

# Spatial Structural Change\*

Fabian Eckert<sup>†</sup>    Michael Peters<sup>‡</sup>

February 2018

## Abstract

This paper studies the spatial implications of structural change. The secular decline in spending on agricultural goods hurts workers in rural locations and increases the return to moving towards non-agricultural labor markets. We combine detailed spatial data for the U.S. between 1880 and 2000 with a novel quantitative theory to understand this process and to quantify its macroeconomic implications. We find that spatial reallocation across labor markets accounts for almost none of the aggregate decline in agricultural employment. Despite ample migration, population net flows were only weakly correlated with agricultural specialization. Spatial reallocation nevertheless had important aggregate effects. Without migration income per capita would have been 15% lower and spatial welfare inequality would have been substantially higher, especially among low-skilled, agricultural workers.

---

\*We thank seminar participants at Boston University, Stanford, Penn State, UNC, the SED, the EEA, the NBER Macroeconomic Across Time and Space Conference, the Edinburgh Workshop on Agricultural Productivity and the Yale International Lunch for very helpful comments. We are particularly thankful to Simon Alder, Timo Boppert, David Lagakos, Doug Gollin, Kjetil Storesletten and Fabrizio Zilibotti for their suggestions. We also thank Costas Arkolakis, Pete Klenow, and Aleh Tsyvinski for their comments. Zara Contractor, Andrés Gvartz and Adam Harris provided outstanding research assistance.

<sup>†</sup>Yale University. [fabian.eckert@yale.edu](mailto:fabian.eckert@yale.edu)

<sup>‡</sup>Yale University and NBER. [m.peters@yale.edu](mailto:m.peters@yale.edu)

# 1 Introduction

Structural change is a key feature of long-run economic growth. As countries grow richer, aggregate spending shifts towards non-agricultural goods and the share of employment in the agricultural sector declines. This sectoral bias of the growth process also implies that economic growth is unbalanced across space. In particular, by shifting expenditure away from the agricultural sector, the structural transformation is biased against regions, that have a comparative advantage in the production of agricultural goods. To what extent this spatial bias affects welfare and allocative efficiency depends crucially on the ease with which resources can be reallocated, in particular on the costs workers face to move towards non-agricultural labor markets. In this paper we use detailed data on the spatial development of the US between 1880 and 2000 and a novel quantitative theory of spatial structural change to analyze how this spatial unbalancedness of the growth process affected the US economy.<sup>1</sup>

We start by documenting a striking - and to the best of our knowledge - new empirical fact: the spatial reallocation of people from agricultural to non-agricultural labor markets accounts for essentially none of the aggregate decline in agricultural employment since 1880. Rather, the entire structural transformation is due to a decline in agricultural employment, which occurs within labor markets.<sup>2</sup> While this is seemingly inconsistent with the secular trend in urbanization, whereby the share of urban dwellers among US workers increased from 25% to 75% between 1880 and 2000, we explicitly show that this is not the case. In fact, like the change in agricultural employment, the process of urbanization was also a predominantly local phenomenon taking place within labor markets.

The most obvious explanation for this pattern is that frictions to spatial mobility were prohibitively large for the majority of workers throughout the 20th century. This, however, is not the case as US Census data shows that throughout the last century, about 30% of every cohort of workers lived and worked in states different from their state of birth. Moving frictions per se, therefore, cannot explain the quantitative insignificance of the spatial reallocation channel. The reason why migration across labor markets cannot account for much of the decline in agricultural employment is rather that the correlation between agricultural employment shares and net migration outflows was essentially zero.

To understand these patterns and to analyze the implications for the spatial distribution of welfare and aggregate productivity, we propose a new quantitative theory of spatial structural change. Our theory combines an otherwise standard, neoclassical model of the structural transformation with an economic geography model with frictional labor mobility. At the spatial level, regions are differentially exposed to the secular decline in the demand for agricultural goods as they differ in their sectoral productivities, the skill composition of their local labor force and the ease with which other, less agricultural labor markets are accessible through migration.

---

<sup>1</sup>By “Spatial Structural Change” we refer to the simultaneous changes in sectoral employment and the spatial organization of economic activity. The father of the study of structural change, Simon Kuznets, was maybe the first to highlight the importance of studying spatial and sectoral reallocation in one unified framework (Lindbeck, ed (1992)).

<sup>2</sup>These patterns hold true regardless of whether we define labor markets at the state, commuting zone or county level. There are roughly 700 commuting zones and 3000 counties. In our quantitative analysis we focus on commuting zones.

To explain why migration flows were only weakly correlated with the initial level of agricultural specialization, our model highlights two forces. First, while the structural transformation indeed put downward pressure on wages in rural labor markets, such shifts were small relative to the equilibrium level of spatial wage differences. And because such wage differences are only imperfectly correlated with the regional agricultural employment share, individuals, in their search for higher earnings, often relocate *towards* agricultural areas. Additionally, we also find that the migration elasticity, i.e. the sensitivity of migration flows to regional wages, is limited. In particular, we show that spatial population *gross* flows are much larger than population *net* flows. This suggests that idiosyncratic, non-monetary preference shocks, which by definition are uncorrelated with local industrial structure, are an important determinant of migration decisions.

We then use the model to study the implications of spatial structural change for aggregate economic performance and the spatial distribution of welfare. We first focus on the role of spatial reallocation for aggregate productivity. Because we estimate that rural, agricultural-intensive regions are - on average - less productive and generate less value added per worker than non-agricultural areas, the lack of spatial reallocation suggests that the US economy potentially missed out on substantial productivity improvements. Quantitatively, we find that such productivity losses were modest. If spatial mobility was costless, aggregate income would only have been 4% higher in the year 2000. In contrast, if moving costs had been prohibitively high, income per capita would have been 15% lower. The observed process of spatial arbitrage in the US therefore seemed to have captured a large share of potential efficiency gains.

Next, we turn to the evolution of spatial welfare inequality. We find that welfare inequality across US commuting zones declined substantially between 1910 and 2000. In 2000 the interquartile range of the distribution of spatial welfare corresponded to a doubling of lifetime income for the average region in the US. In contrast, in 1910 one would have had to increase regional income by 160%. Importantly, spatial mobility was an important driving force behind this reduction in spatial welfare inequality. If spatial mobility had been prohibitively costly, the spatial dispersion of welfare had declined much less. In particular, unskilled workers, who have a comparative advantage in the agricultural sector and are hence particularly exposed to the urban bias of the structural transformation, would have seen no decline in spatial inequality over the 20th century. This highlights the important role of migration to mitigate the distributional consequences of structural shifts in the US economy.

Finally, our theoretical framework might also prove useful for applications beyond the one at hand. Our model combines basic ingredients from an economic geography model (spatial heterogeneity, intra-regional trade, costly labor mobility) with the usual features of neoclassical models of structural change (non-homothetic preferences, unbalanced technological progress, aggregate capital accumulation). Despite this richness, the theory remains highly tractable. Building on recent work by [Boppart \(2014\)](#), we first show that by combining a price independent generalized linear (PIGL) demand system with the commonly-used Frechet distribution of individual skills, one can derive closed-form solutions for most aggregate quantities of interest. We then show how this structure can be embedded in an otherwise

standard overlapping-generation model. Doing so allows us to tractably accommodate both individual savings (and hence aggregate capital accumulation) and costly spatial mobility. In particular, while individuals are forward looking in terms of their savings behavior, we show that the spatial choice problem reduces to a static one as long as goods are freely traded. As a result, we do not have to keep track of individuals' expectations about the entire distribution of future wages across locations - the aggregate interest rate is sufficient. Moreover, because our model essentially nests a version of a canonical aggregate model of structural change as the number of regions collapses to one, we view our framework as a natural spatial extension of neoclassical theories of the structural transformation.

**Related Literature** We combine insights from the macroeconomic literature on the structural transformation with recent advances in spatial economics. The literature on the process of structural change has almost exclusively focused on the time series properties of sectoral employment and value added shares - see [Herrendorf et al. \(2014\)](#) for a survey of this large literature.<sup>3</sup> In contrast, the recent generation of quantitative spatial models in the spirit of [Allen and Arkolakis \(2014\)](#) are mostly static in nature and focus on the spatial allocation of workers across heterogeneous locations.<sup>4</sup> We show that these two aspects interact in a natural way. The structural transformation induces changes in demand, which are non-neutral across space and hence affect the spatial equilibrium of the system. Conversely, the spatial topography, in particular the extent to which individuals are spatially mobile, has macroeconomic implications by determining equilibrium factor prices, capital accumulation and aggregate productivity.

Relatively few existing papers explicitly introduce a spatial dimension into an analysis of the structural transformation. An early contribution is [Caselli and Coleman II \(2001\)](#), who argue that spatial mobility was an important by-product of the process of structural change in the US. [Michaels et al. \(2012\)](#) also study the relationship between agricultural specialization and population growth across US counties. Their analysis, however, is more empirically oriented and does not use a calibrated structural model. More recently, [Desmet and Rossi-Hansberg \(2014\)](#) propose a spatial theory of the US transition from manufacturing to services and [Nagy \(2017\)](#) examines the process of city formation in the United States before 1860.

In allowing for a spatial microstructure, this paper also offers a new interpretation of the so-called “agricultural productivity gap”, i.e. the observation that value added per worker is persistently low in the

---

<sup>3</sup>Authors such as [Kuznets \(1957\)](#) and [Chenery \(1960\)](#) have been early observers of the striking downward trend in the aggregate agricultural employment share and the simultaneous increase in manufacturing employment in the United States. To explain these patterns, two mechanisms have been proposed. Demand side explanations stress the role of non-homotheticities, whereby goods differ in their income elasticity (see e.g. [Kongsamut et al. \(2001\)](#), [Gollin et al. \(2002\)](#), [Comin et al. \(2017\)](#) and [Boppart \(2014\)](#)). Supply-side explanations argue for the importance of unbalanced technological progress across sectors and capital-deepening (see e.g. [Baumol \(1967\)](#), [Ngai and Pissarides \(2007\)](#), [Acemoglu and Guerrieri \(2008\)](#), and [Alvarez-Cuadrado et al. \(2017\)](#)).

<sup>4</sup>This literature has addressed questions of spatial misallocation ([Hsieh and Moretti \(2015\)](#), [Fajgelbaum et al. \(2015\)](#)), the regional effects of trade opening ([Fajgelbaum and Redding \(2014\)](#), [Tombe et al. \(2015\)](#)), the importance of market access ([Redding and Sturm \(2008\)](#)) and the productivity effects of agglomeration economies ([Ahlfeldt et al. \(2015\)](#)). See [Redding and Rossi-Hansberg \(2017\)](#) for a recent survey of this growing literature.

agricultural sector (see e.g. [Gollin et al. \(2014\)](#), [Buera and Kaboski \(2009\)](#) or [Herrendorf and Schoellman \(2015\)](#)). Our model endogenously generates an “agricultural productivity gap”, without resorting to labor market frictions across sectors, as spatial mobility costs keep wages in agricultural areas low. The existence of such spatial gaps and their implications for aggregate productivity and welfare has been subject to a recent active literature (see e.g. [Young \(2013\)](#), [Bryan et al. \(2014\)](#), [Hsieh and Moretti \(2015\)](#) or [Lagakos et al. \(2017\)](#)). [Bryan and Morten \(2017\)](#) and [Hsieh and Moretti \(2015\)](#) use spatial models related to ours to study the aggregate effects of spatial misallocation. In contrast to us, their models are static and they do not focus on the structural transformation.

On the theoretical side, we build on [Boppart \(2014\)](#) and assume a price independent generalized linear (PIGL) demand structure. This demand structure has more potent income effects than the widely-used Stone-Geary specification, a feature which is required to generate declines in agricultural employment of the magnitude observed in the data. At the same time, we show how it can be introduced in a general equilibrium trade model in a tractable way.<sup>5</sup>

The remainder of the paper is structured as follows. In Section 2, we document the empirical fact that spatial reallocation accounts for essentially none of the aggregate decline in agricultural employment over the last 120 years. Section 3 presents our model. In Section 4, we calibrate the model to time-series and spatial data from the US. In Sections 5 and 5.2, we explain why the spatial reallocation component of the structural transformation is small and we quantify the implications for aggregate productivity and spatial inequality. Section 7 concludes. An Appendix contains the majority of our theoretical proofs and further details on our empirical results.

## 2 Spatial Reallocation and Structural Change

The long-run decline in the agricultural employment share in the US has been dramatic: since 1880 it fell from 50% to essentially nil. Naturally, this secular reallocation of resources across sectors has spatial consequences as it is biased against regions, which specialize in the production of agricultural goods. From an accounting perspective, there are two margins through which the economy can accommodate this spatial bias of the structural transformation. Either the process of structural change can induce *spatial reallocation*, whereby labor reallocates from agricultural to non-agricultural labor markets. Or it can lead to *regional transformation*, whereby agricultural employment shares decline within labor markets. Formally, the aggregate decline in the agricultural employment share since 1880 can be decomposed as

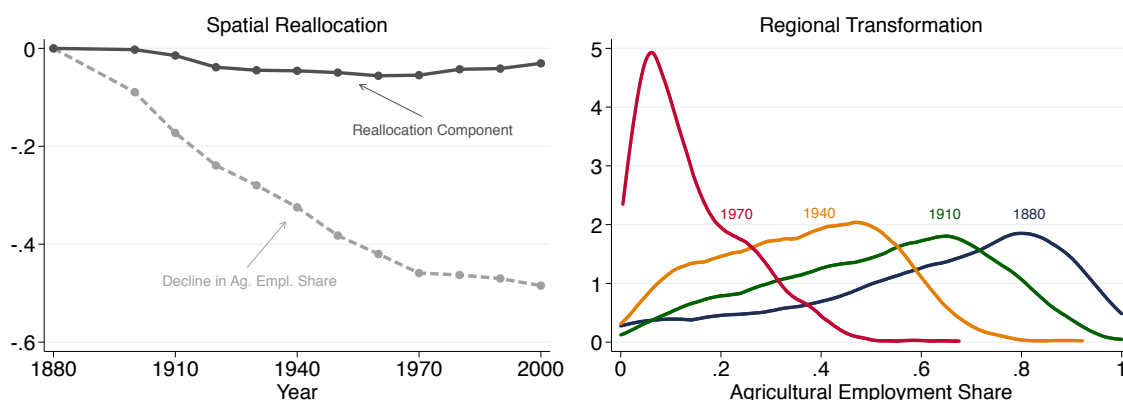
---

<sup>5</sup>Between 1880 and 2000 the aggregate agricultural employment share declines from around 50% to 2%. A model with Stone-Geary preferences can match the post-war data (see e.g. [Herrendorf et al. \(2013\)](#)), but has difficulties at the longer time horizon as income effects vanish asymptotically. [Alder et al. \(2018\)](#) show that the PIGL demand system provides a good fit to the data since 1900. The non-homothetic CES demand system, recently employed by [Comin et al. \(2017\)](#), has similar favorable time-series properties. However, it has less tractable aggregation properties making it harder to embed it in a spatial general equilibrium model.

$$s_{At} - s_{A1880} = \sum_r s_{rAt} l_{rt} - \sum_r s_{rA1880} l_{r1880} = \underbrace{\sum_r s_{rA1880} (l_{rt} - l_{r1880})}_{\text{Spatial Reallocation}} + \underbrace{\sum_r (s_{rAt} - s_{rA1880}) l_{rt}}_{\text{Regional Transformation}}, \quad (1)$$

where  $s_{At}$  is the aggregate agricultural employment share at time  $t$ ,  $l_{rt}$  denotes the share of employment in region  $r$  at time  $t$  and  $s_{rAt}$  is the regional employment share in agriculture. As highlighted by (1), the spatial reallocation margin is important for the decline in agricultural employment, if net population growth,  $l_{rt} - l_{r1880}$ , and the initial agricultural employment share,  $s_{rA1880}$ , are negatively correlated.

For the case of the U.S., the relative importance of the reallocation and transformation margins is striking: in an accounting sense the spatial reallocation of labor accounts for essentially none of the structural transformation observed in the aggregate. To see this, consider Figure 1, where we implement (1) by empirically equating labor markets with US commuting zones.<sup>6</sup> Out of the total decline of about 48%, only 3% is due to the reallocation of workers across commuting zone boundaries.



Notes: In the left panel, the light grey line shows the absolute decline in the aggregate agricultural employment share since 1880,  $s_{At} - s_{A1880}$ , and the across labor market reallocation component highlighted in Equation 1, i.e.  $\sum_r s_{rA1880} (l_{rt} - l_{r1880})$ , where  $s_{rAt}$  and  $l_{rt}$  are the agricultural employment share and the population share of region  $r$  at time  $t$  (dark grey line). The gap between the two lines is accounted for by sectoral reallocation taking place within labor markets. In the right panel we show the cross-sectional distribution of agricultural employment shares across US commuting zones for different years. We omit the 2000 cross-section from the left panel for expositional purposes only: the agricultural employment share does not change much between 1970-2000. For a detailed description of the construction of the regional data we refer to Section 4.

Figure 1: Spatial Structural Change: Spatial Reallocation vs. Regional Transformation

It follows that most of structural change takes place *within* labor markets through a transformation of the local structure of employment. This is seen in the right panel of Figure 1, where we display the distribution of agricultural employment shares across US commuting zones for different years. There is substantial cross-sectional dispersion in regional specialization. While the majority of commuting zones had agricultural employment shares exceeding 75% in 1880, many labor markets were already much less agriculturally specialized and had agricultural employment shares below 25% at the end of the 19th century. Throughout the 20th Century, there is a marked leftwards shift, whereby *all* commuting zones

<sup>6</sup>In the paper, we use the commuting zone definition by Tolbert and Sizer (1996) as our baseline definition of a labor market. There are 741 such commuting zones in the US. We describe our data in more detail in Section 4 below. In Section A in the Appendix, we replicate Figure 1 at the county and state level and it looks effectively identical.

see a decline in agricultural employment. Hence, the structural transformation did not induce regional specialization, but rather features a fractal property whereby all local economies undergo changes in their sectoral structure akin to the aggregate economy.

These patterns might seem surprising as they are seemingly at odds with the process of urbanization, whereby between 1880 and 2000 the fraction of the US population living in cities tripled from about 25% to 75%. This, however, is not the case. While inconsistent with a view of urbanization as a movement of workers towards urbanized commuting zones, Figure 1 is consistent with a world in which workers move towards local urban centers within commuting zones. In Section A of the Appendix, we replicate Figure 1 for the increase in the aggregate urbanization rate, and show that the increase in urbanization in the US was indeed a within labor market phenomenon, ie. was very local in nature. In particular, like for the change in agricultural employment, the reallocation of individuals from rural to urbanized commuting zones explains almost none of the sharp rise in urbanization. Furthermore, this local incidence of the structural transformation is not unique to the US but present in many countries around the world - again we refer the reader to Section A of the Appendix for details.

The patterns in Figure 1 raise two questions. First, why did the spatial bias of the structural transformation not cause a more pronounced migration response towards non-agricultural labor markets? Secondly, does this “missing spatial reallocation” have important consequences for aggregate productivity and spatial welfare differences across labor markets? To answer these questions we need a theory of spatial structural change, which we present next.

### 3 A Quantitative Theory of Spatial Structural Change

In this section we present a novel theory of spatial structural change. Our model rests on two pillars. In particular, we combine an essentially neoclassical model of the structural transformation featuring both non-homothetic preferences and unbalanced sectoral technological progress with a quantitative economic geography model. The latter introduces the spatial dimension by allowing for variations in local industrial structure, intra-regional trade and costly spatial mobility.

#### 3.1 Environment

We consider an economy consisting of  $R$  regions indexed by  $r$ . Each region produces two goods, an agricultural good and a non-agricultural good, indexed by  $s = A, NA$ . We identify a region with a local labor market, i.e. to supply labor in region  $r$ , individuals have to reside there. While population mobility across labor markets is subject to migration costs, the allocation of labor across sectors within a labor market is frictionless. Since the model is dynamic we additionally index most objects by  $t$ . For expositional simplicity, we sometimes omit time subscripts when there is no risk of confusion.

**Technology** Each of the two goods is a CES composite of differentiated regional varieties with a constant elasticity of substitution  $\sigma$ , i.e.

$$Y_s = \left( \sum_{r=1}^R Y_{rs}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \quad (2)$$

where  $Y_{rs}$  is the amount of goods in sector  $s$  produced in region  $r$  and  $Y_s$  is aggregate output in sector  $s$ . As is standard in macroeconomic models of the structural transformation (see e.g. [Herrendorf et al. \(2014\)](#)) we assume that regional production functions are fully neoclassical and given by

$$Y_{rst} = Z_{rst} K_{rst}^\alpha H_{rst}^{1-\alpha},$$

where  $K_{rst}$  and  $H_{rst}$  denote capital and labor (in efficiency units) in region  $r$ , sector  $s$  and time  $t$  respectively, and  $Z_{rst}$  denotes productivity.<sup>7</sup> It is conceptually useful to decompose regional productivity  $Z_{rst}$  as follows

$$Z_{rst} \equiv Z_{st} Q_{rst} \quad \text{with} \quad \sum_r Q_{rst}^{\sigma-1} = 1. \quad (3)$$

Here  $Z_{st}$  is an *aggregate* TFP shifter in sector  $s$ , which affects all regions proportionally. Additionally, there are *region-specific* sources of sectoral productivity denoted by the vector  $\{Q_{rs}\}_{rs}$ . The common component of  $Q_{rs}$  across sectors within region  $r$  captures differences in *absolute* advantage. Regional differences in  $Q_{rs}/Q_{rs'}$  capture differences in *comparative* advantage. Given the normalization embedded in (3), the vector  $\{Q_{rs}\}_{rs}$  can be thought of as parameterizing the heterogeneity in productivity across space.

The aggregate capital stock accumulates according to the usual law of motion

$$K_{t+1} = (1 - \delta) K_t + I_t,$$

where  $I_t$  denotes the amount of investment at time  $t$  and  $\delta$  is the depreciation rate. For simplicity, we assume that the investment good is a Cobb-Douglas composite of the agricultural and non-agricultural good given in (2). Letting  $\phi$  be the share of the agricultural good in the production of investment goods, the price of the investment good is given by  $P_{It} = P_{At}^\phi P_{NA t}^{1-\phi}$ .<sup>8</sup> For the remainder of the paper the investment good will serve as the numeraire of our economy.

Note that our environment imposes some restrictions. First of all, we assume that goods are freely traded so that prices are equalized across locations. As we explain in detail below, this assumption considerably

<sup>7</sup>For simplicity and building on the work of [Herrendorf et al. \(2015\)](#), we assume that capital shares are identical across sectors. These authors find that sectoral differences in the capital shares and elasticities of substitution are of second order importance and conclude that ‘‘Cobb–Douglas sectoral production functions that differ only in technical progress capture the main forces behind postwar US structural transformation that arise on the technology side’’ ([Herrendorf et al., 2015](#), p. 106).

