Spatial Structural Change*

Fabian Eckert† Michael Peters‡

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 analyze this process and to quantify its macroeconomic implications. We show that spatial reallocation across labor markets accounts for almost none of the aggregate decline in agricultural employment. The reason is that population flows, while large in the aggregate, were only weakly correlated with agricultural specialization. Labor mobility nevertheless had important aggregate effects. Without migration income per capita would have been 15% lower. Moreover, spatial welfare inequality would have been substantially higher, especially among low-skilled, agricultural workers, which were particularly exposed to the structural transformation.

*We thank seminar participants at Boston University, Stanford, Penn State, UNC, Maryland, 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 Boppart, 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 Gvirtz and Adam Harris provided outstanding research assistance.

†Yale University. fabian.eckert@yale.edu

‡Yale University and NBER. 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 implies that economic growth is also 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.[1]

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.[2] While this is seemingly inconsistent with the secular trend in urbanization, we explicitly show that this is not the case. In fact, like the change in agricultural employment, the process of urbanization, whereby the share of urban dwellers among US workers increased from 25% to 75% between 1880 and 2000, 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 borne out empirically. In particular, US Census data reveals that throughout the last century, about 30% of workers lived in states different from their state of birth. Hence, the reason why migration across labor markets cannot account for much of the decline in agricultural employment is not the absence of migrants, but rather that the correlation between agricultural employment shares and net population outflows was essentially zero.

To understand why this was the case and to analyze the implications for aggregate productivity and the spatial distribution of welfare, 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.

To explain why migration flows were only weakly correlated with agricultural specialization, our model

[1]By "Spatial Structural Change" we refer to the simultaneous change in the structure of sectoral employment and the spatial organization of the economy. 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)).

[2]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.

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. Moreover, such wage differences were only imperfectly correlated with the regional agricultural employment share, as many locations with a comparative advantage in the agricultural sector were, in fact, quite productive. Individuals, in their search for higher earnings, therefore often relocated *towards* agricultural areas. Secondly, we also find that the migration elasticity, i.e. the sensitivity of migration flows to regional wages, was limited. In particular, spatial population gross flows were much larger than population net flows. This suggests that idiosyncratic, non-monetary preference shocks, by definition uncorrelated with the local industrial structure, were an important determinant of migration decisions. In fact, we explicitly show that a model without these features predicts that spatial reallocation alone accounts for one third of the aggregate decline in agricultural employment.

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 directed spatial reallocation suggests that the US economy potentially missed out on substantial productivity improvements. Quantitatively, we find that such productivity losses were modest. If population outflows and agricultural employment shares had been perfectly negatively correlated, 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 during the structural transformation therefore seemed to have captured a large share of the 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 welfare differences across commuting zones corresponded to a doubling of lifetime income for the average commuting zone in the US. In contrast, in 1910 one would have had to increase income by 160%. Importantly, the process of spatial mobility was a central 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 spatial frictions to mobility disappear, 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.[3] 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.[4] 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 sectoral labor supply and hence equilibrium factor prices 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.

[3] Kuznets (1957) and Chenery (1960) have been early observers of the striking downward trend in the aggregate agricultural employment share in the United States. To explain these patterns, two mechanism 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)).

[4] 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.

In allowing for a spatial microstructure, we find a distinct role for the structural transformation to affect macroeconomic outcomes. This is in contrast to many macroeconomic models, where "structural change is of secondary interest, because a simple aggregate model tells us all we need to know about growth" (Buera and Kaboski, 2009, p. 472). Our model, for example, endogenously generates an "agricultural productivity gap", i.e. the fact that value added per worker is persistently low in the agricultural sector (see e.g. Gollin et al. (2014) or Herrendorf and Schoellman (2015)), if spatial mobility costs keep wages in agricultural areas low. The existence of such spatial gaps and their implications for aggregate productivity and welfare has recently become subject of an active literature (see e.g. Young (2013), Bryan et al. (2014), Hsieh and Moretti (2015), Herkenhoff et al. (2017) 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 integrated into a general equilibrium trade model in a tractable way.[5]

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 theory. In Section 4, we calibrate the model to time-series and spatial data from the US. In Section 5, we explain why the spatial reallocation component is small and we quantify the implications for aggregate productivity and spatial inequality. Section 6 provides an analysis of the robustness of our results and 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, the economy can accommodate this spatial bias of the structural transformation in two ways. Either the process of structural change can induce *spatial reallocation,*

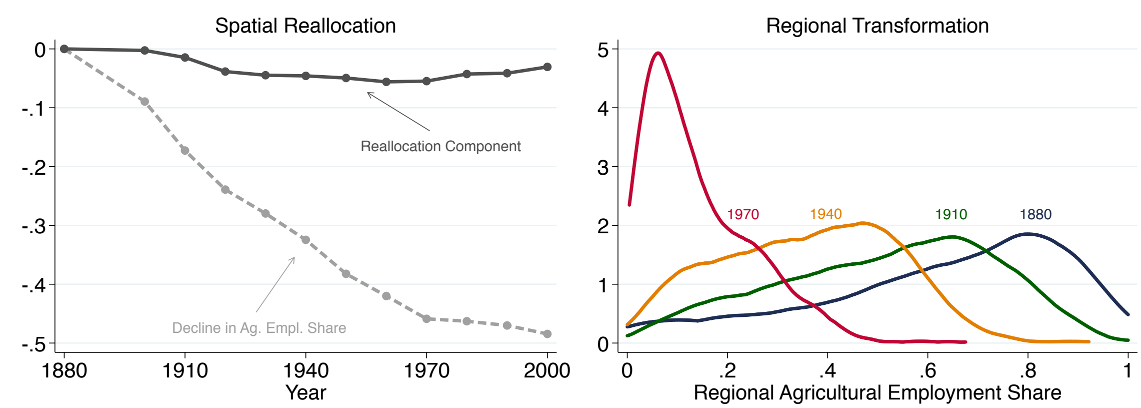
[5] Between 1880 and 2000 the aggregate agricultural employment share declined from around 50% to 2%. A model with Stone-Geary preferences can match the post-war data (see e.g. Herrendorf et al. (2013)), but has difficulties at longer time horizons as income effects vanish asymptotically. Alder et al. (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.

whereby labor reallocates from agricultural to non-agricultural labor markets. Or it can lead to a *regional transformation,* whereby agricultural employment shares decline within local labor markets. Formally, the aggregate decline in the agricultural employment share since 1880 can be decomposed as

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

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 euation (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.[6] Out of the total decline of almost 50%, only 3% are 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, i.e. $s_{At} - s_{A1880}$. The dark line shows the across labor market reallocation component highlighted in equation (1), i.e. $\sum_r s_{rA1880} \left(l_{rt} - l_{r1880} \right)$. In the right panel we show the cross-sectional distribution of agricultural employment shares between 1880 and 1970. We omit the 2000 cross-section from the right panel for expositional purposes as 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* local 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

[6] 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 of the Appendix, we replicate Figure 1 at the county and state level and show that it looks effectively identical.

had agricultural employment shares exceeding 75% in 1880, many labor markets were already much less agriculturally specialized and had agricultural employment shares below 25%. Throughout the 20th Century, there is a marked leftwards shift, whereby *all* commuting zones 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. In fact, in Section A of the Appendix we show that this local incidence of the structural transformation is not unique to the US but present in many countries around the world.

These patterns might seem surprising as they are seemingly at odds with the process of urbanization, whereby the fraction of the US population living in cities tripled from about 25% to 75% between 1880 and 2000. This, however, is not the case. In Section A of the Appendix, we replicate Figure 1 for the increase in the aggregate urbanization rate, and we show that the rise in urbanization was *also* 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 increase in the urban population share.

Figure 1 raises two question. First, why did 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 that 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 a spatial dimension by allowing for heterogeneity in regional productivity, 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 omit time subscripts when there is no risk of confusion.

Technology Each sector s produces a final good by aggregating the differentiated regional varieties with a constant elasticity of substitution σ, i.e.

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

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 the amount of capital and efficiency units of labor employed in sector s in region r at time t. Z_{rst} denotes productivity.[7] It is conceptually useful to decompose regional productivity Z_{rst} as follows

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

Here Z_{st} is an *aggregate* TFP shifter in sector s, which affects all regions proportionally. Additionally, there are *region-specific* sources of 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 equation (3), the vector $\{Q_{rs}\}_{rs}$ can be thought of as parametrizing 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 δ 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 equation (2). Letting ϕ be the share of the agricultural good in the production of investment goods, the price of the investment good is given by $p_{It} = p_{At}^{\phi} p_{NAt}^{1-\phi}$.[8] For the remainder of the paper the investment good will serve as the numeraire of our economy.

