = \alpha w_i, \quad Q_n H_{ni}(\omega) = (1-\alpha) w_i \tag{A.2}$$

where P_n is price of consumption good and Q_n is price of a unit of floor space. We suppose that consumption happens in the residential place. This gives the indirect utility for individuals given commuting choice from residential place n to workplace i:

$$V_{ni}(\omega) = \frac{b_{ni}(\omega)B_n}{d_{ni}} \frac{w_i}{{P_n}^{\alpha} Q_n^{1-\alpha}} \tag{A.3}$$

Suppose that, for every worker, the idiosyncratic term of utility $b_{ni}(\omega)$ is independently drawn from the common Fréchet distribution:

$$\mathcal{F}(b) = \mathbf{Pr}[b_{ni}(\omega) \leq b] = e^{-b^{-\varepsilon}} \tag{A.4}$$

and shocks are independent across any choices of (i, n). The shape parameter ε is greater than one. The cumulative distribution function of the indirect utility is:

$$G_{ni}(v) = \mathbf{Pr}(V_{ni}(\omega) \leq v) = \mathbf{Pr}\left[b_{ni}(\omega) \leq \frac{d_{ni}}{B_n} \frac{P_n^{\alpha} Q_n^{1-\alpha}}{w_i} v\right] = e^{-\Omega_{ni} v^{-\varepsilon}} \tag{A.5}$$

where the location parameter is given by:

$$\Omega_{ni} = \left(\frac{B_n}{d_{ni}} \frac{w_i}{P_n^{\alpha} Q_n^{1-\alpha}}\right)^{\varepsilon} \tag{A.6}$$

The corresponding probability density function is:

$$g_{ni}(v) = \frac{d}{dv} G_{ni}(v) = \varepsilon \Omega_{ni} v^{-\varepsilon-1} e^{-\Omega_{ni} v^{-\varepsilon}} \tag{A.7}$$

The cumulative distribution function of the maximum of indirect utility across pairs is:

$$\begin{aligned} G(v) &= \mathbf{Pr}\left[\max_{(n,i)} V_{ni} \leq v\right] \\ &= \mathbf{Pr}\left[b_{ni}(\omega) \leq \frac{d_{ni}}{B_n} \frac{P_n^{\alpha} Q_n^{1-\alpha}}{w_i} v \quad \forall n, i\right] \\ &= \prod_{n \in \mathcal{S}} \prod_{i \in \mathcal{S}} \mathcal{F}\left(\frac{d_{ni}}{B_n} \frac{P_n^{\alpha} Q_n^{1-\alpha}}{w_i} v\right) = e^{-\Omega v^{-\varepsilon}} \end{aligned} \tag{A.8}$$

where we let $\Omega = \sum_{n \in \mathcal{S}} \sum_{i \in \mathcal{S}} \Omega_{ni}$. Given them, probability that (n, i) pair gives the highest utility becomes:

$$
\begin{aligned}
\lambda_{ni} &= \mathbf{Pr}\Big(V_{ni} \geq \max_{(k,j)\neq(n,i)} V_{kj}\Big) \\
&= \int_0^\infty \mathbf{Pr}(V_{kj} \leq v; \quad \forall (k,j) \neq (n,i)) g_{ni}(v) \mathrm{d}v \\
&= \int_0^\infty e^{-(\Omega-\Omega_{ni})v^{-\varepsilon}} \varepsilon \Omega_{ni} v^{-\varepsilon-1} e^{-\Omega_{ni} v^{-\varepsilon}} dv \\
&= \varepsilon \Omega_{ni} \frac{1}{\varepsilon \Omega} \int_0^\infty e^{-\Omega v^{-\varepsilon}} \varepsilon \Omega v^{-\varepsilon-1} dv = \frac{\Omega_{ni}}{\Omega}
\end{aligned}
$$

By the law of large numbers over the continuum of agents, this probability is identical to the share of people who choose location pair (n, i) in a city. Thus, the residential population in n is given by:

$$
R_n = \sum_{i \in \mathcal{S}} \lambda_{ni} L = \Omega^{-1} L B_n^{\varepsilon} (P_n^{\alpha} {Q_n}^{1-\alpha})^{-\varepsilon} \sum_{i \in \mathcal{S}} d_{ni}^{-\varepsilon} w_i^{\varepsilon} \tag{A.9}
$$

and population in workplace i is:

$$
L_i = \sum_{n \in \mathcal{S}} \lambda_{ni} L = \Omega^{-1} L w_i^{\varepsilon} \sum_{n \in \mathcal{S}} d_{ni}^{-\varepsilon} B_n^{\varepsilon} (P_n^{\alpha} {Q_n}^{1-\alpha})^{-\varepsilon} \tag{A.10}
$$

The share of residents in n who commute to workplace i is given by the conditional probability:

$$
\lambda_{ni|n} \equiv \frac{\lambda_{ni}}{\sum_{j \in \mathcal{S}} \lambda_{nj}} = \frac{(w_i/d_{ni})^{\varepsilon}}{\sum_{j \in \mathcal{S}} (w_j/d_{nj})^{\varepsilon}} \tag{A.11}
$$

Further, we consider the feature of utility conditional on the residential choice. The cumulative distribution function of the indirect utility conditional on the choice of locations (residential place n and workplace i) becomes:

$$
\begin{aligned}
\widetilde{G}_{ni}(v) &= \frac{\mathbf{Pr}\Big(V_{ni} \leq v \text{ and } V_{ni} \geq \max_{(k,j)\neq(n,i)} V_{kj}\Big)}{\mathbf{Pr}\Big(V_{ni} \geq \max_{(k,j)\neq(n,i)} V_{kj}\Big)} \\
&= \frac{1}{\lambda_{ni}} \int_0^v \prod_{j \neq i} G_{nj}(u) \left(\prod_{k \neq n} \prod_{j \in \mathcal{S}} G_{kj}(u) \right) g_{ni}(u) du \\
&= \frac{1}{\lambda_{ni}} \int_0^v e^{-\Omega u^{-\varepsilon}} \varepsilon \Omega_{ni} u^{-\varepsilon-1} du \\
&= G(v)
\end{aligned} \tag{A.12}
$$

This implies that the distribution of indirect utility conditional on the location choice of workers is independent of the location choice. The intuition is the following. When residential place n exhibits lower living costs (Q_n, P_n) and workplace i exhibits higher return to work (w_i), the expected utility of a worker living in n and working in i is high conditional on idiosyncratic shock. On the other hand, when residential place n shows high living costs and workplace i shows a lower wage rate, the expected utility is low. These two opposite effects, together with Fréchet distribution of idiosyncratic shocks, lead to the same average utility level across all location pairs.

Assuming that agents can freely choose the pair of locations, the average utility of an individual is given by:

$$\bar{U} = \int_0^\infty v dG(v) = \int_0^\infty \varepsilon \Omega v^{-\varepsilon-1} e^{-\Omega v^{-\varepsilon}} v dv = \bar{\gamma}\Omega^{1/\varepsilon} = \bar{\gamma}\left(\sum_{n\in\mathcal{S}}\sum_{i\in\mathcal{S}}\left(\frac{B_n}{d_{ni}}\frac{w_i}{P_n^{\alpha}Q_n^{1-\alpha}}\right)^{\varepsilon}\right)^{1/\varepsilon} \quad \text{(A.13)}$$

with $\bar{\gamma} = \Gamma\left(\frac{\varepsilon-1}{\varepsilon}\right)$ is the value of Gamma function. This is the expected utility of agents ex-ante, and agents receive a different level of utility ex-post due to the idiosyncratic shocks. Further, we can express the average utility by:

$$\bar{U} = \bar{\gamma}\frac{B_n}{d_{nn}}\frac{w_n}{P_n^{\alpha}{Q_n}^{1-\alpha}}\lambda_{nn}^{-\frac{1}{\varepsilon}} \quad \text{(A.14)}$$

The right-hand side is defined for any location n, and this is equalized in equilibrium. This implies that a location exhibiting a high value of amenities (B_n) and high real income ($\frac{w_n}{P_n^{\alpha}Q_n^{1-\alpha}}$) is associated with high non-commuting probability (λ_{nn}) since workers have less incentive to commute to other locations. On the other hand, when location n has less attractive characteristics, it must be associated with lower non-commuting probability. Since this holds for any n. Taking the geometric mean of this, we obtain:

$$\bar{U} = \bar{\gamma}\frac{\widetilde{B}}{\widetilde{d}}\frac{\widetilde{w}}{\widetilde{P}^{\alpha}\widetilde{Q}^{1-\alpha}}\widetilde{\lambda}_{\mathrm{N}}^{-1/\varepsilon} \quad \text{(A.15)}$$

where $\widetilde{x} = (\prod_{i\in\mathcal{S}} x_i)^{1/N}$ represents geometric mean of variables $\{x_i\}$. We back to the discussion about this measure in our counterfactual experiments in the section E.

A.2 Production

There are representative firms in each location and they produce consumption goods under perfect competition. Firms produce a consumption good by combining labor, developed land and intermediate numeraire.

Production technology of a representative firm is Cobb-Douglas technology:

$$Y_n = A_n {L_n}^{\beta} {H_n^B}^{\gamma} {X_n}^{1-\beta-\gamma}, \quad \beta,\ \gamma \in (0,1) \quad \text{(A.16)}$$

where A_n is Hicks-neutral productivity in n, L_n is employment, H_n^B is floor spaces used for production, and X_n is intermediate use of final goods. Therefore, final goods are consumed by people in the city and also used in production of the final goods as a material. The cost minimization problem of firms and zero profit condition leads to the price to consumers:

$$P_n = \frac{\psi}{A_n} {w_n}^{\beta} q_n^{\gamma}$$

where q_n is floor space price for commercial use. The consumption goods are freely traded within a city following the traditional assumptions in urban economics, therefore we normalize the price:

$$1 = P = P_n = \frac{\psi}{A_n} w_n^{\beta} q_n^{\gamma} \tag{A.17}$$

Letting Y_n denote the total revenue of firms in location n, zero profit conditions imply:

$$w_n L_n = \beta Y_n, \quad q_n H_n^B = \gamma Y_n, \quad X_n = (1 - \beta - \gamma) Y_n, \tag{A.18}$$

where q_n represents price of floor space for firms.

A.3 Construction

In each location i, there are T_i units of land to be developed. The developed area is exogenous and fixed. There is a large number of developers that can freely enter to develop land in each location. Developers supply h_i^R units of floor space for residential use on one unit of land, and h_i^B units for commercial use. Therefore, total supply of floor spaces per unit of land is $h_i = h_i^R + h_i^B$ and this is equal to height of buildings.

In the baseline, we suppose that construction technology is same between two different uses. This is primary because that we do not distinguish the buildings for different uses in the data in precise. Therefore, floor space price is equalized between two uses: $Q_i = q_i$ in equilibrium. Developers use composite goods (final goods) for their construction. Namely, we suppose that the construction costs of h_i units of floor space per unit of land is

$$\xi(h_i) = \kappa_i h_i^{\nu}, \quad \nu > 1 \tag{A.19}$$

where κ_i is constant cost shifter that can vary across locations. The construction cost features iso-elastic and convex variable costs.

Developers in i decide the amount of floor space such that:

$$h_i = \arg \max_{h \in \mathbb{R}_{++}} \; Q_i h - \xi(h) = Q_i h - \kappa_i h^{\nu} \tag{A.20}$$

This leads to the floor space per unit of land and profits:

$$h_i = \left(\frac{Q_i}{\kappa_i \nu} \right)^{\frac{1}{\nu - 1}}, \quad \pi_i = \kappa_i (\nu - 1) \left(\frac{Q_i}{\kappa_i \nu} \right)^{\frac{\nu}{\nu - 1}} \tag{A.21}$$

Since developers freely enter to the market and perfect competition leads to the land rent per unit of land

$$r_i = \pi_i \tag{A.22}$$

and the total amount of land rent for landlords is $r_i T_i$. The total floor spaces in each location is given by:

$$H_i = h_i T_i = \left(\frac{Q_i}{\kappa_i \nu}\right)^{\frac{1}{\nu-1}} T_i \tag{A.23}$$

Therefore, given the fixed amount of developed land T_i, the supply elasticity of floor spaces is constant $1/(\nu - 1)$.

<u>**Microfoundation for production technology of floor spaces**</u> Suppose that floor spaces in each location (H_i) is produced by a competitive developers using the stock of developed land (T_i) and homogeneous tradable goods (X_i). The production technology takes a form of Cobb-Douglas:

$$H_i = \Lambda_i \left(\frac{T_i}{\mu}\right)^{\mu} \left(\frac{X_i}{1-\mu}\right)^{1-\mu}, \quad \mu \in (0,1) \tag{A.24}$$

where μ is the input share of developed land, $1 - \mu$ is the input share of numeraire, and Λ_i is Hicks-neutral productivity of developers. Given the size of developed area (T_i), the representative construction firm chooses the input of numeraire to maximize the profit:

$$X_i = \arg \max_{x \in \mathbb{R}_{++}} \; Q_i \Lambda_i \left(\frac{T_i}{\mu}\right)^{\mu} \left(\frac{x}{1-\mu}\right)^{1-\mu} - x \tag{A.25}$$

Therefore, $X_i = (1 - \mu) Q_i H_i$. Using this, the total supply of floor spaces by developers is:

$$H_i = \frac{1}{\mu} \Lambda_i^{\frac{1}{\mu}} Q_i^{\frac{1-\mu}{\mu}} T_i \tag{A.26}$$

and the average height – floor spaces per unit of developed land – is given by:

$$h_i = \frac{H_i}{T_i} = \frac{1}{\mu} \Lambda_i^{\frac{1}{\mu}} Q_i^{\frac{1-\mu}{\mu}} \tag{A.27}$$

Therefore, in our specification of baseline model, the parameter of floor space supply elasticity is corresponding to the production technology parameter such that:

$$\frac{1}{\nu - 1} = \frac{1-\mu}{\mu} \;\Leftrightarrow\; 1 - \mu = \frac{1}{\nu} \tag{A.28}$$

A.4 City equilibrium

The average income per capita among workers residing in location n is:

$$\bar{w}_n = \sum_{i \in \mathcal{S}} \lambda_{ni|n} w_i = \sum_{i \in \mathcal{S}} \frac{(w_i/d_{ni})^{\varepsilon}}{\sum_{j \in \mathcal{S}} (w_j/d_{nj})^{\varepsilon}} w_i \tag{A.29}$$

Since landlords consume homogeneous goods, the total expenditure on consumption in location n is given by:

$$E_n = \alpha \bar{w}_n R_n + (1-\alpha)\bar{w}_n R_n + (1-\beta-\gamma) Y_n = \bar{w}_n R_n + \frac{1-\beta-\gamma}{\beta} w_n L_n \tag{A.30}$$

where we use the zero profit condition of firms.

We start with the commuter market clearing condition. The supply of labor in location i is given by:

$$L_i = \sum_{n \in \mathcal{S}} \lambda_{ni|n} R_n \tag{A.31}$$

The demand for labor in location i is:

$$L_i = \frac{\beta Y_i}{w_i} = \frac{\beta}{\gamma} \left(\frac{1}{\kappa_i \nu} \right)^{\frac{1}{\nu-1}} \frac{Q_i^{\frac{\nu}{\nu-1}}}{w_i} T_i^B \tag{A.32}$$

where T_i^B is developed land used for commercial use and we substituted the zero profit condition of firms. Inserting price of floor space for commercial use

$$Q_i = A_i^{\frac{1}{\gamma}} w_i^{-\frac{\beta}{\gamma}} \psi^{-\frac{1}{\gamma}} \tag{A.33}$$

into the above, we obtain the demand for labor:

$$L_i = \frac{\beta}{\gamma} \left(\frac{1}{\kappa_i \nu} \right)^{\frac{1}{\nu-1}} A_i^{\frac{\nu}{\gamma(\nu-1)}} w_i^{-\left(1+\frac{\beta\nu}{\gamma(\nu-1)}\right)} \psi^{-\frac{1}{\gamma}\frac{\nu}{\nu-1}} T_i^B \tag{A.34}$$

Combining them, the labor market clearing condition is:

$$\mathcal{O}_i A_i^{\frac{\nu}{\gamma(\nu-1)}} w_i^{-\left(1+\frac{\beta\nu}{\gamma(\nu-1)}\right)} T_i^B = \sum_{n \in \mathcal{S}} \lambda_{ni|n} R_n, \quad \forall\, i \tag{A.35}$$

where we let constant $\mathcal{O} \equiv \frac{\beta}{\gamma} \left(\frac{1}{\kappa_i \nu} \right)^{\frac{1}{\nu-1}} \psi^{-\frac{1}{\gamma}\frac{\nu}{\nu-1}}$. The left-hand side is a downward sloping demand curve for labor, while the right-hand-side is upward sloping supply curve for labor. Therefore, this relationship pins down unique wage in equilibrium given $\{A_i\}$ and $\{R_n\}$.

Next, we formulate the land market clearing condition. Combining utility maximization of workers and profit maximization of developers, the demand for housing is:

$$(1-\alpha)\bar{w}_n R_n = \left(\frac{1}{\kappa_n \nu}\right)^{\frac{1}{\nu-1}} Q_n^{\frac{\nu}{\nu-1}} T_n^R \tag{A.36}$$

Hence, demand for land by developers to construct housing is:

$$T_n^R = \frac{(1-\alpha)\bar{w}_n R_n}{\left(\frac{1}{\kappa_n \nu}\right)^{\frac{1}{\nu-1}} Q_n^{\frac{\nu}{\nu-1}}} \tag{A.37}$$

Inserting the zero profit condition of developers to this, we derive:

$$T_n^R = \frac{\nu-1}{\nu}\frac{(1-\alpha)\bar{w}_n R_n}{r_n} \tag{A.38}$$

Turning to the land demand for construction of commercial floor space, zero profit condition of developers implies:

$$T_n^B = \frac{\nu-1}{\nu}\frac{\gamma Y_n}{r_n} = \frac{\nu-1}{\nu}\frac{\gamma}{\beta}\frac{w_n L_n}{r_n} = \frac{\nu-1}{\nu}\frac{\gamma}{\beta}\frac{{A_n}^{\frac{1}{\beta}} {Q_n}^{-\frac{\gamma}{\beta}} \psi^{-\frac{1}{\beta}}}{r_n} L_n \tag{A.39}$$

where we inserted the zero profit condition of firms. We again substitute zero profit condition of developers into this, and it gives:

$$T_n^B = \frac{\nu-1}{\nu}\frac{\gamma}{\beta}\psi^{-\frac{1}{\beta}} {A_n}^{\frac{1}{\beta}} {r_n}^{-\left(\frac{\nu-1}{\nu}+1\right)} L_n \tag{A.40}$$

Together them, we derive the (developed) land market clearing conditions:

$$\frac{\nu-1}{\nu}\frac{(1-\alpha)\bar{w}_n R_n}{r_n} + \frac{\nu-1}{\nu}\frac{\gamma}{\beta}\psi^{-\frac{1}{\beta}} {A_n}^{\frac{1}{\beta}} {r_n}^{-\left(\frac{\nu-1}{\nu}+1\right)} L_n = T_n, \quad \forall\, n \tag{A.41}$$

The left-hand side is a downward sloping demand curve for land. This pins down the land rent in equilibrium given $\{R_n\}$ and $\{L_n\}$. Further discussion about solving equilibrium and its existence is in the section B.

A.5 Labor mobility

We consider labor mobility in the wide economy, where workers choose to live in a city or the rest of the economy. For tractability, we consider the Rosen-Roback framework for labor mobility. Therefore, workers in the economy anticipate the average real income $\bar{U}$ from a city and the average real income $\mathbb{U}$ from the rest of the economy. The real income level $\mathbb{U}$ is given exogenously. Besides, we assume that workers differ in additional utility benefit from the choice of a city or the rest of the economy. Specifically, workers choose a city when:

$$\bar{U} \cdot z_{co}(\omega) \geq \mathbb{U} \cdot z_{co}(\omega)$$

where $z_{co}(\omega)$ is a idiosyncratic shock measuring preferences for a city and outside of a city by individual worker ω. A large value of $z_{co}(\omega)$ implies that worker ω is attached to a city or its outside for idiosyncratic reasons.

We assume that workers follow two steps in location choices. First, they draw $z_{co}(\omega)$ and anticipate the average real income $\bar{U}$ in a city, and they decide to live in a city (c) or the rest of the economy (o). This determines the total population of a city. Second, conditional on residing in a city, they decide the residential place and workplace within a city as we described in the subsection A.1. Suppose that the idiosyncratic shock follows Fréchet distribution with shape parameter σ, and the total population in the wide economy is $\bar{L}$ that is large constant.