<sup>8</sup>The accompanying production function for the investment good is given by  $I_t = \phi^\phi (1 - \phi)^{1-\phi} X_A^\phi X_{NA}^{1-\phi}$ , where  $X_s$  is the amount of sector  $s$  goods used in the investment good sector. We abstract from changes in sectoral spending within the investment good sector (see ([Herrendorf et al., 2017](#))).



simplifies workers' spatial choice problem. However, in Section B.1 of the Appendix, we show that differences in productivity  $Q_{rst}$  are isomorphic to some forms of trade costs. Secondly, note that we do not explicitly include land as a factor of production in the agricultural sector. This assumption is in fact not very restrictive and we do mainly to keep the production side comparable to macroeconomic models of the structural transformation. In Section B.1 in the Appendix we formally derive the equilibrium of our model with an explicit role for land and we show that these economic forces can be captured by assuming the production in the agricultural sector is subject to decreasing returns to scale. We cover this case specifically as on our robustness exercises. Finally, we assume that capital is traded on a frictionless spot market. We abstract from capital adjustment costs to highlight spatial frictions in the reallocation of workers.

**Demographics and Labor Supply** We phrase our analysis as an overlapping generations (OLG) economy. Individuals live for two periods. When young, individuals decide on their preferred regional labor market (subject to migration costs), work to earn labor income and save to smooth consumption over their life-cycle. When old, individuals solely consume the receipts of their saving decision and have a single off-spring, who has the option migrate to a new region. The OLG structure is analytically extremely convenient. Crucially, it generates a motive for savings (and hence capital accumulation), while still being sufficiently tractable to allow for spatial mobility subject to migration costs. In addition, it is also empirically attractive in that it captures the importance of cohort effects in accounting for the structural transformation (see e.g. [Hobijn et al. \(2018\)](#) and [Porzio and Santangelo \(2017\)](#)). Because individuals' labor market opportunities depends on their location choice, the migratory response of young workers is an important determinant of the sectoral labor supply function for the aggregate economy.

In order to focus on this novel spatial dimension of our model, we first follow the macroeconomic literature on structural change and assume that individuals' labor supply is perfectly substitutable across sectors within labor markets, i.e. there is effectively only a single dimension of skill. This implies that the sectoral allocation of labor within labor markets is only determined from the labor demand side. In our quantitative analysis, however, we will allow for an upward sloping labor supply function across sectors within labor markets: while migration costs affect the costs of spatial reallocation, the sectoral labor supply elasticity determines the costs of transforming the local labor market through sectoral reallocation.

Individuals are heterogenous in the number of efficiency units they can provide to the market,  $z^i$ , and we assume that these are drawn from a Frechet distribution, i.e.  $F(z) = e^{-z^{-\zeta}}$ .<sup>9</sup> Here, the parameter  $\zeta$  governs the dispersion of skills across individuals and the average level of efficiency units is given by  $E[z] = \Gamma_\zeta$ , where  $\Gamma_\zeta \equiv \Gamma(1 - \frac{1}{\zeta})$  and  $\Gamma(\cdot)$  denotes the Gamma function. The distribution of earnings of individual  $i$  in region  $r$  at time  $t$ ,  $y_{rt}^i = z^i w_{rt}$ , is therefore also Frechet distributed with mean  $\Gamma_\zeta w_{rt}$ .

---

<sup>9</sup>Strictly speaking, the individual heterogeneity in skills is not essential at this point. We introduce it here in anticipation of our two-sector extension in Section 3.3. To generate an upward sloping supply function across sectors, we assume that individuals face an occupational choice problem and draw a vector of *sector*-specific of efficiency units  $(z_{NA}^i, z_A^i)$ . As we will show below, all our expressions seamlessly generalize to this two-sector case. Hence, it is attractive for exposition purposes to introduce skill heterogeneity already at this point.

Hence, in our framework individuals make three economic choices: (i) a spatial decision on where to live and work in the beginning of life, (ii) an inter-temporal choice on how much to consume and save when young and (iii) an intra-temporal choice of how to allocate their spending optimally across the two consumption goods in both periods of life.

**The intra-temporal problem: Non-homothetic preferences and the allocation of spending** To generate the structural transformation at the aggregate level, we follow the existing macroeconomic literature by including two channels: Non-homothetic preferences imply a form of Engel's Law whereby consumers reduce their relative agricultural spending as they grow richer, while sectoral differences in technological progress induce a relative price effect that begets reallocation in consumer spending.<sup>10</sup>

We follow [Boppart \(2014\)](#) and assume that individual preferences can be represented by the indirect utility function<sup>11</sup>

$$V(e, P) = \frac{1}{\eta} \left( \frac{e}{p_A^\phi p_{NA}^{1-\phi}} \right)^\eta - \frac{v}{\gamma} \left( \frac{p_A}{p_{NA}} \right)^\gamma + \frac{v}{\gamma} - \frac{1}{\eta}. \quad (4)$$

This is a slight generalization of the PIGL demand system employed by [Boppart \(2014\)](#).<sup>12</sup> In particular, Roy's Identity implies that the expenditure share on the agricultural good,  $\vartheta_A(e, p)$ , is given by

$$\vartheta_A(e, p) \equiv \frac{x_A(e, p) p_A}{e} = \phi + v \left( \frac{p_A}{p_{NA}} \right)^\gamma e^{-\eta}, \quad (5)$$

and hence incorporates both income effects (governed by  $\eta$ ) and price effects (governed by  $\gamma$ ).<sup>13</sup> For  $\eta > 0$ , the expenditure share on agricultural goods is declining in total expenditure. This captures the income effect of non-homothetic demand, whereby higher spending reduces the relative expenditure share on agricultural goods. Holding real income  $e$  constant, the expenditure share is increasing in the relative price of agriculture if  $\gamma > 0$ . We can also see that this demand system nests important special cases. The case of  $\eta = 0$  corresponds to a homothetic demand system, where expenditure shares only depend on relative prices. The case of  $\eta = \gamma = 0$  is the Cobb Douglas case where expenditure shares are

<sup>10</sup>Both of these mechanisms have been shown to be quantitatively important. See for example [Herrendorf et al. \(2014\)](#), [Alvarez-Cuadrado and Poschke \(2011\)](#), [Boppart \(2014\)](#) or [Comin et al. \(2017\)](#).

<sup>11</sup>The most common choice among non-homothetic preferences is the Stone-Geary specification, which relies on subsistence requirements to generate a less than unitary income elasticity for consumers' spending on agricultural goods (see e.g. [Kongsamut et al. \(2001\)](#)). [Herrendorf et al. \(2013\)](#), for example, show that these preferences do a good job at quantitatively accounting for the decline in agricultural spending in the US in the post-war period. However, [Alder et al. \(2018\)](#) show that the Stone-Geary specification is unable to generate the large decline in agricultural employment since 1880, as the non-homotheticity vanishes asymptotically. In contrast, they find that *Price-Independent Generalized Linear* ("PIGL") Preferences, introduced in the literature on structural change in [Boppart \(2014\)](#), provide a much better fit to the data.

<sup>12</sup>For  $V(e, p)$  to be well-defined, we have to impose additional parametric conditions. In particular, we require that  $\eta < 1$ , that  $\gamma \geq \eta$ . These conditions are satisfied in our empirical application. See Section B.7 of the Appendix for a detailed discussion. [Boppart \(2014\)](#) uses this demand system to study the the evolution of service sector. In terms of (4) he assumes that  $p_A$  is the price of goods and  $p_{NA}$  is the price of services and considers the case of  $\phi = 0$ . In his Appendix, however, [Boppart \(2014\)](#) also discusses the case of (4).

<sup>13</sup>Roy's Identity implies that  $\vartheta_A(e, p) = x_A^M(p, e) \frac{p_A}{e} = -\frac{\partial V(p, e(p, u)) / \partial p_A}{\partial V(p, e(p, u)) / \partial e} \frac{p_A}{e}$ , where  $x_A^M(p, e)$  denotes the Marshallian demand function. For  $V(p, e)$  given in (4), this expression reduces to (5). See Section C.1 in the Online Appendix for details.

constant and equal to  $\phi$ .<sup>14</sup>

The preference specification in (4) also has advantageous aggregation properties in our context. In particular, we show below that these preferences (combined with the Frechet distribution of individuals skills) allow us to derive closed-form expressions for the economy's aggregate demand system, despite the fact that they fall outside the Gorman class.<sup>15</sup>

**The inter-temporal problem: The consumption-saving decision** Given the indirect utility function above, life-time utility of individual  $i$  after having moved to region  $r$ ,  $U_{rt}^i$ , is given by

$$U_{rt}^i = \max_{[e_t, e_{t+1}, s]} \{V(e_t, P_t) + \beta V(e_{t+1}, P_{t+1})\}, \quad (6)$$

subject to

$$\begin{aligned} e_t + s_t &= y_{rt}^i \\ e_{t+1} &= (1 + r_{t+1})s_t. \end{aligned}$$

Here,  $y_{rt}^i = z^i w_{rt}$  is individual  $i$ 's real income in region  $r$ ,  $s_t$  denotes the amount of savings and  $r_t$  is the real interest rate.

**Spatial Mobility** A crucial aspect of our theory are agents' endogenous location choices. Because our focus is on a model of long-run growth, we model individuals' migration decisions as referring to life-time migration. Hence, agents have the option to move once, in the beginning of their lives, before they learn the actual realization of their labor efficiency  $z^i$ .<sup>16</sup> We follow the literature on discrete choice models and assume the value of a bilateral move from  $j$  to  $r$  to agent  $i$  can be summarized by

$$\mathcal{U}_{jr}^i = E[U_{rt}^i] - MC_{jr} + A_r + \kappa v_r^i,$$

where  $E[U_{rt}^i]$  is the *expected* utility of living in region  $r$ ,  $MC_{jr}$  denotes the cost of moving from  $j$  to  $r$ ,  $A_r$  is a location amenity, which summarizes the attractiveness of region  $r$  and is common to all individuals and  $v_r^i$  is an idiosyncratic error term, which is independent across locations and individuals. Furthermore,  $\kappa$  parametrizes the importance of the idiosyncratic shock, i.e. the extent to which individuals sort based on their idiosyncratic tastes relative to the systematic attractiveness of region  $r$ . The higher  $\kappa$ , the less

<sup>14</sup>Note that  $\phi$  is also the agricultural share in the investment good sector. Hence, this case is akin to the neoclassical growth model, where consumption and investment goods are identical.

<sup>15</sup>This is in contrast to the non-homothetic CES demand system, which has recently been analyzed in [Comin et al. \(2017\)](#) and for which no closed form aggregation results exist.

<sup>16</sup>This structure has two convenient analytic properties. First, allowing mobility to depend on the realization of the efficiency bundle  $z^i$  would be less tractable as we would need to keep track of a continuum of ex-ante heterogenous individuals. Secondly, this structure retains the convenient aggregation properties of the Frechet distribution. If workers' spatial choice was conditional on  $z^i$ , the distribution of skills within a location would no longer be of the Frechet form.

responsive are individuals to the fundamental value of a location  $r$ .

As in the standard conditional logit model, we assume that  $v_r^i$  is drawn from a Gumbel distribution. This implies that the share of people moving from  $j$  to  $r$  is given by

$$\rho_{jrt} = \frac{\exp\left(\frac{1}{\kappa} (E[U_{rt}^i] + A_{rt} - MC_{jr})\right)}{\sum_{l=1}^R \exp\left(\frac{1}{\kappa} (E[U_{lt}^i] + A_{lt} - MC_{jl})\right)}. \quad (7)$$

Hence, individual flock towards regions which offer large earnings (and hence promise large future lifetime utility  $E[U_{rt}^i]$ ) and high amenities  $A_{rt}$ . The extent to which population flows are directed towards such regions is moderated by moving costs  $MC_{jr}$  and the importance of idiosyncratic shocks, parametrized by  $\kappa$ .

Importantly, expression (7) highlights the determinants of the spatial reallocation component of the structural transformation: individuals leave agricultural areas if the correlation of the initial agricultural employment share  $s_{rAt-1}$  and *future* utility and amenities is negative and the elasticity of moving flows with respect to such fundamental differences is large, i.e.  $\kappa$  is small.

From Equation (7) we also obtain the law of motion for the spatial reallocation of workers between  $t$  and  $t-1$  as

$$L_{rt} = \sum_{j=1}^R \rho_{jrt} L_{jt-1}, \quad (8)$$

i.e. the number of people in region  $r$  at time  $t$  is given by the total inflows from all other regions (including itself). Since in the model workers only move once, we will discipline the model with data on *lifetime* migration, i.e. the fraction of people who live (and work) in a different location from where they were born (see [Molloy et al. \(2011\)](#)).

## 3.2 Competitive Equilibrium

Given the environment above, we can now characterize the equilibrium of the economy. We proceed in three steps. First we characterize the household problem, i.e. the optimal consumption-saving decision and the spatial choice. We then show that the solution to the household problem together with our distributional assumptions on individuals' skills delivers an analytic solution for the economy's aggregate demand system, despite the fact that our economy does not admit a representative consumer. Finally, we show that individuals' migration choices only depend on the distribution of equilibrium wages and not on any other future equilibrium prices. This implies that the dynamic competitive equilibrium of our economy has a structure akin to the neoclassical growth model: given the sequence of interest rates  $\{r_t\}_t$ , we can solve the entire path of spatial equilibria from static equilibrium conditions. The equilibrium sequence of interest rates can then be calculated from households' savings decisions. The model can

then be solved by iteratively computing the sequence of spatial equilibria and finding a fixed point for the path of interest rates.

### Individual Behavior

First consider the households' consumption-saving decision given in equation (6). Let the optimal level of expenditure when young (old) of the generation that is born at time  $t$  be denoted by  $e_t^Y$  ( $e_{t+1}^O$ ). This two-period OLG structure together with the specification of preferences in equation (4) has a tractable solution for both the optimal allocation of expenditure and the consumers' total utility  $U_r$ . We summarize this solution in the following Proposition.

**Proposition 1.** *Consider the maximization problem in equation (6) where  $V(e, P)$  is given in equation (4). The solution to this problem is given by*

$$e_t^Y(y) = \psi(r_{t+1})y \quad (9)$$

$$e_{t+1}^O(y) = (1 + r_{t+1})(1 - \psi(r_{t+1}))y \quad (10)$$

$$U_r^i = U_t(y) = \frac{1}{\eta} \psi(r_{t+1})^{\eta-1} y^\eta + \Lambda_{t,t+1}$$

where

$$\psi(r_{t+1}) = \left(1 + \beta^{\frac{1}{1-\eta}} (1 + r_{t+1})^{\frac{\eta}{1-\eta}}\right)^{-1} \quad (11)$$

$$\Lambda_{t,t+1} = -\frac{v}{\gamma} \left( \left(\frac{p_{At}}{p_{NA_t}}\right)^\gamma + \beta \left(\frac{p_{A_{t+1}}}{p_{NA_{t+1}}}\right)^\gamma \right) + (1 + \beta) \left( \frac{v}{\gamma} - \frac{1}{\eta} \right).$$

*Proof.* See Section B.2 in the Appendix. □

Proposition 1 characterizes the solution to the household problem. Four properties are noteworthy. First of all, the policy functions for the optimal amount of spending are linear in earnings. This will allow for a tractable aggregation of individuals' demands. Secondly, these expenditure policies resemble the familiar OLG structure, where the individual consumes a share

$$\frac{e_t^Y(y)}{y} = \psi(r_{t+1}) = \frac{1}{1 + \beta^{\frac{1}{1-\eta}} (1 + r_{t+1})^{\frac{\eta}{1-\eta}}}$$

of his income when young and consumes the remainder (and the accrued interest) when old. Because  $\eta < 1$ , the consumption share  $\psi(r_{t+1})$  is decreasing in the interest rate and decreasing in the discount factor  $\beta$  as both increase the value of saving. If  $\eta = 0$ , i.e. if demand is homothetic, we recover the canonical OLG solution for log utility where the consumption share is simply given by  $1/(1 + \beta)$ . Importantly, the consumption share  $\psi(r_{t+1})$  only depends on the interest rate  $r_{t+1}$  and not on relative prices  $P_t$  or  $P_{t+1}$ . This is due to our assumption that nominal income  $e$  is deflated by the same price index as the

investment good. This is convenient for tractability and similar to the single-good neoclassical growth model, where the consumption good and the investment good uses all factors in equal proportions. For our purposes, this ensures that an increase in the price of the investment good,  $p_{It}$ , makes savings more attractive but at the same reduces the marginal utility of spending. Third, lifetime utility  $U_{rt}^i$  only depends on the location  $r$  via individual income  $y_{rt}^i$ . This is due to our assumption that trade is frictionless so that the price indices (which determine  $\Lambda_{t,t+1}$ ) do not vary across space. Finally, lifetime utility is *additively separable* in income  $y_{rt}$  and current and future prices  $P_t$  and  $P_{t+1}$  (which determine  $\Lambda_{t,t+1}$ ).

Moreover, we can calculate individuals' lifetime utility  $E[U_{rt}^i]$  analytically: because life-time utility is a power function of individual income  $y_r^i$  and individual income is Frechet distributed, we get that  $E[y^\eta] = \Gamma_{\eta/\zeta} w_{rt}^\eta$ . Together with equation (7), this delivers closed form expressions for individual migration decisions, which we summarize in the following Proposition.

**Proposition 2.** *Consider the environment above. Define the relative life-time value of location  $r$  at time  $t$ ,  $\mathcal{W}_{rt}$ , by  $E[U_{rt}^i] + A_{rt} = \mathcal{W}_{rt} + \Lambda_{t,t+1}$ . Then*

$$\mathcal{W}_{rt} = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} w_{rt}^\eta + A_{rt}. \quad (12)$$

The share of people moving from  $j$  to  $r$  at time  $t$ ,  $\rho_{jrt}$ , is then given by

$$\rho_{jrt} = \frac{\exp\left(\frac{1}{\kappa} (\mathcal{W}_{rt} - MC_{jr})\right)}{\sum_{l=1}^R \exp\left(\frac{1}{\kappa} (\mathcal{W}_{rt} - MC_{jl})\right)}. \quad (13)$$

In particular,  $\rho_{jrt}$  is fully determined from static equilibrium wages  $\{w_{rt}\}_r$  and exogenous amenities and does not depend on future prices.

Proposition 2 implies that individuals' migration decisions are fully captured by  $\mathcal{W}_{rt}$ , which is a summary measure of regional attractiveness. Note that the cross-sectional variation in  $\mathcal{W}_{rt}$  results from differences in average wages,  $w_{rt}$ , and in amenities  $A_{rt}$  across labor markets. The former is endogenous and depends on the extent of spatial sorting and aggregate demand conditions. The latter is fully exogenous. Importantly, the expression for individuals' spatial choice probabilities in (13) does *not* feature  $\Lambda_{t,t+1}$ , which is constant across locations (because of the absence of trade costs) and hence does not determine spatial labor flows. It is here that the additive separability of future prices embedded in  $\Lambda_{t,t+1}$  is crucial, since it implies that agents' spatial choice problem reduces to a static decision problem, which depends only on current, not future, equilibrium objects (in particular the distribution of wages). This structure allows us to calculate the transitional dynamics in the model with a realistic geography, i.e. with about 700 regions.

## Equilibrium Aggregation and Aggregate Structural Change

The spatial equilibrium of economic activity is shaped by the heterogeneous local incidence of aggregate demand and supply conditions. Our economy does not admit a representative consumer, since the PIGL

preference specification falls outside of the Gorman class. To see this, consider a set of individuals  $i \in \mathcal{S}$ , with spending  $e_i$ . The aggregate demand for agricultural products of this set of consumers is given by

$$PC_{\mathcal{S}}^A = \int_{i \in \mathcal{S}} \vartheta_A(e_i, P) e_i di = \left( \phi + v \left( \frac{p_A}{p_M} \right)^\gamma \int_{i \in \mathcal{S}} e_i^{-\eta} \omega_i di \right) E_{\mathcal{S}},$$

where  $E_{\mathcal{S}} = \int_{i \in \mathcal{S}} e_i di$  denotes aggregate spending and  $\omega_i = e_i/E_{\mathcal{S}}$  is the share of spending of individual  $i$ . Hence, as long as preferences are non-homothetic, i.e. as long as  $\eta > 0$ , aggregate demand does not only depend on aggregate spending  $E_{\mathcal{S}}$  and relative prices, but on the entire distribution of spending  $\{e_i\}_i$ . Characterizing the aggregate demand function in our economy, which features heterogeneity through individuals' location choice (which determines the factor prices they face) and the actual realization of the skill vector  $z^i$ , is therefore in principle non-trivial.

Our model, however, delivers tractable expressions for the economy's aggregate quantities. This is due to three properties of our theory. First of all, the distributional assumption on individual skills implies that individual income  $y^i$  is Frechet distributed. Secondly, Proposition 1 showed that individuals' expenditure policy functions are linear in income  $y^i$  and hence also Frechet distributed. Finally, individual spending shares  $\vartheta_A(e, P)$  are a power function of expenditure and can therefore be calculated explicitly. This allows us to solve for the aggregate demand system explicitly as a function of equilibrium wages.

**Proposition 3.** *Let  $\mathcal{S}_r^Y$  and  $\mathcal{S}_r^O$  be the set of young and old consumers in region  $r$ . The aggregate expenditure share on agricultural good of the set of consumers  $\mathcal{S}_r^S$  is given by*

$$\vartheta_A([e_i]_{i \in \mathcal{S}_r^S}, P) \equiv \frac{PC_{\mathcal{S}_r^S}^A}{E_{\mathcal{S}_r^S}} = \phi + \tilde{v} \left( \frac{p_A}{p_M} \right)^\gamma E_{\mathcal{S}_r^S}^{-\eta}, \quad (14)$$

where  $E_{\mathcal{S}_r^S}$  is mean spending at time  $t$  and given by

$$E_{\mathcal{S}_r^Y} = \psi(r_{t+1}) \Gamma_\zeta w_{rt} \text{ and } E_{\mathcal{S}_r^O} = (1 + r_t) (1 - \psi(r_t)) \Gamma_\zeta w_{rt-1},$$

and  $\tilde{v} = v \Gamma_{\frac{\zeta}{1-\eta}} / \Gamma_\zeta^{1-\eta}$  is a constant. The aggregate share of agriculture in value added is given by

$$\vartheta_t^A \equiv \frac{PC_t^A + \phi I_t}{PY_t} = \phi + \tilde{v} \left( \frac{p_A}{p_M} \right)^\gamma \frac{\sum_{r=1}^R \left( E_{\mathcal{S}_r^Y}^{1-\eta} L_{rt} + E_{\mathcal{S}_r^O}^{1-\eta} L_{rt-1} \right)}{PY_t}. \quad (15)$$

*Proof.* See Section B.3 in the Appendix. □

Proposition 3 is an ‘‘almost-aggregation’’ result. Even though the PIGL preferences fall outside of the Gorman class, equation (14) shows that the aggregate demand of a given set of consumers resembles that of a representative consumer with mean spending  $E_{\mathcal{S}}$  and an adjusted preference parameter  $\tilde{v}$ . Because the linearity of individuals' policy functions allows to express aggregate spending directly as a function of equilibrium spatial wages, aggregate sectoral spending  $PC_t^A$  is then simply the spatial aggregate over

the respective consumer groups and the aggregate value added share of the agricultural sector takes the form in equation (15). Note that  $\vartheta_t^A$  can be directly calculated from current and past wages  $\{w_{rt}, w_{rt-1}\}_r$  and the spatial allocation of factors  $\{L_{rt}, L_{rt-1}\}_r$ . We exploit this “almost-aggregation” property intensely in computing the model. Finally, note that these equations highlight the usual demand side forces of the structural transformation: to the extent that  $\eta > 0$ , i.e. preferences are non-homothetic, the agricultural value added share,  $\vartheta_t^A$  will decline as income rises. Similarly, changes in relative technological progress (and therefore in sectoral prices) will affect agricultural spending as long as  $\gamma \neq 0$ .