We assume that goods are freely traded so that prices are equalized across locations. As we explain in detail below, this assumption considerably 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

[7] We follow the literature and assume that capital shares are identical across sectors. Herrendorf et al. (2015) for example 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).

[8] The accompanying production function for the investment good is given by $I_t = \phi^{\phi} (1-\phi)^{\phi} X_A^{\phi} X_{NA}^{1-\phi}$, 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. For an analysis of investment-specific structural change, see Herrendorf et al. (2017).

(simple) forms of trade costs. We also 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. Finally, note that we do not explicitly include land as a factor of production in the agricultural sector. This assumption turns out to not be very restrictive and we impose it mainly to keep the production side of the economy 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. There we also argue that the implied economic forces can be captured by assuming that production in the agricultural sector is subject to decreasing returns to scale. We cover this case specifically in one of our robustness exercises in Section 6.

Labor Supply In order to focus on the novel spatial dimension of our model, we first follow the macroeconomic literature and assume that efficiency units are perfectly substitutable across sectors. Hence, it is only workers' sorting across space, which makes sectoral labor supply in the aggregate not fully elastic. We assume that individuals are heterogenous in the number of efficiency units they can provide to the market, z^i, which are drawn from a Frechet distribution, i.e. $F(z) = e^{-z^{-\zeta}}$. The parameter ζ governs the dispersion of skills across individuals and the average level of efficiency units is given by $E[z] = \Gamma(1-\zeta^{-1})$, where $\Gamma(.)$ denotes the Gamma function. For the remainder of this paper, we define $\Gamma_\zeta \equiv \Gamma(1-\zeta^{-1})$. In our quantitative analysis we allow for an upward sloping labor supply function across sectors within labor markets, which we explicitly introduce in Section 3.3 below.[9]

Demographics We phrase our analysis as an overlapping generations (OLG) economy. Individuals live for two periods. When young, individuals move to 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 remain where they are, consume the receipts of their saving decisions and have a single off-spring, who in turn has the option to migrate to a new region. For our application we think of a period as lasting 30 years.

The OLG structure is analytically very 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 as recently highlighted by Hobijn et al. (2018) and Porzio and Santangelo (2017).

To summarize, 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

[9]The individual heterogeneity 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 $\left(z^i_{NA}, z^i_A\right)$. As we will show below, all our expressions seamlessly generalize to this two-sector case. Hence, it is attractive for expositional purposes to introduce skill heterogeneity already at this point.

and (iii) an intra-temporal choice of how to allocate their spending optimally across the two consumption goods in both periods of life. To characterize optimal behavior, we solve these decision problems by backward induction and hence present them in reverse chronological order.

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.[10] We build on Boppart (2014) and assume that individual preferences are of *Price-Independent Generalized Linear* ("PIGL") form and can be represented by the indirect utility function

$$V(e,p) = \frac{1}{\eta}\left(\frac{e}{p_A^{\phi} p_{NA}^{1-\phi}}\right)^{\eta} - \frac{\nu}{\gamma}\left(\frac{p_A}{p_{NA}}\right)^{\gamma} + \frac{\nu}{\gamma} - \frac{1}{\eta}, \tag{4}$$

where e denotes total spending and $p = (p_A, p_{NA})$ is the vector of sectoral prices.[11] In particular, Roy's Identity implies that the expenditure share on the agricultural good, $\vartheta_A(e,p)$, is given by

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

and hence incorporates both income effects (governed by η) and price effects (governed by γ). For $\eta > 0$, the expenditure share on agricultural goods is declining in total expenditure. This captures the income effect of non-homothetic demand. Holding real income e constant, the expenditure share is increasing in the relative agricultural price 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 $\nu = 0$ is the Cobb Douglas case where expenditure shares are constant and equal to ϕ.[12]

The preference specification in equation (4) has advantageous aggregation properties. 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.

[10] 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).

[11] The most common choice among non-homothetic preferences is the Stone-Geary specification (see e.g. Kongsamut et al. (2001)). Alder et al. (2018) show that the Stone-Geary specification is unable to generate the large decline in agricultural employment since 1880, as the the non-homotheticity vanishes asymptotically. In contrast, they find that PIGL Preferences provide a much better fit to the data. 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 Sections B.7 and C.1 of the Appendix for a detailed discussion.

[12] Note that ϕ 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.

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

$$U_{rt}^{i} = \max_{[e_t, e_{t+1}, s]} \left\{ V\left(e_t, p_t\right) + \beta V\left(e_{t+1}, p_{t+1}\right) \right\}, \tag{5}$$

subject to

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

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. 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}.[13] We follow the literature on discrete choice models and assume that the value of a bilateral move from j to r to agent i can be summarized by

$$\mathscr{U}_{jr}^{i} = E\left[U_{rt}^{i}\right] - MC_{jr} + A_r + \kappa \nu_r^{i},$$

where $E\left[U_{rt}^{i}\right]$ 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 ν_r^{i} is an idiosyncratic preference shock, which is specific to region r and independent across locations and individuals. Furthermore, κ 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 κ, the less responsive are individuals to the fundamental value of a location r.

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

$$\rho_{jrt} = \frac{\exp\left(\frac{1}{\kappa}\left(E\left[U_{rt}^{i}\right] + A_{rt} - MC_{jr}\right)\right)}{\sum_{l=1}^{R} \exp\left(\frac{1}{\kappa}\left(E\left[U_{lt}^{i}\right] + A_{lt} - MC_{jl}\right)\right)}. \tag{6}$$

Hence, individuals migrate towards regions which offer high earnings (and hence promise large future lifetime utility $E\left[U_{rt}^{i}\right]$) and favorable 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 κ. From equation (6) we also obtain the law of motion for the spatial reallocation of

[13] 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.

workers between t and $t-1$ as

$$L_{rt} = \sum_{j=1}^{R} \rho_{jrt} L_{jt-1}, \tag{7}$$

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 extent to which people live (and work) in a different location from where they were born (see Molloy et al. (2011)).

3.2 Competitive Equilibrium

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. Finally, we show that individuals' migration choices only depend on the distribution of equilibrium wages and not on any other future equilibrium variables. 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 alone. The equilibrium sequence of interest rates can then be calculated from the dynamics of the aggregate capital stock, implied by households' savings decisions. The model can thus 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 (5). Let the optimal level of expenditure when young (old) of the generation that is born at time t be denoted by e_t^Y $\left(e_{t+1}^O\right)$. 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 (5) where $V(e,p)$ is given in equation (4). The solution to this problem is given by*

$$e_t^Y(y) = \psi(r_{t+1})\, y \tag{8}$$

$$e_{t+1}^O(y) = (1+r_{t+1})(1-\psi(r_{t+1}))\, y \tag{9}$$

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

where

$$\psi(r_{t+1}) = \left(1+\beta^{\frac{1}{1-\eta}}(1+r_{t+1})^{\frac{\eta}{1-\eta}}\right)^{-1} \quad (10)$$
$$\Lambda_{t,t+1} = -\frac{\nu}{\gamma}\left(\left(\frac{p_{At}}{p_{NAt}}\right)^{\gamma}+\beta\left(\frac{p_{At+1}}{p_{NAt+1}}\right)^{\gamma}\right)+(1+\beta)\left(\frac{\nu}{\gamma}-\frac{1}{\eta}\right).$$

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

Proposition 1 characterizes the solution to the household problem. Two 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. Moreover, they resemble the familiar OLG structure, where the individual consumes a share $\psi(r_{t+1})$ of his income when young and consumes the remainder (and the accrued interest) when old. Importantly, this share only depends on the interest rate r_{t+1} and not on relative prices p_t or p_{t+1}.[14] 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 $1/(1+\beta)$.

Secondly, lifetime utility U^i_{rt} only depends on the location r *via* individual income y^i_{rt}. 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. Moreover, utility is *additively separable* in income y_{rt} and current and future prices P_t and P_{t+1} (which determine $\Lambda_{t,t+1}$). Both of these properties allow for a tractable solution of individuals' spatial choice problem. To see this, note first that we can calculate individuals' lifetime utility $E\left[U^i_{rt}\right]$ analytically: because life-time utility is a power function of individual income y^i_r and individual income is Frechet distributed, we get that $E\left[y^{\eta}\right] = \Gamma_{\eta/\zeta} w^{\eta}_{rt}$. Together with equation (6), 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, $\mathscr{W}_{rt}$, by $E\left[U^i_{rt}\right] + A_{rt} = \mathscr{W}_{rt} + \Lambda_{t,t+1}$. Then*

$$\mathscr{W}_{rt} = \frac{\Gamma_{\eta/\zeta}}{\eta}\psi(r_{t+1})^{\eta-1} w^{\eta}_{rt} + A_{rt}. \quad (11)$$

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

$$\rho_{jrt} = \frac{\exp\left(\frac{1}{\kappa}\left(\mathscr{W}_{rt} - MC_{jr}\right)\right)}{\sum_{l=1}^{R}\exp\left(\frac{1}{\kappa}\left(\mathscr{W}_{rt} - MC_{jl}\right)\right)}. \quad (12)$$

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

[14] This is due to our assumption that nominal income e is deflated by the same price index as the investment good. This is convenient and similar to the single-good neoclassical growth model, where the consumption good and the investment good use 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.