Then, the total measure of workers that choose to live in a city is given by:

$$L = \frac{\bar{U}^{\sigma}}{\bar{U}^{\sigma} + \mathbb{U}^{\sigma}} \bar{L} = \frac{\left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma}}{\left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma} + 1} \bar{L} \tag{A.42}$$

Notice that in our application, we can consider $\left|\left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma}\right| << 1$, as population in a city is not large part of the total population in a country. Therefore, we can approximate:

$$L = \left(1 - \frac{1}{\left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma} + 1}\right) \bar{L} \approx \left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma} \bar{L} \tag{A.43}$$

Using the expected utility of workers in a city, this becomes:

$$L = \left(\frac{\bar{\gamma} \left(\sum_{n=1}^{N} \sum_{i=1}^{N} \left(\frac{B_n}{d_{ni}} \frac{w_i}{P_n^{\alpha} Q_n^{1-\alpha}} \right)^{\varepsilon} \right)^{1/\varepsilon}}{\mathbb{U}} \right)^{\sigma} \bar{L} \tag{A.44}$$

Therefore, city population (L) relative to wide economy ($\bar{L}$) increases in the expected utility relative to the exogenous level of utility in the wide economy $\mathbb{U}$. The parameter σ captures the elasticity of this adjustment. Note that we posit that workers cannot change their locations after choosing the city or outside of the city. Therefore, a realization of the worker's utility can be lower than the outside utility $\mathbb{U}$ in equilibrium.

B Appendix: Solving the model

In this section, we investigate the details of the constitution of competitive equilibrium. In subsection B.1, we start with the system of equations that characterize the competitive equilibrium. The subsection B.2 shows the existence of the competitive equilibrium, and subsection B.3 extends the discussion to its uniqueness.

B.1 System of equations

We introduce local spillovers in productivity:

$$A_n = \bar{a}_n L_n^{\chi}, \quad \text{with} \quad \bar{a}_n \equiv a_n K_n^{-\chi} \tag{B.1}$$

where $\chi > 0$ is the strength of the spillovers. We keep amenities exogenous. As we derived in the previous section, we have characterized the equilibrium by the following equations. First, population is:

$$R_i = L\bar{U}^{-\varepsilon} B_i^{\varepsilon} Q_i^{-(1-\alpha)\varepsilon} \sum_{n\in\mathcal{S}} d_{in}^{-\varepsilon} w_n^{\varepsilon}, \quad \forall\, i \in \mathcal{S} \tag{B.2}$$

Commuter market clearing condition is:

$$L_n = \sum_{i\in\mathcal{S}} \frac{w_n^{\varepsilon} d_{in}^{-\varepsilon}}{\sum_{k\in\mathcal{S}} w_k^{\varepsilon} d_{ik}^{-\varepsilon}} R_i, \quad \forall\, n \in \mathcal{S} \tag{B.3}$$

Profit maximization and zero profit condition for producers in location n is:

$$w_n = \bar{a}_n^{\frac{1}{\beta}} L_n^{\frac{\chi}{\beta}} Q_n^{-\frac{\gamma}{\beta}}, \quad \forall\, n \in \mathcal{S} \tag{B.4}$$

Using profit maximization and zero profit condition for producers and developers together, the floor space market clearing condition is:

$$\frac{(1-\alpha)R_n}{Q_n} \sum_{k\in\mathcal{S}} \frac{w_k^{\varepsilon} d_{nk}^{-\varepsilon}}{\sum_{\ell\in\mathcal{S}} w_\ell^{\varepsilon} d_{n\ell}^{-\varepsilon}} w_k + \frac{\gamma}{\beta}\frac{w_n L_n}{Q_n} = \left(\frac{Q_n}{\kappa_n \nu}\right)^{\frac{1}{\nu-1}} T_n, \quad \forall\, n \tag{B.5}$$

where the first term on the left-hand-side is demand for floor space by residents and the second term is demand for floor space by producers. The labor mobility is:

$$L = \bar{L}\left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma} \tag{B.6}$$

These $4N+1$ equations (B.2) to (B.6) determine four endogenous vectors of population ($\boldsymbol{R}$), employment ($\boldsymbol{L}$), wage rate ($\boldsymbol{w}$), and floor space price ($\boldsymbol{Q}$) and one scalar of city population L.

We now solve the system of these equations. The parameters of the model are: (α, ε, β, γ, χ, ν, σ). Without loss of generality, we set $L = \bar{L}$ and $\bar{U} = \mathbb{U}$, therefore the average real income $\bar{U}$

is equal to the reservation utility level in outer economy, $\mathbb{U}$. This implies that the free mobility of labor is hold in the spatial equilibrium. Then, combining (B.2) and (B.4) leads to population:

$$R_i = \bar{L}\mathbb{U}^{-\varepsilon} B_i^{\varepsilon} Q_i^{-(1-\alpha)\varepsilon} \sum_{n\in\mathcal{S}} d_{in}^{-\varepsilon} \bar{a}_n^{\frac{\varepsilon}{\beta}} L_n^{\frac{\chi}{\beta}\varepsilon} Q_n^{-\frac{\gamma}{\beta}\varepsilon} \tag{B.7}$$

For employment, commuter market clearing condition (B.3) becomes:

$$L_n = \sum_{i\in\mathcal{S}} \frac{d_{in}^{-\varepsilon}\left(\bar{a}_n^{\frac{1}{\beta}} L_n^{\frac{\chi}{\beta}} Q_n^{-\frac{\gamma}{\beta}}\right)^{\varepsilon}}{\sum_{k\in\mathcal{S}} d_{ik}^{-\varepsilon}\left(\bar{a}_k^{\frac{1}{\beta}} L_k^{\frac{\chi}{\beta}} Q_k^{-\frac{\gamma}{\beta}}\right)^{\varepsilon}} R_i \tag{B.8}$$

Plugging (B.7) into this, we derive:

$$L_n^{1-\frac{\chi}{\beta}\varepsilon} = \bar{L}\mathbb{U}^{-\varepsilon} \bar{a}_n^{\frac{\varepsilon}{\beta}} Q_n^{-\frac{\gamma}{\beta}\varepsilon} \sum_{i\in\mathcal{S}} d_{in}^{-\varepsilon} B_i^{\varepsilon} Q_i^{-(1-\alpha)\varepsilon} \tag{B.9}$$

Given the floor space price ($\boldsymbol{Q}$), we find the unique set of vectors ($\boldsymbol{R}$, $\boldsymbol{L}$) that solve the $2N$ equations (B.7) and (B.9) together. We now turn to the floor space clearing condition. Plugging (B.2) and (B.4) into (B.5),

$$\frac{(1-\alpha)R_n}{Q_n} \sum_{k\in\mathcal{S}} \frac{d_{nk}^{-\varepsilon}\left(\bar{a}_k^{\frac{1}{\beta}} L_k^{\frac{\chi}{\beta}} Q_k^{-\frac{\gamma}{\beta}}\right)^{\varepsilon+1}}{R_n \bar{L}^{-1}\mathbb{U}^{\varepsilon} B_n^{-\varepsilon} Q_n^{(1-\alpha)\varepsilon}} + \frac{\gamma}{\beta}\frac{\bar{a}_n^{\frac{1}{\beta}} L_n^{\frac{\chi}{\beta}} Q_n^{-\frac{\gamma}{\beta}}}{Q_n} L_n = \left(\frac{Q_n}{\kappa\nu}\right)^{\frac{1}{\nu-1}} T_n \tag{B.10}$$

Further, plugging (B.7) into above, floor space market clearing condition is:

$$(1-\alpha)\bar{L}\mathbb{U}^{-\varepsilon} B_n^{\varepsilon} Q_n^{-(1-\alpha)\varepsilon} \sum_{k\in\mathcal{S}} d_{nk}^{-\varepsilon}\left(\bar{a}_k^{\frac{1}{\beta}} L_k^{\frac{\chi}{\beta}} Q_k^{-\frac{\gamma}{\beta}}\right)^{\varepsilon+1} + \frac{\gamma}{\beta}\bar{a}_n^{\frac{1}{\beta}} L_n^{\frac{\chi}{\beta}+1} Q_n^{-\frac{\gamma}{\beta}} = \left(\frac{1}{\kappa\nu}\right)^{\frac{1}{\nu-1}} Q_n^{\frac{\nu}{\nu-1}} T_n. \tag{B.11}$$

We then use profit maximization and zero profit condition to transform the floor space clearing condition to the land market clearing condition. Profit maximization and zero profit condition for developers relates the floor space price and price per unit of developed land uniquely:

$$Q_n = \bar{\kappa}_n r_n^{\frac{\nu-1}{\nu}} \tag{B.12}$$

where $\bar{\kappa}_n = \kappa_n \nu \left(\frac{1}{\kappa_n(\nu-1)}\right)^{\frac{\nu-1}{\nu}}$ is constant. Substituting (B.12) into (B.11),

$$\begin{aligned} &(1-\alpha)\bar{L}\mathbb{U}^{-\varepsilon} B_n^{\varepsilon} \bar{\kappa}_n^{-(1-\alpha)\varepsilon} r_n^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon} \left(\sum_{k\in\mathcal{S}} d_{nk}^{-\varepsilon} \bar{a}_k^{\frac{\varepsilon+1}{\beta}} L_k(\boldsymbol{r})^{\frac{\chi}{\beta}(\varepsilon+1)} \bar{\kappa}_k^{-\frac{\gamma}{\beta}(\varepsilon+1)} r_k^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)}\right) \\ &+ \frac{\gamma}{\beta}\bar{a}_n^{\frac{1}{\beta}} L_n(\boldsymbol{r})^{\frac{\chi}{\beta}+1} \bar{\kappa}_n^{-\frac{\gamma}{\beta}} r_n^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}} = \left(\frac{1}{\kappa\nu}\right)^{\frac{1}{\nu-1}} \bar{\kappa}_n^{\frac{\nu}{\nu-1}} r_n T_n, \quad \forall\, n\in\mathcal{S} \end{aligned} \tag{B.13}$$

This is the market clearing condition for developed land. The right hand side is proportional to the endowment of developed land (T_n). The left hand side summarizes the demand for developed land. The first term in the left hand side is demand for developed land to construct residential floor spaces, and the second term is demand for developed land to construct commercial floor spaces. In (B.13), we denote employment as a function of land price, $L_n(\boldsymbol{r})$, as we can characterize unique mapping from land rent to employment from above discussion. Then, we define:

$$J_n(\boldsymbol{r}) = c_{1,n}\widetilde{\mathbb{U}} r_n^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon-1}\left(\sum_{k\in\mathcal{S}} d_{nk}^{-\varepsilon} c_{2,k} L_k(\boldsymbol{r})^{\frac{\chi}{\beta}(\varepsilon+1)} r_k^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)}\right) + c_{3,n} L_n(\boldsymbol{r})^{\frac{\chi}{\beta}+1} r_n^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}-1}$$
$$- c_{4,n} T_n \tag{B.14}$$

where we let

$$\widetilde{\mathbb{U}} = (1-\alpha)\bar{L}\mathbb{U}^{-\varepsilon},\ c_{1,n} = B_n^{\varepsilon}\bar{\kappa}_n^{-(1-\alpha)\varepsilon-\frac{\gamma}{\beta}(\varepsilon+1)},\ c_{2,n} = \bar{a}_n^{\frac{\varepsilon+1}{\beta}}\bar{\kappa}_n^{-\frac{\gamma}{\beta}(\varepsilon+1)},$$
$$c_{3,n} = \frac{\gamma}{\beta}\bar{a}_n^{\frac{1}{\beta}}\tilde{\kappa}_n^{-\frac{\gamma}{\beta}},\ c_{4,n} = \left(\frac{1}{\kappa\nu}\right)^{\frac{1}{\nu-1}}\bar{\kappa}_n^{\frac{\nu}{\nu-1}}$$

refer positive parameters. This operator $J_n(\boldsymbol{r})$ is excess demand function for developed land. Thus, the equilibrium price per unit of developed land satisfies:

$$J_n(\boldsymbol{r}) = 0 \tag{B.15}$$

for every $n \in \mathcal{S}$.

B.2 Existence of equilibria

We turn to check properties of the excess demand function: (B-I) the excess demand function is continuous, (B-II) the excess demand function satisfies Warlas law, and (B-III) the excess demand function satisfies boundary condition. These three properties establish the existence of equilibrium.

B-I. The excess demand function $J_n(\cdot)$ is continuous in $\boldsymbol{r}$, since $J_n(\boldsymbol{r}^m) \to J_n(\boldsymbol{r})$ for any sequence such that $\boldsymbol{r}^m \to \boldsymbol{r}$.

B-II. The excess demand function satisfies Warlas' law for any land rent $\boldsymbol{r}$, as immediately follow by construction.

B-III. The excess demand function is obviously bounded below for any $\boldsymbol{r} \in \mathbb{R}^N_{++}$. Letting $\bar{T} \equiv \max_n T_n$, the excess demand function must satisfy $J_n(\boldsymbol{r}) \geq -c_{3,n}\bar{T} > -\infty$ for any vector $\boldsymbol{r}$ by its construction. Further, the excess demand function satisfies the boundary behavior: for a sequence $\boldsymbol{r}^m \to \hat{\boldsymbol{r}} \in \mathbb{R}^N_{+}\backslash(\mathbb{R}^N_{++} \cup \boldsymbol{0})$, the excess demand function satisfies $\max_n\ J_n(\boldsymbol{r}^m) \to \infty$.

It is sufficient to show that the demand for land exhibits $\mathcal{D}_n(\boldsymbol{r}^m) \to \infty$ as $m \to \infty$ for some location n, and we show that by contradiction. Suppose that a sequence $\mathcal{D}_n(\boldsymbol{r}^m)$ is bounded above.

Since $\mathcal{D}_n(\boldsymbol{r}^m) \geq 0$ for all m, $\mathcal{D}_n(\boldsymbol{r}^m)$ is bounded below. Thus, the sequence is bounded and it has a convergent subsequence. Without loss of generality, let $\mathcal{D}_n(\boldsymbol{r}^m)$ be the subsequence and D_n^* be its limit. By Warlas law, $\boldsymbol{r}^m \cdot \mathcal{D}_n(\boldsymbol{r}^m) = \boldsymbol{r}^m \cdot \boldsymbol{T}$. Taking its limit, $\boldsymbol{r} \cdot \mathcal{D}_n^* = \boldsymbol{r} \cdot \boldsymbol{T}$. We also consider $\mathcal{D}_n$ such that $\boldsymbol{r} \cdot \mathcal{D}_n = \boldsymbol{r} \cdot \boldsymbol{T}$, and we take

$$\mathcal{D}_n^m = \frac{\boldsymbol{r}^m \cdot \boldsymbol{T}}{\boldsymbol{r}^m \cdot \mathcal{D}_n} \mathcal{D}_n$$

for each m. By its construction, we obtain $\boldsymbol{r}^m \cdot \mathcal{D}_n^m = \boldsymbol{r}^m \cdot \boldsymbol{T}$. Hence, $\mathcal{D}_n^m \leq \mathcal{D}_n(\boldsymbol{r}^m)$. Taking the limit for both sides,

$$\frac{\boldsymbol{r} \cdot \boldsymbol{T}}{\boldsymbol{r} \cdot \mathcal{D}_n} \mathcal{D}_n = \mathcal{D}_n \leq \mathcal{D}_n^*$$

that leads to contradiction to the strong monotonicity of the demand for land.

These properties of the excess demand function **(B-I)-(B-III)** concludes the existence of equilibrium land price vector $(\boldsymbol{r})$. Note that it is not required that the excess demand function is homogeneous of degree zero for the statement of existence of equilibrium. Next, we investigate the uniqueness of equilibrium.

B.3 Uniqueness of equilibrium

In general, with local spillovers, there are potentially multiple equilibria. We start with the special case, and then discuss the general cases.

No floor spaces in production ($\gamma = 0$) Suppose that production does not require floor spaces, $\gamma = 0$. Then, (B.9) becomes:

$$\frac{\mathbb{U}^\varepsilon}{\bar{L}\bar{a}_n^{\varepsilon/\beta}} L_n^{1-\frac{\chi\varepsilon}{\beta}} = \sum_j d_{jn}^{-\varepsilon} B_j^\varepsilon Q_j^{-(1-\alpha)\varepsilon} \tag{B.16}$$

The floor space market clearing condition (B.11) is:

$$\frac{\mathbb{U}^\varepsilon T_n}{(1-\alpha)\bar{L}B_n^\varepsilon} Q_n^{1+(1-\alpha)\varepsilon+\frac{1}{\nu-1}} = (\kappa\nu)^{\frac{1}{\nu-1}} \sum_j d_{nj}^{-\varepsilon} \bar{a}_j^{\frac{\varepsilon+1}{\beta}} L_j^{\frac{\chi(\varepsilon+1)}{\beta}} \tag{B.17}$$

Then, we can apply the results by Allen et al. (2015) to find unique solution ($\{Q_n\}$, $\{L_n\}$). The matrix of coefficients of these equations are:

$$\mathbf{M}_0 = \begin{pmatrix} 1-(1-\alpha)\varepsilon + \frac{1}{\nu-1} & 0 \\ 0 & 1-\frac{\chi}{\beta}\varepsilon \end{pmatrix}, \quad \mathbf{M}_1 = \begin{pmatrix} 0 & \frac{\chi}{\beta}(\varepsilon+1) \\ -(1-\alpha)\varepsilon & 0 \end{pmatrix} \tag{B.18}$$

The matrix $\mathbf{M}_0$ is invertible if $\chi\varepsilon/\beta \neq 1$, and we can obtain:

$$|\mathbf{M}_1\mathbf{M}_0^{-1}| = \begin{pmatrix} 0 & \frac{\chi}{\beta}(\varepsilon+1)|1-\frac{\chi}{\beta}\varepsilon|^{-1} \\ (1-\alpha)\varepsilon\big[1+(1-\alpha)\varepsilon+\frac{1}{\nu-1}\big]^{-1} & 0 \end{pmatrix} \tag{B.19}$$

The system of equations has a unique up-to-scale solution if eigenvalues of this matrix are no larger than one. Therefore, the sufficient conditions for a unique up-to-scale solution are:

$$\frac{\chi}{\beta}(\varepsilon+1)\left(1-\frac{\chi}{\beta}\varepsilon\right)^{-1} \leq 1 \ \Leftrightarrow \ \chi \leq \frac{\beta}{2\varepsilon+1} \tag{B.20}$$

Intuitively, when the local spillover is not strong and taste shocks across locations show large variation, the solution is unique.

No local spillovers ($\chi = 0$) Consider no agglomeration economies in our model. If $\chi = 0$, (B.14) becomes:

$$J_n(\boldsymbol{r}) = c_{1,n}\widetilde{\mathbb{U}} r_n^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon-1}\left(\sum_{k\in\mathcal{S}} d_{nk}^{-\varepsilon} c_{2,k} r_k^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)}\right) + c_{3,n} L_n(\boldsymbol{r}) r_n^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}-1} - c_{4,n} T_n \tag{B.21}$$

and (B.9) gives mass of workers in workplace n:

$$L_n(\boldsymbol{r}) = c_{5,n} \bar{L} \mathbb{U}^{-\varepsilon} r_n^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}\varepsilon} \sum_{i\in\mathcal{S}} d_{in}^{-\varepsilon} c_{6,i} r_i^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon} \tag{B.22}$$

where $c_{5,n} \equiv \bar{a}_n^{\frac{\varepsilon}{\alpha}} \bar{\kappa}_n^{-\frac{\gamma}{\beta}\varepsilon}$ and $c_{6,i} = B_i^{\varepsilon} \bar{\kappa}_i^{-(1-\alpha)\varepsilon}$. Plugging this into (B.21) yields:

$$J_n(\boldsymbol{r}) = J_n^0(\boldsymbol{r}) + J_n^1(\boldsymbol{r}) - \bar{J}_n \tag{B.23}$$

where we let:

$$\begin{aligned}
J_n^0(\boldsymbol{r}) &= c_{1,n}\widetilde{\mathbb{U}} r_n^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon-1}\left(\sum_{i\in\mathcal{S}} d_{ni}^{-\varepsilon} c_{2,i} r_i^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)}\right),\\
J_n^1(\boldsymbol{r}) &= c_{3,n} c_{5,n} \bar{L} \mathbb{U}^{-\varepsilon} r_n^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)-1}\left(\sum_{i\in\mathcal{S}} d_{in}^{-\varepsilon} c_{6,i} r_i^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon}\right),\\
\bar{J}_n &= c_{4,n} T_n.
\end{aligned} \tag{B.24}$$

Then, we have:

$$\frac{dJ_n^0}{dr_j} = \mathbf{1}_{j=n} \times \left(-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon - 1\right)\frac{J_n^0}{r_j} - \frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)\frac{J_n^0}{r_j} p_{nj}^0 \tag{B.25}$$

where $p_{nj}^0 = \frac{d_{nj}^{-\varepsilon} c_{2,j} r_j^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)}}{\sum_{i\in\mathcal{S}} d_{ni}^{-\varepsilon} c_{2,i} r_i^{-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1)}}$, and

$$\frac{dJ_n^1}{dr_j} = \mathbf{1}_{j=n} \times \left(-\frac{\nu-1}{\nu}\frac{\gamma}{\beta}(\varepsilon+1) - 1\right)\frac{J_n^1}{r_j} - \frac{\nu-1}{\nu}(1-\alpha)\varepsilon\frac{J_n^1}{r_j} p_{nj}^1 \tag{B.26}$$

with $p_{nj}^{1} = \frac{d_{jn}^{-\varepsilon} c_{6,j} r_j^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon}}{\sum_{i \in \mathcal{S}} d_{in}^{-\varepsilon} c_{6,i} r_i^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon}}$. Then, we can immediately see that $J_n(\boldsymbol{r})$ exhibits:

$$\frac{dJ_n(\boldsymbol{r})}{dr_n} < 0, \quad \left.\frac{dJ_n(\boldsymbol{r})}{dr_j}\right|_{j \neq n} < 0, \quad \frac{dJ_n(\boldsymbol{r})}{dr_n} < \left.\frac{dJ_n(\boldsymbol{r})}{dr_j}\right|_{j \neq n} \tag{B.27}$$

Then, we need the condition such that:

$$\sum_{j \neq n} \left| \frac{dJ_n(\boldsymbol{r})}{dr_j} \right| \leq \left| \frac{dJ_n(\boldsymbol{r})}{dr_n} \right| \tag{B.28}$$

for every n. Mathematically, this implies the negative of the normalized Jacobian is a diagonally dominant matrix, and economically, this corresponds to the gross substitute properties.