## Equilibrium Conditions

As highlighted above, the key properties of our model are that (i) individual moving decisions are static and (ii) that our economy generates an aggregate demand system as a function of regional wages. This implies that, for a given path of interest rates  $\{r_t\}_t$ , we can calculate the equilibrium by simply solving a sequence of static equilibrium conditions. Consider first the goods market. The market clearing condition for agricultural products is given by

$$L_{rt}\Gamma_\zeta w_{rt}s_{rAt} = (1 - \alpha)\pi_{rAt}\vartheta_t^A PY_t, \quad (16)$$

where  $s_{rAt}$  is the agricultural employment share in region  $r$  at time  $t$ .<sup>17</sup> Hence, total agricultural labor earnings in region  $r$  are equal to a share  $1 - \alpha$  of total agricultural revenue in region  $r$ . This in turn is equal to region  $r$ 's share,  $\pi_{rAt}$ , in aggregate spending on agricultural goods  $\vartheta_t^A PY_t$ . The CES structure of consumers' preferences implies that regional trade shares,  $\pi_{rAt}$ , are given by

$$\pi_{rAt} = \left(\frac{P_{rAt}}{P_{At}}\right)^{1-\sigma} = \frac{(Q_{rAt}w_{rt}^{\alpha-1})^{\sigma-1}}{\sum_{j=1}^R (Q_{jAt}w_{jt}^{\alpha-1})^{\sigma-1}}, \quad (17)$$

i.e. they neither depend on the identity of the sourcing region, nor the equilibrium capital rental rate  $R_t$  or the common component of productivity  $Z_{st}$ . Rather, a region  $r$ 's agricultural competitiveness only depends on its productivity  $Q_{rAt}$  and the equilibrium price of labor. An analogous expression holds for the non-agricultural sector.

To characterize the equilibrium, we find it useful to express the sectoral market clearing conditions in equation (16), as

$$L_{rt}\Gamma_\zeta w_{rt} = (1 - \alpha)\left(\pi_{rAt}\vartheta_t^A + \pi_{rAt}\left(1 - \vartheta_t^A\right)\right) PY_t \quad (18)$$

$$\frac{s_{rAt}}{1 - s_{rAt}} = \frac{\pi_{rAt}}{\pi_{rNA_t}} \frac{\vartheta_t^A}{1 - \vartheta_t^A}. \quad (19)$$

---

<sup>17</sup>Because efficiency units are fully substitutable, the equilibrium only determines the allocation of efficiency units. Once we explicitly introduce sector-specific human capital in Section 3.3, this indeterminacy will be resolved. It is this latter model that we take to the data.



Equation (18) shows that local labor earnings are a *demand-weighted* average of regional sectoral trade shares and hence highlights the urban bias of the structural transformation: a decline of the aggregate agricultural spending share  $\vartheta_t^A$  tends to reduce regional earnings in locations who have a comparative advantage, in the agricultural sector. Moreover, equation (19) illustrates the spatial co-movement of sectoral employment shares. Because regional (scaled) agricultural shares  $\frac{s_{rAt}}{1-s_{rAt}}$  are proportional to the aggregate (scaled) agricultural expenditure share  $\frac{\vartheta_t^A}{1-\vartheta_t^A}$ , a decline in the aggregate spending share  $\vartheta_t^A$  tends to reduce agricultural employment shares in *all* locations.

These equations highlight how the spatial distribution of economic activity is fully determined from *static* equilibrium conditions. Note first that GDP is proportional to aggregate labor earnings, i.e.  $(1-\alpha)PY_t = \Gamma_\zeta \sum_r L_{rt} w_{rt}$  (see equation (18)) and that the spatial labor supply function is *only* a function of the the spatial distribution of wages (see Proposition (2)). Likewise the agricultural value added share  $\vartheta_t^A$  only depends on the vector of current and past wages and population. Together these equilibrium conditions fully determine the equilibrium wages and labor allocations across space.

Finally, since the future capital stock is simply given by the savings of the young generation, we get that capital accumulates according to

$$K_{t+1} = (1 - \psi(r_{t+1})) \sum_r \Gamma_\zeta w_{rt} L_{rt} = (1 - \psi(r_{t+1})) (1 - \alpha) PY_t, \quad (20)$$

i.e. future capital is simply a fraction  $1 - \psi(r_{t+1})$  of aggregate labor earnings. This proportionality between the aggregate capital stock and aggregate GDP is a consequence of the linearity of agents' consumption policy rules. Note that our model retains many features of the baseline neoclassical growth model. In particular, for given initial conditions  $\{K_0, \{L_{r,-1}\}_r, \{w_{r,-1}\}_r\}$  and a path of interest rates  $\{r_t\}_t$ , the equilibrium evolution of wages and people are solutions to the static equilibrium conditions highlighted above. Given these allocations, the model predicts the evolution of the capital stock according to equation (20). A dynamic equilibrium then requires that the set of interest rates  $\{r_t\}_t$  is consistent with the implied evolution of the capital stock. In practice, it is straightforward to calculate this dynamic equilibrium even for a realistic economic geography of hundreds of locations. More formally, a competitive equilibrium in our economy is defined in the following way.

**Definition 4.** *Consider the economy described above. Let the initial capital stock  $K_0$ , the initial spatial allocation of people  $\{L_{r,-1}\}_r$  and the vector of wages  $\{w_{r,-1}\}_r$  be given. A dynamic competitive equilibrium is a set of prices  $\{P_{rst}\}_{rst}$ , wages  $\{w_{rt}\}_{rt}$ , capital rental rates  $\{R_t\}_t$ , labor and capital allocations  $\{L_{rst}, K_{rst}\}_{rst}$ , consumption and saving decisions  $\{e_{rt}^Y, e_{rt}^O, s_{rt}\}_{rt}$ , and demands for regional varieties  $\{c_{rst}\}_{rst}$  such that consumers' choices  $\{e_{rt}^Y, e_{rt}^O, s_{rt}\}_{rt}$  maximize utility, i.e. are given by equations (9) and (10), the demand for regional varieties follows equation (17), firms' factor demands maximize firms' profits, markets clear, the capital stock evolved according to equation (20) and the allocation of people across space  $\{L_{r,t}^Y\}$  is consistent with individuals' migration choices in equation (7).*

One implication of our theory is that the spatial structure of the model naturally generates an explanation

for an “agricultural productivity gap”, i.e. why value added per worker in the agricultural sector is relatively low (see e.g. [Gollin et al. \(2014\)](#)). In our model value added per worker in agriculture *relative* to average value added per worker is given by

$$\frac{VA_t^{Ag}/L_t^A}{VA_t/L_t} = \frac{\vartheta_t^A}{s_{At}} = \frac{\sum_r s_{Ar} \frac{w_r L_r}{\sum_r w_r L_r}}{\sum_r s_{Ar} \frac{L_r}{\sum_r L_r}}.$$

Hence, there is an agricultural productivity gap, i.e.  $\frac{VA_t^{Ag}}{L_t^A} < \frac{VA_t}{L_t}$ , whenever the spatial correlation between wages and agricultural employment share is *negative*. Differences in regional efficiency combined with frictions to spatial mobility are therefore one potential mechanism to explain why agricultural productivity low. Frictions to labor mobility are crucial to generate these patterns - if wages were equalized across locations, value added per worker would be equalized across sectors and our economy would essentially be aggregative. To see this more clearly, it is useful to consider two special cases, where there is no distinct role for space.

**Special Case 1: The Spaceless Economy** The special case of our theory with just a single location shows that our model is essentially a standard, macroeconomic model of the structural transformation augmented by a spatial layer. With a single region, it is easy to show that GDP is given by

$$PY_t = Z_t K_t^\alpha L^{1-\alpha},$$

where  $Z_t \equiv \Gamma_\zeta^{1-\alpha} Z_{At}^\phi Z_{NA_t}^{1-\phi}$ . Moreover, capital still accumulates according to equation (20). Even though consumers’ savings rates  $1 - \psi_{t+1}$  depend on future interest rates (and hence on the future capital stock  $K_{t+1}$ ), it can be shown that - given some initial condition  $K_0$  and processes for productivity  $\{Z_{At}, Z_{NA_t}\}_t$  - there exists a unique dynamic equilibrium path of capital  $\{K_t\}_t$ . Section B.4 in the Appendix provides a formal proof.

Furthermore, as in the baseline macroeconomic model of the structural transformation, this equilibrium path can be characterized *independently* of the sectoral labor allocation (see e.g. [Herrendorf et al. \(2014\)](#)). To see this, suppose that the economy is on a balanced growth path where aggregate income, capital and wages grow at rate  $g$  and the interest rate is constant. In Section B.4 in the Appendix we characterize this path and show that the agricultural share in value added  $\vartheta_t^A$  is given by

$$\vartheta_t^A = \phi + \tilde{\nu} \chi \left( \frac{Z_{NA_t}}{Z_{At}} \right)^\gamma w_t^{-\eta},$$

where  $\chi$  is a constant, which is a simple function of exogenous parameters. This relative demand system again resembles a representative household with PIGL preferences. Finally, note that in the single region case, sectoral employment shares are equal to sectoral value added shares, i.e.  $s_{At} = \vartheta_t^A$  (see (19)). Hence, as in most frictionless models of the structural transformation, value added per worker is equalized across

sectors.

**Special Case 2: Frictionless labor mobility** Another polar case of our model is the case of frictionless labor mobility. In this parametrization, there are no moving costs and individuals have no locational preferences.<sup>18</sup> This implies that equilibrium wages (and individual welfare) are equalized across space at each point in time. The equilibrium trade shares,  $\pi_{rst}$  are therefore given by  $\pi_{rst} = Q_{rst}^{\sigma-1}$ , i.e. are fully exogenous and only depend on regional productivity (see equation (17)). Moreover, equation (18) implies that the size of the local *population* is a demand-weighted average of the exogenous local sectoral productivities. The change in the local population is therefore given by

$$dL_{rt} = \underbrace{(Q_{rAt}^{\sigma-1} - Q_{rNAAt}^{\sigma-1}) d\vartheta_t}_{\text{Urban Bias}} + \underbrace{\sum_s \vartheta_t^s dQ_{rst}^{\sigma-1}}_{\text{Local Productivity Growth}}. \quad (21)$$

The first term on the right is what we refer to as the urban bias: A decline in the spending share on agricultural goods reduces population in region  $r$  if and only if region  $r$  has a *comparative* advantage in the agricultural sector.<sup>19</sup> In particular, if it was not for local productivity shocks, regional population growth and the initial agricultural employment share would be perfectly negatively aligned and spatial reallocation would account for a sizable share of the decline in agricultural employment.

This already suggests that local productivity shocks,  $dQ_{rst}$  play an important role to explain why spatial reallocation plays a minor role for the decline in the agricultural employment share. In our full, quantitative model, three additional channels affecting incentives to spatially reallocation are at work. First, costly spatial mobility prevent wages from being equalized and tends to keep individuals in their location. Second, local amenities are an additional determinant of moving flows. Finally, idiosyncratic shocks reduce the sensitivity of moving flows to wage differentials and generate gross mobility flows, which exceed the extent of net reallocation.

### 3.3 Selection, Human Capital and Labor Supply

So far we assumed that human capital is perfectly substitutable across sectors and that all individuals are ex-ante identical. In preparation for the quantitative exercise, we now extend it to allow for imperfect substitutability of efficiency units across industries and for systematic differences in skill-supply.

We add these ingredients for two reasons. First, the extent of skill substitutability across sectors is an important determinant of the costs of the *regional* transformation. While moving costs are a hurdle for

<sup>18</sup>Formally, this parametrization is nested in our full model by assuming that  $MC_{jr} = 0$ ,  $A_r = 0$  and  $\kappa \rightarrow 0$ .

<sup>19</sup>In fact, it is easy to show that  $\text{sgn}(Q_{rAt}^{\sigma-1} - Q_{rNAAt}^{\sigma-1}) = \text{sgn}(s_{rAt} - s_{NAAt})$ . To see this, note that  $s_{rAt} = \vartheta_t^A$  and

$$Q_{rAt}^{\sigma-1} - Q_{rNAAt}^{\sigma-1} = \pi_{rAt} - \pi_{rNAAt} = \frac{s_{rAt} L_{rt}}{\vartheta_t^A} - \frac{s_{rNAAt} L_{rt}}{1 - \vartheta_t^A} = \frac{s_{rNAAt}}{1 - \vartheta_t^A} \left( \frac{s_{rAt} / \vartheta_t^A}{(1 - s_{rAt}) / (1 - \vartheta_t^A)} - 1 \right) L_{rt}$$

spatial reallocation, an upward sloping relative sectoral supply function within locations makes it costly to reallocate agricultural workers to factories. Second, systematic differences in skills are important in that they generate spatial sorting. If skilled workers have a comparative advantage in manufacturing jobs, they are more likely to move to locations which are productive in the manufacturing sector. This spatial sorting captures a distinct role for the *supply* of agricultural labor. A shift of people towards urban areas will, for example, reduce the relative supply of agricultural goods at given prices.<sup>20</sup>

Our model can easily be extended along these lines. In particular, we assume that individuals draw a two-dimensional vector of skill-specific efficiency units  $z^i = (z_A^i, z_{NA}^i)$  and sort across industries based on their comparative advantage. We also assume that individuals can be of two types - high skilled and low skilled. Their skill type  $h \in \{L, H\}$  determines the distribution of  $z^i$ . As before, we assume that  $z_s^i$  is drawn independently from a Frechet distribution  $F_s^h(z) = e^{-\Psi_s^h z^{-\zeta}}$ , where  $\Psi_s^h$  parametrizes the average level of human capital of individuals of skill type  $h$  in sector  $s$ . Without loss of generality we parameterize  $\Psi_s^h$  as  $\Psi_A^L = \Psi_{NA}^L = 1$ ,  $\Psi_{NA}^H = \mu q$  and  $\Psi_A^H = q$ . Here,  $q$  measures the absolute advantage of skilled individuals and  $\mu$  governs the comparative advantage of skilled workers in the non-agricultural sector. We denote the share of the aggregate labor force that is skilled by  $\lambda$  and assume that it is constant.<sup>21</sup> In contrast, the spatial allocation of human capital, i.e. the share of skilled workers in region  $r$  at time  $t$ ,  $\lambda_{rt}$ , is endogenous and determined by workers' migration decisions.

Because of the properties of the Frechet distribution, these additional ingredients leave the rest of the theoretical analysis almost unchanged.<sup>22</sup> As we show in detail in Section B.5 of the Appendix, the key endogenous object is no longer the vector of regional wages  $w_{rt}$ , but *average* earnings of individuals in skill group  $h$ , which are given by  $E^h [y_r^i] = \Gamma_\zeta \Theta_r^h$ , where

$$\Theta_{rt}^h = \left( \Psi_A^h w_{rAt}^\zeta + \Psi_{NA}^h w_{rNAAt}^\zeta \right)^{1/\zeta}. \quad (22)$$

While  $\Theta_{rt}^h$ , differs across skill-types, it is equalized across sectors within locations and can be directly calculated from regional wages  $(w_{rA}, w_{rNA})$ , which are no longer equalized across industries within locations. Moreover, the regional attractiveness in Proposition 2 is skill-specific and given by  $\mathcal{W}_{rt}^h = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_{rt}^h)^\eta + A_{rt}$ , i.e. high skilled individuals put a higher relative weight on non-agricultural wages  $w_{rNAAt}$  and hence consider locations with a strong manufacturing sector particularly attractive. Finally, the law of motion of the population is now skill-specific and given by  $L_{rt} \lambda_{rt}^h = \sum_{j=1}^R \rho_{jrt}^h \lambda_{rt-1}^h L_{jt-1}$ , where  $\rho_{jrt}^h$  is given in Proposition 2. The addition of imperfect skill substitutability, while begetting extra notation, leaves the tractability of our framework untouched and it is this version of the model that we take to the data in the next section.

<sup>20</sup>Spatial mobility costs are therefore one example of cohort-specific ‘‘sectoral moving costs’’ highlighted by [Hobijn et al. \(2018\)](#) and [Porzio and Santangelo \(2017\)](#).

<sup>21</sup>Hence, we abstract from human capital accumulation and simply assume that skills are fully inherited between parents and children. Our timing assumption therefore implies that individuals know their skill  $h \in \{L, H\}$  prior to migrating but not the realization of their  $z^i$ .

<sup>22</sup>The key property is the fact that the Frechet distribution is max-stable.

## 4 Spatial Structural Change in the US: 1880 - 2000

We now apply this theory to the experience of the United States over the last 120 years. To do so we construct a novel panel data set on the regional development of the United States between 1880 and 2000. We describe the data in Section 4.1. One of the main features of our data collection effort is that we compile measures of average labor earnings at the regional level. In Section 4.2 we use this data to provide direct empirical evidence on the urban bias of the structural transformation. Finally, we calibrate the model in Section 4.3. Our quantitative analysis is contained in Section 5.

### 4.1 Data

To compose our panel we draw on various data sets published by the US Census Bureau. In particular, we use information from the Census of Manufacturing for 1880 and 1910, the Population Census for 1880-2000 and the County and City Data Books for 1940-2000. From these sources we construct a panel data set of total workers  $\{L_{rt}\}_{rt}$ , average manufacturing earnings  $\{w_{rt}\}_{rt}$ , and sectoral employment shares  $\{s_{rAt}\}_{rt}$  for all US counties at 30 year intervals between 1880 and 2000.<sup>23</sup> We define the agricultural sector to comprise agriculture, fishing and mining industries following the 1950 Census Bureau industrial classification system. All remaining employed workers are assigned to the non-agricultural sector. We construct average manufacturing wages from county level total manufacturing payroll data and manufacturing head counts obtained from the same source. Table 17 in Appendix E contains a comprehensive list of all data sources and more details on the construction of the data set.

We aggregate this county-level data to the level of commuting zones, which we take as our definition of a regional labor market (see Tolbert and Sizer (1996)).<sup>24</sup> We do so for two reasons. First, we need stable regional boundaries over time. Second, labor markets should be large enough such that they contain both an agricultural and a non-agricultural sector. Commuting zones partition the territory of the United States into 741 polygons, all of which exhibited non-zero non-agricultural employment shares in 1880. This procedure leaves us with a panel data set, which features sectoral employment shares, total employment and average manufacturing earnings for all continental commuting zones from 1880 to 2000. For the main calibration of the model we employ the cross-sections 1880, 1910, 1940, 1970 and 2000 only and normalize the size of the total US workforce to unity in each period.

---

<sup>23</sup>Twelve states have not obtained statehood in 1880. We list them here and give the year of their accession to the Union: North Dakota (1889), South Dakota (1889), Montana (1889), Washington (1889), Idaho (1890), Wyoming (1890), Utah (1896), Oklahoma (1907), New Mexico (1912), Arizona (1912), Alaska (1959) and Hawaii (1959). We exclude Alaska, Hawaii and Washington D.C. The 1880 Census does report data for counties in all states, even those that had not yet officially obtained statehood in 1880, with two exceptions: Oklahoma and Hawaii. We impute 1880 data for Oklahoma's counties using a procedure described in Appendix (E).

<sup>24</sup>To do so, we construct a crosswalk between counties and 1990 commuting zones for every decade between 1880 and 2000. Using this we re-aggregate the county level data for the various years to commuting zones, employing area weights to allocate workers wherever counties are split. A detailed description of the construction of the county to commuting zone cross-walk for 1790-2000 as well as a panel of the populations of US commuting zones for that period is made available on the authors' website (Eckert et al. (2018)).

In Figure ?? we depict the geography of the US at the commuting zone level. In the left panel, we show the regional agricultural employment shares in 1880. While some regions (in particular, commuting zones in Northeastern states like Massachusetts or New York) already have agricultural employment shares of less than 10%, many commuting zones in the South have more than 75% of their population employed in the agricultural sector. On the right we show local population growth rates between 1880 and 1910. The absence of a strong correlation between population growth and agricultural specialization is apparent. In particular, most population growth is observed in western commuting zones, which - in 1880 - tend to have intermediate agricultural shares. In contrast, consideration other than the structural transformation seemed to have caused the observed migration patterns and we will use our model as a way to measure these alternative mechanisms.

In addition, we rely heavily on the 1940 edition of the decennial Micro Census by the US Census Bureau. This is the most recent Census which contains individual identifiers for all US counties and it is the first Census for which earnings and education variables are available. We use this information to calibrate the spatial distribution of skilled workers.

To estimate the moving cost parameters we exploit the information in the Population Census. In the model, individuals move once in their lifetime to access their preferred labor market. In the data we therefore look at patterns of lifetime mobility. All censuses between 1880 and 2000 contain information on the state of residence and the state of birth.

Finally, we use micro-data on expenditure patterns from the 1930s to estimate consumer preferences. The Consumer Expenditure Survey in 1936 (“Study of Consumer Purchases in the United States, 1935-1936”) contains detailed information on individual expenditure and allows us to calculate the expenditure share of food. We exploit this cross-sectional information on expenditure shares and total expenditure to estimate the extent of non-homotheticities in demand. Our information for the time-series of relative prices is taken from [Alder et al. \(2018\)](#).

## 4.2 Earnings, Agricultural Employment and the Spatial Bias of Structural Change

The spatial bias of the structural transformation and the extent to which individuals respond to such changing demand condition through migration is a central aspect of our analysis. In this section we provide some direct evidence on this mechanism, without relying on the calibrated model.