Proposition 2 implies that individuals' migration decisions are fully captured by $\mathscr{W}_{rt}$, which is a summary measure of regional attractiveness. Note that the cross-sectional variation in $\mathscr{W}_{rt}$ results from differences in average wages, w_{rt}, and in amenities A_{rt} across labor markets. The former is endogenous and depends on the distribution of regional productivity Q_{rst}, the extent of spatial sorting and aggregate demand conditions. In particular, the structural transformation will lower relative wages in agricultural areas and hence raise the return to relocate towards non-agricultural areas. In contrast, regional amenities, while time-varying, are fully exogenous. This highlights that the spatial reallocation component of the structural transformation is larger, if the correlation of the initial agricultural employment share s_{rAt-1} and *future* wages w_{rt} and amenities A_{rt} is negative and the elasticity of moving flows with respect to such fundamental differences is large, i.e. κ is small.

Importantly, this expression does *not* feature $\Lambda_{t,t+1}$, which is constant across locations and hence does not determine spatial labor flows. This implies that agents' spatial choice problem reduces to a static decision problem, which depends only on current, not future, equilibrium objects. This 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 $i \in \mathscr{S}$, with spending e_i. The demand for agricultural products by this set of consumers is given by

$$PC^A_{\mathscr{S}} = \int_{i\in\mathscr{S}} \vartheta_A(e_i,p)\, e_i di = \left(\phi + \nu\left(\frac{p_A}{p_{NA}}\right)^{\gamma} \int_{i\in\mathscr{S}} e_i^{-\eta}\omega_i di\right) E_{\mathscr{S}},$$

where $E_{\mathscr{S}} = \int_{i\in\mathscr{S}} e_i di$ denotes aggregate spending and $\omega_i = e_i/E_{\mathscr{S}}$ 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 $E_{\mathscr{S}}$ 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 choices (and hence in 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 $\mathscr{S}_r^g$ for $g = Y, O$ be the set of young and old consumers in region r respectively. The* *aggregate expenditure share on agricultural good of the set of consumers $\mathscr{S}_r^g$ is given by*

$$\vartheta_A\left([e_i]_{i\in\mathscr{S}_r^g}, p\right) \equiv \frac{PC^A_{\mathscr{S}_r^g}}{E_{\mathscr{S}_r^g}} = \phi + \tilde{\nu}\left(\frac{p_A}{p_{NA}}\right)^{\gamma} E_{\mathscr{S}_r^g}^{-\eta}, \tag{13}$$

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

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

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

$$\vartheta_t^A \equiv \frac{PC_t^A + \phi I_t}{PY_t} = \phi + \tilde{\nu}\left(\frac{p_A}{p_{NA}}\right)^{\gamma} \frac{\sum_{r=1}^R \left(E_{\mathscr{S}_r^Y}^{1-\eta} L_{rt} + E_{\mathscr{S}_r^O}^{1-\eta} L_{rt-1}\right)}{PY_t}. \tag{14}$$

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 (13) shows that the aggregate demand of a given set of consumers resembles that of a representative consumer with mean spending $E_{\mathscr{S}}$ and an adjusted preference parameter $\tilde{\nu}$. Because the linearity of individuals' policy functions allows to express aggregate spending directly as a function of equilibrium local wages, aggregate sectoral spending 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 (14). Note that ϑ_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, ϑ_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

Two central 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

$$L_{rt}\Gamma_\zeta w_{rt} s_{rAt} = (1-\alpha)\pi_{rAt}\vartheta_t^A PY_t. \tag{15}$$

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 in aggregate spending on agricultural goods $\vartheta_t^A PY_t$, π_{rAt}. The CES structure of consumers' preferences implies that these regional trade shares are given by

$$\pi_{rAt} = \left(\frac{p_{rAt}}{p_{At}}\right)^{1-\sigma} = \frac{\left(Q_{rAt} w_{rt}^{\alpha-1}\right)^{\sigma-1}}{\sum_{j=1}^{R}\left(Q_{jAt} w_{jt}^{\alpha-1}\right)^{\sigma-1}}, \tag{16}$$

i.e. they neither depend on the identity of the sourcing region, nor on the equilibrium capital rental rate R_t or the common component of productivity Z_{At}. Rather, 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 (15), as

$$L_{rt}\Gamma_{\zeta} w_{rt} = (1-\alpha)\left(\pi_{rAt}\vartheta_t^A + \pi_{rAt}\left(1-\vartheta_t^A\right)\right) PY_t \tag{17}$$

$$\frac{s_{rAt}}{1-s_{rAt}} = \frac{\pi_{rAt}}{\pi_{rNAt}}\frac{\vartheta_t^A}{1-\vartheta_t^A}. \tag{18}$$

Equation (17) shows that total earnings in region r 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 ϑ_t^A tends to reduce regional earnings in locations who have a comparative advantage, in the agricultural sector. Moreover, equation (18) illustrates the spatial co-movement of sectoral employment shares. Because regional (scaled) agricultural shares $s_{rAt}/(1-s_{rAt})$ are proportional to the aggregate (scaled) agricultural expenditure share $\vartheta_t^A/(1-\vartheta_t^A)$, a decline in the aggregate spending share ϑ_t^A tends to reduce agricultural employment shares in *all* locations.

These equations also 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. $(1-\alpha)PY_t = \Gamma_{\zeta}\sum_r L_{rt}w_{rt}$ and that the spatial labor supply function is only a function of the spatial distribution of wages. Likewise the agricultural value added share ϑ_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

$$K_{t+1} = (1-\psi(r_{t+1}))\sum_r \Gamma_{\zeta} w_{rt} L_{rt} = (1-\psi(r_{t+1}))(1-\alpha)PY_t, \tag{19}$$

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. A dynamic equilibrium then requires the sequence of interest rates, $\{r_t\}_t$, to be consistent with the evolution of the capital stock implied by equation (19).

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 $\{e^Y_{rt}, e^O_{rt}, s_{rt}\}_{rt}$, and demands for regional varieties $\{c_{rst}\}_{rst}$ such that consumers' choices $\{e^Y_{rt}, e^O_{rt}, s_{rt}\}_{rt}$ maximize utility, i.e. are given by equations (8) and (9), the demand for regional varieties follows equation (16), firms' factor demands maximize firms' profits, markets clear, the capital stock evolved according to equation (19) and the allocation of people across space $\{L^Y_{r,t}\}$ is consistent with individuals' migration choices in equation (6).*

To see how "space" and the process of structural change interact, it is instructive to consider a special case of our model, where space does *not* play any interesting role. In particular, consider a parametrization, where (i) there are no moving costs $\left(MC_{jr} = 0\right)$, (ii) spatial productivities are constant $(Q_{rst} = Q_{rs})$, (iii) there are no amenity differences $(A_{rt} = 0)$ and (iv) individuals have no idiosyncratic preferences for particular locations $(\kappa \to 0)$. These assumptions imply that equilibrium wages (and individual welfare) are equalized across space at each point in time. In this case, our model reduces to a standard, macroeconomic model of the structural transformation augmented by a spatial layer. While the macroeconomic aggregates affect the spatial allocations, there is no feedback from space to the macroeconomy.

To see this, we show in Section B.4 in the Appendix that aggregate GDP in this model is given by an aggregate production function

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

where $Z_t \equiv \Gamma_{\zeta}^{1-\alpha} Z_{At}^{\phi} Z_{NAt}^{1-\phi}$ and $L = \sum_r L_{rt}$. Moreover, it can be shown that - given some initial condition K_0 and processes for productivity $\{Z_{At}, Z_{NAt}\}_t$ - there exists a unique dynamic equilibrium path of capital $\{K_t\}_t$. Furthermore, as in the baseline, aggregate 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)). In particular, 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. The agricultural share in value added ϑ_t^A is then given by

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

where χ is a constant, which is a simple function of exogenous parameters. This relative demand system again resembles a representative household with PIGL preferences. Finally, sectoral employment shares are equal to sectoral value added shares, i.e. $s_{At} = \vartheta_t^A$, so that value added per worker is equalized across sectors.