With local spillovers ($\chi > 0$) Next, we allow agglomeration economies in production place, $\chi \neq 0$. Employment (B.9) becomes:

$$L_n(\boldsymbol{r}) = \left(\bar{L} \mathbb{U}^{-\varepsilon} c_{2,n}^{\frac{\varepsilon}{\varepsilon+1}} r_n^{-\frac{\gamma}{\beta}\frac{\nu-1}{\nu}\varepsilon} \sum_{i \in \mathcal{S}} d_{in}^{-\varepsilon} B_i^{\varepsilon} \bar{\kappa}_i^{-(1-\alpha)\varepsilon} r_i^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon} \right)^{\frac{\beta}{\beta - \chi\varepsilon}} \tag{B.29}$$

When we substitute this into (B.14), the excess demand function becomes:

$$\begin{aligned} J_n(\boldsymbol{r}) = {} & c_{7,n} \widetilde{\mathbb{U}}^{\frac{\chi+\beta}{\beta-\chi\varepsilon}} r_n^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon - 1} \left(\sum_{i \in \mathcal{S}} d_{ni}^{-\varepsilon} c_{2,i}^{\frac{\beta}{\beta-\chi\varepsilon}} r_i^{-\frac{\nu-1}{\nu}\frac{\varepsilon+1}{\beta-\chi\varepsilon}} \left(\sum_{j \in \mathcal{S}} d_{ji}^{-\varepsilon} c_{6,j} r_j^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon} \right)^{\frac{\chi(\varepsilon+1)}{\beta-\chi\varepsilon}} \right) \\ & + c_{3,n} \left(\bar{L} \mathbb{U}^{-\varepsilon} c_{2,n}^{\frac{\varepsilon}{\varepsilon+1}} \right)^{\frac{\beta+\chi}{\beta-\chi\varepsilon}} r_n^{-\gamma\frac{\nu-1}{\nu}\left(\frac{\varepsilon}{\beta-\chi\varepsilon}+\frac{1}{\beta}\right)-1} \left(\sum_{i \in \mathcal{S}} d_{in}^{-\varepsilon} B_i^{\varepsilon} \bar{\kappa}_i^{-(1-\alpha)\varepsilon} r_i^{-\frac{\nu-1}{\nu}(1-\alpha)\varepsilon} \right)^{\frac{\beta+\chi}{\beta-\chi\varepsilon}} - c_{4,n} T_n \end{aligned} \tag{B.30}$$

where $c_{7,n} = c_{1,n}(1-\alpha)^{-\frac{\chi(\varepsilon+1)}{\beta-\chi\varepsilon}}$. Therefore, if $\chi\varepsilon < \beta$, the excess demand function exhibits (B.27). The intuition for this condition is followings: first, the idiosyncratic shocks in commuting choices exhibit large variation as it prevents the perfect concentration of employment (i.e., small value of ε); second, the agglomeration force in production place, χ, is not too strong to dominate the dispersion forces.

When we compare this condition to the case of $\gamma = 0$, we see the condition becomes tight. Intuitively, without land in production, an increase in employment directly increases the wage of workers, which leads to further agglomeration forces and an increase in demand for housing. This effect is captured by $\chi(\varepsilon + 1)$. Therefore, we have this part in the condition for the case with $\gamma = 0$. However, when production requires land, the dispersion force from the land offsets such a wage increase. Yet, we need the condition of gross substitute properties for the uniqueness of equilibrium.

C Appendix: Calibration of the model

C.1 Gravity equation for commuting

We parametrize the commuting costs (loss of utility) by a function of commuting time, $d_{ni} = t_{ni}^{\delta}$ with positive parameter δ. The commuting pattern within a city shows:

$$\log \lambda_{ni} = -\varepsilon\delta \ln t_{ni} + \mu_n + \phi_i + \lambda_0 \tag{C.1}$$

where

$$\mu_n \equiv \varepsilon \log B_n - (1-\alpha)\varepsilon \log Q_n, \quad \phi_i \equiv \varepsilon \log w_i \tag{C.2}$$

where μ_n is the residential fixed effect that compounds residential amenity (B_n) and price of residential floor spaces (Q_n), and ϕ_i is the workplace fixed effect that reflects wage rate (w_i) in equilibrium. λ_0 is a constant that corresponds to the denominator of commuting probability. Therefore the combination of two parameters, the shape parameter of Fréchet distribution (ε) and the parameter for commuting cost (δ), governs the elasticity of commuting flows to travel time.

Estimating (C.1) using all positive commuting flow in $\log \lambda_{ni}$ and travel time t_{ni} yields estimates for the combination of two parameters: $\hat{\theta} = \hat{\varepsilon}\hat{\delta}$. Kreindler and Miyauchi (2021) have estimated the gravity equation for Dhaka and obtain the coefficient of 2.5.

To decompose the coefficient, our benchmark value of ε is the estimated value in Kreindler and Miyauchi (2021). They estimated the Fréchet shape parameter around 8.0 by using the variation of workplace labor earnings for the fixed effects ϕ_i. This implies that the parameter in the commuting costs, δ, is equal to $\delta = 2.5/8.0 \approx 0.31$. We use this value ($\delta$) as a benchmark throughout our quantification.

There are various estimates for the commuting elasticity in the literature; therefore, we also do some robustness checks for different values for the elasticity ε.

C.2 Citywide parameters

We feed some parameters based on the descriptive statistics. In particular, we look into the data on expenditure share on housing and other consumer goods for the target city to determine the parameter in the preference. For the production technology, we set cost share from the literature.

Our data source for Dhaka is the Household Income and Expenditure Survey (HIES) in 2016. It reports average expenditure share across different consumption segments: food and beverage (42.59 %), cloth and footwear (6.42 %), housing and house rent (17.25 %), fuel and lighting (5.02 %), household effect (3.03 %), medical and education (10.69 %) and others including transportation and recreation (15.0 %). Given these numbers, we set α to be 0.75.

We also set 0.60 for input share of labor in production (β) and 0.20 for input share of commercial floor spaces in production (γ) based on Valentinyi and Berthold (2008).

C.3 Model inversion and estimation of key parameters

Our aim in the next step is to recover the location-specific fundamentals (A_i, B_i) and estimate key parameters in the determinants in the supply of floor spaces and spillovers. Conditional on city-wide parameters $(\alpha, \varepsilon, \beta, \gamma, \nu, \kappa)$ and observations in data about workers in residential places ($\boldsymbol{R}$), workers in workplaces ($\boldsymbol{L}$), average height of buildings ($\boldsymbol{h}$), size of built up area ($\boldsymbol{T}$) and commuting costs ($\boldsymbol{d}$), we recover unobserved vector of productivity ($\boldsymbol{A}$) and amenities ($\boldsymbol{B}$) to make the equilibrium allocation consistent with observations. Furthermore, given parameters for local externalities in productivity (χ) and size of the area ($\boldsymbol{K}$), we can obtain the unobserved location fundamentals in workplaces ($\boldsymbol{a}$), and they are unique. We use this process to estimate the strength of spillover.

The process is sequential, and therefore we explain it step by step. In step 1, we calibrate wages across locations by exploiting commuting market clearing conditions. Intuitively, we can infer wage by plugging employment, population and commuting costs into the commuting market clearing condition. In step 2, we estimate the parameter related to floor supply elasticity. Specifically, we estimate cost elasticity in building ν by using the data on height and land market clearing conditions. We consider the moment conditions for the residuals of floor space supply to estimate the parameter. At the end of this step, we can also compute the floor space prices and land rents across locations to be consistent with market clearing conditions. In step 3, we back out productivity ($\boldsymbol{A}$) and amenities ($\boldsymbol{B}$). The zero profit condition of firms allows us to recover productivity, while population distribution can be used to back out amenities. Finally, in step 4, we decompose the estimated productivity into the exogenous part and spillovers. We consider the moment conditions for the fundamental advantages to estimate the strength of spillover (χ).

Back out wages (STEP 1)

First, the model leads to a system of equations that relate workplace and residential population:

$$L_n = \sum_{i \in \mathcal{S}} \frac{(w_n/d_{in})^{\varepsilon}}{\sum_{k \in \mathcal{S}} (w_k/d_{ik})^{\varepsilon}} R_i \tag{C.3}$$

where we can plug travel time into the commuting cost, $d_{in} = t_{in}^{\delta}$. Given the information of $(\boldsymbol{L}, \boldsymbol{R})$, commuting cost ($\boldsymbol{d}$) and parameter of idiosyncratic shocks ε, we solve the system of N equations for wages $\boldsymbol{w} = \{w_i\}_{i=1}^{N}$.

Conditional on the observation of $(\boldsymbol{L}, \boldsymbol{R})$, this system of equations is homogeneous degree zero in $\boldsymbol{w}$, and the solution for (C.3) is unique up to scale. To see this, suppose that there are two linearly independent vectors $\boldsymbol{w}$ and $\boldsymbol{w}'$ solving the system of equations. We let $\nabla_k \equiv w'_k/w_k$ for every $k \in \mathcal{S}$ and without loss of generality we set $n \in \arg\max_{k \in \mathcal{S}} \nabla_k$. Using this, we define $\tilde{\boldsymbol{w}} = \nabla_n \cdot \boldsymbol{w}$. By construction, $\tilde{w}_n = w'_n$ and $\tilde{w}_k = \nabla_n \cdot w_k > \nabla_k \cdot w_k = w'_k$ for other elements $k \neq n$.

Therefore, for such n, we have:

$$0 = L_n - \sum_{i\in\mathcal{S}} \frac{(w'_n/d_{in})^\varepsilon}{\sum_{k\in\mathcal{S}}(w'_k/d_{ik})^\varepsilon} R_i < L_n - \sum_{i\in\mathcal{S}} \frac{(w_n/d_{in})^\varepsilon}{\sum_{k\in\mathcal{S}}(w_k/d_{ik})^\varepsilon} R_i \tag{C.4}$$

and this leads to contradiction. Therefore, the solution for the system of equations (C.3) is unique up to scale. We normalize the vector by its geometric mean, $\left(\prod_n w_n\right)^{1/N} = 1$.

Given the inverted wage vector $\boldsymbol{w} = \{w_n\}$, we compute the wage earnings per capita at location i:

$$\bar{w}_i = \sum_{n\in\mathcal{S}} \frac{(w_n/d_{in})^\varepsilon}{\sum_{k\in\mathcal{S}}(w_k/d_{ik})^\varepsilon} w_n \tag{C.5}$$

Using this, we infer the aggregate income of workers living in location i by $\overline{w}_i R_i$.

Estimate parameter of elasticity ν, floor space prices and land rent (STEP 2)

Next, we consider the estimation of the elasticity in the construction of floor spaces, ν, together with back up of relevant variables in equilibrium: floor space price ($\boldsymbol{Q}$), average height ($\boldsymbol{h}$) and price of developed land ($\boldsymbol{r}$). First, given the parameter ν, we compute land price and floor space price by using land market clearing conditions and profit maximization in the construction of floor spaces. Then, we predict the average height in each location and define the target of matching the predicted height and observations. Using the moment condition, we search the parameter value ν. In our data, we cannot distinguish the average height of housing and commercial floors. We only observe the average of all buildings in a neighborhood which includes any different uses. Hence, we cannot identify different parameters for each use, and we only estimate the single parameter ν.

Given the elasticity in construction of floor spaces, ν, we use land market clearing condition for developed land to back out the inferred unit land price ($\boldsymbol{r}$). We compute unit land price by:

$$r_i = \frac{1}{T_i}\left(\frac{\nu-1}{\nu}(1-\alpha)\bar{w}_i R_i + \frac{\gamma}{\beta}\frac{\nu-1}{\nu} w_i L_i\right) \tag{C.6}$$

where T_i is the area of developed land. In the bracket, the first term is the demand for developed land to be used for residential places, and the second term is the demand for developed land to build commercial floors.

Further, the profit maximization of developers in the supply of floor spaces to compute their unit prices:

$$Q_i = \kappa_i \nu \left(\frac{r_i}{\kappa_i(\nu-1)}\right)^{(\nu-1)/\nu} \tag{C.7}$$

The elasticity of floor space price $\boldsymbol{Q}$ to land price $\boldsymbol{r}$ is given by $1 - 1/\nu \in (0, 1]$. The height of buildings in each location is:

$$h_i = \left(\frac{Q_i}{\kappa_i \nu}\right)^{1/(\nu-1)} \tag{C.8}$$

Combining them, we can express the height by:

$$h_i = \left(\frac{r_i}{\kappa_i(\nu - 1)} \right)^{1/\nu}. \tag{C.9}$$

Inserting the land market cleating condition into this, we obtain:

$$\nu \ln h_i = -\ln T_i - \ln \nu - \ln \kappa_i + \ln \left((1-\alpha)\bar{w}_i R_i + \frac{\gamma}{\beta} w_i L_i \right) \tag{C.10}$$

The left-hand side is the log of height. On the right-hand side, the first term is the log of the size of developed land, the second term is constant, the third term is the supply shifter, and the last term is the demand for floor spaces. Among them, we can use the data on height for the left-hand side, data on developed area, and demand for floor spaces is computed by using wages and income we calculated in step 1.

The supply shifter can be seen as structural residuals. Therefore, we use the residuals to estimate the parameter ν. Namely, we compute the structural residuals:

$$\ln \kappa_i = -\nu \ln h_i - \ln T_i - \ln \nu + \ln \left((1-\alpha)\bar{w}_i R_i + \frac{\gamma}{\beta} w_i L_i \right) \tag{C.11}$$

given parameter ν and observation of heights (h_i) and developed land T_i and demand. Using them, we define the moment conditions for the structural residuals:

$$\mathbb{E}[\mathcal{M}_h(X;\nu)] \equiv \mathbb{E}\left[\mathbb{I}_i(g) \times \left(\ln \kappa_i - \overline{\ln \kappa}\right)\right] = 0 \tag{C.12}$$

where we group locations into G different bins based on the total population density (the sum of population and employment) and $\mathbb{I}_i(g)$ is an indicator function for location i that is in the g-th bin. $\ln \kappa_i$ is the computed residuals and $\overline{\ln \kappa}$ is their average. The key identifying assumption for this is: that any differences between the structural residuals in the average height are unrelated to the level of total population density among the same group.

In our baseline, we use total density, which sums population density and employment density to define $G = 10$ bins. Therefore, we sort all locations by the level of population density and employment density, and we group locations for different 10 bins based on their total density. Our target of parameters is:

$$\widehat{\boldsymbol{\nu}} = \arg\min \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathcal{M}_h(X_i;\boldsymbol{\nu}) \right)^{\top} \mathbb{V}^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathcal{M}_h(X_i;\boldsymbol{\nu}) \right) \tag{C.13}$$

where $\mathbb{V}$ is positive semidefinite matrix.

The reason that we use total density, which sums population density and employment density, to define bins in the baseline is that the variation of floor spaces in the city is mainly driven by the total demand for both housing and business use. We also consider other definitions for the grouping.

Floor space prices and land rent. Once we estimated ν, we can back out floor space prices from:

$$\ln Q_i = (\nu - 1)\ln h_i + \ln \kappa_i + \ln \nu \tag{C.14}$$

On the right-hand side, we plug the data on average height and residuals κ_i. In practice, we smooth the residuals since we find that the structural residuals are noisy in the periphery of the city. Based on the grouping of locations used for the moment condition, we compute the average of the structural residuals within the group. Then, we use such averages instead of the original values of κ_i. Plugging such average values into the definition of the structural residuals, we adjust the size of developed land area (T_i) such that zero profit conditions for developers hold.

Finally, we substitute the adjusted land area into the land market cleating conditions to estimate land rents. The land market clearing condition implies the developed land used for housings and that for commercial floor spaces are given by:

$$T_i^R = \frac{\nu - 1}{\nu}\frac{(1-\alpha)\overline{w}_i R_i}{r_i}, \quad T_i^B = \frac{\gamma}{\beta}\frac{\nu - 1}{\nu}\frac{w_i L_i}{r_i} \tag{C.15}$$

Back out productivity and amenities (STEP 3)

Productivity We use the zero profit condition for producers of homogeneous tradable goods to back out the productivity in each location:

$$A_n \propto w_n^{\beta} Q_n^{\gamma} \tag{C.16}$$

Note that trade costs are negligible in a city, therefore we normalize the price to be one (numeraire). Given parameter of cost share for labor (β) and floor spaces (γ), we can compute the overall productivity for each location. $1 - \beta - \gamma$ is the share of numeraire for intermediate use in production.

Amenities The model states that residential population in i is:

$$R_i = \mathbb{U}^{-\varepsilon}\bar{\gamma}^{\varepsilon} L B_i^{\varepsilon} {Q_i}^{-(1-\alpha)\varepsilon} W_i \quad W_i \equiv \sum_{n \in \mathcal{S}} d_{in}^{-\varepsilon} w_n^{\varepsilon} \tag{C.17}$$

where $\mathbb{U}$ is utility level in the large economy, $\bar{\gamma}$ is constant value of Gamma function, L is total number of workers in a city, and B_i is overall amenity. W_i is the proximity to the labor market potential of location i which summarizes the accessibility to labor returns in a city.

Since the probability of location choice is homogeneous of degree zero in overall amenity, we can back out the amenity vector up to scale. We take geometric mean for relevant variables:

$$\tilde{B} = \left(\prod_{i \in \mathcal{S}} B_i\right)^{\frac{1}{N}}, \ \tilde{R} = \left(\prod_{i \in \mathcal{S}} R_i\right)^{\frac{1}{N}}, \ \tilde{Q} = \left(\prod_{i \in \mathcal{S}} Q_i\right)^{\frac{1}{N}}, \ \tilde{W} = \left(\prod_{i \in \mathcal{S}} W_i\right)^{\frac{1}{N}}, \tag{C.18}$$

Using them, we back out the amenity vector normalized by its geometric mean:

$$\frac{B_i}{\tilde{B}} = \left(\frac{R_i}{\tilde{R}}\right)^{1/\varepsilon} \left(\frac{Q_i}{\tilde{Q}}\right)^{1-\alpha} \left(\frac{W_i}{\tilde{W}}\right)^{-1/\varepsilon} \tag{C.19}$$

The vector of overall amenity ($\boldsymbol{B}$) is uniquely determined with this normalization. We finally set $\tilde{B} = 1$.

Estimate the agglomeration parameters (STEP 4)

Finally, we consider the parameter of agglomeration economy. In this paper, we allow that the local productivity increases in employment density in the place and the strength of agglomeration economy is governed by a single parameter χ. In this last step, we consider estimation of the parameter and exogenous fundamental productivity ($\boldsymbol{a} = \{a_n\}_{n=1}^N$). Using our specification the local agglomeration economy, the fundamental productivity is:

$$a_n = A_n \left(\frac{L_n}{K_n}\right)^{-\chi} \tag{C.20}$$

where K_n is size of location and L_n/K_n is employment density. Given information of employment ($\boldsymbol{L}$) and area size ($\boldsymbol{K}$), we can compute fundamentals ($\boldsymbol{a}$) conditional on parameter values χ.

We search the best fitted values for the parameter $\hat{\chi}$ that minimizes the variation of fundamentals within each bin of locations relative to the global (city-level) average of fundamentals. Specifically, we define the moment conditions for fundamental productivity:

$$\mathbb{E}[\mathcal{M}_A(X;\boldsymbol{\chi})] \equiv \mathbb{E}\left[\mathbb{I}_i^A(g) \times (\ln \bar{a}_i - \ln \bar{a})\right] = 0 \tag{C.21}$$

Where $\overline{a}$ is geometric mean of fundamentals at the city level.

We define the bins in a similar manner as in (C.12) of Step 2. Namely, we define $G = 10$ bins based on the information of the total density, which is the sum of employment density and population density. The key identifying assumption here is that any differences between the average of exogenous fundamentals in a group and their overall average are unrelated to the level of total density.