The key implication of the spatial bias is a negative relationship between initial agricultural specialization and subsequent economic performance. In particular, if spatial moving costs are important, the process of structural change will reduce factor prices relatively more in agricultural areas. To test this relationship, we consider the regression

$$\ln \bar{w}_{rt+1}^M = \delta_{t,State} + \alpha \ln s_{rAt} + \beta \ln s_{rAt} \times \Delta s_{At+1} + \gamma \ln \bar{w}_{rt}^M + u_{rt+1},$$

where  $\bar{w}_{rt}^M$  denotes average manufacturing earnings in region  $r$  at time  $t$  (which are directly observed

in the data),  $\delta_{t,State}$  contains year and state fixed effects,  $\ln s_{rAt}$  is the agricultural employment share in region  $r$  and  $\Delta s_{At+1}$  denotes the change in the *aggregate* agricultural employment share between  $t$  and  $t + 1$ . The coefficient  $\alpha$  captures the direct effect of agricultural specialization on manufacturing earnings growth. The coefficient  $\beta$  captures the urban bias. In particular, we expect  $\beta$  to be positive: the larger the decline in the agricultural share, the more adversely will regions with a comparative advantage in agriculture be affected.

	Dep. variable: ln Manufacturing Earnings			
	1880-1910	Full Sample		
$\ln s_{rAt-1}$	-0.040*** (0.015)	-0.106*** (0.006)	-0.087*** (0.007)	-0.063*** (0.008)
$\ln s_{rAt-1} \times \Delta s_{At}$			0.243*** (0.094)	0.247** (0.099)
lagged ln man earnings	0.118*** (0.022)	0.120*** (0.018)	0.117*** (0.018)	0.154*** (0.021)
Year FE		✓	✓	✓
State FE	✓	✓	✓	✓
State $\times$ Year FE				✓
Observations	717	2868	2868	2868
$R^2$	0.594	0.983	0.983	0.985

Notes: Robust standard errors in parentheses with \*\*\*, \*\* and \* respectively denoting significance at the 1%, 5% and 10% levels.

Table 1: The Urban Bias of the Structural Transformation

The results are reported in Table 1. In the first column we focus on the time period from 1880 to 1910. In this time period, the aggregate agricultural employment share declined from 50% to around 35%. Cross-sectionally, column one shows that regions with a higher agricultural share in 1880 experienced significantly lower earnings growth in the subsequent 30 years. Column two shows that this relationship is not confined to the 1880-1910 period, but holds true in the entire sample. In the last two columns we directly exploit the panel structure of our data, to estimate the coefficient  $\beta$  from the interaction between the initial regional agricultural share and the change in aggregate agricultural share. Consistent with an urban bias, a faster decline in agricultural employment is particularly harmful for regions, which have a larger agricultural share.

The results in Table 1 highlight why the weak correlation between agricultural employment shares and subsequent population outflows is surprising: the structural transformation does indeed reduce relative wages in agricultural locations. The quantitative impact is, however, limited. To see this, consider Figure 2, where we depict the cross-sectional correlation between agricultural employment shares in 1880 and average manufacturing earnings in 1910. While there is a strong negative relationship, there is also ample variation in future wages holding the agricultural employment share fixed. This residual variation is important for the correlation between agricultural employment shares and population outflows. In our theory, individuals do not have particular preferences for non-agricultural locations - they care about the sectoral composition of a location only in as far as it offers higher life-time utility through favorable



Notes: The figure shows the correlation between the agricultural employment share in 1880 and the log of average manufacturing earnings in 1910. The size of the markers in the right panel indicates the size of commuting zones as measured by their population in 1880.

Figure 2: Agricultural Employment and Future Earnings in 1880

factor prices. And Figure 2 suggests that the cohort born in 1880 does not necessarily have to move towards non-agricultural places to increase their life-time earnings.

## 4.3 Calibration

In this section we describe the calibration of our model. First we discuss our calibration strategy, then the fit of our model with respect to both targeted and non-targeted moments.

### 4.3.1 Calibration Strategy

In this section, we outline the general features of our calibration strategy. In Section D in the Online Appendix we provide considerably more detail. Even though our parameters are calibrated jointly, we organize the discussion of our calibration strategy around the structural parameters and the respective moments, which are most informative.

**The evolution of aggregate productivity:**  $\{Z_{At}, Z_{NA_t}\}_t$  We calibrate the model such that *aggregate* income per capita grows at a constant rate and that the capital-output ratio is constant. In Section B.6 in the Appendix, we show that this implies that interest rates are constant and have a closed-form expression.<sup>25</sup> We then calibrate the time series of aggregate sectoral productivities,  $\{Z_{At}, Z_{NA_t}\}_t$ , to match the evolution of relative prices and a GDP growth rate of 2%.

**Moving costs and idiosyncratic location preferences:**  $MC_{jr}$  and  $\kappa$  We specify the cost of spatial reallocation as having both a fixed and a variable component and we allow the latter to depend on the migration distance in a flexible way. In particular, we assume that

<sup>25</sup>In the counterfactual analysis, interest rates will of course be free to vary over time.



$$MC_{jr} = \tau + \delta_1 d_{jr} + \delta_2 d_{jr}^2$$

whenever  $j \neq r$  and zero otherwise. Here,  $\tau > 0$  parameterizes the fixed cost of moving and  $\delta_1$  and  $\delta_2$  govern how mobility costs vary with distance. We normalize  $d_{jr}$  so that the maximum distance in the US is 1. Because in the theory workers move only once (at the beginning of their working life), for the remainder of the paper we all mention of migration and reallocation is understood to refer to lifetime migration. Following [Molloy et al. \(2011\)](#) we measure the aggregate lifetime migration rate as the fraction of people who live in a different location than where they were born.

In the Census data we only observe individuals' region of birth at the state level. We therefore aggregate the commuting-zone level migration flows in the model to the level of US states and choose  $(\kappa, \tau, \delta_1, \delta_2)$  so as to make the model best fit the observed state level flows. We focus on workers between 26 and 50 years of age and chose  $\tau$  so as to match their aggregate lifetime interstate migration rate between 1910 and 1940 exactly. We then chose  $(\kappa, \delta_1, \delta_2)$  to minimize the distance between spatial mobility rates in the data and the model.<sup>26</sup>

**Spatial productivities and amenities:**  $\{Q_{rst}, A_{rt}\}_{rst}$  We follow the recent quantitative spatial economics literature and calibrate local productivities and amenities  $\{Q_{rst}, A_{rt}\}_{rst}$  as structural residuals [Redding and Rossi-Hansberg \(2017\)](#). More specifically, we chose  $\{Q_{rst}, A_{rt}\}_{rst}$  so that the model matches the spatial data on agricultural employment shares, local employment and average manufacturing earnings exactly, conditional on all other structural parameters. In [Appendix D.1](#) we formally show that there is a unique mapping from the observed spatial data on agricultural employment shares, populations and average manufacturing earnings to the vector of local productivity  $\{Q_{rst}\}_{rst}$ , conditional on a set of calibrated parameters. Intuitively, local sectoral employment shares contain information on  $\frac{Q_{rAt}}{Q_{rNA_t}}$  while the level of wages along with the total number of workers informs the level of  $Q_{rNA_t}$ . The vector of amenities  $\{A_{rt}\}_{rt}$  can then be inferred from the observed net population flows.

**Preference parameters:**  $\eta, \gamma, \phi$  We estimate the strength of the non-homotheticity in the demand system ( $\eta$ ) from the cross-sectional relationship between sectoral spending shares and the level of expenditure. To do so, we use historical micro data from the Consumer Expenditure Survey in 1936, the ‘‘Study of Consumer Purchases in the United States, 1935-1936’’. As  $\gamma$  determines the price elasticity of demand, we discipline  $\gamma$  with the elasticity of substitution. [Comin et al. \(2017\)](#) estimate this elasticity to be around 0.7 in post-war data for the US and we calibrate our model to be consistent with this number. Finally, we use the time-series of the aggregate agricultural employment share to identify the remaining parameters  $\phi$  and  $v$ . Given that our model matches the joint distribution of agricultural employment

---

<sup>26</sup>More specifically, we chose  $(\kappa, \delta_1, \delta_2)$  to minimize  $\sum_i \sum_{j \neq i} L_{j,1940} \left( \log \rho_{ij,1940}^{DATA} - \log \rho_{ij,1940}^{MODEL}(\kappa, \delta_1, \delta_2) \right)^2$  conditional on always exactly matching the aggregate interstate migration rate through the choice of  $\tau$ . As shown in [Section D.6](#) in the Appendix, the number of stayers in a commuting zone is a monotone function in  $\tau$  given  $(\kappa, \delta_1, \delta_2)$ .

shares and population size perfectly, internal consistency requires us to also match the time series of the aggregate agricultural employment share exactly. As we discuss in detail in Section C.6 in the Appendix, the income effects as implied from the cross-sectional spending-food relationship are not strong enough to explain the entire decline in agricultural employment in the time-series.<sup>27</sup> We therefore allow the parameter  $\phi$  to be time-specific to fully account for the residual decline in agricultural employment and choose  $\mathbf{v}$  to minimize the required time-variation in  $\phi_t$ . Intuitively,  $\mathbf{v}$  is chosen for the model to explain as much of the aggregate process of structural change as possible, given the income and price elasticities  $\eta$  and  $\gamma$ . Recall that  $\phi$  does not enter the household's decision problem directly.

**Skill supply:  $\zeta, \mu, q$  and  $\{\lambda_{r1880}\}_r$**  To parametrize the skill supply, we need values for the supply elasticity ( $\zeta$ ), the comparative and absolute advantage of skilled workers ( $q$  and  $\mu$ ) and the initial distribution of skilled workers across space in 1880,  $\{\lambda_{r1880}\}_r$ . We define skilled individuals as workers who completed at least a high school education in 1940 and hold the aggregate share of skilled workers fixed.<sup>28</sup> This choice yields an aggregate skilled employment share of about 0.3. We then calibrate  $\{\lambda_{r1880}\}_r$  for the model to exactly replicate the the spatial skill distribution 1940, which is the only year for which educational attainment at the commuting zone level is directly observable. We calibrate the parameter  $\zeta$ , i.e. the dispersion of individual productivity, to match the dispersion of earnings in the 1940 Census data. The model implies that the variance of log earnings within region-skill cells is given by  $\frac{\pi^2}{6}\zeta^{-2}$ . We therefore identify  $\zeta$  from  $\zeta = \frac{\pi}{6^{1/2}}\text{var}(\hat{u}_{rsh}^i)^{-1/2}$ , where  $\text{var}(\hat{u}_{rsh}^i)$  is the variance of the estimated residuals from a regression of log earnings on commuting zone, sector and skill-group fixed effects in the 1940 Census Data.

The two parameters  $q$  and  $\mu$  in turn are chosen to match the aggregate skill premium and the aggregate relative manufacturing employment share of skilled workers in 1940. We calculate the skill premium as the ratio of average labor earnings of skilled relative to unskilled individuals in the 1940 Census data. Similarly, we compute the relative manufacturing employment share of skilled workers as the non-agricultural employment share of skilled workers relative to the one of unskilled workers. Note that these measures already incorporate the unbalanced spatial sorting of skilled and unskilled individuals, i.e. they take into account that skilled workers live in high-wage and manufacturing intensive localities.

**Other parameters:  $\beta, \delta, \sigma$  and  $\alpha$**  We estimate the rate of time preference  $\beta$  from aggregate macro relationships. In particular, we chose  $\beta$  to be consistent with the aggregate rate of investment. The capital share  $\alpha$  is set to match an aggregate capital share of 1/3. The elasticity of substitution between regional varieties,  $\sigma$ , is set to 4. In Section 6 below and in the Online Appendix we provide extensive robustness

<sup>27</sup>This discrepancy between the cross-section and time-series is not particular to our application. For example, the results reported in Comin et al. (2017) also imply different estimates for the income elasticity stemming from the cross-section and the time-series. While reconciling this discrepancy between the cross-section and the time-series is an important open research question, it is not the main focus of our paper.

<sup>28</sup>Because we focus on the spatial aspects of the structural transformation, we abstract from skill deepening. The model could easily be extended to allow for changes in the aggregate supply of human capital.

Parameter		Target	Value	Moments	
				Data	Model
<i>Skill Supply</i>					
$\zeta$	Skill heterogeneity	Residual Earnings variance in 1940	1.62	0.62	0.62
$\xi$	Share of skilled individuals	Share with more than high school degree in 1940	0.3	0.3	0.3
$\mu$	Comparative advantage	Rel. non-ag. share of skilled workers in 1940	3.41	1.21	1.21
$q$	Absolute advantage	Skill premium in 1940	0.68	1.62	1.61
$[\lambda_{r,1880}]$	Initial distribution of skilled individuals	Spatial skill distribution in 1940	.	$\{\lambda_{r,1940}\}$	Accounting
<i>Regional Fundamentals</i>					
$[Q_{rAt}]_{r,t}$	Agricultural productivity	Regional empl. shares and earnings	See Appendix	$\{e_{r,t}^{Man}, s_{A,r,t}\}$	Accounting
$[Q_{rNA}]_{r,t}$	Non-agricultural productivity				
$[A_{r,t}]_{r,t}$	Amenities	Net migration flows		$\{L_{r,t}^h\}$	
<i>Time Series Implications</i>					
$[Z_{NA}]_{r,t}$	Non-agricultural productivity	Aggregate growth rate of GDP pc	See Appendix	2%	2%
$[Z_A]_{r,t}$	Agricultural productivity	Relative price of ag. goods	See Appendix	$\{P_{A,t}/P_{NA,t}\}$	Accounting
<i>Preference Parameters</i>					
$\beta$	Discount rate	Investment rate along the BGP	0.29	0.15	0.15
$\phi$	Ag. share in price index	Time series of ag. empl. share	See Appendix	$\{s_{A,t}\}$	Accounting
$\nu$	PIGL Preference parameter	Time series of ag. empl. share	0.017	0.20	0.20
$\eta$	Non-homotheticity	Ag. share - expenditure relationship	0.35	0.04	0.04
$\gamma$	Price sensitivity	Elasticity of substitution in 2000	0.35	0.7	0.7
<i>Moving Costs and Mobility</i>					
$\tau$	Fixed costs of moving	Lifetime interstate migration rate	1.63	0.32	0.32
$\kappa$	Dispersion of idiosyncratic tastes	Observed state-to-state flows	0.42	Estimated with NLS	
$[\delta_1, \delta_2]$	Distance elasticity of moving costs		[8.44, -6.39]	Estimated with NLS	
<i>Other parameters</i>					
$\delta$	Depreciation rate (over 30 years)	set exogenously	0.91 (0.08 annually)	.	.
$\alpha$	Capital share in production function	set exogenously	1/3	.	.
$\sigma$	Elasticity of substitution	set exogenously	4	.	.

Notes: The table contains the calibrated parameters. See Section 4.3 for the calibration strategy. In Section D in the Online Appendix we provide additional details.

Table 2: Calibrated structural parameters

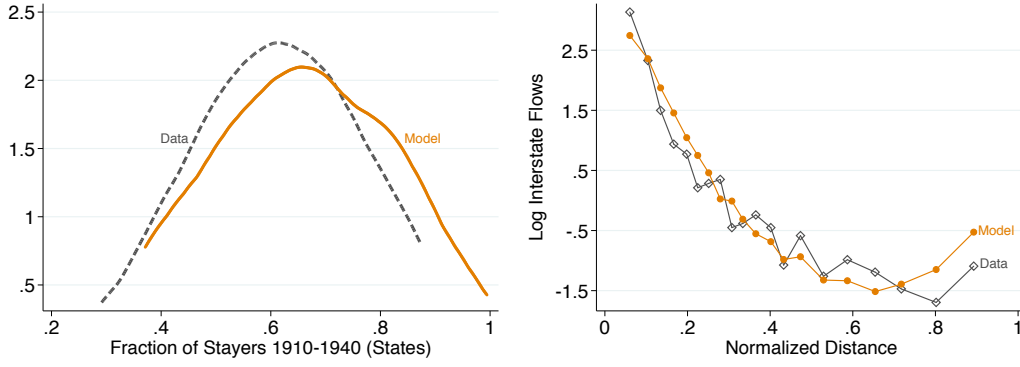
checks for this parameter.<sup>29</sup> Finally, the depreciation rate  $\delta$  is set to a 0.08, which is a central value in the literature.

### 4.3.2 Calibration Results and Model Fit

In Table 2 we report the calibrated parameters and the main targeted moment, both in the data and the model. Naturally, the parameters are calibrated jointly.

The presence of *spatial mobility costs* is an important component of our theory. To see that our model matches important features of the lifetime migration data well, we display the distribution of stayers across states and the relationship between moving flows and distance in Figure 3. In the left panel we depict the cross-sectional distribution of the share of “stayers”, i.e.  $\rho_{rr}$ , both in the data and the model. Because regions differ both in attractiveness, i.e. the utility they provide to their residents, and in their distance to more attractive places, there is sizable heterogeneity in the extent to which regions are able to retain workers. Figure 3 shows that the model matches this cross-sectional heterogeneity, even though

<sup>29</sup> Allen and Arkolakis (2014) use  $\sigma = 9$  for a model calibrated to US counties, while Monte et al. (2015) use  $\sigma = 4$  for the same purpose. Since we calibrate our model to the more aggregated commuting zones,  $\sigma$  would be expected to be lower. We consider values for  $\sigma \in [3, 9]$ .



Notes: In the left panel we plot the distribution of the share of people staying in their home state between 1910 and 1940 in the data (grey dotted line) and model (orange solid line). To construct the right panel, we run a gravity equation of the form  $\log(\rho_{jr}/(1-\rho_{jr})) = \alpha_j + \beta_r + u_{jr}$  where  $\rho_{jr}$  denote the share of people moving from  $j$  to  $r$  and  $\alpha_j$  and  $\beta_r$  denote origin and destination fixed effects. We run the regression both in the model (orange dots) and in the data (grey diamonds) and then plot the average  $\hat{u}_{jr}$  by distance percentile.

Figure 3: Lifetime state-to-state migration: Model vs Data

it is only calibrated to match the average rate of mobility. The simple correlation between the share of stayers in the model and in the data is 0.3. In the right panel, we show that the model also matches the distance gradient of moving flows. In particular, we run a gravity-type regression of moving flows and compare the model outcomes to those of the data. We consider the specification

$$\log \frac{\rho_{jr}}{1 - \rho_{jj}} = \alpha_j + \beta_r + u_{jr},$$

where  $\alpha_j$  and  $\beta_r$  are origin and destination fixed effects. We then plot the estimated residual  $\hat{u}_{jr}$  as a function of distance for the actual data and the data stemming from the model. The right panel of Figure 3 shows that the model captures this systematic pattern of lifetime migration flows well. In particular, spatial mobility is very local, i.e. spatial flows are steeply decreasing in distance.

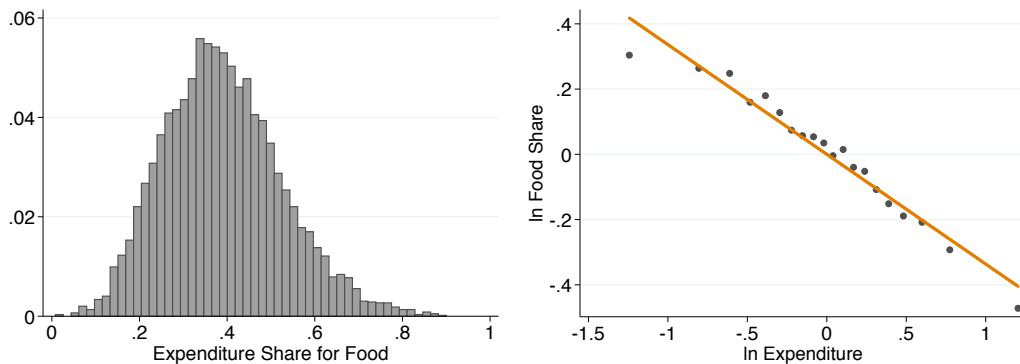
Bivariate correlations		
$\rho(\ln Q_{NA,r}, \ln Q_{A,r})$	$\rho(\ln Q_{NA,r}, A_r)$	$\rho(\ln Q_{A,r}, A_r)$
0.225	0.492	0.152

Notes: This table reports the cross-sectional correlation between the spatial fundamentals. We report the average of the correlations for the years 1880, 1910, 1940, 1970 and 2000.

Table 3: Correlation of spatial fundamentals

An integral part of our calibration strategy is that we calibrate the cross-sectional distribution of sectoral productivities  $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$  and amenities  $\{A_{rt}\}_{rt}$  as structural residuals of the model in line with the new quantitative spatial economics literature (Redding and Rossi-Hansberg, 2017). In Table 3 we report the cross-sectional bivariate correlations between these fundamentals. We find that productive regions have an absolute advantage in *both* sectors - the cross-sectional correlation between  $\ln Q_{rA}$  and  $\ln Q_{rNA}$  is about 0.225. Similarly, the correlation of productivity and amenities is also positive, in particular for non-agricultural productivity. Hence, productive places are *also* relatively pleasant places to live. This

is in line with recent direct evidence for developing countries by [Gollin et al. \(2017\)](#). In [Section B.8](#) in the Appendix, we provide additional details about these estimates of spatial fundamentals. In particular, we also relate these model-based measures of regional productivity,  $Q_{r,t}$ , to direct empirical measures of local productivity growth. More specifically, we use data on changes in regional market access due to the expansion of the railroad network from [Donaldson and Hornbeck \(2016\)](#) and show these are highly correlated with *changes* in  $\{Q_{rA}, Q_{rNA}\}_r$  as inferred from our model. See [Section B.9](#) in the Appendix for details.



Notes: The figure shows the cross-sectional distribution of the individual expenditures shares on food (left panel) and the bin scattered relationship between the (log) expenditure share on food and (log) total expenditure (right panel). The relationship in the right panel is conditional on a set of location and family size fixed effects. The underlying data stems from the Data Source: Study of Consumer Purchases in the United States, 1935-1936.

Figure 4: Expenditure Shares on Food in 1936

Our model of consumer preferences also does a good job at replicating the non-homothetic structure of consumer demand. Recall that the PIGL demand system implies that expenditure shares at the individual level are given by  $\vartheta_A(e, p) = \phi + \nu \left(\frac{p_A}{p_M}\right)^\gamma e^{-\eta}$ . For  $\phi \approx 0$ , this implies that there is a log-linear relationship between expenditure shares and total expenditure. In the left panel of [Figure 4](#) we depict the cross-sectional distribution of the expenditure share for food across US households in 1935. It is apparent there is substantial heterogeneity and that a large fraction of households has food shares exceeding 40%. In the right panel we depict the binned scatter plot between (the log of) expenditures and expenditure shares after taking out a set of regional fixed effects, to control for relative prices. The slope of the regression line is exactly the extent of the demand non-homotheticity  $\eta$ . The expenditure share is not only systematically declining in the level of expenditure but the cross-sectional relationship is essentially log-linear, as predicted by the theory. The slope coefficient implies that  $\eta = 0.32$ .<sup>30</sup>

<sup>30</sup>For our estimate, we of course do not impose the restriction that  $\phi = 0$  and estimate the demand function using non-linear least squares. The parameter  $\eta$  is precisely estimated and - depending on the specification - between 0.3 and 0.34.