While its equilibrium path is independent of the spatial microstructure, this model nevertheless has strong implications for the allocation of labor across space. In particular, equilibrium trade shares are given by $\pi_{rst} = Q_{rs}^{\sigma-1}$, i.e. are fully exogenous and only depend on regional productivity (see equation (16)). Moreover, equation (17) implies that the size of the local *population* is a demand-weighted average of the constant local sectoral productivities

$$L_{rt} = Q_{rA}^{\sigma-1} \vartheta_t^A + Q_{rNA}^{\sigma-1} \left(1 - \vartheta_t^A\right). \tag{20}$$

Equation (20) concisely summarizes the *urban bias of the structural transformation*: A decline in the spending share on agricultural goods reduces population in all regions that have a comparative advantage in agricultural goods and increases the size of non-agricultural localities.[15] Moreover, because the structural transformation is the only reason for individuals to relocate, regional population growth and the initial agricultural employment share are predicted to be perfectly negatively correlated.

Importantly, without spatial frictions, the macroeconomic forces of structural change affect the spatial allocation of factors, but all aggregate allocations can be characterized independently of the spatial microstructure. Like the allocation of resources across sectors, space becomes an inconsequential layer that is determined residually from the macroeconomic dynamics which are akin to the single-sector neoclassical growth model. Both the structural transformation across sectors and its spatial implications are of secondary interest for our understanding of the process of economic growth. We therefore refer to this parametrization of our model also as the *Quasi-Spaceless Economy.*

This separability between the macroeconomic and spatial allocations breaks down if moving costs or regional amenities generate spatial dispersion in the marginal product of labor. In this case, aggregate structural change and the spatial allocation of factors are jointly determined. In particular, *spatial* considerations also have implications for the *sectoral* allocation of resources as average products are no longer equalized. Relative value added per worker in agriculture is for example given by

$$\frac{VA_t^A / L_t^A}{VA_t / L_t} = \frac{\vartheta_t^A}{s_{At}} = \frac{\sum_r s_{rAt} \frac{w_{rt} L_{rt}}{\sum_r w_{rt} L_{rt}}}{\sum_r s_{rAt} \frac{L_{rt}}{\sum_{rt} L_{rt}}}.$$

Hence, the agricultural sector has low productivity (i.e. the economy suffers from an "agricultural productivity gap" as in Gollin et al. (2014)), whenever the spatial correlation between wages and agricultural employment share is *negative.* And because the structural transformation puts continuous downward pressure on wages in agricultural areas, value added per worker in the agricultural areas might remain low despite the large extent of reallocation out of the agricultural sector. In Section 5.2 we will explicitly compare the aggregate implications of our calibrated model of spatial structural change with the Quasi-Spaceless Economy.

[15] In fact, it is easy to show that $\operatorname{sgn}\left(Q_{rAt}^{\sigma-1} - Q_{rNAt}^{\sigma-1}\right) = \operatorname{sgn}\left(s_{rAt} - s_{At}\right)$. To see this, note that $s_{At} = \vartheta_t^A$ and $Q_{rst}^{\sigma-1} = \pi_{rst} = s_{rst} L_{rt} / \vartheta_t^s$.

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 our model 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 crucially determines the costs of the regional transformation. While moving costs are a hurdle for 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 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. Such sorting behavior is a strong feature of the data.[16] Allowing for such skill differences in the theory allows us to explicitly calibrate our model to be consistent with the spatial distribution of sectoral employment patterns and skill shares.

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 $z^i = \left(z^i_A, z^i_{NA}\right)$ 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^i_s is drawn independently from a Frechet distribution $F^h_s(z) = e^{-\Psi^h_s z^{-\zeta}}$, where Ψ^h_s parametrizes the average level of human capital of individuals of skill type h in sector s. Without loss of generality we parametrize Ψ^h_s as $\Psi^L_A = \Psi^L_{NA} = 1$, $\Psi^H_{NA} = \mu q$ and $\Psi^H_A = q$. Here, q measures the absolute advantage of skilled individuals and μ 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 λ and assume that it is constant.[17] In contrast, the spatial allocation of human capital, i.e. the share of skilled workers in region r at time t, λ_{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. 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 rather *average* earnings of individuals in skill group h, which are given by

$$E^h\left[y^i_r\right] = \Gamma_\zeta \Theta^h_r \quad \text{where} \quad \Theta^h_{rt} = \left(\Psi^h_A w^\zeta_{rAt} + \Psi^h_{NA} w^\zeta_{rNAt}\right)^{1/\zeta}. \tag{21}$$

While Θ^h_{rt}, 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 $\mathscr{W}^h_{rt} =$

[16] See Figure 13 in the Appendix

[17] 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.

$\frac{\Gamma_{\eta/\zeta}}{\eta}\psi\left(r_{t+1}\right)^{\eta-1}\left(\Theta_{rt}^{h}\right)^{\eta}+A_{rt}$. Hence, high skilled individuals put a higher relative weight on non-agricultural wages w_{rNAt} 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 ρ_{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.

4 Spatial Structural Change in the US: 1880 - 2000

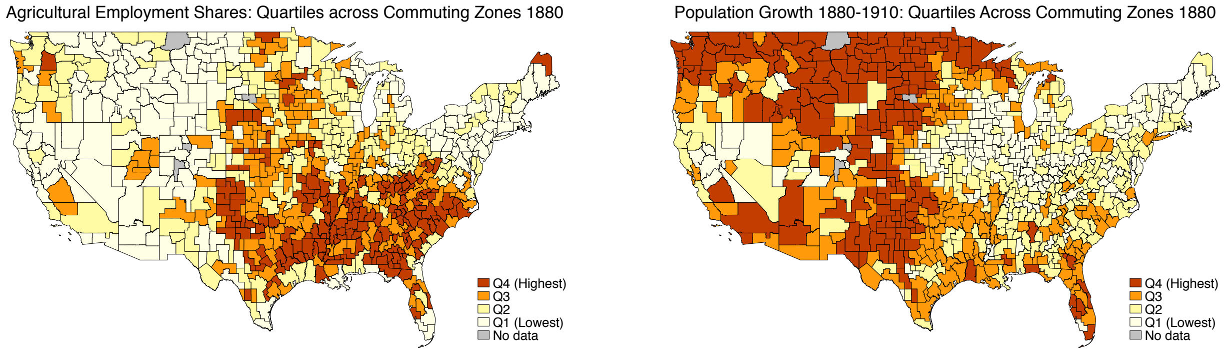
We now apply this theory to the experience of the United States over the last 120 years. To do so we construct a novel panel data set on the regional development of the United States between 1880 and 2000. We describe the data in Section 4.1. One of the main features of our data collection effort is 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, i.e. that agricultural labor markets saw their relative earnings decline. Finally, we calibrate the model in Section 4.3. Our quantitative analysis is contained in Section 5.

4.1 Data

We combine 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 employment $\{L_{rt}\}_{rt}$, average manufacturing earnings $\{w_{rt}\}_{rt}$, and sectoral employment shares $\{s_{rAt}\}_{rt}$ for all US counties between 1880 and 2000.[18] We define the agricultural sector to comprise agriculture, fishing and mining 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 data on total manufacturing payrolls and manufacturing head counts. To estimate the moving cost parameters we exploit information on lifetime mobility as, in the model, individuals move once in their lifetime to access their preferred labor market. All censuses between 1880 and 2000 contain information on the state of residence and the state of birth. Section E in the Online Appendix 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

[18] Twelve states have not obtained statehood in 1880. These are (with the year of their accession in parentheses): 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).



Notes: The figure shows the agricultural employment shares in 1880 (left panel) and regional population growth between 1880 and 1910 (right panel) across US commuting zones.

Figure 2: Agricultural Specialization and Population Growth in 1880

a regional labor market (see Tolbert and Sizer (1996)).[19] We choose commuting zones 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 continental territory of the United States into 712 polygons. For the main calibration of the model we assume that a period is 30 years and hence employ the cross-sections 1880, 1910, 1940, 1970 and 2000. We normalize the size of the total US workforce to unity in each period.

In Figure 2 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 Northeastern states like Massachusetts or New York already had agricultural employment shares of less than 10%, many commuting zones in the South had more than 75% of their population employed in the agricultural sector. In the right panel, 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. This already suggests that considerations other than the structural transformation were important for the observed migration patterns. Below we will use our model 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 information on earnings and education is available. We use this information to calibrate the spatial distribution of skilled workers.

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-

[19] To do so, we construct a crosswalk between counties and 1990 commuting zones for every decade between 1880 and 2000. We 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)).