D Appendix: Data for Dhaka

This part presents additional figures about the data for Dhaka. Subsection D.1 presents basic data for the city. The data is based on the earlier work by Bird et al. (2018). Subsection D.2 shows the results of calibration.

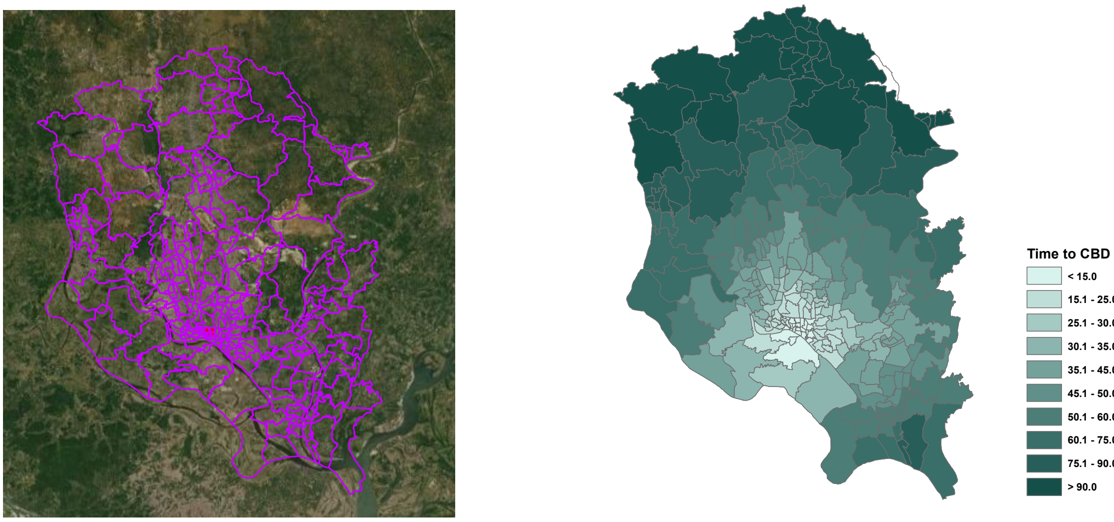
D.1 Data

Dhaka We study the City of Dhaka, Bangladesh, which is a typical developing country city in many ways. Dhaka has experienced rapid population growth since the partition of India in 1947, and it continues to attract new populations at a rapid pace. Figure D.1 (a) shows the area of the city that we use in our quantitative analysis. Each partition on the map represents a union, which defines a location in our quantification. The number of unions in our analysis is 264. As we see on the map, the size of unions is significantly different. The central area of the city shows small partitions while the peripheral areas are large.

Figure D.1 (b) shows how each union is far from the central business district (CBD). The CBD is defined as a union with the highest employment density. On the map, dark areas in the North of the city show the longest travel time to the CBD, and travel time for people living in these places takes more than 90 minutes if they commute to the CBD. We observe the large heterogeneity across wards in the city area, which shows the importance of the commuting costs for workers in the city. In our analysis, we use the average travel time instead of the travel time with congestion. The transport network is mapped in 2011 using data from the Revised Strategic Transport Plan (DTCA 2015) and OpenStreetMap (2015). Travel times are estimated for both pedestrians traveling at 5 km/h and motorists traveling at different speeds according to different types of roads: 15km/h on one-lane roads, 20 km/h on two-lane roads, and 60km/h for four-lane or wider roads. Then, travel time between unions is computed on the shortest path. See Bird et al. (2018) for further details.

Table D.4 reports the basic statistics for 264 unions used in our analysis: population, employment, population density, employment density, size of unions, size of the developed area within unions, the share of the developed area within unions, average height of buildings in unions and travel time to the CBD. Employment density shows significant variation relative to population density. We also find remarkable differences in the degree of developed area in the city.

Figure D.1 : The Area of Dhaka and Travel Time to CBD



(a) Dhaka　　**(b)** Travel time to CBD

Note: The left panel shows the area of Dhaka and ward used in our analysis. The right panel shows the travel time (minutes) to CBD in Dhaka.

Table D.4 : Summary Statistics

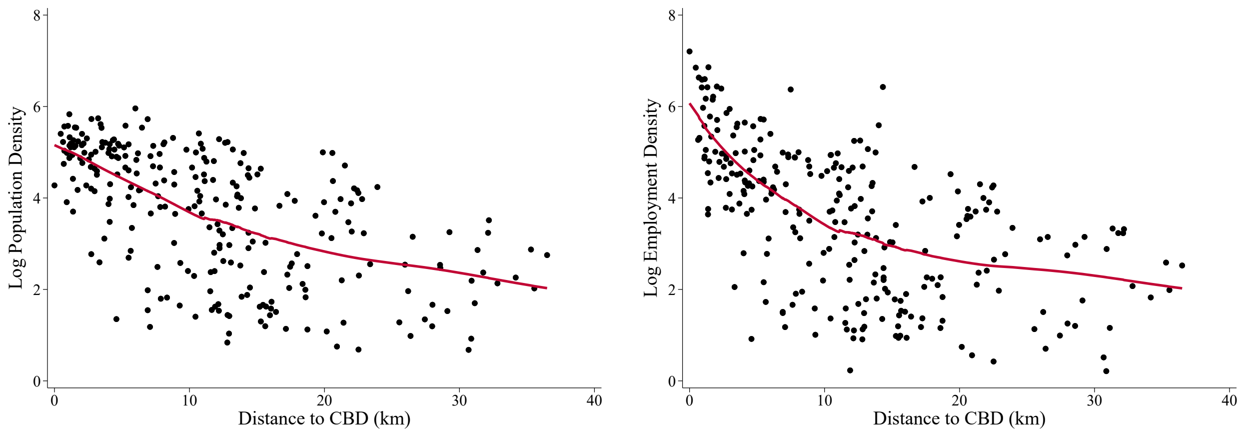
Variable in data	Average	1 percentile	25 percentile	50 percentile	75 percentile	99 percentile
Population	11,883	495	4,742	8,113	14,770	58,717
Employment	11,883	328	3,726	7,581	14,417	64,342
Population density (per ha)	79.80	1.983	12.22	53.26	136.7	311.9
Employment density (per ha)	102.5	1.527	8.620	41.18	112.6	944.1
Area size (ha)	552.8	5.403	74.25	180.6	550.8	3,669
Area of developed land (ha)	91.58	1.858	28.66	55.08	110.5	621.7
Share of developed land area (percent)	36.4	0.809	13.7	37.2	52.8	91.2
Average height of buildings (metres)	7.620	2.007	5.384	7.085	10.30	12.90
Travel time to CBD (minutes)	48.00	5.556	27.73	46.92	61.85	112.6

Note: This table reports summary statistics about the main data for 264 unions in Dhaka.

Density The population and residential data are based on the Population and Housing Census of 2010 (BBS 2011) at the union level. Following Bird and Venables (2019), we take households as the unit of observation and re-scale the resulting spatial distribution of households by average household size. Our data for employment is built up from employment data at the establishment location taken from the Economic Census 2013 (BBS2014). The Economic Census reports 4-digit industry codes (BSIC 2009) and allows us to establish employment in different sectors (manufacturing, nontradable services and tradable services). In the baseline, we do not distinguish different sectors and use the sum of total employment in each union.

The city shows the concentration of population and employment in the central area. Figure D.2 show population density and employment density across different unions by their distances from the CBD. Both panels show the highest density around the CBD, consistent with the monocentric city framework. The employment density quickly drops as locations are far from the CBD. In the peripheral areas, population density and employment density show a similar level, implying less commuting in these areas.

Figure D.2 : Workplace Density and Residential Density



(a) Worker Density at Residents

(b) Worker Density at Workplace (Employment)

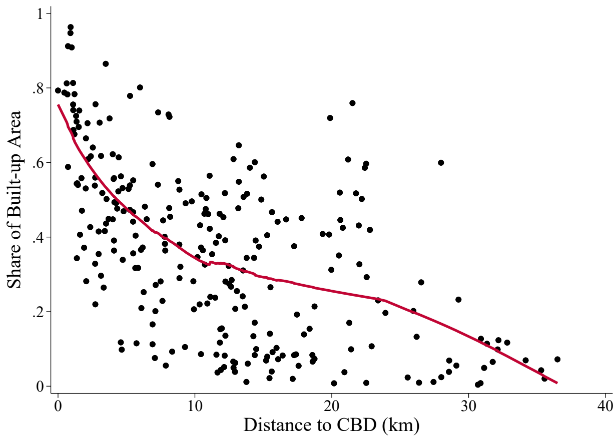
Note: The left panel shows the population density in residential places, and the right panel shows the employment density. Each point shows a ward in Dhaka, and red lines are fitted lines.

Satellite data We use the data on the build-up area and the average height of buildings in all wards of Dhaka from the German Aerospace Center (DLR). These data are produced by DLR, using their World Settlement Footprint 3D (WSF 3D). The methodologies of WSF 3D are described in Marconcini et al. (2020) and Esch et al. (2020).

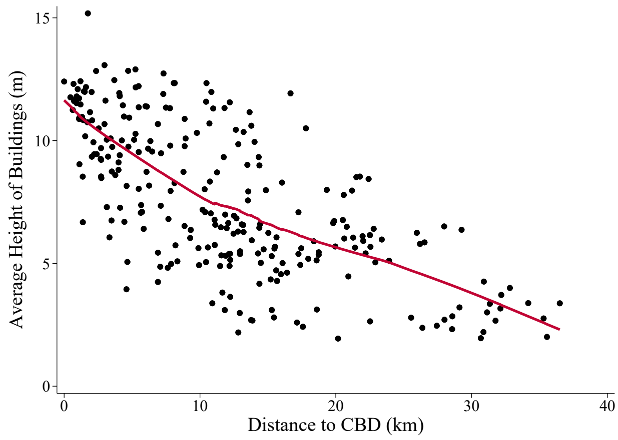
The WSF 3D is a three-dimensional model of the built environment worldwide. The WSF 3D estimates building area, height, and density at an aggregated 90 m spatial resolution. The original satellite data, 12m WSF-3D layer, provides information about the height of all buildings in 12 meters $\times$ 12 meters pixels. Using the height data of all 12m WSD-3D pixels that belong to a building structure in each union, the average height of buildings at the union level is computed. We also obtain the total built-up area within the union. Based on these two measures, the total volume of all buildings within the ward results from multiplying the total building footprint area per ward by the average building height per ward.

Figure D.3 shows this data for Dhaka. Panel (a) presents the variation of the share of built-up area in the data. The variation is significant. The central area of the city exhibits more than 80 percent of the developed area share, while many locations show less than 10 percent. Panel (b) shows the average height of buildings that decline from the center of the city toward the periphery. Together with these figures, the total volume of floor spaces is relatively large in the central area that is used for both firms and households. In contrast, housings in the peripheral area show lower height on the small share of developed area.

Figure D.3 : Built-up Area and Average Height



(a) Share of Built-up area in Data



(b) Average Height in Data

Note: The left panel shows the share of built-up areas. The measure is computed by: the estimated developed area divided by the total area size in each ward. The right panel shows estimated average height of buildings in each ward. Each point shows a ward and the red lines are fitted lines.

D.2 Calibration results

This subsection presents additional results for our calibration omitted in the main text. Following the sequential steps in the calibration we described in the section C, we show the results step by step here. We start with calibration of wages and commuting patterns in Dhaka, then estimate floor space supply elasticity. Given the results, we estimate floor space prices and land rent. Lastly, we estimate spillovers in productivity. In the last step, we also show the results when we shut down the spillover components.

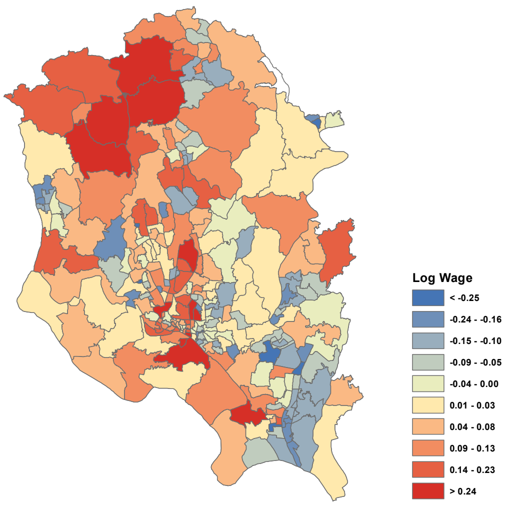
Calibration of wage and commuting

Figure D.4 display the results for the estimated wage rate in step 1. Panel (a) shows the spatial variation of the wage rate. The red-colored areas show high wage rates, and the blue-colored areas show lower wage rates. We can identify the high wage rate around the center of the city (Tejgaon, Hazaribagh), the north part of the city (Savar), where the Bangladesh Export Processing Zone is located, and the south part of the portage (Narayanganj). By construction, our estimated wages are correlated to the employment size. Panel (b) confirms a positive relationship between high wage rate and employment density. Intuitively, people commute to the unions with high wage rates as their return from commuting is solely high wages. Table D.5 shows the summary statistics. In this process, we can also compute the estimated commuting patterns. Figure D.5 confirms the gravity pattern of commuting, which relates the predicted number of bilateral commuters between unions and the commuting time between them.

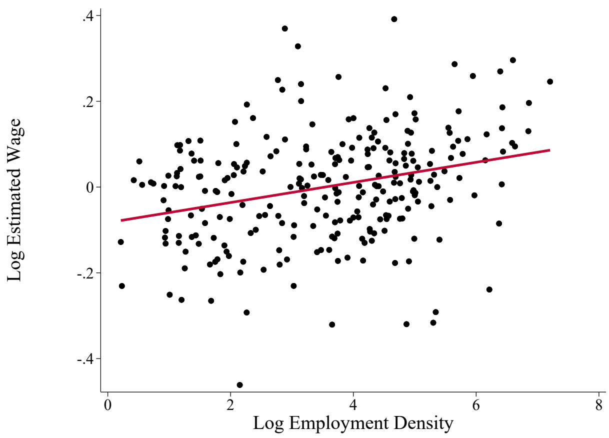
Once we obtain the estimated wages across unions and predicted patterns of commuting, we can compute the estimated wage income for workers. Figure D.6 displays the variation of the estimated income per worker. Since workers residing in unions around the central area have high accessibility to the places exhibiting high wage rates, workers in these places show high income per capita. In contrast, in the peripheral areas with a small number of employment, less accessibility to jobs leads to lower income per capita for workers residing there.

Figure D.7 presents that the variation of the volume of buildings in DLR data is strongly correlated with the total population (i.e., sum of employment and residents). Figure D.8 confirms that the variation in the demand for floor spaces reflects the total population. In this figure, we compare the total demand floor spaces based on the model and total population across wards in the city.

Figure D.4 : Estimated Wage Rate



(a) Estimated Wages in Dhaka



(b) Estimated Wage and Employment Density

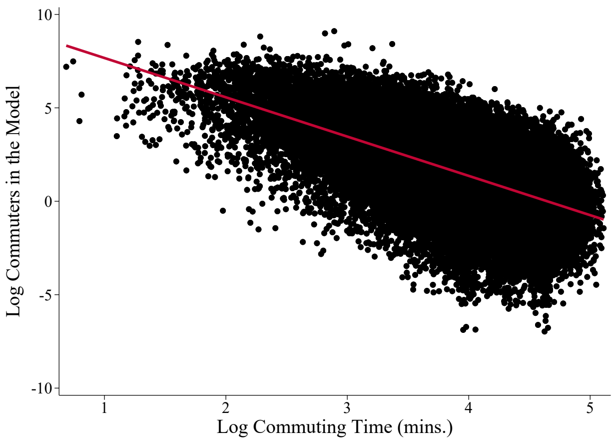
Note: The left panel shows mapping of calibrated wages. The right panel shows the relationship of calibrated wages and employment density.

Table D.5 : Summary Statistics for Estimated Wage Rate

Variable in calibration	Average	1 percentile	25 percentile	50 percentile	75 percentile	99 percentile
Wage rate	1.008	0.726	0.923	1.007	1.083	1.389
Income per capita	1.108	0.966	1.076	1.101	1.130	1.373

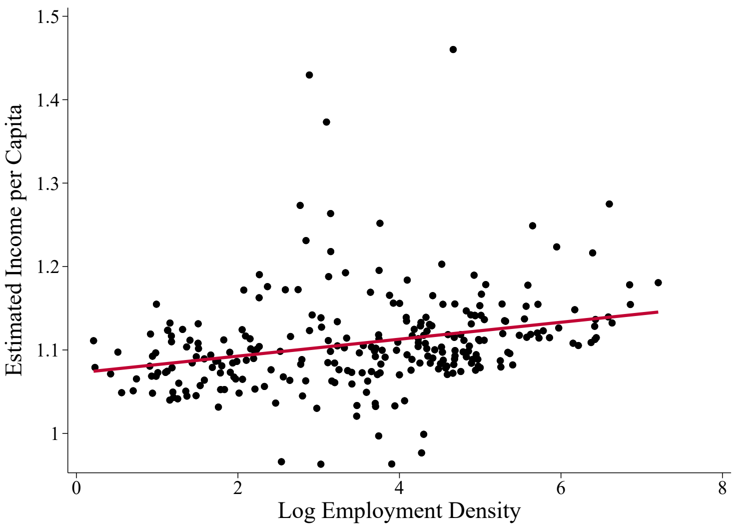
Note: This table reports summary statistics about the calibrated wages and income for 264 unions in Dhaka.

Figure D.5 : Gravity for Commuting



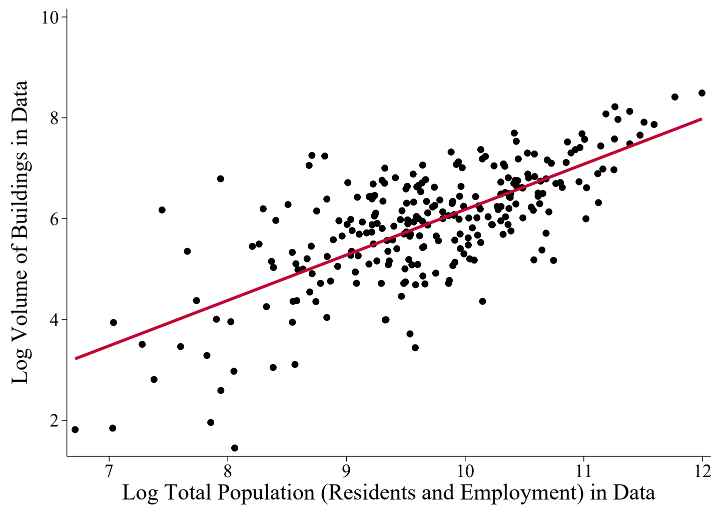
Note: The graph shows the gravity relationship between number of commuters predicted in the model and commuting time for each pair of different wards.

Figure D.6 : Estimated Income



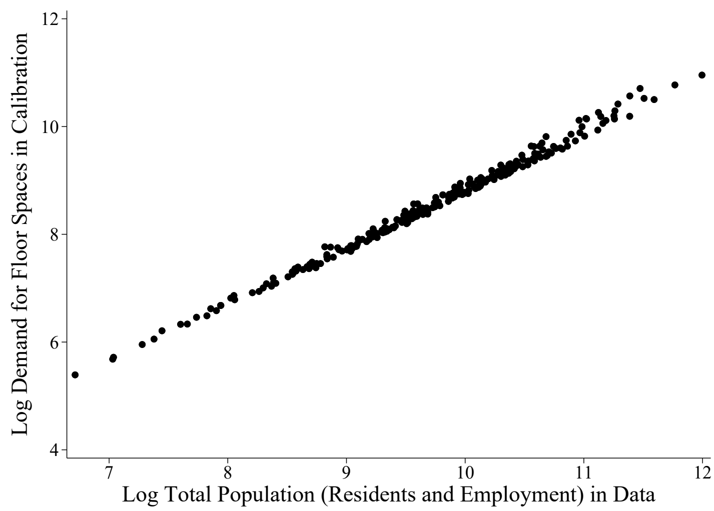
Note: The graph shows the income per capita calibrated in the model and employment density for different wards.

Figure D.7 : Volume of Buildings and Total Population



Note: Each observation in this figure is logarithm of the volume of buildings estimated by the DLR data from satellite images and logarithm of total population (sum of residential population and employment) for each ward in Dhaka. The red line is fitted line.