## 5 Causes and Consequences of Spatial Structural Change

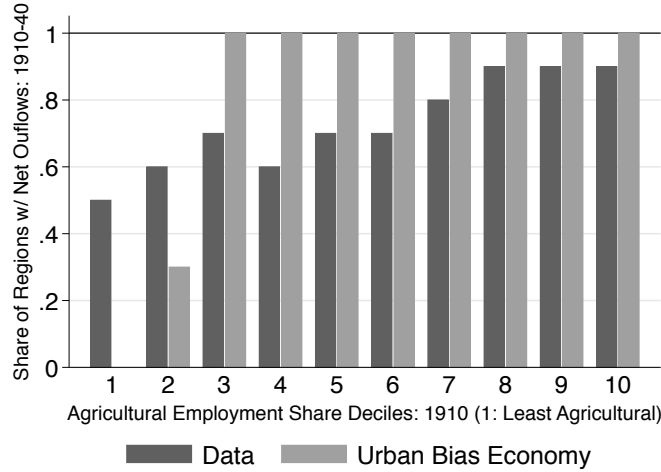
Using our calibrated model we can now turn to the causes and consequences of spatial structural change. In Section 5.1 we discuss why the urban bias of the structural transformation did not cause more outmigration from agricultural areas. In Section 5.2 we ask whether the observed patterns of spatial mobility had important implications for aggregate productivity and the spatial distribution of welfare.

### 5.1 Causes: The Urban Bias and the Pattern of Spatial Reallocation

In Section 4.2 we showed that the mechanism of the urban bias has empirical content: the process of structural change did reduce relative wages in agriculturally specialized labor markets and hence was a secular force towards spatial reallocation. However, we also showed that agricultural employment shares and future earnings were only imperfectly correlated. This tends to weaken the link between population outflows and agricultural specialization as migrants might move *towards* agricultural localities in their search for higher wages. Hence, there were important offsetting factors counteracting the urban bias of the structural transformation.

Our model suggests four main margins that could counteract the force of the urban bias channel for the process of spatial reallocation. At the individual level, moving costs between labor markets reduce the level of migration and idiosyncratic motives in individual moving decisions weaken the correlation between population net flows and regional factor prices. Because agricultural intensive regions have (on average) low factor prices, both channels reduce the importance of spatial reallocation for the aggregate decline in agricultural employment. At the regional level, both shocks to spatial fundamentals (i.e. productivities and amenities) and a positive correlation between agricultural comparative advantage ( $Q_{rA}/Q_{rNA}$ ) and the level of productivity ( $Q_{rNA}$ ) will reduce the correlation between population outflows and agricultural employment, as they counteract the secular decline of labor demand in agricultural areas. To understand the relative importance of these channels, we compare the reallocation patterns of the calibrated model, which naturally features all channels, with an economy, where *only* the the urban bias channel is present. We refer to this parametrization as the “Urban Bias Economy”. We then quantify the relative importance of individual counteracting channels by considering various alternative parametrization which shut them off one at a time.

**The Urban Bias Economy** We define the Urban Bias Economy as being characterized by (i) the absence of moving frictions ( $MC_{jr} = 0$ ), (ii) no *changes* in regional fundamentals ( $Q_{rst} = Q_{rs1880}$ ), (iii) no amenities ( $A_{rt} = 0$ ) and (iv) no idiosyncratic preferences for particular locations ( $\kappa \rightarrow 0$ ). Note that this is the frictionless parametrization of our model introduced in Section 3 with the additional restriction that there are *no* changes in regional productivities. This makes the urban bias of the structural transformation the only driver of spatial mobility. While this model implies the same decline in the aggregate agricultural share as our baseline model, we find that it has strikingly different implications for the link between



Notes: The figure reports the share of regions within the different deciles of the agricultural employment share in 1910, which experience net population outflows between 1910 and 1940. The first (last) decile refers the commuting zones with lowest (highest) agricultural employment share. We report the results for the data (dark grey) and the urban bias economy, which has no moving costs ( $MC_{jr} = 0$ ), (ii) no changes in regional fundamentals ( $Q_{rst} = Q_{rs1880}$ ), (iii) no amenities ( $A_{rt} = 0$ ) and (iv) no idiosyncratic preferences for particular locations ( $\kappa \rightarrow 0$ ).

Figure 5: Agricultural specialization and Population Outflows

initial agricultural employment shares and subsequent population flows.

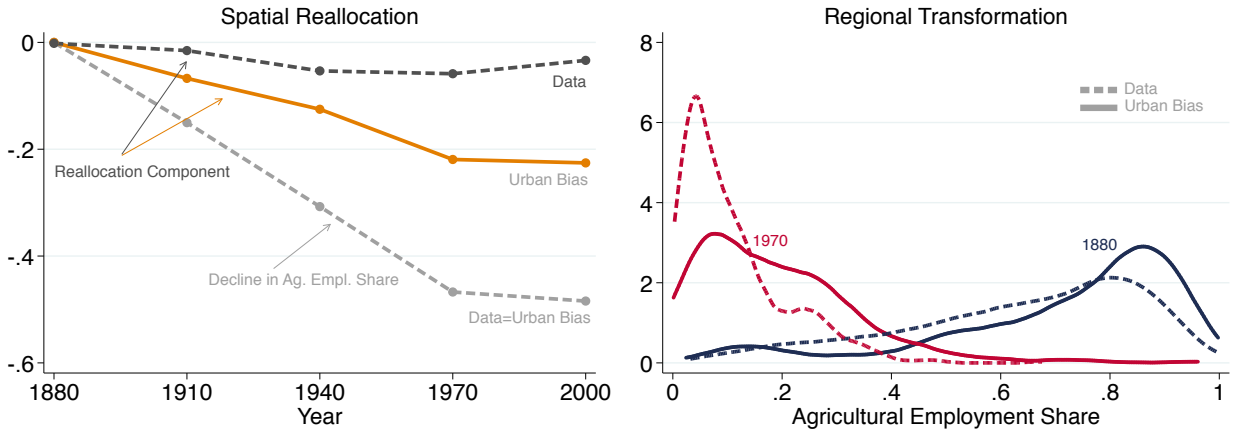
In Figure 5 we depict the share of commuting zones experiencing population outflows between 1910 and 1940 within different deciles of their agricultural employment share in 1910. While the data shows a negative correlation between outflows and initial agricultural specialization, the relationship is noisy: even among the set of the regions with lowest agricultural employment shares in 1910, about 50% experience net population outflows.<sup>31</sup> This is very different for the “urban bias” economy, where agricultural specialization and population outflows are perfectly aligned. In particular, this model implies that the *only* regions experiencing population inflows are the 17% of commuting zones with the lowest agricultural share in 1910.<sup>32</sup> Hence, if only the urban bias mechanism had been at play, the structural transformation would have induced much more population growth in non-agricultural, urban localities.<sup>33</sup>

As a result, in the Urban Bias Economy, the spatial reallocation of individuals explains about one third of the aggregate decline in agricultural employment - see Figure 6. Conversely, while the Urban Bias Economy *overestimates* the role of spatial reallocation, it *underestimates* the extent of the structural transformation at the local level. In particular, there are more regions which remain dominated by the agricultural sector throughout the 20th century. Consider, for example, the two red densities on the far left, which corresponds to 1970. In the data (the dashed line), the vast majority of commuting zones have

<sup>31</sup>Note that Figure 5 reports the share of regions with net outflows, i.e. the extensive margin of population flows. In the aggregate, net population flows are by construction zero.

<sup>32</sup>This can be seen in Figure 5: all regions in the first decile and about 70% of the second decile experience net population outflows, implying 17% of all regions do so.

<sup>33</sup>Note that this strong form of sorting is implied by (21): only counties with  $s_{rAt} < s_{At}$  see their population increase. And because population size and agricultural employment shares are strongly negatively correlated, far less than half of the regions are predicted to experience population inflows.



Notes: In the left panel we show the aggregate agricultural employment share (light grey dashed line) and the predicted agricultural share holdings regional agricultural shares at their 1880 level in the urban bias economy (solid orange line). This line is calculated as  $\sum_r s_{rA1880}^M l_{rt}^M$ , where  $s_{rAt}^M$  and  $l_{rt}^M$  are the agricultural employment share and the population share of region  $r$  at time  $t$  in the urban bias economy. For comparison we also depict this object from the data (dark grey dashed line). In the right panel we show the cross-sectional distribution of agricultural employment shares between 1910 and 1970, both in the urban bias economy (solid lines) and in the data (dashed lines).

Figure 6: Spatial Reallocation and the Urban Bias

an agricultural employment share below 20%. In the model, there is a substantial number of regions, where 30%-40% of the workforce is still employed in the agricultural sector. The reason for this is that worker mobility and the local structural transformation are substitutes - the easier it is to reallocate people across space, the more regional specialization can be sustained throughout the structural transformation. We summarize these differences between the Urban Bias Economy and our baseline calibration in the first two rows in Table 4. As seen in Figure 6, the spatial reallocation component is equal to 15% and hence five times as large if only the urban bias is at play. Importantly, the net reallocation rate across labor markets is not substantially different. If only the urban bias channel had been at play, the net migration rate would have been 14%, which is almost identical to the one observed in the data. Hence, in principle, the level of net spatial reallocation seen in the US was sufficient to account for a large share of the aggregate decline in agricultural employment. This suggests that the presence of moving frictions cannot be the main counteracting force of the urban bias channel in the data; this is exactly what we find quantitatively in the next subsection.

One important difference between our baseline model and the Urban Bias Economy is the existence of idiosyncratic preference shocks for particular locations. The reason why our model infers that idiosyncratic shocks are empirically important is the prevalence of *bi-directional flows* in the data. If idiosyncratic shocks were absent, all individuals would agree on the ranking of potential destinations and the gross and net flows between any pair of locations would coincide. The “turnover” rate, i.e. the ratio of the gross and net migration rate, is therefore a measure of the importance of idiosyncratic migration motives. It can be interpreted as the number of bodies that have to be moved across space to reallocate one person “on net” between regions. The last column of Table 4 shows that the turnover rate in the data and in our baseline model is around 3. This is in sharp contrast to the Urban Bias Economy, where this ration is 1



	<i>Model ingredients</i>				<i>Migration patterns</i>		
	Spatial productivity shocks	Moving costs	Amenities	Idiosync. location preferences	Spatial Reallocation Component	Net migration rate	Turnover (gross/net)
<i>The Urban Bias of Structural Change</i>							
Urban Bias Economy	✗	✗	✗	✗	-15.2%	14%	1
Baseline model	✓	✓	✓	✓	-3.2%	12.9%	2.8
<i>Decomposition: The partial effect of ...</i>							
Spatial productivity ( $Q_{rst}$ )	✓	✗	✗	✗	-3.6%	22.0%	1
	✗	✓	✓	✓	-6.5%	10.7%	3.2
Moving Costs ( $MC_{jr}$ )	✗	✓	✗	✗	0%	0%	1
	✓	✗	✓	✓	-0.6%	41.1%	2.4
Amenities ( $A_{rt}$ )	✗	✗	✓	✗	-7.0%	60.8%	1
	✓	✓	✗	✓	2.6%	6%	6.2

Notes: The table reports three outcomes for various parametrizations of our model. The spatial reallocation component is calculated as in (1), i.e. is given by  $\sum_r s_{rA1880} (l_{r2000} - l_{r1880})$ . The net migration rate is the average net migration rate across US states for the years 1880 - 2000. The turn over rate is the gross migration rates relative to the net migration rate. Both rates are measured at the state level. The first two rows contain the results for the Urban Bias Economy and the baseline model. The Urban Bias Economy abstracts from spatial productivity shocks ( $Q_{rst} = Q_{rs1880}$ ), moving costs ( $MC_{jr} = 0$ ), regional amenities ( $A_{rt} = 0$ ) and idiosyncratic taste shocks (i.e.  $\kappa \rightarrow 0$ ). In the remaining rows we measure the partial effect spatial productivity shocks (rows 3 and 4), moving costs (rows 5 and 6) and regional amenities (rows 7 and 8).

Table 4: Decomposing the Spatial Reallocation Component

as gross and net flows coincide.<sup>34</sup>

**Offsetting Factors** To better understand which aspect of our theory was the most important to counteract the force of the urban bias, we now provide a formal, model-based decomposition of the partial effects of regional productivity shocks, moving frictions and regional amenities. Because these ingredients interact non-linearly, we quantify the partial effects by calculating both the consequences of adding the respective ingredient to the Urban Bias Economy and of removing it from our baseline model. The results are contained in the lower part of Table 4.

Rows three and four show that the evolution of spatial productivity was the main reason for the low correlation between agricultural specialization and population outflows. The frictionless model introduced in Section 3, which is exactly the Urban Bias Economy augmented by the observed process of spatial productivity shocks, implies a spatial reallocation component of -3.6%. This is almost exactly the same number as observed in the data. Conversely, if the structure of regional absolute and comparative advantage,  $Q_{rst}$ , had been fixed at its level in 1880, the impact of spatial reallocation would have doubled from 3.2% to 6.5%. The reason is that - as seen in equation (21) - regional productivity shocks introduce noise in the relationship between earnings growth and initial agricultural specialization and hence weaken the quantitative importance of the urban bias.

In contrast, moving costs and regional amenities play less of a role. While adding moving costs to the Urban Bias Economy does reduce the role of spatial reallocation, it does so for the wrong reason: the

<sup>34</sup>To be precise, in the absence of moving costs, the equilibrium allocation of net flows is unique, but the gross flows are not determined. This indeterminacy vanishes for an arbitrarily small moving cost. In that case, the net migration rate is equal to the gross migration rate.

implied *level* of migration is very small. More interestingly, row six shows that without moving costs the spatial reallocation component would be even *lower* - 0.6% compared to 3.2% in the data. The reason is that a reduction in moving costs would not only induce people to move towards regions with high earnings but also trigger more mobility for idiosyncratic reasons. The latter moves are by construction uncorrelated with agricultural specialization and hence do not contribute to the spatial reallocation component of structural change. Similarly, regional amenities cannot explain the quasi-absence of the spatial reallocation component in the data. In fact, the last row of Table 4 shows that the role of spatial reallocation would have been even smaller in the absence of amenities. Since rural regions have (on average) low future amenities (see Table 3), the amenity channel provides an additional push factor out of agricultural labor markets. This is consistent with the direct empirical evidence reported in Gollin et al. (2017) for developing countries today.

## 5.2 Consequences: Aggregate Productivity and Spatial Welfare

We now turn to the macroeconomic implications of spatial structural change. We focus on two aggregate outcomes: aggregate productivity and the spatial distribution of welfare. The urban bias of the structural transformation systematically changes the marginal product of labor across labor markets. If and how fast workers reallocate spatially therefore has important implications for allocative efficiency and hence aggregate productivity of the US economy. Furthermore, in the presence of spatial frictions, this secular demand shift also has distributional consequence by lowering relative wages in agricultural regions. Labor mobility therefore tends to both increase allocative efficiency and reduce spatial inequality.

To quantify the importance of spatial mobility, we compare our baseline model with three alternative parameterizations. First, we consider a model without any labor mobility by making the costs of moving prohibitively high. Secondly, we consider the frictionless model introduced in Section 3. In that model spatial sorting is costless and based only on local productivity and aggregate demand conditions. More specifically, we assume that both moving costs are absent ( $MC_{jr} = 0$ ) and agents do not derive utility from non-pecuniary features of particular locations ( $A_r = 0$  and  $\kappa = 0$ ), so that spatial sorting is only based on expected earnings (see also row 3 in Table 4). Finally, we consider the Urban Bias Economy, which further abstracts from regional productivity changes, i.e. assumes that  $Q_{rst} = Q_{rs1880}$ .

**Spatial Structural Change and Aggregate Productivity** We report the outcomes of these three different models in Table 5. In the first row, we verify that all models generate a similar decline in the aggregate agricultural share. Hence, the aggregate “size” of the structural transformation is constant across specifications. For convenience we again report the average amount of spatial reallocation (as measured by the net migration rate) and the spatial reallocation component of the structural transformation in rows two and three.

	<i>Baseline</i>	<i>Counterfactuals</i>		
		No Spatial Mobility	Frictionless Mobility	Urban Bias
Decline in Agricultural Employment Share	-48.1%	-47.8%	-48.2%	
Net migration rate	13%	0	22%	14%
Spatial Reallocation Component	-3.2%	0	-3.6	-15.2%
Increase in GDP pc (rel. to 1880)	10.78	8.89	11.09	
<i>Difference to baseline</i>		-17.57%	2.87%	
Increase in TFP (rel. to 1880)	4.88	4.27	4.99	
<i>Difference to baseline</i>		-12.59%	2.22%	
Agricultural productivity Gap	0.83	0.73	0.83	

Notes: The model with no spatial mobility corresponds to the baseline model except that individuals are not allowed to move (i.e.  $MC_{jr} \rightarrow \infty$ ). The model without spatial frictions corresponds to the baseline model except that moving costs are zero (i.e.  $MC_{jr} = 0$ ), that there are no amenities ( $A_r = 0$ ) and no idiosyncratic tastes (i.e.  $\kappa = 0$ ). The Urban Bias Economy is like the frictionless model but also assumes the regional technologies are constant, i.e.  $Q_{rst} = Q_{rs1880}$ . The net migration rate is the average migration rate between 1910 and 2000. The spatial reallocation component is calculated as in (1), i.e. is given by  $\sum_r s_{rA1880} (l_{r2000} - l_{r1880})$ . TFP is measured as  $TFP_t = Y_t / (K_t^\alpha L_t^{1-\alpha})$ . The ‘‘Agricultural Productivity Gap’’ is measured as aggregate value added per worker in agriculture relative the rest of the economy, i.e.  $(VA_t^A / L_t^A) / (VA_t / L_t)$ , in the year 2000

Table 5: Spatial Reallocation and Aggregate Productivity during the Structural Transformation

In the lower panel, we report the implications for GDP growth and aggregate TFP.<sup>35</sup> Spatial mobility contributed heavily to aggregate productivity growth in the US during the 20th century. Without labor mobility, income per capita would have been lower by almost 17.5% and this decline is entirely accounted for by lower TFP, i.e. more misallocation.<sup>36</sup> If, on the other hand, spatial mobility was only based on regional earnings and mobility was free, aggregate productivity and income per capita would have been higher by about 2.2% and 3% respectively. Moreover, as long as mobility is free, these gains do not crucially depend on whether we allow for idiosyncratic spatial productivity growth. Compared to the losses from the absence of spatial mobility, these numbers are relatively small.<sup>37</sup>

As discussed at the end of Section 3.2, these spatial productivity gains are also reflected in sectoral productivity gaps. Both the baseline model and the models without spatial frictions, for example, imply an agricultural productivity gap of around 17% (Gollin et al., 2014). Hence, value added per worker in agriculture is about 17% lower than for the average worker even though there are no sectoral frictions. These productivity differences are due to skill-based sorting, whereby high-skilled individuals both work in the non-agricultural sector within regions and move towards urban locations spatially. Without spatial mobility this gap would increase to 27% as the structural transformation amplifies wage differences

<sup>35</sup>We calculate TFP from the perspective of an aggregate production function, i.e.  $TFP_t = Y_t / (K_t^\alpha L_t^{1-\alpha})$ .

<sup>36</sup>Income per capita is not proportional to TFP because of capital accumulation. As the capital-output ratio is essentially constant in all parametrization, we have that  $Y \propto TFP^{1-\alpha}$ .

<sup>37</sup>To put these results in perspective, Bryan and Morten (2017), using a static spatial equilibrium model, estimate higher productivity losses from limited labor mobility for Indonesia. In their specifications without agglomeration and endogenous amenities, which we for simplicity abstract from in our model, they find that productivity would increase by about 17% if moving costs were zero and there were no amenity differences. There are two main reasons, why our results are small. First of all, by explicitly considering capital as a factor of production (which can be traded without frictions) our economy allows for some factor other than labor to adjust. Secondly, Bryan and Morten (2017) argue that the US economy has lower mobility costs than Indonesia.

between agricultural and non-agricultural regions.

**The Distributional Effects of Spatial Structural Change** We now turn to the evolution of spatial welfare inequality. We focus on the expected life-time value of being in location  $r$ ,  $\mathcal{W}_{rt}^h = E^h [U_r] + A_{rt}$ , as our measure of regional welfare for skill group  $h$  (see (12)). Our model implies that both regional income  $\Theta_r^h$  and regional amenities are negatively correlated with the agricultural employment share so that welfare is systematically lower in agriculturally intensive places. Moreover, the urban bias of the structural transformation is an additional inequality-enhancing force. To quantify the evolution of spatial welfare inequality during the 20th century, we convert utility differences into “life-time-income” equivalents. Specifically, let  $\Delta_t^h$  denote the interquartile range of  $\mathcal{W}_{rt}^h$  at time  $t$ . Let  $T_{rt}^h$  be the increase of expected lifetime income, an individual with skill  $h$  in region  $r$  requires to increase the average utility of living in region  $r$  by  $\Delta_t^h$ , i.e.

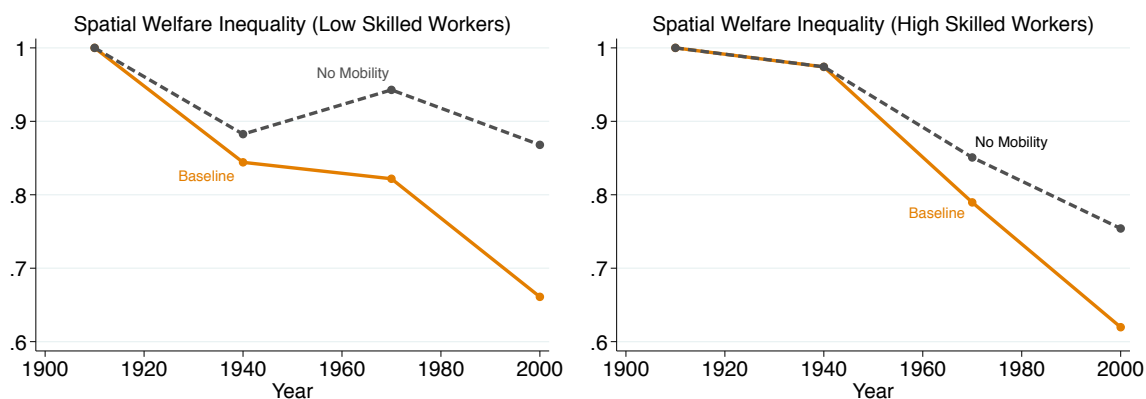
$$\mathcal{W}_{rt}^h (T_{rt}^h y) = \mathcal{W}_{rt}^h (y) + \Delta_t^h.$$

Our measure of spatial inequality for individuals of skill  $h$  is then given as the cross-regional average of  $T_{rt}^h$ , i.e.  $T_t^h = \frac{1}{R} \sum T_{rt}^h$ .