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 The Urban Bias of Structural Change: Direct Evidence

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

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

where $\overline{w}_{rt}^{M}$ 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 log agricultural employment share in region r and Δs_{At+1} denotes the change in the *aggregate* agricultural employment share between t and $t+1$. The coefficient α captures the direct effect of agricultural specialization on manufacturing earnings growth. The coefficient β captures the urban bias. In particular, we expect β to be positive: the larger the decline in the agricultural share, the more adversely will regions with a comparative advantage in agriculture be affected.

The results are reported in Table 1. In the first column, for consistency with Figure 2, 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 1 shows that regions with a higher agricultural share in 1880 experienced lower earnings growth in the subsequent 30 years. Column 2 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 β 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 the aggregate agricultural employment share 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.

	Dep. variable: ln manufacturing earnings			
	1880-1910	Full Sample		
$\ln s_{rAt1}$	-0.040***	-0.106***	-0.087***	-0.063***
	(0.015)	(0.006)	(0.007)	(0.008)
$\ln s_{rAt1} \times \Delta s_{At+1}$			0.243***	0.247**
			(0.094)	(0.099)
lagged ln man earnings	0.118***	0.120***	0.117***	0.154***
	(0.022)	(0.018)	(0.018)	(0.021)
Year FE		✓	✓	✓
State FE	✓	✓	✓	✓
State × 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. $s_{rA,t}$ is the agricultural share of region r at time t. Δs_{At+1} is the change in aggregate agricultural employment share between t and $t+1$.

Table 1: The Urban Bias of Structural Change

The quantitative impact is, however, limited. To see this, consider Figure 3, where we depict the cross-sectional correlation between agricultural employment shares in 1880 and average manufacturing earnings in 1910. While there is a 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 care about the sectoral composition of a location only in as far as it offers higher life-time utility through favorable factor prices. And Figure 3 suggests that the cohort born in rural areas in 1880 did not necessarily have to move towards non-agricultural places to increase their life-time earnings.

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 indicates the size of commuting zones as measured by their population in 1880.

Figure 3: Agricultural Specialization and Future Earnings

4.3 Calibration

In this section we describe the calibration of our model. First we discuss our calibration strategy. We then turn to 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: $\{Z_{At}, Z_{NAt}\}_t$ 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.[20] We calibrate the time series of aggregate sectoral productivities, $\{Z_{At}, Z_{NAt}\}_t$, to match the evolution of relative prices and a GDP growth rate of 2%.

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). 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 Q_{rAt}/Q_{rNAt} while the level of wages along with the total number of workers informs the level of Q_{rNAt}. The vector of amenities $\{A_{rt}\}_{rt}$ can then be inferred from the observed net population flows.

Moving costs and idiosyncratic location preferences: MC_{jr} **and** κ 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, letting d_{jr} be the distance between j and r, we assume that

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

whenever $j \neq r$ and zero otherwise. Here, $\tau > 0$ parametrizes the fixed cost of moving and δ_1 and δ_2 govern how mobility costs vary with distance. We normalize d_{jr} so that the maximum distance in the US is 1. We calibrate these parameters by matching salient features of the data on 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 from where they were born.

[20] Note that Z_{At} and Z_{NAt} do not grow at a constant rate. Because the reallocation of labor across locations has direct effects on aggregate productivity, a constant rate of GDP growth and a constant interest rate is inconsistent with constant aggregate productivity growth. See also Herrendorf et al. (2017) for a related discussion. In our counterfactual analysis, interest rates will of course be free to vary over time.

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 the importance of locational preferences (κ) and the parameters of mobility costs (τ, δ_1, δ_2) so as to make the model best fit the observed state level flows. To capture employment-related mobility, we focus on workers between 26 and 50 years of age and chose τ 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.[21]

Preference parameters: η, γ, ϕ We estimate the strength of the non-homotheticity in the demand system (η) 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 γ determines the price elasticity of demand, we discipline γ with the elasticity of substitution between agricultural and non-agricultural goods. 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 ϕ and ν. Given that our model matches the joint distribution of agricultural employment shares and population size across space perfectly, internal consistency requires us to also match the time series of the aggregate agricultural employment share exactly. 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.[22] We therefore allow the parameter ϕ to be time-specific to fully account for the residual decline in agricultural employment and choose ν to minimize the required time-variation in ϕ_t. Intuitively, ν is chosen for the model to explain as much of the aggregate process of structural change as possible, given the income and price elasticities η and γ. Recall that ϕ does not enter the household's decision problem directly.

Skill supply: ζ, μ, q and $\{\lambda_{r1880}\}_r$ To parametrize the skill supply, we need values for the supply elasticity (ζ), the comparative and absolute advantage of skilled workers (q and μ) and the initial distribution of skilled workers across space in 1880, $\{\lambda_{r1880}\}_r$. We define skilled individuals as workers who completed at least high school in 1940 and hold the aggregate share of skilled workers fixed.[23] This choice yields an aggregate skilled employment share of about 0.3. We then calibrate $\{\lambda_{r1880}\}_r$ for the

[21] 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 τ. As shown in Section D.6 in the Appendix, the number of stayers in a commuting zone is a monotone function in τ given $(\kappa, \delta_1, \delta_2)$.

[22] 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.

[23] 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.

model to exactly replicate the spatial skill distribution in 1940, which is the only year for which educational attainment at the commuting zone level is directly observable. We calibrate the dispersion of efficiency units ζ, 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 $(\pi^2/6)\zeta^{-2}$. We therefore identify ζ from $\zeta = (\pi/6^{1/2})\text{var}\left(\hat{u}^i_{rsh}\right)^{-1/2}$, where var $\left(\hat{u}^i_{rsh}\right)$ 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 μ 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 non-agricultural intensive localities.

Parameter		**Target**	**Value**	**Moments**	
				Data	Model
		Skill Supply			
ζ	Skill heterogeneity	Residual Earnings variance in 1940	1.62	0.62	0.62
ξ	Share of skilled individuals	Share with at least high school in 1940	0.3	0.3	0.3
μ	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
$[\lambda_{r1880}]$	Initial distr. of skilled individuals	Spatial skill distribution in 1940	.	$\{\lambda_{r1940}\}$	Accounting
		Regional Fundamentals			
$[Q_{rAt}]_{rst}$	Agricultural productivity	Regional empl. shares and earnings	See Appendix	$\{e^{Man}_{rt}, s_{rAt}\}$	Accounting
$[Q_{rNAt}]_{rst}$	Non-agricultural productivity				
$[A_{rt}]_{rt}$	Amenities	Net migration flows		$\{L^h_{rt}\}$	
		Time Series Implications			
$[Z_{NAt}]$	Non-agricultural productivity	Aggregate growth rate of GDP pc	See Appendix	2%	2%
$[Z_{At}]$	Agricultural productivity	Relative price of ag. goods	See Appendix	$\{p_{At}/p_{NAt}\}$	Accounting
		Preference Parameters			
β	Discount rate	Investment rate along the BGP	0.29	0.15	0.15
ϕ	Ag. share in price index	Time series of ag. empl. share	See Appendix	$\{s_{At}\}$	$\{s_{At}\}$
ν	PIGL Preference parameter	Time series of ag. empl. share	See Appendix	$\{s_{At}\}$	$\{s_{At}\}$
η	Non-homotheticity	Ag.share - expenditure relationship	0.32	Estimated with NLS	
γ	Price sensitivity	Elasticity of substitution in 2000	0.35	0.7	0.7
		Moving Costs and Mobility			
τ	Fixed costs of moving	Lifetime interstate migration rate	1.63	0.32	0.32
κ	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	
		Other parameters			
δ	Depreciation rate (over 30 years)	set exogenously	0.91 (0.08 pa)	.	
α	Capital share in production function	Aggregate capital share	0.33	.	.
σ	Elasticity of substitution	set exogenously	4	.	.

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

Table 2: Structural Parameters

Other parameters: β, δ, σ and α We estimate the rate of time preference β from aggregate macro relationships. In particular, we chose β to be consistent with the aggregate rate of investment. The capital share α is set to match an aggregate capital share of 0.33. The elasticity of substitution between regional varieties, σ, is set to 4. In Section 6 below and in the Online Appendix we provide extensive robustness checks for this parameter.[24] Finally, the depreciation rate δ is set to a 0.08 at annual frequency, 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.