Figure D.8 : Demand for Floor Spaces and Total Population



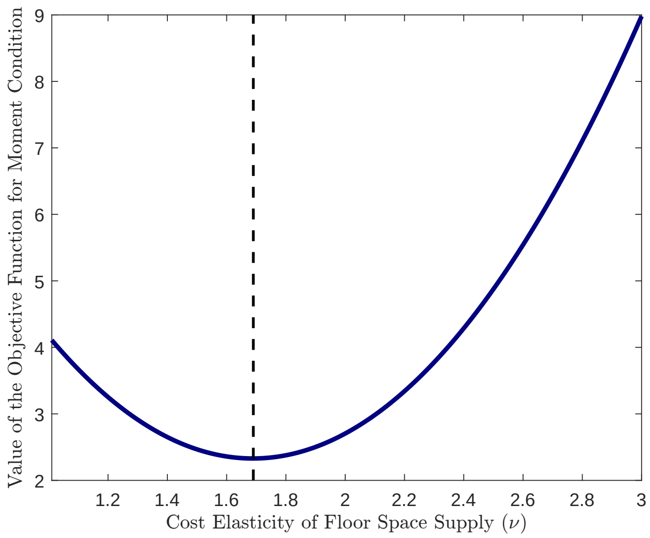
Note: Each observation in this figure shows the logarithm of demand for floor spaces computed by using estimated wages and the logarithm of total population that is sum of residential population and employment.

Estimation of floor space supply elasticity

Next, we estimate the parameter ν. Figure D.9 shows the value of the objective function for the moment condition. The dashed line corresponds to the value of ν, which achieves the minimum of the objective function and we use it as a baseline estimate, $\nu = 1.69$. The objective function has a unique minimum.

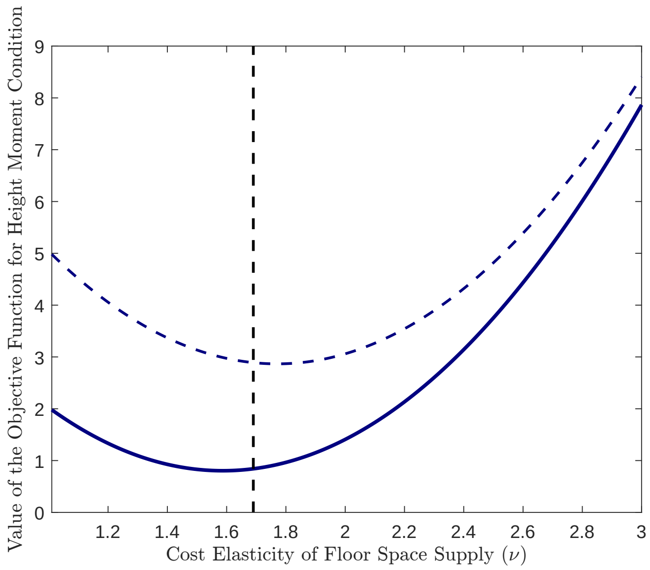
In the baseline, we define 10 different bins of unions based on the sum of population density and employment density. This is based on the idea that the variation of demand for floor spaces stems from both usages of business and residence. However, we can also define such bins based on different measures. In Figure D.10 shows the value of objective functions for two different definitions. The line shows the values when we define 10 bins based on the population density. Therefore, we weigh more on the demand for housing. This gives the unique minimum and the estimate is $\nu = 1.59$. In contrast, the dashed line shows the values when we define 10 bins based on the employment density. Then, we obtain the estimate $\nu = 1.77$. Intuitively, when we estimate the cost elasticity by weighting more on employment density, we obtain higher cost elasticity for the development of floor spaces. Our baseline estimate is in the middle of these two estimates.

Figure D.9 : Objective function for the moment condition of ν



Note: The graph shows the value of the objective function of the moment condition for housing supply elasticity (ν). We define the 10 different bins based on the total population density.

Figure D.10 : Objective function for the moment condition of ν based on different bins



Note: The graph shows the value of the objective function of the moment condition for housing supply elasticity (ν) when we use different definitions of bins. The line shows the objective function when we define 10 bins solely based on population density (i.e., density of workers in residential places). The dotted line shows that when we define 10 bins solely based on employment density (i.e., density of workers in workplaces).

Calibration of floor space prices and land rent

Table D.6 shows the summary statistics for the estimated land rent and floor space prices for 264 unions. Figure D.11 presents the variation of the estimated land rent and floor space prices across the distance from the CBD.

The estimated land rent shows a huge difference across locations in the city. Unions in the central area show relatively high land rent, and the land rent curve shows a negative slope from the city center. This is consistent with the monocentric city structure of Dhaka, and the demand for floor spaces in the central area for both production and housing drives high land rent and floor space prices.

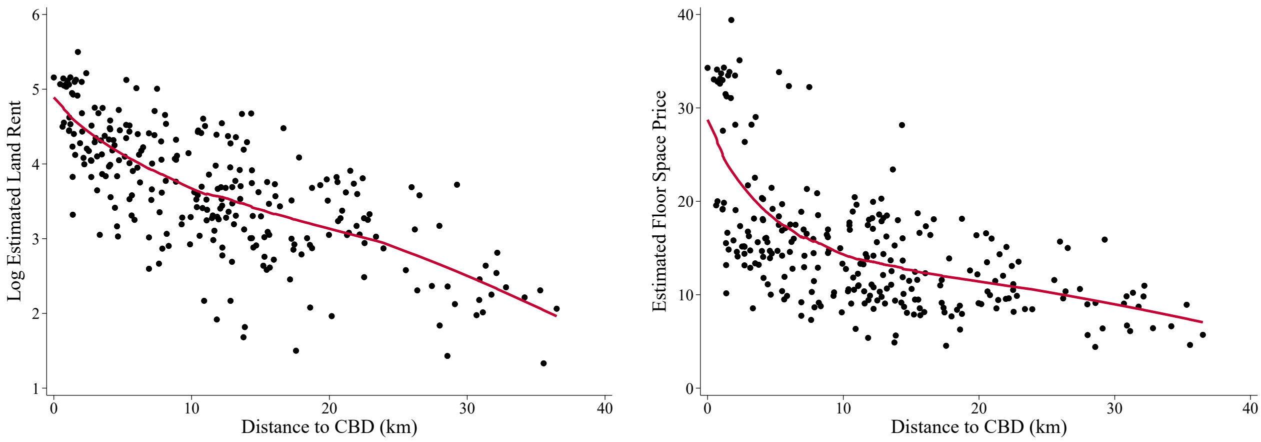
Figure D.12 shows the relationship between the estimated share of floor spaces used by firms and employment density. In the central district of the city, almost all of the floor spaces are used for production instead of housing. The concentration of firms in these areas is the primary source of the significant difference in floor space prices between the city center and the outer area.

Table D.6 : Summary Statistics for Estimated Land Rent and Floor Space Price

Variable in calibration	Average	1 percentile	25 percentile	50 percentile	75 percentile	99 percentile
Land rent	52.64	4.486	21.19	39.82	74.31	174.0
Floor space price	14.91	4.623	9.874	13.78	17.48	34.31

Note: This table reports summary statistics about the calibrated land rent and floor space prices for 264 unions in Dhaka.

Figure D.11 : Estimated Land Rent and Floor Prices



(a) Estimated Land Rent

(b) Estimated Floor Space Price

Note: The left panel shows the land rent calibrated in the model and distance to CBD for different wards. The right panel shows the floor space prices (Q_i) calibrated in the model and distance to CBD for different wards.

Figure D.12 : Estimated Floor Space Usage

Note: The graph shows the share of floor space used for busienss calibrated in the model and employment density for different wards.

Estimation of productivity spillovers

Next, we back out (overall) productivity and amenities. Using the estimated wages and floor space prices above, we compute the overall productivity by using the zero profit conditions in the production of tradable goods. We also exploit the population distribution across locations to infer the amenities.

Table D.7 reports the summary of the estimated productivity and amenities. As we show in the main text, the overall productivity is positively correlated with the employment density. The correlation is 0.6741. In contrast, the value of amenities is not significantly correlated with population density. Figure D.13 shows the variation of amenities and population density and the correlation is 0.2644. Given this, we decompose the overall productivity into the spillover term and exogenous term by estimating the spillover parameter (χ). As we discuss in the main text, we consider the moment condition for the fundamental advantage in productivity. Figure D.14 displays the values of objective function. The objective function shows the unique minimum at $\chi = 0.045$, and we use this estimate as a benchmark value for the local spillover. Using the value, we can compute the exogenous productivity (a_i). The third row in Table D.7 shows its summary statistics.

Our assumption for the identification for estimating the parameter of local spillover suppose that there is no systematic relationship between exogenous productivity and distance from the center. Figure D.15 shows the variation of the fundamental advantage in productivity and distance from the CBD. The correlation between the log of fundamental advantage and total population density (sum of employment density and residential population density) is -0.0182. This confirms that there is no systematic relationship between fundamental productivity advantages and total population density or distance from CBD.

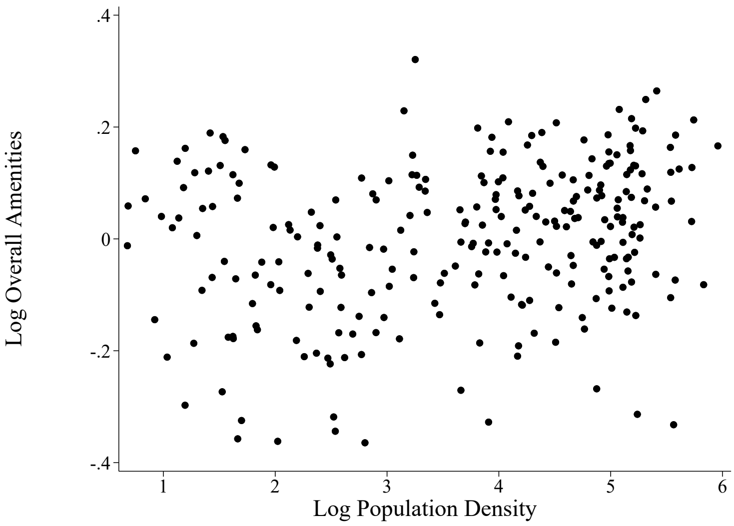
To gauge the role of spillovers, we also run the model by shutting down the spillovers. Namely, using the calibrated amenities, fundamental advantages and estimated parameters, we compute the equilibrium when setting $\chi = 0$. We compare such equilibrium and the baseline calibration. Figure D.16 shows the variation of floor space prices when we abstract the spillovers. The red line shows the fitted line for that case, and the dashed line is the fitted line for the baseline case. When there are no spillovers, the floor space prices in the central area of the city become lower than the baseline, but it does not show a dramatic drop. The difference reflects the gains from the agglomeration of workers. Figure D.17 displays the results for height of buildings. The central area shows the drop of height compared to the baseline in our data.

Table D.7 : Summary Statistics for Productivity and Amenities

Variable in calibration	Average	1 percentile	25 percentile	50 percentile	75 percentile	99 percentile
Productivity	1.007	0.750	0.927	0.995	1.073	1.377
Amenities	1.009	0.699	0.921	1.022	1.107	1.283
Exogenous productivity	0.856	0.687	0.803	0.852	0.897	1.066

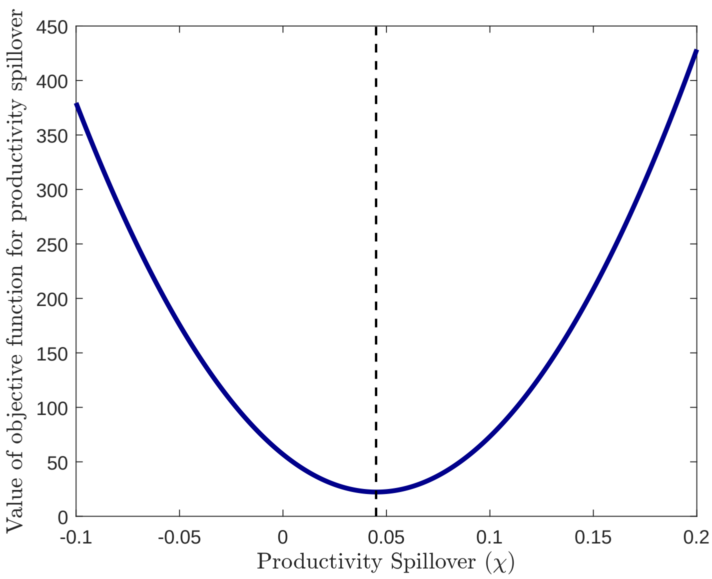
Note: This table reports summary statistics about the calibrated productivity (A_i), amenities (B_i) and fundamental advantages (a_i) for 264 unions in Dhaka.

Figure D.13 : Amenities and Population Density



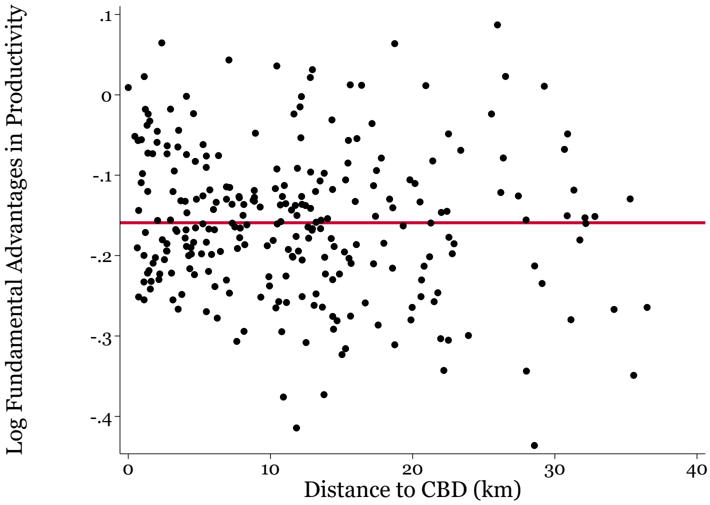
Note: The graph shows logarithm of the value of estimated amenities and logarithm of population density in each ward.

Figure D.14 : Objective function for moment condition of productivity spillovers



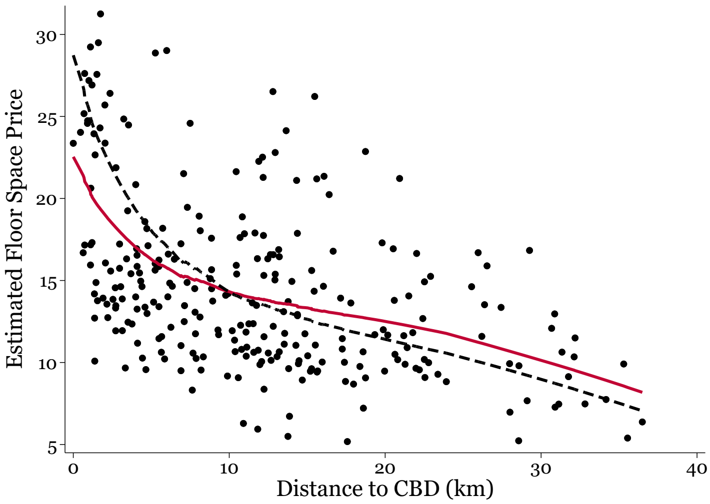
Note: The graph shows the value of objective functions for the moment condition of productivity spillovers.

Figure D.15 : Examining the productivity moment condition



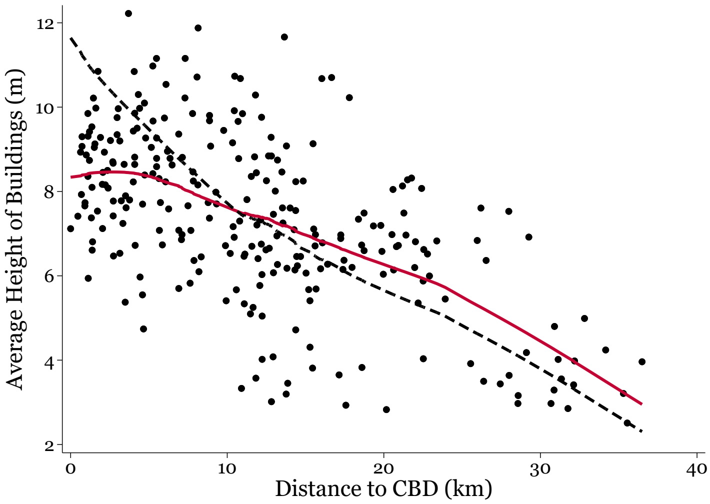
Note: This figure shows the variation in fundamental advantage in productivity. Each point shows the logarithm of fundamental advantage and distance to CBD for each location. The red horizontal line is the average log of fundamental productivity in the city.

Figure D.16 : Floor space price when $\chi = 0$



Note: This figure shows the variation in floor space prices in equilibrium when $\chi = 0$. Each point shows the floor space prices and distance to CBD for each location. The red line is the fitted line for the equilibrium floor space prices, while the dashed line is the fitted line for the baseline estimates.

Figure D.17 : Average Height when $\chi = 0$



Note: This figure shows the variation in heights in equilibrium when $\chi = 0$. Each point shows the average height and distance to CBD for each location. The red line is the fitted line for the equilibrium height, while the dashed line is for the average height in the baseline (data).

E Appendix: Counterfactuals

The model allows us to do counterfactual policy experiments in the city. This section provides the details about the counterfactual experiments. Subsection E.1 and E.2 explains the procedure of computing counterfactual change of the equilibrium for different types of counterfactual experiments. The former considers the change in the fundamentals, while the latter considers the change in the parameter. Subsection E.3 provides the additional counterfactual results.

E.1 Computing counterfactuals when fundamentals are changed

Suppose that any exogenous variables of location characteristics in the model have changed. For instance, we can consider change in commuting cost (d_{ni}), fundamental productivity (a_i), amenities (B_n) and development of land (T_i). For these changes, we can use the exact hat algebra. For any variable $\boldsymbol{x}$, we let $\hat{\boldsymbol{x}}$ denote the change of the variable relative to the baseline. We can solve the equilibrium conditions for the relative change of endogenous variables $(\hat{\boldsymbol{w}}, \hat{\boldsymbol{R}}, \hat{\boldsymbol{L}}, \hat{\boldsymbol{q}}, \hat{\boldsymbol{Q}}, \hat{\boldsymbol{H}})$ in the counterfactual equilibrium relative to their levels in the baseline equilibrium. We describe how to compute these counterfactual results in the followings.

Closed city As an example, we consider the proportional change in travel time in a city. That is, we have $\hat{d}_{in} \neq 1$ for some i and n. We guess the proportional change in employment ($\hat{\boldsymbol{L}}^0$) and population ($\hat{\boldsymbol{R}}^0$). In a closed city case, total population in the city is fixed. Therefore, we must have $\sum_i \lambda_i^L \hat{L}_i = 1$ and $\sum_n \lambda_n^R \hat{R}_n = 1$ where λ_i^L is employment share in i and λ_n^R is population share in n for the baseline (i.e., observation).

1. Given the guess of proportional change in employment $\hat{\boldsymbol{L}}^0$, we compute the proportional change in productivity
$$\hat{A}_i = \left(\hat{L}_i^0\right)^{\chi} \tag{E.1}$$

2. We solve the commuting market clearing conditions for the change in wage. Namely, we solve the system of N equations:
$$\hat{L}_i L_i = \sum_{n \in \mathcal{S}} \frac{(\hat{w}_i w_i)^{\varepsilon} (\hat{d}_{in} d_{in})^{-\varepsilon}}{\sum_{j \in \mathcal{S}} (\hat{w}_j w_j)^{\varepsilon} (\hat{d}_{jn} d_{jn})^{-\varepsilon}} \hat{R}_n R_n \tag{E.2}$$
for $\hat{w}_i$, and it is unique. Using the proportional change in wages, we compute the proportional change in income per worker residing in location n is:
$$\hat{\bar{w}}_n \bar{w}_n = \sum_{i \in \mathcal{S}} \frac{(\hat{w}_i w_i)^{\varepsilon} (\hat{d}_{in} d_{in})^{-\varepsilon}}{\sum_{j \in \mathcal{S}} (\hat{w}_j w_j)^{\varepsilon} (\hat{d}_{nj} d_{nj})^{-\varepsilon}} \hat{w}_i w_i \tag{E.3}$$

3. We substitute these changes into the land market clearing conditions to obtain the propor-

tional change in land rent:

$$\hat{r}_n = \hat{\bar{w}}_n \hat{R}_n^0 \left(\frac{\frac{\nu-1}{\nu}(1-\alpha)\bar{w}_n R_n}{r_n T_n} \right) + \hat{w}_n \hat{L}_n^0 \left(\frac{\frac{\nu-1}{\nu}\frac{\gamma}{\beta} w_n L_n}{r_n T_n} \right) \tag{E.4}$$

where on the right-hand side, the first parenthesis is share of developed land used for housing in the baseline and the second one is that for production.

4. Using the condition of profit maximization for developers, we obtain:

$$\begin{aligned} \hat{Q}_n &= (\hat{r}_n)^{\frac{\nu-1}{\nu}}, \\ \hat{H}_n &= (\hat{Q}_n)^{\frac{1}{\nu-1}} \end{aligned} \tag{E.5}$$

5. we compute the proportional change of commuting probabilities:

$$\hat{\lambda}_{ni} = \frac{(\hat{w}_i)^{\varepsilon} \hat{d}_{ni}^{-\varepsilon} (\hat{Q}_n)^{-(1-\alpha)\varepsilon}}{\sum_{k \in \mathcal{S}} \sum_{j \in \mathcal{S}} \lambda_{kj} (\hat{w}_j)^{\varepsilon} \hat{d}_{kj}^{-\varepsilon} (\hat{Q}_k)^{-(1-\alpha)\varepsilon}} \tag{E.6}$$

where λ_{kj} is the probability in baseline.