In Figure 7, we depict the evolution of spatial welfare inequality, relative to the level in 1910 for both low skilled and high skilled individuals. The solid lines refer to our baseline calibration, the dashed lines correspond to the economy without spatial mobility.<sup>38</sup> Figure 7 shows that spatial welfare inequality decreased substantially during the 20th century. The required increase in lifetime income to raise utility from the 25% quantile to the 75% quantile of the spatial distribution of welfare decreased by about 35% since 1910. In terms of levels, in 1910 the average region would have required an increase in average lifetime income by about 160% to increase utility by the interquartile range of regional welfare differences. In 2000, lifetime income would only have to be doubled.<sup>39</sup>

<sup>38</sup>Naturally, we do not display the two specifications without moving costs, as welfare inequality is by construction zero.

<sup>39</sup>The decline in inequality is not driven by changes in the distribution of regional amenities. The dispersion in amenities is roughly constant between 1910 and 2000 and spatial welfare differences had declined even in the absence of amenities.



Notes: The figure shows the average increase in life-time income required to increase utility from the 25% quantile to the 75% quantile of the cross-sectional welfare distribution at time  $t$  relative to 1910. In the left (right) panel we depict the results for low (high) skilled workers. See Section B.10 in the Appendix for details.

Figure 7: The Evolution of Spatial Welfare Inequality

Figure 7 also shows that spatial mobility was a crucial factor for this decline in welfare inequality. Without labor mobility to prevent the regional distribution of income from fanning out, spatial welfare would have declined much less. This margin was a particularly important adjustment mechanism for low-skilled workers as these workers have a comparative advantage in the agricultural sector and hence are particularly exposed to the structural transformation. Spatial mobility was therefore a crucial adjustment mechanism for low-skilled workers to weather the first structural transformation away from agriculture. Without it, spatial welfare inequality would have been substantially higher.

## 6 Robustness

In this section we demonstrate the robustness of our results to a range of alternative parameterizations of the model. We focus on the spatial reallocation component of the structural transformation (Table 4) and the effects of spatial mobility on aggregate GDP (Table 5). In Section D.8 of the Online Appendix we provide more details.

The results are summarized in Table 6. For the spatial reallocation components we again report the implications of the Urban Bias Economy and the decomposition into the individual margins by removing the respective ingredient from our baseline calibration. For the implications for aggregate GDP, we report the change of GDP in 2000 for the economy without mobility and the frictionless economy relative to the baseline. For all specifications, we always recalibrate the spatial structural residuals and aggregate technology series.<sup>40</sup>

<sup>40</sup>Specifically, we always match the time series data on relative prices and aggregate GDP and the spatial data on earnings, population size and agricultural employment shares.

	Baseline	Elasticity of Subst.		Skill heterogeneity		Land	Agglomeration		Congestion
		$\sigma = 3$	$\sigma = 8$	$\zeta = 1.4$	$\zeta = 4$	$\omega_A = -0.1$	$\omega_{NA} = 0.03$	$\omega_{NA} = \omega_A = 0.03$	$\iota_1 = 0.86$ $\iota_2 = 0.19$
<i>Spatial Reallocation Component</i>									
Calibrated	-3.16%	-3.16%	-3.15%	-3.13%	-3.20%	-3.16%	-3.15%	-3.15%	-3.18%
Urban Bias	-15.22%	-14.02%	-17.21%	-15.41%	-13.36%	-15.17%	-15.71%	-15.73%	-14.13%
Productivity	-6.55%	-6.34%	-6.98%	-6.64%	-5.50%	-6.55%	-6.59%	-6.59%	-6.42%
Moving Costs	-0.65%	-0.37%	-1.58%	-0.76%	-0.47%	-0.64%	-0.78%	-0.79%	0.132%
Amenities	2.63%	2.05%	3.86%	3.02%	1.97%	2.61%	2.84%	2.85%	2.13%
<i>Change in GDP in 2000</i>									
No Mobility	-17.56%	-21.49%	-11.32%	-17.62%	-17.20%	-17.56%	-17.56%	-17.56%	-17.44%
Frictionless	2.86%	2.29%	5.01%	2.69%	3.53%	2.85%	3.20%	3.20%	2.82%

Notes: The table contains various robustness exercises. For all specifications, we always recalibrate our baseline model. For the “congestion” case we follow [Adao et al. \(2018\)](#) in using a model implied optimal IV approach to estimate  $\iota_1$  and  $\iota_2$ , using the regional incidence of aggregate trends as an instrument for local population inflows. See Section D.9 of the Online Appendix for details on this strategy. For more additional robustness results and their discussion in turn, see Section D.8 of the Online Appendix.

Table 6: Robustness

In the first four columns we focus on the elasticity of substitution  $\sigma$  and the dispersion of individual skills  $\zeta$ . These parameters determine the extent to which demand declines if regional prices change ( $\sigma$ ) and the elasticity of sectoral labor supply within locations ( $\zeta$ ). While the value of  $\sigma$  is quantitatively important, in particular for the implications for aggregate productivity, the results are qualitatively similar to our baseline results. In contrast, our results are essentially insensitive to the precise value of  $\zeta$ .

In the last four columns, we consider the case of endogenous spatial fundamentals  $Q_{rst}$  and  $A_{rst}$ . In particular, we assume that  $Q_{rst} = \tilde{Q}_{rst} L_{rt}^{\omega_s}$  and  $A_{rt} = \tilde{A}_{rt} - \iota_1 L_{rt}^{\iota_2}$ , where  $\tilde{Q}_{rst}$  and  $\tilde{A}_{rt}$  are exogenous. As we discuss more formally in Section D.8 of the Online Appendix, the case of  $\omega_s < 0$  can be thought of capturing decreasing returns in sector  $s$ . This would for example be the case if land was a factor of production and in fixed supply. Similarly,  $\omega_s > 0$  captures the existence of agglomeration benefits. Finally,  $\iota > 0$  allows for congestion forces in location amenities. In terms of agglomeration, we assume an elasticity of 0.03, which is the preferred value in [Bryan and Morten \(2017\)](#). To estimate the congestion parameters  $(\iota_1, \iota_2)$ , we follow [Adao et al. \(2018\)](#) and use a model implied optimal IV approach by exploiting the urban bias as an instrument for local population inflows. See Section D.9 of the Online Appendix for details on this strategy. The last four columns of Table 6 again show that our results are quantitatively robust to such considerations.

## 7 Conclusion

The structural transformation, i.e. the systematic reallocation of employment out of the agricultural sector, is a key feature of long-run economic growth. This sectoral bias of the growth process naturally affects the spatial allocation of economic activity. In particular, by shifting expenditure away from the agricultural sector, the structural transformation benefits urban, non-agricultural labor markets and hurts rural ones. This urban bias of the structural transformation raised the return to relocate workers spatially. In this paper, we use a novel theory of spatial structural change and detailed regional data to analyze the spatial nature of the structural transformation of the US between 1880 and 2000.

We first document empirically that the process of the spatial reallocation of workers towards non-agricultural labor markets explained essentially none of the aggregate decline in the agricultural employment share from 50% to nearly zero over the last 120 years. In contrast, the entire decline is accounted for by within labor market changes, whereby agricultural employment declines in each locality. While these patterns seem to contradict the large increase in urbanization over the same time-period, we show that this is not the case: like the change in agricultural employment, the increase in the share of urban dwellers was also very local in nature.

To explain this fact and to understand whether this mode of adjustment had important aggregate implications, we construct a new quantitative theory of the structural transformation that explicitly incorporates a spatial layer. The model combines the basic features of an economic geography model featuring costly labor mobility with the canonical ingredients of neoclassical models of structural change, i.e. non-homothetic preferences, unbalanced technological progress and aggregate capital accumulation. Despite this richness, we show that the analysis remains highly tractable and can be applied to a realistic geography.

Our analysis yields two main results. First, we show that the evolution of spatial productivity was the main reason for the insignificance of the spatial reallocation channel. Because local productivity is subject to shocks and regional absolute and comparative advantage are imperfectly correlated, the relationship between agricultural specialization and future earnings (and hence net population outflows) is only weakly negative - *despite* the urban bias of the structural transformation. Secondly, we show that the possibility of spatial reallocation had important macroeconomic consequences. On the one hand, the process of spatial arbitrage was an important contributor of aggregate productivity growth in the US. Without labor mobility, aggregate income would have been 17% smaller. On the other hand, it played a crucial role for the evolution of regional welfare inequality during the structural transformation of the US. In the absence of migration across labor markets, welfare inequality would have been markedly higher, in particular among unskilled workers, which were especially exposed to the secular demand shift away from agriculture.

## References

- Acemoglu, Daron and Veronica Guerrieri**, “Capital Deepening and Nonbalanced Economic Growth,” *Journal of Political Economy*, 2008, 116 (3), 467–498.
- Adao, Rodrigo, Costas Arkolakis, and Federico Esposito**, “Trade, Agglomeration Effects, and Labor Markets: Theory and Evidence,” 2018.
- Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf**, “The Economics of Density: Evidence From the Berlin Wall,” *Econometrica*, 2015, 83 (6), 2127–2189.

- Alder, Simon, Timo Boppart, and Andreas Muller**, “A Theory of Structural Change that Can Fit the Data,” 2018.
- Allen, Treb and Costas Arkolakis**, “Trade and the Topography of the Spatial Economy,” *The Quarterly Journal of Economics*, 2014, 129 (3), 1085–1140.
- Alvarez-Cuadrado, Francisco and Markus Poschke**, “Structural Change out of Agriculture: Labor Push versus Labor Pull,” *American Economic Journal: Macroeconomics*, 2011, 3 (3), 127–158.
- , **Ngo Van Long, and Markus Poschke**, “Capital Labor Substitution, Structural Change, and Growth,” *Theoretical Economics*, 2017, 12 (3), 1229–1266.
- Baumol, William J**, “Macroeconomics of Unbalanced Growth: the Anatomy of Urban Crisis,” *American Economic Review*, 1967, 57 (3), 415–426.
- Boppart, Timo**, “Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non-Gorman Preferences,” *Econometrica*, 2014, 82 (6), 2167–2196.
- Bryan, Gharad and Melanie Morten**, “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia,” 2017.
- , **Shyamal Chowdhury, and Ahmed Mushfiq Mobarak**, “Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh,” *Econometrica*, 2014, 82 (5), 1671–1748.
- Buera, Francisco J and Joseph P Kaboski**, “Can Traditional Theories of Structural Change Fit the Data?,” *Journal of the European Economic Association*, 2009, 7 (2-3), 469–477.
- Caselli, Francesco and Wilbur John Coleman II**, “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation,” *Journal of Political Economy*, 2001, 109 (3), 584–616.
- Chenery, Hollis B**, “Patterns of Industrial Growth,” *American Economic Review*, 1960, 50 (4), 624–654.
- Comin, Diego, Danial Lashkari, and Marti Mestieri**, “Structural Change with Long-Run Income and Price Effects,” 2017.
- Desmet, Klaus and Esteban Rossi-Hansberg**, “Spatial Development,” *American Economic Review*, 2014, 104 (4), 1211–43.
- Diamond, Rebecca**, “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000,” *American Economic Review*, March 2016, 106 (3), 479–524.
- Donaldson, Dave and Richard Hornbeck**, “Railroads and American Economic Growth: A Market Access Approach,” *The Quarterly Journal of Economics*, 2016, 131 (2), 799–858.



**Eckert, Fabian, Andres Gvirtz, and Michael Peters**, “Constructing a Population Panel for US Commuting Zones 1790-2000,” 2018.

**Fajgelbaum, Pablo and Stephen J. Redding**, “External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914,” Working Paper 20217, National Bureau of Economic Research June 2014.

**Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suarez Serrato, and Owen M. Zidar**, “State Taxes and Spatial Misallocation,” Working Paper 21760, National Bureau of Economic Research November 2015.

**Gollin, Douglas, David Lagakos, and Michael E. Waugh**, “The Agricultural Productivity Gap,” *The Quarterly Journal of Economics*, 2014, 129 (2), 939–993.

—, **Martina Kirchberger, and David Lagakos**, “In Search of a Spatial Equilibrium in the Developing World,” Working Paper 23916, National Bureau of Economic Research October 2017.

—, **Stephen Parente, and Richard Rogerson**, “The Role of Agriculture in Development,” *American Economic Review*, May 2002, 92 (2), 160–164.

**Herrendorf, Berthold and Todd Schoellman**, “Why is Measured Productivity so Low in Agriculture?,” *Review of Economic Dynamics*, 2015, 18 (4), 1003–1022.

—, **Christopher Herrington, and Akos Valentinyi**, “Sectoral Technology and Structural Transformation,” *American Economic Journal: Macroeconomics*, October 2015, 7 (4), 104–33.

—, **Richard Rogerson, and Akos Valentinyi**, “Two Perspectives on Preferences and Structural Transformation,” *American Economic Review*, December 2013, 103 (7), 2752–89.

—, —, **and** —, “Chapter 6 - Growth and Structural Transformation,” in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 2 of *Handbook of Economic Growth*, Elsevier, 2014, pp. 855 – 941.

—, —, **and** —, “Structural Change in Investment and Consumption: A Unified Approach,” 2017. Working Paper.

**Hobijn, Bart, Todd Schoellman, and Alberto J. Vindas Q.**, “Structural Transformation: A Race between Demographics and Technology,” 2018. Mimeo, Arizona State University.

**Hsieh, Chang-Tai and Enrico Moretti**, “Housing Constraints and Spatial Misallocation,” Working Paper 21154, National Bureau of Economic Research May 2015.

**Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie**, “Beyond Balanced Growth,” *The Review of Economic Studies*, 2001, 68 (4), 869–882.

- Kuznets, Simon**, “Quantitative Aspects of the Economic Growth of Nations: II. Industrial Distribution of National Product and Labor Force,” *Economic Development and Cultural Change*, 1957, pp. 1–111.
- Lagakos, David and Michael E. Waugh**, “Selection, Agriculture, and Cross-Country Productivity Differences,” *American Economic Review*, April 2013, *103* (2), 948–80.
- , **Mushfiq Mobarak, and Michael E Waugh**, “The Welfare Effects of Encouraging Rural-Urban Migration,” 2017.
- Lindbeck, Assar, ed.**, *Nobel Lectures, Economics 1969-1980*, World Scientific Publishing Co., Singapore, 1992.
- Michaels, Guy, Ferdinand Rauch, and Stephen J Redding**, “Urbanization and Structural Transformation,” *The Quarterly Journal of Economics*, 2012, *127* (2), 535–586.
- Molloy, Raven, Christopher L. Smith, and Abigail Wozniak**, “Internal Migration in the United States,” *Journal of Economic Perspectives*, September 2011, *25* (3), 173–96.
- Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg**, “Commuting, Migration and Local Employment Elasticities,” Working Paper 21706, National Bureau of Economic Research November 2015.
- Nagy, Dávid Krisztián**, “City Location and Economic Development,” 2017.
- Ngai, L. Rachel and Christopher A. Pissarides**, “Structural Change in a Multisector Model of Growth,” *American Economic Review*, March 2007, *97* (1), 429–443.
- Porzio, Tommaso and Gabriella V. Santangelo**, “Structural Change and the Supply of Agricultural Workers,” 2017.
- Redding, Stephen J. and Daniel M. Sturm**, “The Costs of Remoteness: Evidence from German Division and Reunification,” *American Economic Review*, December 2008, *98* (5), 1766–97.
- and **Esteban Rossi-Hansberg**, “Quantitative Spatial Economics,” *Annual Review of Economics*, 2017, *9* (1), 21–58.
- Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander**, “Integrated Public Use Microdata Series: Version 6.0 [dataset]. Minneapolis: University of Minnesota, 2015,” 2015.
- Tolbert, Charles M and Molly Sizer**, “US commuting zones and labor market areas: A 1990 update,” 1996.
- Tombe, Trevor, Xiaodong Zhu et al.**, “Trade, Migration and Productivity: A Quantitative Analysis of China,” 2015.

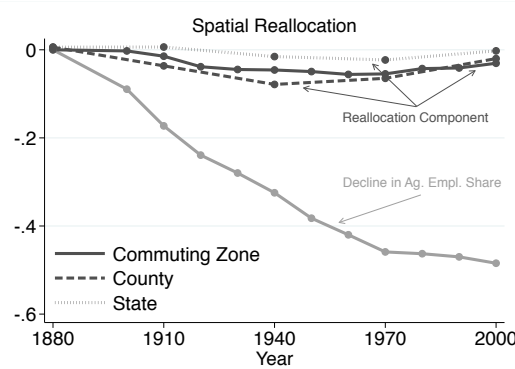
**Young, Alwyn**, “Inequality, the Urban-Rural Gap, and Migration,” *The Quarterly Journal of Economics*, 2013, 128 (4), 1727–1785.

# APPENDIX

# A Spatial Reallocation: Additional Results

In Section 2 we showed that the spatial reallocation of individuals across commuting zones cannot account for the observed decline of the agricultural employment share in the US. In this section we provide additional details for this empirical result.

**Robustness to Labor Market Definition** We first consider definitions of labor markets other than that of a commuting zone, which we use throughout the main body of the paper. Figure 8 replicates Figure 1 on the county and state level. For comparison, we also display the results at the commuting zone level. We use the same underlying data and simply aggregate it differently. Figure 8 shows that the reallocation component is quantitatively unimportant regardless which of these three labor market definitions we chose. Note that the reallocation component is zero by construction if we consider the entire US as one region. The ordering of the lines is indicative of that: for most years the reallocation component is largest for counties (the smallest level of aggregation we consider) and smallest for state (the largest level of aggregation we consider). All in all Figure 8 reinforces our result of spatial reallocation as a highly local phenomenon, which operates predominantly at the intra-county level.

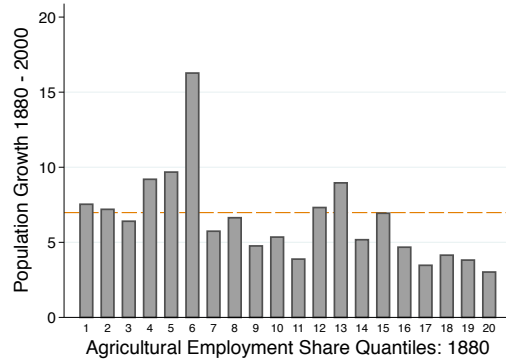


Notes: We show the aggregate agricultural employment share (light grey line) and the predicted agricultural share holding regional agricultural shares at their 1880 level, i.e.  $\sum_r s_{rA1880}(l_{r1880} - l_{rt})$ , where  $s_{rAt}$  and  $l_{rt}$  are the agricultural employment share and the population share of region  $r$  at time  $t$  (dark grey line). We do this for three different definitions of  $r$ : state, commuting zone or county. The underlying data is the same as described in the data appendix and simply aggregated differently for each of the three scenarios.

Figure 8: Spatial Reallocation Across States, Commuting Zones and Counties

**Direct Evidence for Agricultural Specialization and Population Growth** The patterns in Figure reflect the low correlation between agricultural specialization and population growth. We can test this relationship directly in a regression format. In particular, we consider a regression of the form

$$g_{rL}^{1880-2000} = \alpha + \beta s_{rA}^{1880} + u_r,$$



Notes: The figure shows the average rate of population growth between 1880 and 2000 for 20 quantiles of the agricultural employment share in 1880.

Figure 9: Agricultural Specialization in 1880 and Population growth 1880-2000

where  $g_{rL}^{1880-2000}$  denotes regional population growth between 1880 and 2000 and  $s_{rA}^{1880}$  denotes the agricultural employment share in 1880. The results are contained in Table 7.

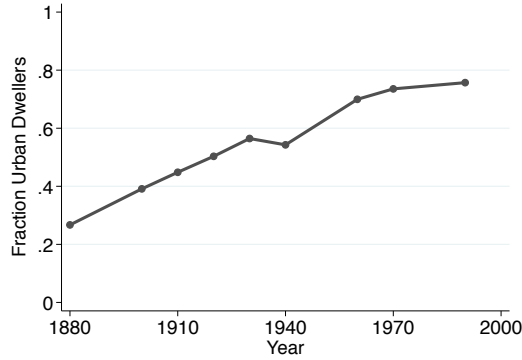
	Dep. variable: Population growth 1880 - 2000			
Agricultural share 1880	-26.241 (40.829)	-3.702* (1.917)		
log Agricultural share 1880			-0.548 (0.658)	
Ag quantile FE	No	No	No	Yes
Weights	No	Yes	Yes	Yes
Observations	717	717	717	717
$R^2$	0.000	0.002	0.000	0.014

Notes: Robust standard errors in parentheses with \*\*\*, \*\* and \* respectively denoting significance at the 1%, 5% and 10% levels. Column 4 contains a whole set of 20 fixed effects for the different quantiles of agricultural employment shares in 1880.

Table 7: Agricultural Specialization in 1880 and Population growth 1880-2000

Columns 1 to 3 show that there is no significant relationship between agricultural specialization in 1880 and population growth between 1880 and 2000. Columns 2 and 3 weigh each regression by their initial population in 1880. In column 4 we include a full set of twenty fixed effects of the initial agricultural share quantiles. While these fixed effects are jointly statistically significant, their explanatory power is still very small. Figure 9 shows this relationship graphically. More specifically, we report average population growth between 1880 and 2000 for twenty quantiles of the agricultural employment share. While population growth tends to be slightly smaller in regions with a high agricultural employment share in 1880, the relationship is not particularly strong and certainly not monotone.

**Spatial Structural Change and Urbanization** The absence of the spatial reallocation channel seems to be inconsistent with sharp increase in the rate of urbanization from 20% to almost 80% shown in Figure 10. We now show that this is not the case. In particular, we show that (like the the secular decrease in



Notes: The graph is constructed from the decennial US micro census available on IPUMS. The line graphs the fraction of Americans who live in an urban environment rather than a rural one over time. An urban environment is defined as a locality with more than 2500 inhabitants.

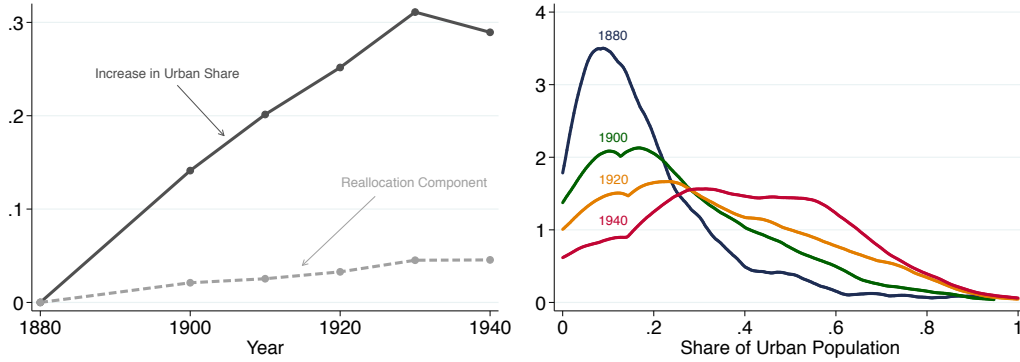
Figure 10: Fraction of US population living in urban environment

agricultural employment) the increase in urbanization is also a very local phenomenon. The share of people living in urban areas (i.e. cities with more than 2500 inhabitants) increases from just shy of 20% in 1850 to more than 50% of the population in 1940. In Figure 11, we again decompose this time-series into a within and across commuting zone component. In particular, we decompose the increase in the rate of urbanization into a spatial reallocation and a regional transformation component as in Equation 1 above, as follows:

$$u_t - u_{1880} = \underbrace{\sum_r u_{r1880}(l_{rt} - l_{r1880})}_{\text{Spatial Reallocation}} + \underbrace{\sum_r (u_{rt} - u_{r1880})l_{rt}}_{\text{Regional Transformation}}, \quad (23)$$

where  $u_{r1880}$  is the urbanization rate in commuting zone  $r$  in 1880.<sup>41</sup> If the increase in urbanization resulted from individuals migrating into highly-urbanized commuting zones, the spatial reallocation component would be close to the actual time series.