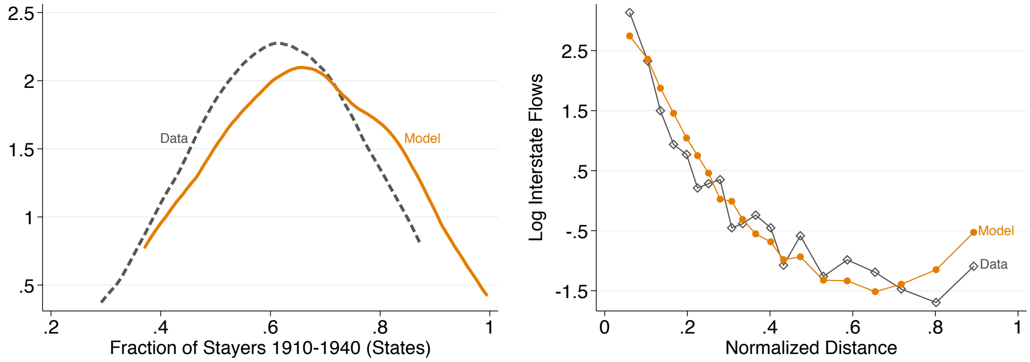
The presence of spatial mobility costs is an important component of our theory. To see that our model matches important features of the lifetime migration data well, we display the distribution of stayers across states and the relationship between moving flows and distance in Figure 4. In the left panel we depict the cross-sectional distribution of the share of "stayers", i.e. ρ_{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 4 shows that the model matches this cross-sectional heterogeneity, even though 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 α_j and β_r are origin and destination fixed effects. We then plot the estimated residual $\hat{u}_{jr}$ as a function of distance for the data and the model. The right panel of Figure 4 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.[25]

[24] 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, σ would be expected to be lower. We consider values for $\sigma \in [3, 9]$.

[25] The increase in moving flows at large distances stems from the fact that there are sizable coast-to-coast flows.



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 $\log\left(\rho_{jr}/\left(1-\rho_{jr}\right)\right) = \alpha_j + \beta_r + u_{jr}$, where ρ_{jr} denotes the share of people moving from j to r and α_j and β_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 4: Lifetime State-to-State Migration: Model vs Data

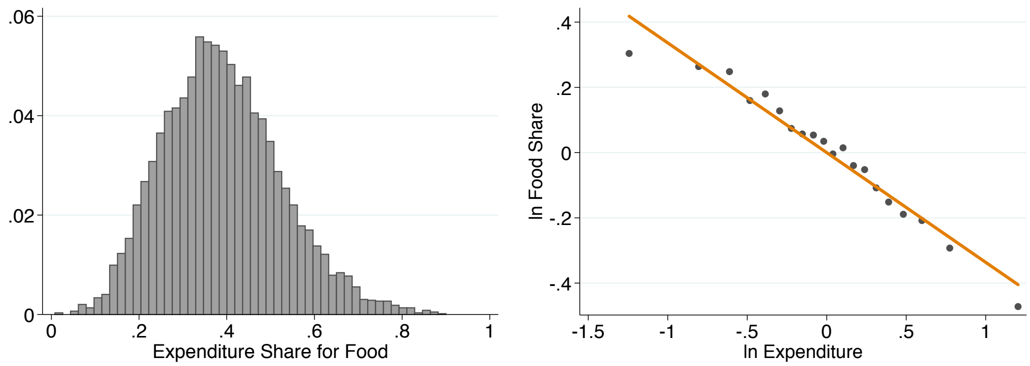
Bivariate correlations		
$\rho\left(\ln Q_{rNAt}, \ln Q_{rAt}\right)$	$\rho\left(\ln Q_{rNAt}, A_{rt}\right)$	$\rho\left(\ln Q_{rAt}, A_{rt}\right)$
0.225	0.492	0.152

Notes: The 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 $\{Q_{rAt}, Q_{rNAt}\}_{rt}$ and amenities $\{A_{rt}\}_{rt}$ as structural residuals (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_{rst}, 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 to be highly correlated with *changes* in $\{Q_{rA}, Q_{rNA}\}_r$ as inferred from our model. See Section B.9 in the Appendix for details.

Our model of consumer preferences 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 5 we depict the cross-sectional distribution of the expenditure share for food across US households in 1935. It is apparent that 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 η. 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 $\eta = 0.32$.[26]



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.

Figure 5: Expenditure Shares on Food in 1936

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. This suggests the presence of important offsetting factors counteracting the urban bias of the structural transformation.

[26] 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 η is precisely estimated and - depending on the specification - between 0.3 and 0.34.

Our model suggests four main margins. 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 the Quasi-Spaceless Model, characterized in Section 3, where *only* the urban bias channel is present.[27]

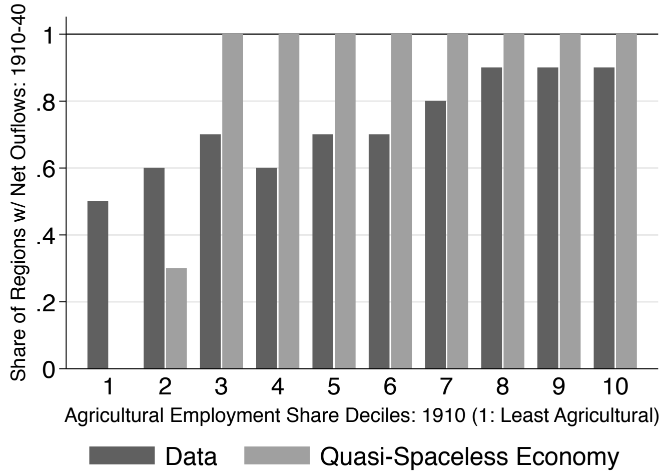
Spatial Reallocation in the Quasi-Spaceless Model In the Quasi-Spaceless Model, population mobility is only generated by the sectoral shift of spending. Hence, while implying the same decline in the aggregate agricultural share as our baseline model, this model has strikingly different implications for the link between initial agricultural employment shares and subsequent population flows. In Figure 6 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 the lowest agricultural employment shares in 1910, about 50% experience net population outflows.[28] This is very different for the Quasi-Spaceless Model, where agricultural specialization and population outflows are perfectly aligned. In particular, this model implies that the *only* regions experiencing population inflows are the roughly 17% of commuting zones with the lowest agricultural share in 1910.[29] Hence, if only the urban bias had been at play, the structural transformation would have induced much more population growth in non-agricultural, urban localities.

As a result, the spatial reallocation of individuals explains about one third of the aggregate decline in agricultural employment. This is seen in the left panel in Figure 7. Conversely, while the Quasi-Spaceless Model *overestimates* the role of spatial reallocation, it *underestimates* the extent of the structural transformation at the local level. In particular, in this counterfactual economy, 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 the year 1970. In the data (the dashed line), the vast majority of commuting zones have an agricultural employment share below 20%. In the Quasi-Spaceless

[27] Recall that this model was characterized by the absence of spatial frictions (i.e. $MC_{jr} = \kappa = A_r = 0$) and no changes in spatial fundamentals $(Q_{rst} = Q_{rs1880})$. In practice we set κ to a very small positive number for computational reasons.

[28] Note that Figure 6 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.

[29] Note that this strong form of sorting is implied by (20): 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: The figure reports the share of regions within 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 Quasi-Spaceless Model, which has no moving costs $\left(MC_{jr}=0\right)$, no changes in regional fundamentals $\left(Q_{rst}=Q_{rs1880}\right)$, no amenities $\left(A_{rt}=0\right)$ and no idiosyncratic preferences for particular locations ($\kappa \rightarrow 0$).

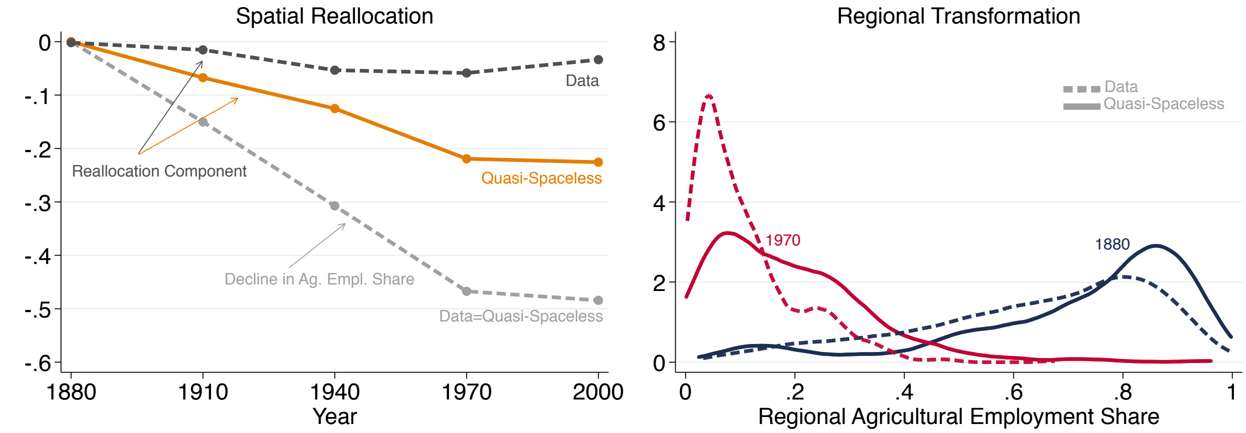
Figure 6: Agricultural Specialization and Population Outflows

Model, there is a substantial number of regions, where 30%-40% of the workforce are still employed in the agricultural sector. The reason 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 Quasi-Spaceless Model and our baseline calibration in the first two rows in Table 4. As seen in Figure 7, the spatial reallocation component is equal to 15% and hence five times as large as in the data. 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.