6. Using this change in commuting probabilities, $\hat{\boldsymbol{\lambda}}$, we compute proportional changes in population and employment:

$$\begin{aligned} \hat{R}_n^1 &= \sum_{i \in \mathcal{S}} \hat{\lambda}_{ni} \frac{\lambda_{ni} L}{R_n}, \\ \hat{L}_i^1 &= \sum_{n \in \mathcal{S}} \hat{\lambda}_{ni} \frac{\lambda_{ni} L}{L_i} \end{aligned} \tag{E.7}$$

7. We update the vector of proportional change in employment and residents:

$$\begin{aligned} \hat{\boldsymbol{L}} &= (1-\varsigma)\hat{\boldsymbol{L}}^0 + \varsigma \hat{\boldsymbol{L}}^1, \\ \hat{\boldsymbol{R}} &= (1-\varsigma)\hat{\boldsymbol{R}}^0 + \varsigma \hat{\boldsymbol{R}}^1 \end{aligned} \tag{E.8}$$

with some update scalar $\varsigma \in (0,1)$. We continue this until $|\hat{L}_i^1 - \hat{L}_i^0|$ and $|\hat{R}_n^1 - \hat{R}_n^0|$ converge to zero for appropriate norm.

Welfare change Using the proportional change in these variables, the change in expected utility of workers in the city can be computed by:

$$\hat{\bar{U}} = \frac{\hat{B}_n}{\hat{d}_{nn}} \frac{\hat{w}_n}{\hat{P}^{\alpha} \hat{Q}_n^{1-\alpha}} \hat{\lambda}_{nn}^{-1/\varepsilon} \tag{E.9}$$

This holds for every location n as expected utility ex-ante is equalized across location choices. Therefore, given the change in amenities ($\hat{B}_n$) and own commuting costs ($\hat{d}_{nn}$), the welfare gains in the counterfactuals are determined by change in wage ($\hat{w}_n$), change in floor space prices ($\hat{Q}_n$) and change of noncommuting probability ($\hat{\lambda}_{nn}$). Each component of welfare change may differ across locations n, but their net effect becomes same in terms of expected utility at the city level. This formula is close to the idea of gains from trade in the context of international trade (Arkolakis et al. (2012)). To investigate this, we consider each part in (E.9) such that:

$$\hat{\bar{U}} = \underbrace{\frac{\hat{B}_n}{\hat{d}_{nn}}}_{\text{Exogenous}} \times \underbrace{\frac{\hat{w}_n}{\hat{P}^{\alpha}\hat{Q}_n^{1-\alpha}}}_{\text{Real income}} \times \underbrace{\hat{\lambda}_{nn}^{-1/\varepsilon}}_{\text{No commuting}} \tag{E.10}$$

The first term is a change in an exogenous component in the utility of workers. The second term is a change in the real income of workers residing in n and working in n (i.e., no commuters). If we do not allow change in location choices of workers in response to the shock, the change in average utility of workers residing in n depends on the exogenous change and real income changes. However, that cannot be equilibrium since workers do not maximize their utility, taking account of taste shocks. The real income changes can differ across locations n. The last term counteracts such real income changes. For locations with an increase in real income for no commuters, their no-commuting probabilities increase in response to the shocks. Therefore, these workers are less benefit from commuting. In contrast, workers residing in a place with lower real income changes exhibit a small probability of non-commuting. They largely gain from commuting through the third term in (E.10). In the equilibrium after the re-optimization of workers in their location choices, the net effect is equalized across locations. The last term in (E.10) depends on the commuting elasticity ε. A small value of ε leads to large effects of this part of gains from commuting since workers with large heterogeneity in tastes across locations are able to choose location pairs more freely.

In particular, change in non-commuting probability is:

$$\hat{\lambda}_{nn} = \frac{(\hat{B}_n\hat{w}_n)^{\varepsilon}\hat{d}_{nn}^{-\varepsilon}(\hat{Q}_n)^{-(1-\alpha)\varepsilon}}{\sum_{k\in\mathcal{S}}\sum_{j\in\mathcal{S}}\lambda_{kj}(\hat{B}_k\hat{w}_j)^{\varepsilon}\hat{d}_{kj}^{-\varepsilon}(\hat{Q}_k)^{-(1-\alpha)\varepsilon}} \tag{E.11}$$

and $\hat{\lambda}_{nn} < 1$ when bilateral commuting costs are reduced with keeping the own commuting cost unchanged ($\hat{d}_{nn} = 1$). This leads to welfare gains through the last term of $\hat{\lambda}_{nn}^{-1/\varepsilon}$.

Taking the geometric mean for the welfare changes, we derive:

$$\hat{\bar{U}} = \hat{\bar{B}} \times \hat{\bar{d}}^{-1} \times \hat{\bar{w}} \times \hat{\bar{Q}}^{-(1-\alpha)} \times \hat{\bar{\lambda}}^{-1/\varepsilon} \tag{E.12}$$

This equation allows us to decompose the welfare change into the different margins: (i) average change in the benefit from amenity, (ii) average change of own commuting time, (iii) average change in wage rate, (iv) average change in floor space prices, and (v) average changes in the residential place advantage.

Open city with total population change When we depart from the closed city framework, the total city population (L) relative to the population in the wider economy ($\bar{L}$) increases as follows:

$$L = \left(\frac{\bar{U}}{\mathbb{U}}\right)^{\sigma} \bar{L} \tag{E.13}$$

where $\bar{L}$ and $\mathbb{U}$ are exogenous. Suppose that there is a positive shock to the exogenous location characteristics in a city, which leads to a higher average utility $\bar{U}$. The increase in the attractiveness of the city induces more inflow of workers to the city; therefore, the total population L increases. This creates an additional effect in general equilibrium. An increase in workers is associated with higher productivity through agglomeration economies. In contrast, a larger population may lead to increase congestion in a city through higher floor space prices and it offsets the part of the utility gain of workers. This mechanism generates the difference in welfare effect compared to the closed city framework.

The procedure to compute the counterfactual equilibrium is similar to the closed city case.

1. Same as in (E.1) to compute proportional change of productivity.

2. We compute proportional change in wage by solving (E.2) and compute change in income per capita (E.3).

3. We compute the proportional change for floor space production (E.5) and commuting probabilities (E.6).

4. Given them, we compute the proportional change of expected utility:

$$\hat{\bar{U}} = \left(\sum_{n\in\mathcal{S}}\sum_{i\in\mathcal{S}} \lambda_{ni}(\hat{B}_n\hat{w}_i)^{\varepsilon}\hat{d}_{ni}^{-\varepsilon}(\hat{Q}_n)^{-(1-\alpha)\varepsilon}\right)^{1/\varepsilon} \tag{E.14}$$

   Then, the labor mobility condition leads to the proportional change in total city population:

$$\hat{L} = \left(\hat{\bar{U}}\right)^{\sigma} \tag{E.15}$$

5. We compute proportional changes in population and employment:

$$\begin{aligned} \hat{R}_n^1 &= \sum_{i\in\mathcal{S}} \hat{\lambda}_{ni}\hat{L}\frac{\lambda_{ni}L}{R_n}, \\ \hat{L}_i^1 &= \sum_{n\in\mathcal{S}} \hat{\lambda}_{ni}\hat{L}\frac{\lambda_{ni}L}{L_i} \end{aligned} \tag{E.16}$$

6. We update the vector of proportional change in employment and residents as in (E.8).

E.2 Computing counterfactuals when structural parameter is changed

OtherOther counterfactual experiments entail the change in the structural parameter. For instance, we consider the different values of parameter ν related to floor supply elasticity. For such counterfactual experiments, we cannot use the exact hat algebra. Therefore, we solve the whole system of equations for endogenous variables after replacing the parameter.

We start from the guess of employment and population, (L_i^0, R_n^0).

1. Given commuting costs (d_{in}) and the shape parameter of Fréchet distribution (ε), we solve the commuting market clearing conditions for wages $\{w_i\}$ and compute average income $\{\bar{w}_n\}$.

2. Given parameters (α, β, γ), new parameter ν' and size of developed land (T_n), land market clearing conditions lead to land rent:

$$r_n = \frac{\nu' - 1}{\nu'} \frac{1}{T_n} \left((1-\alpha)\bar{w}_n R_n^0 + \frac{\gamma}{\beta} w_n L_n^0 \right) \tag{E.17}$$

3. Given fundamentals in floor space supply (κ_n) and new parameter nu', we compute floor space price:

$$Q_n = \kappa_n \nu' \left(\frac{r_n}{\kappa_n(\nu' - 1)} \right)^{(\nu'-1)/\nu'} \tag{E.18}$$

4. Plugging wage and floor space prices into commuting probabilities yield λ_{ni}.

5. Using these probabilities, we compute the population and employment:

$$\begin{aligned} R_n^1 &= \sum_i \lambda_{ni} L, \\ L_i^1 &= \sum_n \lambda_{ni} L \end{aligned} \tag{E.19}$$

   and we update population and employment vectors until they converge.

We start with the baseline equilibrium, and this process leads to the new equilibrium.

Note that the formula (E.9) also holds for this case as long as we fix the values of parameters in preference (α) and commuting elasticity (ε). Letting $\bar{U}'$ refer to the average utility under the new parameter of ν', the welfare change for workers in the city relative to the baseline is:

$$\frac{\bar{U}'}{\bar{U}} = \left(\frac{w_n'}{w_n} \right) \left(\frac{P'}{P} \right)^{-\alpha} \left(\frac{Q_n'}{Q_n} \right)^{-(1-\alpha)} \left(\frac{\lambda_{nn}'}{\lambda_{nn}} \right)^{-1/\varepsilon} \tag{E.20}$$

where we have used that amenities and commuting times are unchanged.

E.3 Additional results for counterfactuals

This subsection includes additional results for the counterfactual experiments.

Changes in floor space supply elasticity

Our first counterfactual experiment is about the floor space supply elasticity. As a baseline scenario, we consider an increase in floor space supply elasticity by 25 percent, which implies a change in parameter ν.

Table E.1 summarizes the relative change in variables of interest from the baseline. There are large variation in the change in population density and employment density. The average height of buildings increases by 33 percent on average, and more than 50 percent in 99 percentile.

Figure E.1 show mapping of the results about change in log of wages and change in log of floor space prices from the baseline equilibrium. In response to an increase in floor supply elasticity, more floor spaces for production are supplied in the central area and workers are concentrated in these area. This leads to increase in wages and drop in price of floor spaces.

Figure E.2 shows the change in average commuting time of workers. For each union, we compute the average commuting time of workers residing there, and we find decline in commuting time in the central area, while increase in the peripheral area. Since we fix travel time as in the baseline, the source of these changes is workers' location changes response to concentration of jobs in the central area and further supply of housing in peripheral areas. Workers in peripheral area face longer commuting time compared to the baseline because their workplaces are more concentrated in the central area.

Table E.2 shows the results when we increase floor supply elasticity by 50 percent, while Table E.3 shows when we decrease floor supply elasticity by 25 percent. We find large effects in the real income changes of workers and total population changes, while change in non-commuting probabilities are modest.

Table E.4 shows the counterfactual results for different values of labor supply, σ. When we set large value $\sigma = 6.0$, total population in the city increase by 27 percent and it offset the decline in land revenue. The total land revenue increases by more than 10 percent. If we set $\sigma = 1.0$, population increases by around 5 percent. The small increase leads to large welfare gains for workers in the city and loss of revenue for landlords.

Lastly, we describe the mathematical implications for the counterfactual effects with focus on the first counterfactual experiment. Considering the change in the parameter ν in the first counterfactual experiment. Using profit maximization of developers, changes in heights and changes in floor space prices response to change in the parameter satisfy:

$$\ln h_n + \frac{\nu - 1}{\nu}\frac{d\ln h_n}{d\ln \nu} = \frac{1}{\nu}\frac{d\ln Q_n}{d\ln \nu} - \frac{1}{\nu}\frac{d\ln \kappa_n}{d\ln \nu} - \frac{1}{\nu} \tag{E.21}$$

Assuming that $\frac{d\ln \kappa_n}{d\ln \nu} = 0$ and $\frac{d\ln T_n}{d\ln \nu} = 0$, we obtain:

$$\frac{d\ln Q_n}{d\ln \nu} = \nu \ln h_n + (\nu - 1)\frac{d\ln h_n}{d\ln \nu} + 1 \tag{E.22}$$

This presents the relationship between the spatial variation in the effects on floor space prices and that in heights. The important feature in this relationship is that the change in floor space price with respect to the small change in ν depends on the heights in the baseline ($\ln h_n$). In our main counterfactual analysis, we set smaller value of ν from the baseline value, and we find:

$$\frac{d \ln h_n}{d \ln \nu} < 0 \tag{E.23}$$

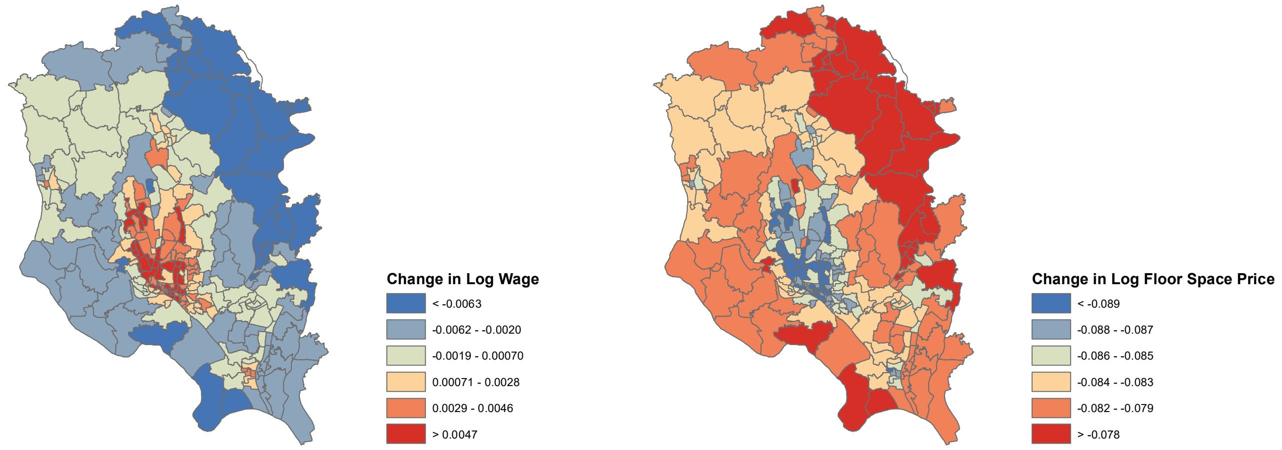
When we compare the central area of the city and periphery, we find very large difference in the response of heights. The central area shows further negative values. However, in the city center, the average height of buildings in the baseline is high and therefore the variation in the floor space prices across locations becomes small. Comparing Figure E.3 and Figure E.4 , the change in average height of buildings shows larger variation relative to the change in floor space prices.

Table E.1 : Summary Statistics for Impact of an 25 % Increase of Floor Supply Elasticity

Variable in change	Average	1 percentile	25 percentile	50 percentile	75 percentile	99 percentile
Population density	1.104	0.915	1.062	1.107	1.171	1.212
Employment density	1.084	0.840	1.017	1.088	1.173	1.230
Wage rate	1.000	0.972	0.994	1.001	1.009	1.015
Floor space prices	0.823	0.811	0.816	0.823	0.828	0.845
Land rent	0.958	0.753	0.913	0.959	1.030 1.072	
Average height	1.337	1.023	1.263	1.338	1.448	1.516
Productivity	1.000	0.989	0.997	1.000	1.004	1.006

Note: This table reports summary statistics about the counterfactual changes when elasticity of floor space supply increases by 25 percent for 264 unions.

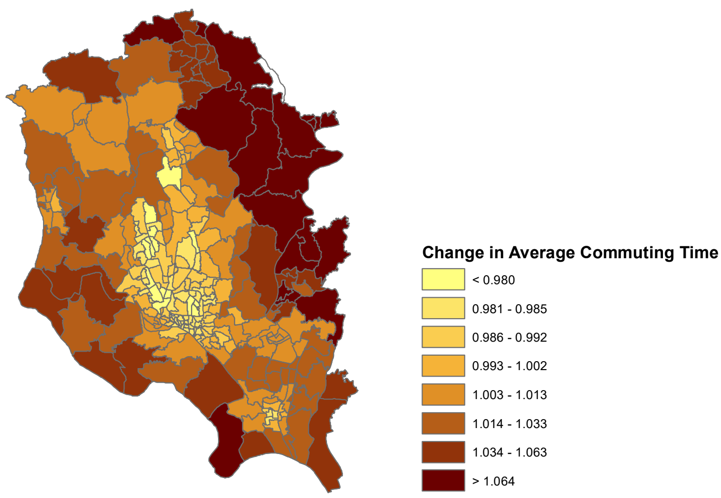
Figure E.1 : Impact of an 25% Increase of Floor Supply Elasticity in the City: Wage and Floor Space Prices



(a) Change in Log Wage **(b)** Change in Log Floor Space Price

Note: The left hand panel shows the counterfactual changes in log of wages (w_i) and the right hand panel shows those in floor space prices (Q_i) to the baseline data.

Figure E.2 : Impact of an 25% Increase of Floor Supply Elasticity in the City: Commuting Time



Note: This shows the change in average commuting time relative to the baseline.

Table E.2 : Counterfactuals for Building Supply Elasticity: 50 percent increase

New parameter: $\nu = 1.46$	(1) Closed City	(2)	(3) Open City	(4)
Counterfactuals (relative to baseline)	Fix Productivity	With Spillovers	Fix Productivity	With Spillovers
Average Welfare of Workers ($\bar{U}$)	1.1143	1.1262	1.0971	1.1003
– Average Non-Commuting Probability ($\widetilde{\lambda_{\mathrm{N}}}^{-1/\varepsilon}$)	1.0034	1.0130	1.0042	1.0072
– Average Real Income	1.1106	1.1118	1.0924	1.0925
Total Land Revenue ($\sum_n r_n T_n$)	0.7644	0.7717	0.9326	0.9425
Total Population ($\bar{L}$)			1.2035	1.2107

Note: These counterfactual exercises assume that building supply elasticity increases. For each counterfactual, the numbers are relative values to the baseline equilibrium. In counterfactuals (1) and (2), we assume the closed city. Counterfactual (1) supposes that overall productivity (A_i) is fixed at the baseline level. In counterfactual (2), we also allow spillovers in productivity so that overall productivity changes with employment density. In counterfactual (3) and (4), we consider the open city where workers can move into and out of the city. Counterfactual (3) fixes overall productivity as in (1), and counterfactual (4) allows productivity spillovers as in (2).

Table E.3 : Counterfactuals for Building Supply Elasticity: 25 percent decline

New parameter: $\nu = 1.92$	(1) Closed City	(2)	(3) Open City	(4)
Counterfactuals (relative to baseline)	Fix Productivity	With Spillovers	Fix Productivity	With Spillovers
Average Welfare of Workers ($\bar{U}$)	0.9266	0.9319	0.9395	0.9388
– Average Non-Commuting Probability ($\widetilde{\lambda_{\mathrm{N}}}^{-1/\varepsilon}$)	0.9966	1.0015	0.9983	0.9975
– Average Real Income	0.9297	0.9305	0.9412	0.9412
Total Land Revenue ($\sum_n r_n T_n$)	1.1533	1.1600	1.0343	1.0316
Total Population ($\bar{L}$)			0.8827	0.8814

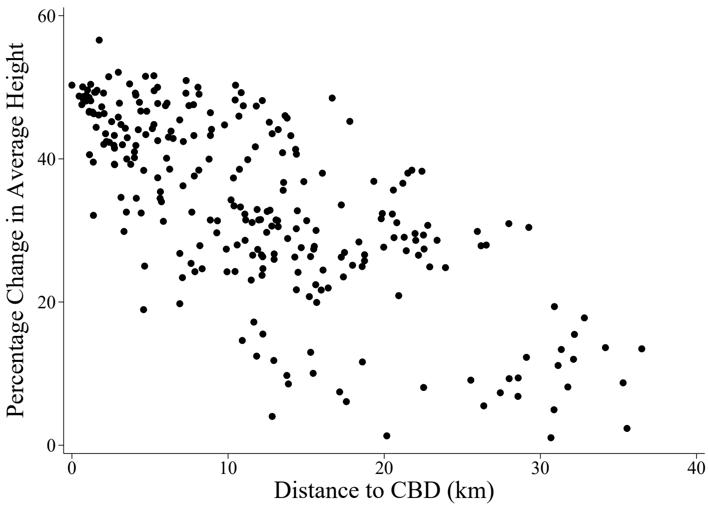
Note: These counterfactual exercises assume that building supply elasticity decreases. For each counterfactual, the numbers are relative values to the baseline equilibrium. In counterfactuals (1) and (2), we assume the closed city. Counterfactual (1) supposes that overall productivity (A_i) is fixed at the baseline level. In counterfactual (2), we also allow spillovers in productivity so that overall productivity changes with employment density. In counterfactual (3) and (4), we consider the open city where workers can move into and out of the city. Counterfactual (3) fixes overall productivity as in (1), and counterfactual (4) allows productivity spillovers as in (2).