<sup>41</sup>Note that this is strictly speaking the negative of Equation 1, this is more convenient here, since, unlike for the aggregate agricultural employment share, the aggregate urbanization share follows a secular positive trend.



Notes: In the left panel, we again show the absolute increase in the share of people living in urban areas and its reallocation component calculated according to Equation 23 as  $\sum_r u_{r1880}(l_{rt} - l_{r1880})$ . In the right panel we show the cross-sectional distribution of the share of the urban population across commuting zones. This figure is based on US Micro Census data from Ruggles et al. (2015). County-level data is only publicly available for the 1880-1940 period and restricting the computation of the counterfactual urbanization share to this period. Figure 10 shows the full time series of the aggregate urban share.

Figure 11: Urbanization within commuting zones

The left panel of Figure 11 shows that this is not the case: as for the agricultural employment share, the cross-commuting zone population flows explain only a minor share of actual increase observed in the data. Similarly to the right panel of 1, the right panel of Figure 11 shows the distribution of the share of the urban population across commuting zones conditional on this share being positive. As for the patterns of agricultural employment depicted in Figure 1 these densities shift to the right, indicating that urbanization takes place within all counties in the US.

**Spatial Reallocation and Structural Change around the World** Finally, we ask whether the insignificance of the across labor market reallocation component of structural change is particular to the United States or a more general feature of the structural transformation. Indeed a look at other countries around the world suggests that the highly localized nature of reallocation and urbanization is a feature of the structural transformation in many countries around the world. To see this, we used data from IPUMS (see Ruggles et al. (2015)) to compute the reallocation component of the decline in the aggregate agricultural share for seven additional countries. The labor market regions available for these countries are generally larger than commuting zones in the United States, but smaller than US states. The results are reported in Table 8. In general we find that the spatial reallocation component is very small. Together with Figure 8 for the United States, we view this quasi-absence of the reallocation component at the sub-state resolution as suggestive evidence for the local nature of the structural transformation more generally.



Country	USA	Argentina	China	India	Mali	Mexico	Spain	Venezuela
Time Period	1880-2000	1970-2001	1982-2000	1987-2009	1987-2009	1970-2010	1991-2011	1981-2001
Number of Regions	712	312	198	413	47	2316	52	157
$\Delta$ Ag. Empl. Share	-0.48	-.076	-.10	-.09	-.12	-.29	-.12	-.02
$\Delta$ implied by reallocation	-0.03	0	0	-0.01	-0.05	-0.03	-0.01	0.01

Notes: These regions are in general larger than counties in the United States. All numbers rounded to two decimal places. The US numbers are based only on continental commuting zones. Source: for all countries except the US we use IPUMS international as a data source. Sources for the US data are discussed in detail below.

Table 8: The Reallocation Component of the Structural Transformation: International Evidence

## B Theory: Proofs and Derivations

### B.1 Trade Costs and Land as a Factor of Production

For our main analysis we abstract from land as an explicit factor of production and from trade costs. In this section, we discuss the importance of these restriction. Consider first the absence of land as a fixed factor in the production function. For expositional simplicity we abstract from capital. Suppose that the production function in region  $r$  and sector  $s$  was given by  $Y_{rs} = Z_{rs}L_{rs}^{1-\gamma_s}T_{rs}^{\gamma_s}$ , where  $T_{rs}$  is the amount of land employed in sector  $s$ . Suppose that  $\gamma_A > \gamma_{NA} = 0$ , i.e. land is only employed in the agricultural sector. Let  $V_r$  be the land rental rate in region  $r$ . Cost minimization requires that  $V_r T_{rs} = \frac{\gamma_s}{1-\gamma_s} w_r L_{rs}$ . Hence, the production function is given by

$$Y_{rs} = Z_{rs}L_{rs}^{1-\gamma_s} \left( \frac{\gamma_s}{1-\gamma_s} \frac{w_r}{V_r} L_{rs} \right)^{\gamma_s} = \tilde{Z}_{rs}L_{rs},$$

where  $\tilde{Z}_{rs} = Z_{rs} \left( \frac{\gamma_s}{1-\gamma_s} \frac{w_r}{V_r} \right)^{\gamma_s}$ . Similarly, the price of sector  $s$  goods in region  $r$  is given by

$$P_{sr} = \frac{1}{Z_{rs}} \left( \frac{w_r}{1-\gamma_s} \right)^{1-\gamma_s} \left( \frac{V_r}{\gamma_s} \right)^{\gamma_s} = \frac{1}{\tilde{Z}_{rs}} \frac{w_r}{1-\gamma_s}.$$

The equilibrium conditions in (16) are then given by

$$L_{rt} \Gamma \zeta w_{rt} s_{rst} = (1-\gamma_s) \pi_{rst} \vartheta_t^s P Y_t$$

where  $\pi_{rst} = \frac{P_{sr}^{1-\sigma}}{(\sum_j P_{sj}^{1-\sigma})^{1/(1-\sigma)}} = \frac{\tilde{Z}_{sr}^{-1} w_{sr}^{1-\sigma}}{(\sum_j \tilde{Z}_{sj}^{-1} w_{sj}^{1-\sigma})^{1/(1-\sigma)}}$ . Given  $\tilde{Z}_{rs}$ , this is the same set of equations as in our baseline economy. Given  $L_{rt}$ , let  $(w_{rt}^*, s_{rst}^*)$  be the equilibrium allocation. Equilibrium land prices,  $V_r^*$ , are then given by  $\frac{V_r^*}{w_r^*} = \frac{\gamma_A}{1-\gamma_A} \frac{L_{rA}^*}{T_r}$ . Hence, the endogenous productivity  $\tilde{Z}_{rs}$  is related to the actual productivity  $Z_{rs}$  by  $\tilde{Z}_{rs} = Z_{rs} (L_{rA}^*/T_r)^{-\gamma_s}$ . Because we infer  $Z_{rs}$  as structural residuals, our *calibrated* model is isomorphic to a model with land in the production function.

To make this model consistent with individuals' spatial labor supply, note also that we require assumptions on land-ownership, i.e. who will receive the returns to land. To map this economy to our baseline

economy, we follow [Redding and Rossi-Hansberg \(2017\)](#) and assume that the return to land is received by all workers in region  $r$  as a proportional subsidy to their wage income. In particular, suppose worker  $i$  receives income  $y_r^i = (1 + \nu) w_r z_r^i$ , where  $\nu$  is the proportional subsidy. We then require that

$$w_r \nu \int z_{rs} di = \nu L_{rt} \Gamma_\zeta w_{rt} = V_r T_r = \frac{\gamma_A}{1 - \gamma_A} L_{rt} \Gamma_\zeta w_{rt} s_{rAt}.$$

Hence,  $\nu = \frac{\gamma_A}{1 - \gamma_A} s_{rst}$ . Letting  $\tilde{w}_{rt} \equiv \left(1 + \frac{\gamma_A}{1 - \gamma_A} s_{rAt}\right) w_{rt}$ , the spatial labor supply function still takes the same form as characterized in [Proposition 2](#), except that  $\tilde{w}_r$  takes the role of  $w_r$ . Intuitively, the agricultural share now determines the attractiveness of a location conditional on the wage as it encapsulates the effect on land prices. However, given that regional amenities  $A_{rt}$  are also determined as structural residuals, we can always find  $A_{rt}$  for the implied choice probabilities to coincide with the initial equilibrium. Note that the spatial equilibrium can still be calculated as a function of static equilibrium variables, i.e. allows for land as a fixed factor does not increase the computational complexity.

While the absence of land is inconsequential for the calibrated economy, our counterfactual exercise will, of course, depend on this restriction. By keeping fundamental productivity  $A_{rs}$  constant but say changing trade costs, the equilibrium population  $L_{rA}^*$  will change. With land in the production function, this will tend to reduce (increase) productivity in growing (declining) regions.

The assumed absence of trade costs is more substantial. Because trade costs in general imply that goods prices vary across space, consumers' spatial choice problem will no longer be static. In particular, the spatial choice probabilities are given by (see [Propositions 1 and 2](#) and [equation \(7\)](#))

$$\rho_{jrt} = \frac{\exp\left(\frac{1}{\kappa} (\mathcal{W}_{rt} + \Lambda_{r,t,t+1} - MC_{jr})\right)}{\sum_{l=1}^R \exp\left(\frac{1}{\kappa} (\mathcal{W}_{lt} + \Lambda_{l,t,t+1} - MC_{jl})\right)}, \quad (24)$$

where  $\Lambda_{r,t,t+1} = -\frac{\nu}{\gamma} \left( \left(\frac{p_{At}^r}{p_{NA_t}^r}\right)^\gamma + \beta \left(\frac{p_{A_{t+1}}^r}{p_{NA_{t+1}}^r}\right)^\gamma \right) + (1 + \beta) \left(\frac{\nu}{\gamma} - \frac{1}{\eta}\right)$  and  $p_{st}^r = \left(\sum_j (\tau_{jr} p_{jst})^{1-\sigma}\right)^{1/(1-\sigma)}$  and  $\tau_{jr}$  denotes the costs to ship goods from region  $j$  to region  $r$ . Crucially, the terms  $\Lambda_{r,t,t+1}$  now depend on  $r$  and hence no longer drop out of [expression \(24\)](#). This implies that moving flows at time  $t$ , depend on the distribution of *future* regional prices. And as *future* prices depend on future equilibrium allocations, computing the equilibrium becomes more involved. While in our baseline model, we only need to guess the time path of interest rates, we would now guess and iterate over the entire sequence of future distributions of regional factor prices.

Note however, that there are some special case of moving costs, which *are* covered in our baseline model. In particular, suppose that the final consumption good in [\(2\)](#) is the output of a national retailer, who sells to consumers irrespective of their location of residences. For the retailer to procure sector  $s$  goods from region  $r$ , however, is subject to trade costs. i.e.  $\tau_{sr} \geq 1$  units have to be shipped for a single unit to reach the retailer. This model is isomorphic to our baseline economy, where effective regional productivity  $Q_{rst}$  is given by  $Q_{rst} = \tilde{Q}_{rst} \tau_{sr}^{-1}$ , i.e. lower trade costs are isomorphic to higher productivity. In [Section B.9](#) below we show that the implied productivities from our model are systematically correlated with

measures of market access from [Donaldson and Hornbeck \(2016\)](#).

## B.2 Proof of Proposition 1

Suppose that the indirect utility function falls in the PIGL class, i.e.  $V(e, P) = \frac{1}{\eta} \left( \frac{e}{B(p)} \right)^\eta + C(P) - \frac{1}{\eta}$ .

The maximization problem is

$$U_r^i = \max_{[e_t, e_{t+1}, s]} \{V(e_t, P_t) + \beta V(e_{t+1}, P_{t+1})\},$$

subject to

$$\begin{aligned} e_t + s_t p_{I,t} &= y_{rt}^i \\ e_{t+1} &= (1 + r_{t+1}) s_t p_{I,t+1}. \end{aligned}$$

Substituting for  $e_{t+1}$  yields

$$U_r^i = \max_{e_t} \left\{ V(e_t, P_t) + \beta V \left( (1 + r_{t+1}) (y_{rt}^i - e_t) \frac{p_{I,t+1}}{p_{I,t}}, P_{t+1} \right) \right\}.$$

The optimal allocation of spending is determined from the Euler equation

$$\frac{\partial V(e_t, P_t)}{\partial e} = \beta (1 + r_{t+1}) \frac{p_{I,t+1}}{p_{I,t}} \frac{\partial V(e_{t+1}, P_{t+1})}{\partial e}.$$

From above (37) we get that this equation is

$$e_t^{\eta-1} B(p_t)^{-\eta} = \beta \left( (1 + r_{t+1}) \frac{p_{I,t+1}}{p_{I,t}} \right)^\eta ((y_{rt}^i - e_t))^{\eta-1} B(p_{t+1})^{-\eta}.$$

This yields

$$\frac{y_{rt}^i - e_t}{e_t} = \left( \left( \frac{1}{1 + r_{t+1}} \frac{B(p_{t+1})/B(p_t)}{p_{I,t+1}/p_{I,t}} \right)^\eta \frac{1}{\beta} \right)^{\frac{1}{\eta-1}},$$

so that

$$\begin{aligned} e_t &= \frac{1}{1 + \left( \phi_{t,t+1}^\eta \frac{1}{\beta} \right)^{\frac{1}{\eta-1}}} y_{rt}^i \\ e_{t+1} &= \frac{(1 + r_{t+1}) p_{I,t+1}}{p_{I,t}} \frac{\left( \phi_{t,t+1}^\eta \frac{1}{\beta} \right)^{\frac{1}{\eta-1}}}{1 + \left( \phi_{t,t+1}^\eta \frac{1}{\beta} \right)^{\frac{1}{\eta-1}}} y_{rt}^i \end{aligned}$$

where  $\phi_{t,t+1} \equiv \frac{1}{1+r_{t+1}} \frac{B(p_{t+1})/B(p_t)}{p_{t,t+1}/p_{t,t}}$ . Hence, (6) implies that

$$\begin{aligned} U_{jr} &= V_{jr}(e_t, P_t) + \beta V_{jr}(e_{t+1}, P_{t+1}) \\ &= A_{jr} \frac{1}{\eta} w_{rt}^\eta \left[ B(p_t)^{-\eta} \left( 1 + \left( \frac{1}{\beta} \phi_{t,t+1}^\eta \right)^{\frac{1}{\eta-1}} \right)^{1-\eta} \right] + C(P_t) + \beta C(P_{t+1}) - \frac{1+\beta}{\eta}. \end{aligned}$$

This can be written as  $U_{jr} = A_{jr} w_{rt}^\eta \Phi_{t,t+1} + \Lambda_{t,t+1}$ , where

$$\begin{aligned} \Phi_{t,t+1} &= \frac{1}{\eta} B(p_t)^{-\eta} \left( 1 + \left( \frac{1}{\beta} \phi_{t,t+1}^\eta \right)^{\frac{1}{\eta-1}} \right)^{1-\eta} \\ \Lambda_{t,t+1} &= C(P_t) + \beta C(P_{t+1}) - \frac{1+\beta}{\eta}. \end{aligned}$$

For our specification we have that  $B(p_t) = p_{A,t}^\phi p_{NA,t}^{1-\phi} = 1$ . Hence,

$$\phi_{t,t+1} \equiv \frac{B(p_{t+1})}{B(p_t)(1+r_{t+1})} = \frac{1}{1+r_{t+1}} = \phi_{t+1}$$

and

$$\Phi_{t,t+1} = \frac{1}{\eta} \left( 1 + \left( \frac{1}{\beta} \left( \frac{1}{1+r_{t+1}} \right)^\eta \right)^{\frac{1}{\eta-1}} \right)^{1-\eta} = \frac{1}{\eta} \left( 1 + \beta^{\frac{1}{1-\eta}} (1+r_{t+1})^{\frac{\eta}{1-\eta}} \right)^{1-\eta}.$$

Note also that

$$e_t = \frac{1}{1 + \left( \phi_{t,t+1}^\eta \frac{1}{\beta} \right)^{\frac{1}{\eta-1}}} w_t = \frac{1}{1 + \beta^{\frac{1}{1-\eta}} (1+r_{t+1})^{\frac{\eta}{1-\eta}}} w_t.$$

This proves the results for Proposition 1.

### B.3 Earning, Expected Earnings and Aggregate Demand

Consider individual  $i$  in region  $r$ . Given her optimal occupational choice, the earnings of individual  $i$  are given by  $y^i = \max_s \{ w_{rs} z_s^i \}$ . We assumed that individual productivities are Frechet Distributed, i.e.  $F_s^h(z) = e^{-\Psi_s^h z^{-\zeta}}$ , where  $\Psi_s^h$  parametrizes the average level of productivity of individuals of skill  $h$  in region  $r$  in sector  $s$  and  $\zeta$  the dispersion of skills. Hence, the distribution of sectoral earning  $y_{sr}^i \equiv w_{s,r} z_s^i$  is also Frechet and given by  $F_{y_{rs}}(y) = P\left(z \leq \frac{y}{w_{rs}}\right) = e^{-\Psi_s^h w_{rs}^\zeta y^{-\zeta}}$ . Using standard arguments about the max stability of the Frechet, the distribution of total earnings  $y^i$  is also Frechet and given by

$$F_{y_r}^h(y) = e^{-(\Theta_r^h)^\zeta y^{-\zeta}} = e^{-\left(\frac{y}{\Theta_r^h}\right)^{-\zeta}} \quad (25)$$

where  $\Theta_r^h = \left( \sum_s \Psi_s^h w_{rs}^\zeta \right)^{1/\zeta} = \left( \Psi_A^h w_{rA}^\zeta + \Psi_{NA}^h w_{rNA}^\zeta \right)^{1/\zeta}$ . Hence, average earnings of individual  $i$  with skill type  $h$  in region  $r$  are given by  $E \left[ y_{r,h}^i \right] = \Gamma \left( 1 - \frac{1}{\zeta} \right) \Theta_r^h$ . From (25) we can derive the distribution of  $y^{1-\eta}$ . As  $\eta < 1$ , we have that

$$F_{y^{1-\eta}}(q) = P(y^{1-\eta} \leq q) = P\left(y \leq q^{1/(1-\eta)}\right) = e^{-\Theta_r^h \left(q^{1/(1-\eta)}\right)^{-\zeta}} = e^{-\left(\frac{q}{\Theta_r^h}\right)^{-\frac{\zeta}{1-\eta}}}.$$

Hence,  $y^{1-\eta}$  is still Frechet. Therefore

$$\int_i y_i^{1-\eta} di = L_r^h E \left[ y_i^{1-\eta} \right] = L_r^h \Gamma \left( 1 - \frac{1-\eta}{\zeta} \right) \times \left( \Theta_r^h \right) = L_r^h \Gamma \left( 1 + \frac{\eta-1}{\zeta} \right) \left( \Theta_r^h \right)^{1-\eta}.$$

To derive the aggregate value added share  $\vartheta_t^A$ , note that

$$\begin{aligned} PY_t^A &= \sum_{r=1}^R \left( \vartheta_A \left( E_{\mathcal{J}_r^Y}, P \right) E_{\mathcal{J}_r^Y} L_{rt} + \vartheta_A \left( E_{\mathcal{J}_r^O}, P \right) E_{\mathcal{J}_r^O} L_{rt-1} \right) + \phi I_t \\ &= \phi \left[ I_t + \sum_{r=1}^R \left( E_{\mathcal{J}_r^Y} L_{rt} + E_{\mathcal{J}_r^O} L_{rt-1} \right) \right] + \tilde{v} \left( \frac{PA}{PM} \right)^\gamma \sum_{r=1}^R \left( E_{\mathcal{J}_r^Y}^{1-\eta} L_{rt} + E_{\mathcal{J}_r^O}^{1-\eta} L_{rt-1} \right). \end{aligned}$$

Now note that total spending of the old equals the capital return plus the non-depreciated stock of capital, i.e.

$$\sum_{r=1}^R E_{\mathcal{J}_r^O} L_{rt-1} = R_t K_t + (1 - \delta) K_t = \alpha PY_t + (1 - \delta) K_t.$$

And because the future capital stock is given by the savings of the young generation (see (20)), we get that

$$I_t + \sum_{r=1}^R \left( E_{\mathcal{J}_r^Y} L_{rt} + E_{\mathcal{J}_r^O} L_{rt-1} \right) = K_{t+1} + \sum_{r=1}^R \left( E_{\mathcal{J}_r^Y} L_{rt} \right) + \alpha PY_t = PY_t.$$

This yields (15).

## B.4 The Spaceless Economy: Equilibrium Existence and Uniqueness

Consider the economy with single location and let  $K_t$  be given. The choice of numeraire implies that

$$1 = P_{At}^\phi P_{NA_t}^{1-\phi} = \left( \frac{1}{Z_{At}} \right)^\phi \left( \frac{1}{Z_{NA_t}} \right)^{1-\phi} \left( \frac{PY_t}{K_t} \right)^\alpha \left( \frac{PY_t}{L \Gamma_\zeta} \right)^{1-\alpha}.$$

Hence,  $PY_t = Z_t K_t^\alpha L^{1-\alpha}$  where  $Z_t = \Gamma_\zeta^\alpha Z_{At}^\phi Z_{NA,t}^{1-\phi}$ . Using that  $K_{t+1} = (1 - \psi_{t+1})(1 - \alpha)PY_t$ , the expression for  $\psi_{t+1}$ ,  $r_t = R_t - \delta$  and  $k_t = K_t/L$ , we get that

$$k_{t+1} \times \left( \left( \frac{1}{\beta} \right)^{\frac{1}{1-\eta}} \left( 1 + \alpha Z_{t+1} k_{t+1}^{-(1-\alpha)} - \delta \right)^{-\frac{\eta}{1-\eta}} + 1 \right) = (1 - \alpha) Z_t k_t^\alpha.$$

Note that the LHS is increasing and continuous in  $k_{t+1}$  and satisfies  $\lim_{k_{t+1} \rightarrow \infty} LHS = \infty$  and  $\lim_{k_{t+1} \rightarrow 0} LHS = 0$ . Hence, there is a unique mapping  $k_{t+1} = m(k_t, Z_t, Z_{t+1})$ . Furthermore,  $m(\cdot)$  is increasing in all arguments.

## B.5 The Equilibrium with Heterogeneous Skills

The equilibrium characterization for the case of heterogenous skills is very similar to our baseline case. In that case, the spatial equilibrium, i.e. the set of equilibrium wages and population levels  $\{w_{rt}, L_{rt}\}_r$ , is determined from labor market clearing condition (18)

$$L_{rt} \Gamma_\zeta w_{rt} = (1 - \alpha) \left( \pi_{rAt} \vartheta_t^A + \pi_{rAt} (1 - \vartheta_t^A) \right) PY_t$$

and the spatial labor supply condition  $L_{rt} = \sum_{j=1}^R \rho_{jrt} L_{jt-1}$ , where  $\rho_{jrt}$ ,  $\vartheta_t^A$  and  $\pi_{rst}$  are given in (13), (15) and (17) and also only depend on  $\{w_{rt}, L_{rt}\}_r$ . Given  $\{w_{rt}, L_{rt}\}_r$ , the future capital stock is then determined from (20) and the regional agricultural employment shares from (19).