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 Quasi-Spaceless Model and of removing it from our baseline model. The results are contained in the lower part of Table 4.

Rows 3 and 4 show that the evolution of spatial productivity was the main reason for the low correlation between agricultural specialization and population outflows. If we augment the Quasi-Spaceless Model



Notes: In the left panel we show the aggregate agricultural employment share (light grey dashed line) and the across labor market reallocation component highlighted in equation (1), i.e. $\sum_r s_{rA1880}(l_{rt} - l_{r1880})$, for the Quasi-Spaceless Model (solid orange line) and the data (dark grey dashed line). In the right panel we show the cross-sectional distribution of agricultural employment shares in 1910 and 1970, both in the Quasi-Spaceless Model (solid lines) and in the data (dashed lines).

Figure 7: Spatial Reallocation and Regional Transformation in the Quasi-Spaceless Model

	Model ingredients				*Migration patterns*		
	Spatial productivity shocks	Moving costs	Amenities	Idiosync. location preferences	Spatial Reallocation Component	Net migration rate	Turnover (gross/net)
Quasi-Spaceless Model	✗	✗	✗	✗	-15.2%	14%	1
Baseline model	✓	✓	✓	✓	-3.2%	12.9%	2.8
Decomposition: The partial effect of ...							
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 parametrization 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 rate relative to the net migration rate. Both rates are measured at the state level. The first two rows contain the results for the Quasi-Spaceless Model and the baseline model. The Quasi-Spaceless Model 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. $\kappa \to 0$). 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

with the observed process of spatial productivity, the implied spatial reallocation component would be -3.6%. This is almost exactly the same number as observed in the data. Similarly, if we start from the baseline model but abstract from regional productivity changes, the impact of spatial reallocation would have doubled from 3.2% to 6.5%. The reason is that 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 Quasi-Spaceless Model does reduce the role of spatial reallocation, it does so for the wrong reason: the implied *level* of migration is essentially zero. More interestingly, row 6 shows that without moving costs our baseline model would predict the spatial reallocation component to 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. If anything, the correlation between agricultural shares and population growth had been positive! 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.

One final important difference between our baseline model and the Quasi-Spaceless Model 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 Quasi-Spaceless Model, where this ratio is 1, as gross and net flows coincide.[30]

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 struc-

[30]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.

		Baseline Model	Counterfactuals	
			Quasi-Spaceless Model	No Spatial Mobility
Decline in Agricultural Employment Share		-48.1%	-48%	-47.8%
Net migration rate		13%	14%	0
Spatial Reallocation Component		-3.2%	-15.2%	0
			Aggregate Productivity	
Increase in GDP pc (rel. to 1880)		10.78	11.04	8.89
Difference to baseline			2.4%	-17.5%
Agricultural productivity Gap		0.83	0.79	0.73
			Spatial Welfare Inequality	
Low Skilled Workers	1910	157%	0	202%
	Δ 1910 - 2000	-35%	-	-13%
High Skilled Workers	1910	149%	0	205%
	Δ 1910 - 2000	-38%	-	-24%

Notes: The model with no spatial mobility corresponds to the baseline model except that individuals are not allowed to move (i.e. $MC_{jr} \to \infty$). The Quasi-Spaceless Model assumes that moving costs are zero (i.e. $MC_{jr} = 0$), that there are no amenities ($A_{rt} = 0$) and no idiosyncratic tastes (i.e. $\kappa = 0$) and that 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 $\sum_r s_{rA1880} (l_{r2000} - l_{r1880})$. The "Agricultural Productivity Gap" is measured as aggregate value added per worker in agriculture relative the rest of the economy, i.e. $\left(VA_t^A / L_t^A\right) / \left(VA_t / L_t\right)$, in the year 2000. For the calculation of the spatial welfare inequality measure we refer to the main text.

Table 5: Aggregate Implications of Spatial Structural Change

tural 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 these consequences of spatial structural change, we therefore compare our model with the Quasi-Spaceless Economy, where space and the structural transformation do not interact. For comparison we also consider a model without any labor mobility by making the costs of moving prohibitively high.

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 2 and 3.

In the lower panel, we report the implications for aggregate GDP growth. If spatial mobility had been only based on regional earnings and mobility was free, income per capita would have been roughly 2.5% higher. Hence, the local nature of the structural transformation did interfere with allocative efficiency. Nevertheless, quantitatively, these efficiency gains seem modest.[31] This, however, does not imply that

[31]To put these results into perspective, Bryan and Morten (2017), using a static spatial equilibrium model, estimate higher

migration was unimportant for the US economy during the structural transformation.[32] Without labor mobility, income per capita would have been lower by almost 17.5%. Hence, spatial mobility contributed heavily to aggregate productivity growth during the transition away from agriculture.

These differences in allocative efficiency are also reflected in *sectoral* productivity gaps. Both the baseline model and the model without spatial frictions, for example, imply an agricultural productivity gap of around 20% even though there are no sectoral frictions. These productivity differences are to a large extent 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 between agricultural and non-agricultural regions and workers sort less effectively.

The Distributional Effects of Spatial Structural Change We now turn to the evolution of spatial welfare inequality. As in Proposition 2 we focus on the expected life-time value of being in location r, $\mathscr{W}_{rt}^{h}=E^{h}[U_{r}]+A_{rt}$, as our measure of regional welfare for skill group h. Our model implies that both regional income Θ_{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 spatial bias of the structural transformation is an additional inequality-enhancing force. At the same time, labor mobility tends to keep wage disparities in check and changing regional productivities are inequality-reducing as they show mean-reversion.

To quantify the evolution of spatial welfare inequality during the 20th century, we convert utility differences into "life-time-income" equivalents. Specifically, let Δ_{t}^{h} denote the interquartile range of $\mathscr{W}_{rt}^{h}$ 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 utility of living in region r by Δ_{t}^{h}, i.e. $\mathscr{W}_{rt}^{h}\left(T_{rt}^{h}y\right)\equiv\mathscr{W}_{rt}^{h}(y)+\Delta_{t}^{h}$. Our measure of spatial inequality for individuals of skill h is then given as the cross-regional average of T_{rt}^{h}, i.e. $T_{t}^{h}=\frac{1}{R}\sum T_{rt}^{h}$.

We report the results in the lower panel of Table 5. Spatial welfare inequality decreased substantially during the structural transformation. In 1910, low skilled workers in 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. Hence, the dispersion in spatial welfare declined by about 35% since 1910. The corresponding numbers for high skilled individuals are quantitatively quite similar.[33]

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. If we recalibrate our model under the assumption that $\alpha=0.05$, i.e. that labor is (essentially) the sole factor of production, we find that relative GDP the Quasi-Spaceless Model is 13% higher. This is comparable to the results reported in Bryan and Morten (2017).

[32] Note that agricultural productivity gap is slightly higher in the Quasi-Spaceless Model. The reason is that model implies more spatial sorting of skilled workers.

[33] The decline in inequality is not driven by changes in the distribution of regional amenities. The dispersion in amenities is

Spatial mobility was a crucial factor for this decline in welfare inequality. Without any spatial frictions, welfare would obviously be equalized at each point in time. This is the case in the Quasi-Spaceless Economy. Conversely, without any spatial mobility since 1880, both the level of welfare inequality would have been higher and the decline would have been much smaller. 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 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 parametrizations 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 Quasi-Spaceless 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.[34] Hence, the spatial reallocation component and the evolution of GDP is identical across specifications.

roughly constant between 1910 and 2000 and spatial welfare differences had declined even in the absence of amenities.