Table E.4 : Counterfactuals for Building Supply Elasticity for different values of σ

Increase floor supply elasticity by 25 %	(1) $\sigma = 6.0$	(2)	(3) $\sigma = 1.0$	(4)
Counterfactuals (relative to baseline)	Fix Productivity	With Spillovers	Fix Productivity	With Spillovers
Average Welfare of Workers ($\bar{U}$)	1.0397	1.0407	1.0563	1.0578
– Average Non-Commuting Probability ($\widetilde{\lambda_{\mathrm{N}}}^{-1/\varepsilon}$)	1.0020	1.0033	1.0020	1.0033
– Average Real Income	1.0376	1.0373	1.0542	1.0543
Total Land Revenue ($\sum_n r_n T_n$)	1.1025	1.1111	0.9221	0.9251
Total Population ($\bar{L}$)	1.2631	1.2705	1.0563	1.0578

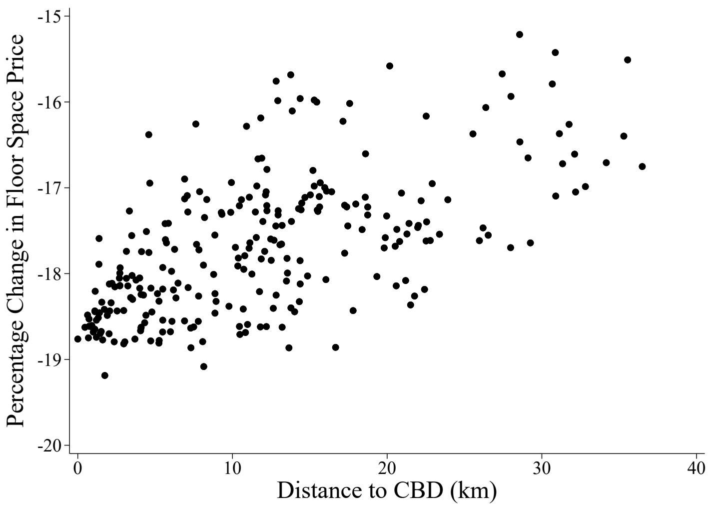
Note: This table shows the results for different values σ. Column (1) and (2) set $\sigma = 6.0$, and Column (3) and (4) set $\sigma = 1.0$. The counterfactual scenario is the same as in the main text. The floor space supply elasticity is increased by 25 percent.

Figure E.3 : Changes in Average Height of Buildings



Note: Each observation in this figure is percentage change in average height of buildings and distance from the CBD. The counterfactual experiment is a 25 percent increase in floor supply elasticity.

Figure E.4 : Changes in Floor Space Prices



Note: Each observation in this figure is percentage change in floor space prices and distance from the CBD. The counterfactual experiment is a 25 percent increase in floor supply elasticity.

E.4 Changes in travel time

Our second counterfactual experiment undertakes the change in travel time. In the main text, we focus on building a highway through the city, which decreases travel time through treated areas. In this subsection, we consider different scenarios.

Table E.5 shows the results for the scenario where all bilateral travel time between any different unions decreases by 25 percent. Travel time within own union is unchanged. Welfare increases around 6 percent for closed city case and 4 percent for the case where we allow population growth (population grows around 9 percent). Real income slightly changes, and the source of welfare gains for workers in the city depends on the change in commuting patterns.

Table E.5 : Counterfactuals for Travel Time: Uniform Change

Uniform change in travel time by -25 %	(1) Closed City	(2)	(3) Open City	(4)
Counterfactuals (relative to baseline)	Fix Productivity	With Spillovers	Fix Productivity	With Spillovers
Average Welfare of Workers ($\bar{U}$)	1.0631	1.0632	1.0423	1.0473
– Average Non-Commuting Probability ($\widetilde{\lambda_{\mathrm{N}}}^{-1/\varepsilon}$)	1.0707	1.0726	1.0707	1.0726
– Average Real Income	0.9930	0.9912	0.9736	0.9764
Total Land Revenue ($\sum_n r_n T_n$)	1.0049	1.006	1.0918	1.1044
Total Population ($\bar{L}$)			1.0865	1.0968

Note: These counterfactual exercises assume that travel time between different unions uniformly decrease by 25 percent. For each counterfactual, the numbers are relative values to the baseline equilibrium. In counterfactuals (1) and (2), we assume the closed city. Counterfactual (1) supposes that overall productivity (A_i) is fixed at the baseline level. In counterfactual (2), we also allow spillovers in productivity so that overall productivity changes with employment density. In counterfactual (3) and (4), we consider the open city where workers can move into and out of the city. Counterfactual (3) fixes overall productivity as in (1), and counterfactual (4) allows productivity spillovers as in (2).

F Appendix: Extensions

As an example of the extension of our baseline model, we describe (i) how amenities at workplaces change the process of our calibration (subsection F.1), and (ii) how we can introduce home production (subsection F.2).

F.1 Amenities at workplaces

Suppose that workers receive utility benefits from amenities in the workplace. This reflects the idea that workers' location choice for the workplace partially depends on amenities they can enjoy in the workplace. This may include the consumption of amenities in the workplace.

Set up The indirect utility of individual worker ω becomes:

$$U_{ni}(\omega) = \frac{b_{ni}(\omega)\mathcal{B}_{ni}}{d_{ni}} \frac{w_i}{P_n^{\alpha} Q_n^{1-\alpha}} \tag{F.1}$$

where worker commuting from n to i receives the benefit from residential amenities (B_n^R) and workplace amenities (B_i^W):

$$\mathcal{B}_{ni} = B_n^R \times B_i^W \tag{F.2}$$

This alters commuting probability of workers:

$$\lambda_{ni} = \frac{d_{ni}^{-\varepsilon} \left(B_n^R B_i^W\right)^{\varepsilon} w_i^{\varepsilon} P_n^{-\alpha\varepsilon} Q_n^{-(1-\alpha)\varepsilon}}{\sum_{k\in\mathcal{S}} \sum_{\ell\in\mathcal{S}} d_{k\ell}^{-\varepsilon} \left(B_k^R B_\ell^W\right)^{\varepsilon} w_\ell^{\varepsilon} P_k^{-\alpha\varepsilon} Q_k^{-(1-\alpha)\varepsilon}} \tag{F.3}$$

This share of population is determined by wage distribution across workplaces (w_i), cost of the residential place (P_n and Q_n), amenities of the residential place (B_n^R) and workplace (B_i^W), and commuting disutility (d_{ni}).

Let $\lambda_{ni|n}$ be the conditional probability of commuting to i conditional on living in location n. This conditional probability is equal to:

$$\lambda_{ni|n} = \frac{\lambda_{ni}}{\sum_{k\in\mathcal{S}} \lambda_{nk}} = \frac{d_{ni}^{-\varepsilon} \left(B_i^W w_i\right)^{\varepsilon}}{\sum_{k\in\mathcal{S}} d_{nk}^{-\varepsilon} \left(B_k^W w_k\right)^{\varepsilon}} \tag{F.4}$$

Other parts of the model remain the same as in baseline model. In particular, we consider only homogeneous tradable goods that are treated as numeraire. Therefore, amenities do not include any endogenous values.

Model inversion We consider the model inversion when there are workplace amenities. When we introduce the workplace amenities, we cannot identify wage rate (w_i) itself, but we can solve the commuting market clearing condition for amenity adjusted wage, $\widetilde{w}_i$:

$$L_i = \sum_{n\in\mathcal{S}} \frac{d_{ni}^{-\varepsilon} \widetilde{w}_i^{\varepsilon}}{\sum_{k\in\mathcal{S}} d_{nk}^{-\varepsilon} \widetilde{w}_k^{\varepsilon}} R_n, \quad \widetilde{w}_i = B_i^W w_i \tag{F.5}$$

The only difference to the baseline model is that we have amenity adjustment. This is unique up to scale, therefore we normalize the vector such that $\left(\prod_{i\in\mathcal{S}} \widetilde{w}_i\right)^{1/N} = 1$.

The residential choice probability in location n is:

$$\lambda_n^R = \frac{R_n}{L} = \frac{\sum_{i \in \mathcal{S}} \left[\frac{B_n^R B_i^W}{d_{ni}} \frac{w_i}{Q_n^{1-\alpha}} \right]^{\varepsilon}}{(\bar{U}/\bar{\gamma})^{\varepsilon}} = \frac{\left[\frac{B_n^R}{Q_n^{1-\alpha}} \right]^{\varepsilon} \sum_{i \in \mathcal{S}} d_{ni}^{-\varepsilon} \widetilde{w}_i^{\varepsilon}}{(\bar{U}/\bar{\gamma})^{\varepsilon}} \tag{F.6}$$

When we observe the average height in location n, h_n, profit maximization and zero profit condition states:

$$Q_n = \kappa_n \nu h_n^{\nu - 1} \tag{F.7}$$

Therefore, we have:

$$\kappa_n \nu^{(1-\alpha)\varepsilon} \left(\frac{\bar{U}}{\bar{\gamma}} \right)^{\varepsilon} \frac{R_n}{L} = (B_n^R)^{\varepsilon} h_n^{-(\nu-1)(1-\alpha)\varepsilon} \widetilde{W}_n, \quad \widetilde{W}_n = \sum_{i \in \mathcal{S}} d_{ni}^{-\varepsilon} \widetilde{w}_i^{\varepsilon} \tag{F.8}$$

Taking the geometric means for variables, we obtain:

$$\frac{R_n}{\bar{R}} = \left[\frac{B_n^R}{\bar{B}^R} \right]^{\varepsilon} \left[\frac{h_n}{\bar{h}} \right]^{-(\nu-1)(1-\alpha)\varepsilon} \left[\frac{\widetilde{W}_n}{\bar{\widetilde{W}}} \right] \tag{F.9}$$

Plugging the population ($R_n/\bar{R}$), average height of buildings ($h_n/\bar{h}$) and calibrated workers' return ($\widetilde{W}_n/\bar{\widetilde{W}}$) into this equation, we can back out the value of amenities in residential place (B_n^R) up to scale.

Next, we consider productivity. The profit maximization and zero profit condition implies:

$$A_i = w_i^{\beta} Q_i^{\gamma} = (\kappa_i \nu)^{\gamma} \left(\frac{\widetilde{w}_i}{B_i^W} \right)^{\beta} h_i^{\gamma(\nu-1)} \tag{F.10}$$

Then, we define the amenity adjusted productivity in workplace:

$$\widetilde{A}_i = A_i (B_i^W)^{\beta} = (\kappa_i \nu)^{\gamma} \widetilde{w}_i^{\beta} h_i^{\gamma(\nu-1)} \tag{F.11}$$

Taking its geometric mean,

$$\frac{\widetilde{A}_i}{\bar{\widetilde{A}}} = \left[\frac{\widetilde{w}_i}{\bar{\widetilde{w}}} \right]^{\beta} \left[\frac{h_i}{\bar{h}} \right]^{\gamma(\nu-1)} \tag{F.12}$$

Plugging adjusted wages ($\widetilde{w}_i/\bar{\widetilde{w}}$) and average height of buildings ($h_i/\bar{h}$) into this, we back out the amenity adjusted productivity.

Summary In sum, conditional on the set of citywide parameters $(\alpha, \varepsilon, \beta, \gamma, \nu)$ and observations in data about commuting cost ($\boldsymbol{d}$), workers in residential places ($\boldsymbol{R}$), workers in workplaces ($\boldsymbol{L}$), and the average height of buildings ($\boldsymbol{h}$), we can obtain the unique unobserved vector of residential place amenity ($\boldsymbol{B}^R$) and amenity adjusted productivity in the workplace ($\widetilde{\boldsymbol{A}}$) that are consistent with the observed data to be equilibrium. The only difference to the baseline model is that wage rate and productivity are adjusted with the value of workplace amenities since their combination defines the return for workers, and they cannot be separately identified.

In practice, our calibrated wages ($\boldsymbol{w}$) in the baseline analysis become exactly the same as workplace amenity adjusted wages ($\widetilde{\boldsymbol{w}}$) in this extension. Therefore, $\boldsymbol{W}$ in the baseline and $\widetilde{\boldsymbol{W}}$ in this extension are also the same, and our estimated residential amenities must be the same between the baseline and the extension: $B_n = B_n^R$. Furthermore, we use the same conditions to back out the productivity and estimated productivity ($\boldsymbol{A}$) in the baseline and that in the extension ($\widetilde{\boldsymbol{A}}$) are also the same. These results conclude that all results in our counterfactuals are unchanged in this extension, as long as amenities are exogenous and fixed.

F.2 Employment effects with home production

We only consider workers who participate in the labor market in our baseline. When we take into account the population in general, there is a large number of people in home production. This is important when we do a counterfactual analysis of a policy. A policy change may alter the value of working, and more people choose to work or quit their job. It changes the labor participation and production in a city.

To see this effect in our framework, we consider two stages in workers' decisions. First, every worker ω in a city draws productivity for home production $z_H(\omega)$ and productivity in the formal sector of homogeneous tradable goods $z_T(\omega)$. The vector of productivity follows Fréchet distribution: $G(z) = e^{z^{-\theta}}$ with $\theta > 1$. Given this productivity draw, people choose whether they work in the formal sector or stay home production. Next, workers who choose to work in the formal sector draw idiosyncratic shocks for commuting, $b_{in}(\omega)$, from Fréchet distribution as in the baseline model, and they decide where they live and where they work. This nested Fréchet structure allows us to characterize the employment effect in a city. The one thing to be noted is that we need the assumption for the shape parameter of Fréchet distribution such that $\varepsilon \leq \theta$: where the elasticity of substitution between home production and employment is lower than that of substitution between locations.

<u>**Set up**</u> We solve the two steps decision by backward. Conditional on the choice of workplaces, individuals with productivity $z_T(\omega)$ for formal sector anticipate the average real income:

$$\mathbb{W}_T(\omega) = \bar{U} \cdot z_T(\omega) \tag{F.13}$$

with

$$\bar{U} = \bar{\gamma} \left[\sum_{n \in \mathcal{S}} \sum_{i \in \mathcal{S}} d_{ni}^{-\varepsilon} B_n^{\varepsilon} (w_i)^{\varepsilon} (P_n^{\alpha} Q_n^{1-\alpha})^{-\varepsilon} \right]^{1/\varepsilon} \tag{F.14}$$

Suppose an individual chooses home production after the draw of productivity $z_O(\omega)$. In that case, they operate a potentially less productive home production technology, with a return given by $z_O(\omega) \cdot w_O$. The wage rate per efficient unit of labor is w_O and it is an exogenous parameter in a city. Voluntary unemployment has the same preference as workers. Their preference is characterized by Cobb-Douglas function over tradable homogeneous goods and residential floor spaces. In addition, they receive utility benefits from the residential place and draw idiosyncratic benefits from residential choice. In particular, their indirect utility conditional on location choice is given by:

$$V_n^O(\omega) = B_n b_n(\omega) \frac{w_O z_O(\omega)}{P_n^\alpha Q_n^{1-\alpha}} \tag{F.15}$$

where $b_n(\omega)$ follows the same distribution as workers: Fréchet distribution with shape parameter ε. Therefore, expected real income from home production for individual worker ω becomes:

$$\mathbb{W}_O(\omega) = \bar{U}^O \cdot z_O(\omega) \tag{F.16}$$

with

$$\bar{U}^O = \bar{\gamma} \left[\sum_{n \in \mathcal{S}} B_n^\varepsilon (w_O)^\varepsilon (P_n^\alpha Q_n^{1-\alpha})^{-\varepsilon} \right]^{1/\varepsilon} \tag{F.17}$$

Now, we derive the probability of working in formal sector and home production. Since $z_T(\omega)$ and $z_O(\omega)$ are following independent Fréchet distribution with shape parameter θ, we have:

$$\pi_T = \left(\frac{\bar{U}}{\mathbb{W}} \right)^\theta, \quad \pi_O = \left(\frac{\bar{U}^O}{\mathbb{W}} \right)^\theta, \tag{F.18}$$

with

$$\mathbb{W} = \left\{ (\bar{U})^\theta + (\bar{U}^O)^\theta \right\}^{1/\theta} \tag{F.19}$$

Using this, we obtain the measure of workers and home production in a city:

$$L = \pi_T \mathcal{M}, \quad L^O = \pi_O \mathcal{M} \tag{F.20}$$

where $\mathcal{M}$ is total population in a city. Once people observe the realization of productivity $z_T(\omega)$ and decide to work in the formal sector, the commuting pattern is given by

$$\lambda_{ni} = \frac{d_{ni}^{-\varepsilon} B_n^\varepsilon w_i^\varepsilon P_n^{-\alpha\varepsilon} Q_n^{-(1-\alpha)\varepsilon}}{\sum_{k \in \mathcal{S}} \sum_{\ell \in \mathcal{S}} d_{k\ell}^{-\varepsilon} B_k^\varepsilon w_\ell^\varepsilon P_k^{-\alpha\varepsilon} Q_k^{-(1-\alpha)\varepsilon}} \tag{F.21}$$

Note that formal workers' probability of location choices is independent of their productivity since their indirect utility is proportional to productivity. For home production, the probability that they choose to live in location n is:

$$\lambda_n^O = \frac{B_n^\varepsilon P_n^{-\alpha\varepsilon} Q_n^{-(1-\alpha)\varepsilon}}{\sum_{k \in \mathcal{S}} B_k^\varepsilon P_k^{-\alpha\varepsilon} Q_k^{-(1-\alpha)\varepsilon}} \tag{F.22}$$

We let

$$\Phi_z \equiv \{ \boldsymbol{z} \ : \ \bar{U} \cdot z_T \geq \bar{U}^O \cdot z_O \} \tag{F.23}$$

and people in a city choose to work in the formal sector if and only if their productivity draw is $\boldsymbol{z} \in \Phi_z$. We also let $G(\boldsymbol{z})$ refer to the joint distribution of productivity $\boldsymbol{z}$ for people in the city. Then, the total units of efficient labor in a city is:

$$E = \mathcal{M} \int_{\Phi_z} z dG(\boldsymbol{z}) = \bar{\zeta} \pi_T^{1-\frac{1}{\theta}} \mathcal{M} = \bar{\zeta} \pi_T^{-1/\theta} L \tag{F.24}$$

where $\bar{\zeta}$ is value of Gamma function $\Gamma\left(\frac{\theta-1}{\theta}\right)$. Therefore average units of efficient labor per worker is:

$$\bar{e} = \frac{E}{L} = \bar{\zeta} \pi_T^{-1/\theta} \tag{F.25}$$

The residential population in n is the sum of workers residing in n and total home production residing in n:

$$R_n = \sum_{i \in \mathcal{S}} \lambda_{ni} L + \lambda_n^O L^O = \left[\pi_T \sum_{i \in \mathcal{S}} \lambda_{ni} + (1 - \pi_T) \lambda_n^O \right] \mathcal{M} \tag{F.26}$$

Production and developers The production technology in the formal sector is similar to the baseline model and constant return to scale. Therefore, the wage rate per unit of efficient labor and commercial floor space price satisfies:

$$P_n = \frac{w_n^\beta Q_n^\gamma}{A_n}, \quad \forall\, n \in \mathcal{S} \tag{F.27}$$

The developers are same as in the baseline model. Height of buildings and land rent are given by:

$$h_i = \left(\frac{Q_i}{\kappa_i \nu} \right)^{\frac{1}{\nu-1}}, \quad r_i = \kappa_i (\nu - 1) \left(\frac{Q_i}{\kappa \nu} \right)^{\frac{\nu}{\nu-1}} \tag{F.28}$$

The floor spaces are used for both housing and production.

General equilibrium The total income in location n is the sum of the total labor income and total return from home production:

$$W_n = \sum_{i \in \mathcal{S}} w_i E_{ni} + w_O E_n^O \tag{F.29}$$

where E_{ni} is the total efficient units of labor along the commuting from n to i, and E_n^O is the efficient units of labor for hompe production in n. The former is given by:

$$E_{ni} = \int_{\Phi_z} L_{ni} z \, \mathrm{d}G(\boldsymbol{z}) = \bar{\zeta} \lambda_{ni} L \pi_T^{-1/\theta} = \bar{\zeta} \pi_T^{-1/\theta} \lambda_{ni|n} \lambda_n^R L \tag{F.30}$$

where we use the independence of two Fréchet distributions in work choice (first stage) and location choices (second stage). We also obtain:

$$E_n^O = \bar{\zeta} \pi_H^{-1/\theta} \lambda_n^O L^O \tag{F.31}$$

Therefore, total income in location n becomes:

$$W_n = \bar{\zeta}\left(\pi_T^{-1/\theta}\bar{w}_n R_n + \pi_H^{-1/\theta} w_O R_n^O\right) \tag{F.32}$$

where R_n is the measure of workers of the formal sector living in n, and R_n^O is that in home production. $\bar{w}_n$ is the average wage rate for workers living in n. The total efficient labor in the formal sector in location i is:

$$E_i = \bar{\zeta}\pi_T^{-1/\theta} L_i \tag{F.33}$$

where L_i is measure of workers working in location i. Lastly, the floor space market clearing condition becomes:

$$H_n = \frac{(1-\alpha)\bar{\zeta}(\pi_T^{-1/\theta}\bar{w}_n R_n + \pi_H^{-1/\theta} w_O R_n^O)}{Q_n} + \frac{\gamma}{\beta}\frac{\bar{\zeta} w_n \pi_T^{-1/\theta} L_n}{Q_n} \tag{F.34}$$

where the first terms is the aggregate demand for residential floor spaces and the second terms is the demand for commercial floor spaces in the formal production sector.