In the case with sector-specific skills and imperfect substitutability, the spatial equilibrium consists of sector-specific wages  $\{w_{rAt}, w_{rNA,t}\}_r$  and skill-specific populations  $\{L_{rt}^L, L_{rt}^H\}_r$ . The corresponding equilibrium conditions are as follows. The regional labor market clearing condition is given by

$$L_{rt} \Gamma_\zeta (\lambda_{rt} \Theta_{rt}^H + (1 - \lambda_{rt}) \Theta_{rt}^L) = (1 - \alpha) \left( \pi_{rAt} \vartheta_t^A + \pi_{rAt} (1 - \vartheta_t^A) \right) PY_t,$$

where  $\Theta_r^h$ , given in (22), denotes regional income for skill group  $h$  and  $\lambda_{rt}$  is the share of skilled individuals.<sup>42</sup> The labor supply equations are now skill-specific and given by  $L_{rt}^h = \sum_{j=1}^R \rho_{jrt}^h L_{jt-1}^h$ , where  $\rho_{jrt}^h$  is still given in (13) with  $\mathcal{W}_{rt}^h = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_r^h)^\eta + A_{rt}$ . Finally, the within-region allocation of factors across sectors, i.e. the counterpart to (19), is given by

$$\frac{\pi_{rAt}}{\pi_{rNA,t}} \frac{\vartheta_t^A}{1 - \vartheta_t^A} = \frac{w_{rA} (H_{rA}^L + H_{rA}^H)}{w_{rNA} (H_{rNA}^L + H_{rNA}^H)} = \left( \frac{w_{rAt}}{w_{rNA,t}} \right)^\zeta \frac{(1 - \lambda_r) \Psi_A^L (\Theta_r^h)^{1-\zeta} + \lambda_r \Psi_A^H (\Theta_r^H)^{1-\zeta}}{(1 - \lambda_r) \Psi_{NA}^L (\Theta_r^h)^{1-\zeta} + \lambda_r \Psi_{NA}^H (\Theta_r^H)^{1-\zeta}}. \quad (26)$$

These equations determine  $\{w_{rAt}, w_{rNA,t}, L_{rt}^L, L_{rt}^H\}_r$ . The skill specific employment shares can then be

<sup>42</sup>To see that, note that labor earnings in region  $r$ , for skill group  $h$  in sector  $s$  are given by  $w_{rst} H_{rst}^h = L_{rt} \Gamma_\zeta \lambda_r^h s_r^h \Theta_r^h$ . Summing over sectors  $s$  and skill groups  $h$  yields the expression above.

calculated as  $s_{rst}^h = \Psi_s^h (w_{rs}/\Theta_r^h)^\zeta$ . The capital accumulation is still given by (20), i.e.

$$K_{t+1} = (1 - \psi(r_{t+1})) \sum_r L_{rt} \Gamma_\zeta (\lambda_{rt} \Theta_{rt}^H + (1 - \lambda_{rt}) \Theta_{rt}^L) = (1 - \psi(r_{t+1})) (1 - \alpha) PY_t.$$

The main difference to the case with substitutable skills is the within-region sectoral supply equation (26). In the baseline model,  $s_{rAt}$  is determined residually from (19). With an upward sloping supply curve, the relative wages have to be consistent with sectoral labor supplies.

## B.6 Balanced Growth Path Relationships

Consider a dynamic allocation where GDP grows at a constant rate  $g$  and the capital output ratio is constant. Static optimality requires that  $R_t = \frac{\alpha PY_t}{K_t}$ . Hence, a constant capital output ratio directly implies that the return to capital  $R_t$  has to be constant. Hence, the real interest on saving is also constant and given by  $r = R - \delta$ . This also implies that the consumption rate in (11) is constant and given by

$$\psi = \left( 1 + \beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}} \right)^{-1}. \quad (27)$$

To solve for the interest rate, note that capital grows at rate  $g$ , so that (20) implies that

$$\frac{K_{t+1}}{K_t} = \frac{(1 - \psi)(1 - \alpha) PY_t}{\alpha PY_t / R} = (1 - \psi) \frac{(1 - \alpha)}{\alpha} (r + \delta) = 1 + g.$$

Hence,

$$1 + g = (1 - \psi) \frac{(1 - \alpha)}{\alpha} (r + \delta) = \left( \frac{\beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}}}{1 + \beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}}} \right) \frac{(1 - \alpha)}{\alpha} (r + \delta). \quad (28)$$

This equation determines  $r$  as a function of parameters. Along the BGP the consumption and investment rate is equal to

$$\begin{aligned} PC_t &= \psi(1 - \alpha) PY_t + \alpha PY_t + (1 - \delta) \frac{\alpha}{R} PY_t = \left[ \psi(1 - \alpha) + \alpha + (1 - \delta) \frac{\alpha}{R} \right] PY_t \\ PI_t &= (1 - \psi)(1 - \alpha) PY_t - (1 - \delta) \frac{\alpha}{R} PY_t = \left[ (1 - \psi)(1 - \alpha) - (1 - \delta) \frac{\alpha}{R} \right] PY_t \end{aligned}$$

Using (28) yields  $\frac{PI_t}{PY_t} = (g + \delta) \frac{\alpha}{R}$  and  $\frac{PC_t}{PY_t} = 1 - (g + \delta) \frac{\alpha}{R}$ . From (28) we also get that  $\psi = 1 - \frac{\alpha}{1-\alpha} \frac{1+g}{R}$ .

## B.7 Regularity conditions for PIGL preferences

In our model, expenditure share on the two goods are given by

$$\begin{aligned}\vartheta_A(e, p) &= \phi + v \left( \frac{p_A}{p_M} \right)^\gamma e^{-\eta} \\ \vartheta_{NA}(e, p) &= 1 - \phi - v \left( \frac{p_A}{p_M} \right)^\gamma e^{-\eta}.\end{aligned}$$

For these expenditure shares to be positive, we need that

$$\vartheta_A(e, p) \geq 0 \Rightarrow e^\eta \geq -\frac{v}{\phi} \left( \frac{p_A}{p_M} \right)^\gamma, \quad (29)$$

and  $\vartheta_{NA}(e, p) \geq 0 \Rightarrow e^\eta \geq \frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma$ . Note first that (29) is trivially satisfied as long as  $v > 0$ . Also note that satisfying both of these restrictions automatically implies that  $\vartheta_s(e, p) \leq 1$ . In addition, as we show in Section C.3 in the Online Appendix, for the Slutsky matrix to be negative semi definite, we need that

$$v(1-\eta) \left( \frac{p_A}{p_M} \right)^\gamma - \frac{(1-\phi)\phi}{v} \left( \frac{p_A}{p_M} \right)^{-\gamma} e^{2\eta} \leq (1-2\phi-\gamma)e^\eta.$$

Hence, for our preferences to be well-defined, we require that

$$e^\eta \geq \frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma \quad (30)$$

$$(1-2\phi-\gamma)e^\eta + \frac{(1-\phi)\phi}{v} \left( \frac{p_A}{p_M} \right)^{-\gamma} e^{2\eta} \geq v(1-\eta) \left( \frac{p_A}{p_M} \right)^\gamma. \quad (31)$$

**Lemma 5.** *A sufficient condition for (30) to be satisfied is that (31) holds and that*

$$\gamma > (1-\phi)\eta. \quad (32)$$

*Proof.* Equation (31) can be written as

$$\left( \frac{e^\eta}{\frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma} - 1 \right) + \phi \frac{e^\eta}{\frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma} \frac{e^\eta}{\frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma} \geq [(1-\phi)(1-\eta) - 1] + (2\phi + \gamma) \frac{e^\eta}{\frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma}.$$

Letting  $x = \frac{e^\eta}{\frac{v}{1-\phi} \left( \frac{p_A}{p_M} \right)^\gamma}$ , this yields

$$(x-1) + (\phi x - (2\phi + \gamma))x \geq -(1 - (1-\phi)(1-\eta)).$$



Now let  $h(x) = (x-1) + (\phi x - (2\phi + \gamma))x$ . For (31) to be satisfied we need that

$$h(x) \geq -(1 - (1 - \phi)(1 - \eta)).$$

Note that  $h$  is strictly concave with a minimum at

$$h'(x^*) = 1 + \phi x^* - (2\phi + \gamma) + \phi x^* = 0.$$

Hence,

$$x^* = 1 - \frac{1 - \gamma}{2\phi} < 1.$$

Also note that

$$h(0) = -1 < -(1 - (1 - \phi)(1 - \eta)).$$

Hence, for (31) to be satisfied, it has to be the case that  $x > x^* = 1 - \frac{1 - \gamma}{2\phi}$  as  $h(x^*) < h(0)$ . Hence, condition (31) implies (30) if

$$h(1) = \phi - 2\phi - \gamma < -(1 - (1 - \phi)(1 - \eta)).$$

Rearranging terms yields  $(1 - \phi)\eta < \gamma$ , which is (32). □

Hence, the preferences are well defined as long (31) is satisfied and (32) holds. Because the RHS of (31) is increasing in  $e$  in the relevant range, i.e. as long as (31) is satisfied, this implies that the preferences are well defined as long as  $e$  is high enough.

Now note that  $e_{it} = \psi_{t+1}y_{it}$ , where  $y_{it}$  denotes total earnings of individual  $i$ . From (25) we know that

$$P(e_{it} \leq \kappa) = P\left(y_{it} \leq \frac{\kappa}{\psi_{t+1}}\right) = e^{-(\Theta_{it}^h \psi_{t+1})^\zeta \kappa^{-\zeta}},$$

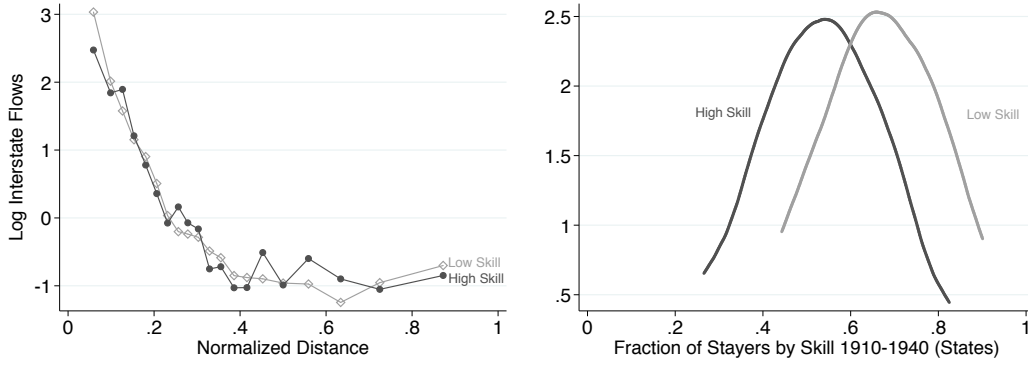
where

$$\left(\Theta_{it}^h \psi_{t+1}\right)^\zeta = \left(\Psi_A^h w_{rAt}^\zeta + \Psi_{NA}^h w_{rNAAt}^\zeta\right) \psi_{t+1}^\zeta > \left(\Psi_A^h w_{rAt}^\zeta + \Psi_{NA}^h w_{rNAAt}^\zeta\right) \left(\frac{1}{\eta} \left(1 + \beta^{\frac{1}{1-\eta}}\right)^{1-\eta}\right)^\zeta.$$

Hence, as long as aggregate productivity in 1880 (and hence  $w_{r,1880}$ ) is high enough, we can make  $P(e_{it} \leq \kappa)$  arbitrarily small.

## B.8 Additional Properties of the Calibrated Model

In this section we provide additional results for the fit of the calibrated model.



Notes: To construct the left panel, we run a gravity equation of the form where origin and destination fixed effects are skill specific:  $\ln \rho_{jr} / (1 - \rho_{jj}) = \alpha_j^h + \alpha_r^h + \hat{u}_{jr}^h$ . The red diamonds line plots the resulting  $\hat{u}_{jr}^h$  by distance percentile, the blue line does the same for low skill types. The right panel we plot the distribution of the share of people staying in their home state between 1910 and 1940 for low and high skilled workers.

Figure 12: Lifetime interstate Migration by Skill in 1940 in the Data

**Migration by skill type** In the model low and high skill workers are subject to the same distance costs, which yields an elasticity of moving flows to distance that is very similar for both groups. As the left panel of the below figure shows this is in line with the data. In the model high skill types are slightly more likely to move but not a lot. In the data low skilled workers are substantially less likely to leave their birth state. The model can match if we allowed  $\tau$  to differ by skill type, since as outlined in the calibration section, the total number of stayers is monotone in  $\tau$ . We chose to abstract from this for simplicity in the main part of the paper.

**Moving costs** To quantify the economic magnitude of our estimated fixed costs, we calculate

$$\Delta_{jt}^h \equiv \frac{\mathcal{U}_{jt}^{h,Mov} - \mathcal{U}_{jt}^{h,Stay}}{\tau}, \quad (33)$$

i.e. the average increase in utility by moving *relative* to the fixed cost of moving as a norm for the economic magnitude of fixed costs. Because utility is not equalized, these gains differ by commuting zone. In Table 9 we report some statistics of the distribution of  $\Delta_{jt}^h$ . In the top row we report these statistics for low skilled individuals, in the bottom row for high skilled individuals. Table 9 shows that for the median commuting zone the expected value of moving is slightly positive and amounts to roughly 5% (10% for high skilled individuals) of the estimated fixed cost of moving.

**Regional Fundamentals** In this section, we describe additional details for the estimated spatial productivities  $\{Q_{rAt}, Q_{rNA_t}\}_{r,t}$  and amenities  $\{A_{rt}\}_{r,t}$ . First we study the fundamental determinants of agricultural specialization by projecting the endogenous agricultural employment share on regional fundamentals and the two state variables of the system, namely the distribution of skills and population size in 1880. We do so in Table 10.

	Expected value of moving relative to fixed costs $\Delta_{jt}^h$				
	10%	25%	50%	75%	90%
Low Skilled	-0.2543	-0.0937	0.0571	0.1900	0.2871
High Skilled	-0.2464	-0.0654	0.1065	0.2543	0.3619

Notes: The table reports different quantiles of the distribution of  $\Delta_{jt}^h$  calculated as (33).

Table 9: The Economic Magnitude of Moving Costs

	Dep. variable: $\ln s_{Art}$			
Non Ag. Productivity ( $\ln Q_{rNA}$ )	-0.331*** (0.020)			-0.344*** (0.007)
Ag. Productivity ( $\ln Q_{rA}$ )		0.404*** (0.022)		0.408*** (0.006)
Amenities ( $A_r$ )			-0.430*** (0.075)	-0.055*** (0.013)
Skill share in 1880 ( $\lambda_{r1880}$ )	0.139 (0.123)	-1.594*** (0.150)	-0.466** (0.196)	-0.117*** (0.032)
In population 1880	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
$N$	3225	3225	2580	2580
$R^2$	0.891	0.902	0.810	0.989

Notes: Robust standard errors in parentheses with \*\*\*, \*\* and \* respectively denoting significance at the 1%, 5% and 10% levels. All specification control for the population in 1880 and for year fixed effects.

Table 10: Fundamental Determinants of Agricultural Specialization

In particular, we report the results from the specification

$$\ln s_{rAt} = \delta_t + \beta_{NA} \ln Q_{rNA} + \beta_A \ln Q_{rA} + \gamma A_{rt} + \zeta \lambda_{r1880} + \eta \ln \text{pop}_{r1880} + u_{rt},$$

where  $\delta_t$  is a year fixed effect. In columns one to three we report the bivariate partial correlations. Rural, agricultural regions are regions with low productivity in the non-agricultural sector, a comparative advantage in the production of agricultural goods and low amenities. The last column reports the respective partial correlations. In particular, the coefficient on regional amenities and the regional skill share drops by a factor of ten. This reflects the existing cross-sectional correlation with regional productivity, in particular non-agricultural productivity  $Q_{rNA}$ .

In Table 11 we focus directly on the dynamics of spatial productivity and amenities. In particular, we consider a simple autoregressive specification

$$x_{rt} = \delta_t + \delta_r + \beta y_{r,t-1} + u_{rt}$$

where  $x$  denotes either log sectoral productivity,  $\ln Q_{rst}$ , or the level of amenities  $A_{rt}$  and  $\delta_t$  and  $\delta_r$  are

	$\ln Q_{rNA,t}$		$\ln Q_{rA,t}$		$A_{r,t}$	
$\ln Q_{rNA,t-1}$	0.832*** (0.011)	0.416*** (0.039)				
$\ln Q_{rA,t-1}$			0.725*** (0.050)	0.045 (0.043)		
$A_{r,t-1}$					0.897*** (0.063)	0.175 (0.133)
Year FE	✓	✓	✓	✓	✓	✓
CZ FE		✓		✓		✓
$N$	2580	2580	2580	2580	1935	1935
$R^2$	0.907	0.958	0.486	0.789	0.749	0.890

Notes: Robust standard errors in parentheses with \*\*\*, \*\* and \* respectively denoting significance at the 1%, 5% and 10% levels.

Table 11: The Process of Spatial Fundamentals

region and year fixed effects. The first four columns show that productivities are mean-reverting and that there is an important fixed, region-specific component determining spatial productivity between 1880 and 2000. It is also interesting to note that these patterns differ slightly across sectors. In particular, conditional on commuting zone fixed effects, past agricultural productivity is uncorrelated contemporaneous agricultural productivity. The last two columns show the same result for regional amenities. Amenities follow a stochastic process, which is similar to agricultural productivity.

## B.9 Local Productivities and Market Access

We calibrate the cross-sectional distribution of sectoral productivities  $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$  and amenities  $\{A_{rt}\}_{rt}$  as structural residuals of the model as is commonly done in the new quantitative spatial economics literature (see (Redding and Rossi-Hansberg, 2017) for a recent and excellent review of this literature). The local sectoral productivities calibrated this way,  $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$ , are residuals necessary to fit the data on wages and sectoral employment conditional on the calibrated parameters of the model. While they reflect a variety factors that make one region more productive than another, some measurable others not, there is one factor that determines the productivity of a location in a very direct way: its integration into the national transportation network. In this section we use direct measures on the extent of regional “market access” by Donaldson and Hornbeck (2016) for each county in the US for 1880 and 1910 to corroborate our model-based measures of regional productivity.

Donaldson and Hornbeck (2016) provide data on county-to-county transport cost for 1880 and 1910. We use this data to construct a commuting zone level index for market access costs for these time periods. In particular we average the market access cost across all county-to-county pairs within two given commuting zones to obtain a measure for the ease of transporting goods between these two zones. Then we take the destination population share weighted sum of market access cost for each origin commuting zone to obtain our market access index at the commuting zone level.

We now relate this index of market access costs to our measured regional productivity residuals  $Q_{rst}$ .

Letting  $MAC_{rt}$  be this index of market access costs of county  $r$  at time  $t$ , we consider a specifications of

$$\ln Q_{rst} = \delta_r + \beta \times MAC_{rt} + u_{rst}, \quad (34)$$

where  $\delta_r$  denotes a fixed effect at different levels of regional aggregation. We expect  $\beta < 0$ , as higher transport costs should reduce a county's earnings potential, i.e. productivity. We estimate 34 separately for the agricultural and the non-agricultural sector. The results are contained in Table 12.

In the first two columns we show that our model infers low productivity in places that have a high market access cost. Columns 3 and 4 show that this relationship is if anything stronger within states. In columns 5 and 6 we exploit the time-variation within commuting zones and show that regions who see their access costs decrease indeed experience faster productivity growth. Finally, in the last two column, we estimate 34 in first differences and explicitly include a whole set of state fixed effects, i.e. allowing for systematic differences in productivity growth across states. Again, we find a significant relationship between (changes in ) market access costs and (changes in) regional productivity. We take this as evidence that transportation costs are one of the directly measurable ingredients in regional productivity shifters that our framework infers as  $\{Q_{rAt}, Q_{rNAAt}\}_{rt}$ .

	$\ln Q_{NA}$	$\ln Q_A$	$\ln Q_{NA}$	$\ln Q_A$	$\ln Q_{NA}$	$\ln Q_A$	$\Delta \ln Q_{NA}$	$\Delta \ln Q_{NA}$
Market Access Costs	-1.897*** (0.224)	-2.077*** (0.125)	-5.205*** (0.290)	-4.291*** (0.195)	-3.937*** (0.353)	-3.760*** (0.309)		
Change in Market Access Costs							-4.005*** (0.388)	-2.395*** (0.328)
Year FE	✓	✓	✓	✓				
CZ FE					✓	✓		
State FE			✓	✓			✓	✓
$N$	1242	1242	1242	1242	1242	1242	621	621
adj. $R^2$	0.133	0.286	0.538	0.455	0.789	0.576	0.337	0.399

Notes: Robust standard errors in parentheses with \*\*\*, \*\* and \* respectively denoting significance at the 1%, 5% and 10% levels. We measure Market Access Costs using the data from Donaldson and Hornbeck (2016). We use data on the cost of reaching any other county from a given county, take the destination population weighted sum and aggregate them to the commuting zone level. Our measure of market access costs is the log of this index. The change in market access costs is the log difference between 1880 and 1910.

Table 12: Spatial Productivity and Market Access Cost from Donaldson and Hornbeck (2016)

## B.10 Spatial Welfare Inequality

In Figure 7 in the main text we depict the evolution of spatial inequality. We construct this figure in the following way. From Proposition 2 we have that expected utility in region  $r$  is given by

$$\mathcal{W}_{rt}^h(\Theta_r^h) = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_r^h)^\eta + \Lambda_{t,t+1} + A_{rt}. \quad (35)$$

Let  $T_{rt}^h(\Delta)$  be the increase in average income  $\Theta_r^h$  required to increase utility by  $\Delta$ , i.e.  $\mathcal{W}_{rt}^h(\Theta_r^h T_{rt}^h(\Delta)) = \mathcal{W}_{rt}^h(\Theta_r^h) + \Delta$ . (35) implies that

$$T_{rt}^h(\Delta) = \left( 1 + \frac{\Delta}{\frac{\Gamma_{\eta/\xi}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_r^h)^\eta} \right)^{1/\eta}.$$

Let  $\Delta_t^{h,IQR}$  be the interquartile range in regional welfare  $\mathcal{W}_{rt}^h$  at time  $t$ , i.e.  $\Delta_t^{h,IQR} = \mathcal{W}_t^{h,75} - \mathcal{W}_t^{h,25}$ , where  $\mathcal{W}_t^{h,x}$  is the  $x$ -quantile of the distribution of  $\mathcal{W}_{rt}^h$ . Figure 7 shows the time series evolution of  $\bar{T}_t^h(\Delta_t^{h,IQR}) = \frac{1}{R} \sum_r T_{rt}^h(\Delta)$  for both high and low skilled individuals.