Inversion We consider how the process of model inversion is similar and different from the baseline model. Only workers in the formal sector commute. First, we only use the data for workers in the formal sector about their population (R_n) and employment (L_i). Their commuting market clearing condition is exactly the same as in the baseline; therefore, we back out exactly the same wage rate (w_i) for them to the baseline calibration. Using this, we can also compute the average wage per unit of efficient labor $\bar{w}_n$ that is also the same as in the baseline model.

Next, we let $\mathcal{M}$ be the total population in a city, including both workers in the formal sector and home production. We compute the share of workers in the formal sector $\pi_T = L/\mathcal{M}$ and home production $\pi_H = 1 - \pi_T$. Then, we consider how to pin down the return from home production, w_O. By definition, we have:

$$\frac{\bar{U}^O}{\bar{U}} = w_O \left[\sum_n \frac{\lambda_{nn}}{d_{nn}^{-\varepsilon} w_n^{\varepsilon}}\right]^{1/\varepsilon} \tag{F.35}$$

Therefore, using the share of employment in the formal sector and outside, we obtain:

$$w_O = \frac{\pi_O}{\pi_T}\left[\sum_n \frac{\lambda_{nn}}{d_{nn}^{-\varepsilon} w_n^{\varepsilon}}\right]^{-1/\varepsilon} \tag{F.36}$$

The right-hand side can be computed by using the share of workers in the formal sector (π_T), that in home production (π_O), commuting of workers (λ_{nn}), and estimated wage (w_n).

In turn, we compute the total income in each location, W_n, based on (π_T, π_H, $\bar{w}_n$, w_O, R_n, R_n^O). Then, we solve the floor space market clearing condition:

$$H_n = \frac{(1-\alpha)W_n}{Q_n} + \frac{\gamma}{\beta}\frac{\bar{w}_n\bar{\zeta}\pi_T^{-1/\theta} L_n}{Q_n} \tag{F.37}$$

for floor space price, Q_n. On the right-hand side, calibrated total income W_n determines the first term of housing demand, and average wage rate $\bar{w}_n$, share of workers in the formal sector π_T and total number of the workers in the formal sector L_n determine the total demand for floor spaces used in production. This second step introduces a change in the calibrated floor space prices compared to the baseline.

In the last step, we use the zero profit condition in the formal sector to back out the productivity (A_n). Finally, we use the residential distribution of workers in the formal sector to back out the residential amenity (B_n) with normalization. Since estimated floor space prices are different from the baseline, these estimates of productivity and amenities also show differences.

Welfare effects The welfare measure of people in a city is exante average utility. In this extension, we can define:

$$\begin{aligned} \mathbb{W} &= \bar{\zeta} \left\{ \left(\bar{U} \right)^{\theta} + \left(\bar{U}^{O} \right)^{\theta} \right\}^{1/\theta} \\ &= \bar{\zeta} \bar{\gamma} \pi_T^{-1/\theta} (\lambda_{nn})^{-1/\varepsilon} B_n w_n P_n^{-\alpha} Q_n^{-(1-\alpha)} \end{aligned} \tag{F.38}$$

Therefore, fixing value of amenities, change in welfare of people in a city is:

$$\widehat{\mathbb{W}} = (\widehat{\pi}_T)^{-1/\theta} \times (\widehat{\bar{\lambda}})^{-1/\varepsilon} \times \widehat{\bar{w}} \times (\widehat{\bar{P}})^{-\alpha} \times (\widehat{\bar{Q}})^{-(1-\alpha)} \tag{F.39}$$

The first term on the right-hand side is the additional term in this extension. Any policy changes can change the share of voluntary unemployment in the city ($\widehat{\pi}_H$). Note that we fix the value of hope production w_O in the counterfactuals.

Calibration The main difference to the baseline calibration is that we allow a difference between total employment and total population in the city. Our Census data finds such difference and a positive number of excess population. Aggregating 264 wards in Dhaka, the Population Census shows a total population of 3,294,420, while the total employment in Employment Census is 3,137,200. We suppose that their difference captures the population outside the formal sector. Following the same step in our calibration in Section 4, we first estimate the wage of the formal workers using commuting market clearing conditions. The challenge in the next step is estimating the return of home production, w_O. The present model, however, allows to compute it to match the observed commuting pattern of workers in the formal sector and estimated wages such that:

$$w_O = \frac{\pi_O}{\pi_T} \left(\sum_{n \in \mathcal{S}} \frac{\lambda_{nn}}{d_{nn}^{-\varepsilon} w_n^{\varepsilon}} \right)^{-1/\varepsilon} \tag{F.40}$$

On the right-hand side, π_O/π_T is the ratio of the total number of workers in the home production to those in the formal sector. In Dhaka, we see $\pi_O/\pi_T = 0.050$. For reference, the unemployment rate in Bangladesh in 2013 was 4.43 percent. Once we obtain wage rates in both formal sector and home production, we use land market clearing conditions to estimate land rent and floor space prices, and other procedures are the same as the baseline. We posit the parameter value for the substitution between formal sector and home production, $\theta = 8.0$ that is the same value as the commuting elasticity, ε.

The estimated value of housing supply elasticity is 1.492, which is close to the baseline value. The estimated return for home production is $w_O = 0.037$, which is significantly low compared to the wage rate in the formal sector.[20] The parameter for productivity spillovers is 0.045, which is the same value as the baseline estimate.

Quantitative implications Based on the estimation results, we now explore the quantitative implications of the extended model with home production. We focus on the counterfactual analysis of housing supply elasticity for the comparison to the baseline results in our main counterfctual experiment. Housing supply elasticity increases by 25 percent, from 1.49 to 1.86. Table F.1 report results.

As expected from the theory, an improvement in building supply elasticity leads to workers' concentration in productive places, and wage rates of workers in these areas increase. Therefore, the value of expected return from working in the formal sector becomes high, and the total number of workers in the formal sector would increase. The fourth row of F.1 confirms that total employment in the formal sector would grow by around 5 percent. At the same time, this composition effects increase the demand for land and prices for floor spaces, and it partially offsets the gains from an increase in housing supply elasticity. As seen in the third row of Table F.1 , we see a slightly small loss of landlords compared to the baseline.

Turning to welfare, the existence of home production adds the composition effect and its consequences on price changes. The welfare gains when we see Column (4) in Table F.1 , the impact on aggregate welfare of people is 4.93 percent, slightly below the baseline results for workers' welfare.

[20]For comparison, wage rates of the formal sector are normalized such that their geometric mean is one.

Table F.1 : Counterfactuals for Building Supply Elasticity: When Home production exists

	(1)	(2)	(3)	(4)
	Closed City		Open City	
Counterfactuals (relative to baseline)	Fix Productivity	With Spillovers	Fix Productivity	With Spillovers
Average Welfare of Workers ($\bar{U}$)	1.0574	1.0663	1.0480	1.0493
– Average Real Income of Workers in the Formal Sector	1.0566	1.0576	1.0459	1.0459
Total Land Revenue ($\sum_n r_n T_n$)	0.8731	0.8796	0.9721	0.9764
Total Number of Workers in the Formal Sector	1.0501	1.0501	1.0501	1.0501
Total Population ($\bar{L}$)			1.0983	1.1011

Note: These counterfactual exercises assume that building supply elasticity increases by 25 percent. For each counterfactual, the numbers are relative values to the baseline equilibrium. In counterfactuals (1) and (2), we assume the closed city. Counterfactual (1) supposes that overall productivity (A_i) is fixed at the baseline level. In counterfactual (2), we also allow spillovers in productivity so that overall productivity changes with employment density. In counterfactual (3) and (4), we consider the open city where workers can move into and out of the city. Counterfactual (3) fixes overall productivity as in (1), and counterfactual (4) allows productivity spillovers as in (2).

Bibliography

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2016). Import competition and the great us employment sag of the 2000s. *Journal of Labor Economics 34*(S1), S141–S198.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy 128*(6), 2188–2244.

Ahlfeldt, G. M. and J. Barr (2022). The economics of skyscrapers: A synthesis. *Journal of Urban Economics 129*, 103419.

Ahlfeldt, G. M. and D. P. McMillen (2018). Tall buildings and land values: Height and construction cost elasticities in chicago, 1870–2010. *Review of Economics and Statistics 100*(5), 861–875.

Ahlfeldt, G. M. and E. Pietrostefani (2019). The economic effects of density: A synthesis. *Journal of Urban Economics 111*, 93–107.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the Berlin Wall. *Econometrica 83*(6), 2127–2189.

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and C. Arkolakis (2018). Modern spatial economics: a primer. In *World Trade Evolution*, pp. 435–472. Routledge.

Allen, T., C. Arkolakis, and X. Li (2015). Optimal city structure. *Yale University, mimeograph.*

Allen, T. and D. Donaldson (2019). The geography of path dependence. *Unpublished manuscript.*

Allen, T., S. Fuchs, S. Ganapati, A. Graziano, R. Madera, and J. Montoriol-Garriga (2020). Is tourism good for locals? Evidence from Barcelona. *Dartmouth College, mimeograph.*

Anderson, J. E., C. A. Milot, and Y. V. Yotov (2014). How much does geography deflect services trade? canadian answers. *International Economic Review 55*(3), 791–818.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Artuç, E., S. Chaudhuri, and J. McLaren (2010). Trade shocks and labor adjustment: A structural empirical approach. *American economic review 100*(3), 1008–45.

Baldwin, R. E. (2001). Core-periphery model with forward-looking expectations. *Regional science and urban economics 31*(1), 21–49.

Barro, R. J., X. Sala-i Martin, O. J. Blanchard, and R. E. Hall (1991). Convergence across states and regions. *Brookings papers on economic activity*, 107–182.

Barro, R. T. and X. Sala-I-Martin (1992). Regional growth and migration: A japan-united states comparison. *Journal of the Japanese and International Economies 6*(4), 312–346.

Bartelme, D. G., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2021). The textbook case for industrial policy: Theory meets data. Technical report, National Bureau of Economic Research.

Baum-Snow, N. and L. Han (2021). The microgeography of housing supply. *Work in progress, University of Toronto.*

Bird, J., Y. Li, H. Z. Rahman, M. Rama, and A. J. Venables (2018). *Toward great Dhaka: A new urban development paradigm eastward.* The World Bank.

Bird, J. H. and A. J. Venables (2019). *Growing a Developing City: A Computable Spatial General Equilibrium Model Applied to Dhaka.* The World Bank.

Boar, C. and D. Lashkari (2021). Occupational choice and the intergenerational mobility of welfare.

Brezis, E. S. and P. R. Krugman (1997). Technology and the life cycle of cities. *Journal of Economic Growth 2*(4), 369–383.

Bryan, G. and M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Buera, F. J. and J. P. Kaboski (2012). The rise of the service economy. *American Economic Review 102*(6), 2540–69.

Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics 121*(2), 587–633.

Bütikofer, A., A. Dalla-Zuanna, and K. G. Salvanes (2019). Breaking the links: Natural resource booms and intergenerational mobility.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica 87*(3), 741–835.

Caliendo, L. and F. Parro (2021). Trade policy. Technical report, National Bureau of Economic Research.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018). The impact of regional and sectoral productivity changes on the us economy. *The Review of economic studies 85*(4), 2042–2096.

Caron, J., T. Fally, and J. R. Markusen (2014). International trade puzzles: A solution linking production and preferences. *The Quarterly Journal of Economics 129*(3), 1501–1552.

Caselli, F. and W. J. Coleman II (2001). The us structural transformation and regional convergence: A reinterpretation. *Journal of political Economy 109*(3), 584–616.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Colantone, I., G. I. Ottaviano, and K. Takeda (2021). Trade and intergenerational income mobility: Theory and evidence from the us.

Combes, P.-P., G. Duranton, and L. Gobillon (2021). The production function for housing: Evidence from france. *Journal of Political Economy 129*(10), 2766–2816.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica 80*(6), 2543–2594.

Comin, D., D. Lashkari, and M. Mestieri (2020). Structural change with long-run income and price effects. *Econometrica*.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives 27*(3), 79–102.

Corak, M. and P. Piraino (2011). The intergenerational transmission of employers. *Journal of Labor Economics 29*(1), 37–68.

David, H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–68.

Davis, M. A. and J. Heathcote (2005). Housing and the business cycle. *International Economic Review 46*(3), 751–784.

Desmet, K. and E. Rossi-Hansberg (2009). Spatial growth and industry age. *Journal of Economic Theory 144*(6), 2477–2502.

Desmet, K. and E. Rossi-Hansberg (2014). Spatial development. *American Economic Review 104*(4), 1211–43.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica 82*(3), 825–885.

Dix-Carneiro, R. and B. K. Kovak (2017). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–46.

Doepke, M. and F. Zilibotti (2008). Occupational choice and the spirit of capitalism. *The Quarterly Journal of Economics 123*(2), 747–793.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2012). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies 79*(2), 645–677.

Donaldson, D. and A. Storeygard (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives 30*(4), 171–98.

Duranton, G. and D. Puga (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review 91*(5), 1454–1477.

Duranton, G. and M. A. Turner (2012). Urban growth and transportation. *Review of Economic Studies 79*(4), 1407–1440.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Eckert, F. and M. Peters (2018). Spatial structural change. *Unpublished Manuscript*.

Esch, T., J. Zeidler, D. Palacios-Lopez, M. Marconcini, A. Roth, M. Mönks, B. Leutner, E. Brzoska, A. Metz-Marconcini, F. Bachofer, et al. (2020). Towards a large-scale 3d modeling of the built environment—joint analysis of tandem-x, sentinel-2 and open street map data. *Remote Sensing 12*(15), 2391.

Fajgelbaum, P. D., E. Morales, J. C. Suárez Serrato, and O. Zidar (2019). State taxes and spatial misallocation. *The Review of Economic Studies 86*(1), 333–376.

Fan, T., M. Peters, and F. Zilibotti (2021). Service-led or service-biased growth? equilibrium development accounting across indian districts. Technical report, National Bureau of Economic Research.

Feigenbaum, J. J. (2015). Intergenerational mobility during the great depression.

Fernandez, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics 1*(1), 146–77.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics 119*(4), 1249–1299.

Ferrie, J. P. (2005). History lessons: The end of american exceptionalism? mobility in the united states since 1850. *Journal of Economic Perspectives 19*(3), 199–215.

Fieler, A. C. (2011). Nonhomotheticity and bilateral trade: Evidence and a quantitative explanation. *Econometrica 79*(4), 1069–1101.

Fogli, A. and V. Guerrieri (2019). The end of the american dream? inequality and segregation in us cities. Technical report, National Bureau of Economic Research.

Gervais, A. and J. B. Jensen (2019). The tradability of services: Geographic concentration and trade costs. *Journal of International Economics 118*, 331–350.

Glaeser, E. L., J. D. Gottlieb, and O. Ziv (2016). Unhappy cities. *Journal of labor economics 34*(S2), S129–S182.

Glaeser, E. L., M. Resseger, and K. Tobio (2009). Inequality in cities. *Journal of Regional Science 49*(4), 617–646.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 501–522.

Handbury, J. (2019). Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities. Technical report, National Bureau of Economic Research.

Hanoch, G. (1975). Production and demand models with direct or indirect implicit additivity. *Econometrica: Journal of the Econometric Society*, 395–419.

Harding, J. P., S. S. Rosenthal, and C. Sirmans (2007). Depreciation of housing capital, maintenance, and house price inflation: Estimates from a repeat sales model. *Journal of urban Economics 61*(2), 193–217.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, Volume 4, pp. 131–195. Elsevier.

Heblich, S., S. J. Redding, and D. M. Sturm (2020). The making of the modern metropolis: evidence from London. *Quarterly Journal of Economics 135*(4).

Henderson, J. V., T. Regan, and A. J. Venables (2021). Building the city: from slums to a modern metropolis. *The Review of Economic Studies 88*(3), 1157–1192.

Henderson, V., A. Storeygard, and D. N. Weil (2011). A bright idea for measuring economic growth. *American Economic Review 101*(3), 194–99.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2014). Growth and structural transformation. In *Handbook of economic growth*, Volume 2, pp. 855–941. Elsevier.

Hornbeck, R. and E. Moretti (2020). Who benefits from productivity growth? direct and indirect effects of local tfp growth on wages, rents, and inequality.

Hummels, D. (2001). Toward a geography of trade costs.

Kleinman, B., E. Liu, and S. J. Redding (2021). Suficient statistics for dynamic spatial economics.

Kramarz, F. and O. N. Skans (2014). When strong ties are strong: Networks and youth labour market entry. *Review of Economic Studies 81*(3), 1164–1200.

Kreindler, G. E. and Y. Miyauchi (2021). Measuring commuting and economic activity inside cities with cell phone records. *The Review of Economics and Statistics*, 1–48.

Krugman, P. (1991). History versus expectations. *The Quarterly Journal of Economics 106*(2), 651–667.

Long, J. and J. Ferrie (2013). Intergenerational occupational mobility in great britain and the united states since 1850. *American Economic Review 103*(4), 1109–37.

Loury, L. D. (2006). Some contacts are more equal than others: Informal networks, job tenure, and wages. *Journal of Labor Economics 24*(2), 299–318.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of economic perspectives 14*(3), 115–136.

Marconcini, M., A. Metz-Marconcini, S. Üreyen, D. Palacios-Lopez, W. Hanke, F. Bachofer, J. Zeidler, T. Esch, N. Gorelick, A. Kakarla, et al. (2020). Outlining where humans live, the world settlement footprint 2015. *Scientific Data 7*(1), 1–14.

Matsuyama, K. (1991). Increasing returns, industrialization, and indeterminacy of equilibrium. *The Quarterly Journal of Economics 106*(2), 617–650.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of economic theory 58*(2), 317–334.

Matsuyama, K. (2009). Structural change in an interdependent world: A global view of manufacturing decline. *Journal of the European Economic Association 7*(2-3), 478–486.

Matsuyama, K. (2019). Engel's law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica 87*(2), 497–528.

McLaren, J. (2017). Globalization and labor market dynamics. *Annual Review of Economics 9*, 177–200.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica 71*(6), 1695–1725.

Michaels, G., F. Rauch, and S. J. Redding (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics 127*(2), 535–586.

Miyauchi, Y., K. Nakajima, and S. J. Redding (2021). Consumption access and agglomeration: evidence from smartphone data. Technical report, National Bureau of Economic Research.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–90.

Ngai, L. R. and C. A. Pissarides (2007). Structural change in a multisector model of growth. *American economic review 97*(1), 429–443.

Ottaviano, G. I. (1999). Integration, geography and the burden of history. *Regional Science and Urban Economics 29*(2), 245–256.

Owens III, R., E. Rossi-Hansberg, and P.-D. Sarte (2020). Rethinking detroit. *American Economic Journal: Economic Policy 12*(2), 258–305.

Pellegrina, H. S. and S. Sotelo (2021). Migration, specialization, and trade: Evidence from brazil's march to the west. Technical report, National Bureau of Economic Research.

Polanyi, M. (1958). Personal knowledge: Towards a post-critical philosophy. *University Of Chicago Press.*

Porcher, C. (2020). Migration with costly information.

Puga, D. and A. J. Venables (1996). The spread of industry: spatial agglomeration in economic development. *Journal of the Japanese and International Economies 10*(4), 440–464.

Ramondo, N., A. Rodríguez-Clare, and M. Saborío-Rodríguez (2016). Trade, domestic frictions, and scale effects. *American Economic Review 106*(10), 3159–84.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics 37*(3), 375–398.

Redding, S. and A. J. Venables (2004). Economic geography and international inequality. *Journal of international Economics 62*(1), 53–82.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*, 21–58.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics 125*(3), 1253–1296.

Simonovska, I. and M. E. Waugh (2014). Trade models, trade elasticities, and the gains from trade. Technical report, National Bureau of Economic Research.

Tsivanidis, N. (2020). Evaluating the impact of urban transit infrastructure: evidence from bogota's transmilenio. Technical report, University of California at Berkeley mimeo.

Valentinyi, A. and H. Berthold (2008). Measuring factor income shares at the sectoral level. *Review of Economic Dynamics 11*, 820–835.

Willis, R. J. and S. Rosen (1979). Education and self-selection. *Journal of political Economy 87*(5, Part 2), S7–S36.