[34] 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
		$\sigma = 3$	$\sigma = 8$	$\zeta = 1.4$	$\zeta = 4$	$\omega_A = -0.1$	$\omega_{NA} = 0.03$	$\omega_{NA} = \omega_A = 0.03$	$\iota_1 = 0.86$ $\iota_2 = 0.19$
Spatial Reallocation Component									
Calibrated	-3.2%	-3.2%	-3.2%	-3.1%	-3.2%	-3.2%	-3.2%	-3.2%	-3.2%
Quasi-Spaceless	-15.2%	-14%	-17.2%	-15.4%	-13.4%	-15.2%	-15.7%	-15.7%	-14.1%
Decomposition: The partial effect of									
Productivity	-6.6%	-6.3%	-7%	-6.6%	-5.5%	-6.6%	-6.6%	-6.6%	-6.4%
Moving Costs	-0.7%	-0.4%	-1.6%	-0.8%	-0.5%	-0.6%	-0.8%	-0.8%	0.1%
Amenities	2.6%	2.1%	3.9%	3%	2%	2.6%	2.8%	2.9%	2.1%
Change in GDP in 2000									
No Mobility	-17.6%	-21.5%	-11.3%	-17.6%	-17.2%	-17.6%	-17.6%	-17.6%	-17.4%
Quasi-Spaceless	2.9%	2.3%	5%	2.7%	3.5%	2.9%	3.2%	3.2%	2.8%

Notes: The table contains various robustness exercises. For all specifications, we always recalibrate our baseline model. We report the spatial reallocation component in the calibrated model and the Quasi-Spaceless Model in rows 1 and 2. In rows 3 - 5 we report the partial effect of abstracting from spatial productivity Q_{rst}, moving costs MC_{jr} and amenities A_{rt} in the baseline calibration. The last two rows report the implications for GDP per capita. For the "congestion" case we follow Adao et al. (2018) in using a model implied optimal IV approach to estimate ι_1 and ι_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 σ and the dispersion of individual skills ζ. These parameters determine the extent to which demand declines if regional prices change (σ) and the elasticity of sectoral labor supply within locations (ζ). While the value of σ 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 ζ.

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, $\iota_j > 0$ 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 (ι_1, ι_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 therefore raised the return for workers to relocate 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 in the US between 1880 and 2000.

We first document empirically that the spatial reallocation of workers towards non-agricultural labor markets explained essentially none of the decline in the aggregate 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 declined 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. First, 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. Additionally, it played a crucial role for the evolution of spatial 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 Sructural 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 Martı 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.

Herkenhoff, Kyle F, Lee E Ohanian, and Edward C Prescott, "Tarnishing the golden and empire states: Land-use restrictions and the US economic slowdown," *Journal of Monetary Economics*, 2017.

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.

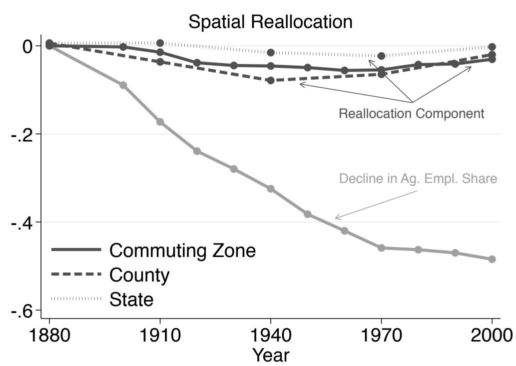
Young, Alwyn, "Inequality, the Urban-Rural Gap, and Migration," *The Quarterly Journal of Economics*, 2013, *128* (4), 1727–1785.

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). Figure 8 reinforces our result of spatial reallocation as a highly local phenomenon, which operates predominantly at the intra-county level.



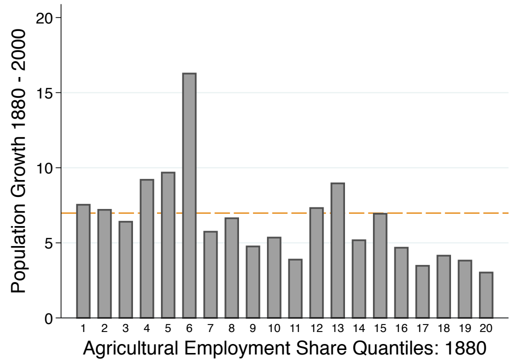
Notes: The figure shows the absolute decline in the aggregate agricultural employment share since 1880, i.e. $s_{At} - s_{A1880}$. The remaining lines show the across labor market reallocation component highlighted in equation (1), i.e. $\sum_r s_{rA1880} (l_{rt} - l_{r1880})$, when we define a labor market at the county, commuting zone and state level.

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

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

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.



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

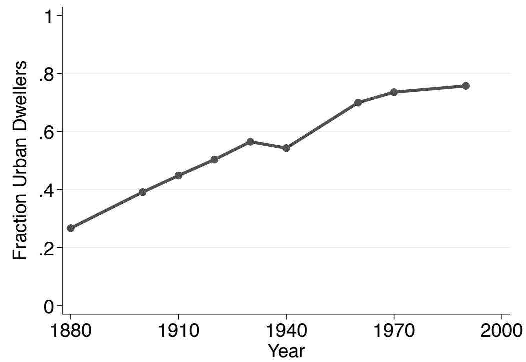
	Dep. variable: Population growth 1880 - 2000			
Agricultural share 1880	-26.241	-3.702*		
	(40.829)	(1.917)		
log Agricultural share 1880			-0.548	
			(0.658)	
Ag quantile FE	No	No	No	Yes
Weights	No	Yes	Yes	Yes
Observations	717	717	717	717
R^2	0.000	0.002	0.000	0.014

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 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%



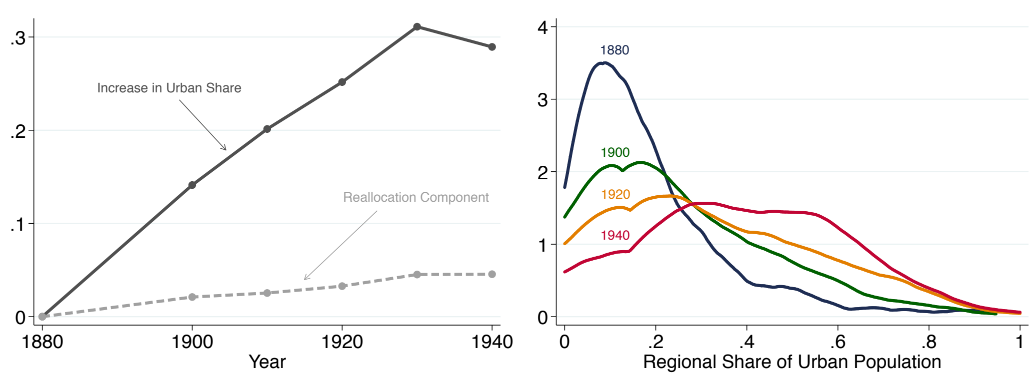
Notes: The graph is constructed from the decennial US micro census available on IPUMS. It shows the fraction of Americans who live in an urban environment, i.e. a locality with more than 2500 inhabitants.

Figure 10: Urbanization in the US from 1880 to 2000

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:

$$u_t - u_{1880} = \underbrace{\sum_r u_{r1880}(l_{rt} - l_{r1880})}_{\text{Spatial Reallocation}} + \underbrace{\sum_r (u_{rt} - u_{r1880}) l_{rt}}_{\text{Regional Transformation}}, \tag{22}$$

where u_{r1880} is the urbanization rate in commuting zone r in 1880.[35] 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.



Notes: In the left panel, we show the absolute increase in the share of people living in urban areas and its reallocation component calculated according to equation (22) 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.

Figure 11: Urbanization within Commuting Zones

[35]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.

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 Figure 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. Note that the county-level data is only publicly available until 1940.

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
Δ Ag. Empl. Share	-0.48	-.076	-.10	-.09	-.12	-.29	-.12	-.02
Δ 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 $V_rT_{rs} = \frac{\gamma_s}{1-\gamma_s}w_rL_{rs}$. 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 equation (15) are then given by

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

where $\pi_{rst} = p_{sr}^{1-\sigma} / \left(\sum_j p_{sj}^{1-\sigma} \right)^{1/(1-\sigma)} = \tilde{Z}_{sr}^{\sigma-1} w_{sr}^{1-\sigma} / \left(\sum_j \tilde{Z}_{sj}^{\sigma-1} w_{sj}^{1-\sigma} \right)^{1/(1-\sigma)}$. 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 $V_r^*/w_r^* = (\gamma_A/(1-\gamma_A))(L_{rA}^*/T_r)$. 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 $y_r^i = (1+v) w_r z_r^i$, where $v$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, $\nu = (\gamma_A/(1-\gamma_A)) s_{rst}$. Letting $\tilde{w}_{rt} \equiv (1 + (\gamma_A/(1-\gamma_A)) s_{rAt}) w_{rt}$, 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 (6))

$$\rho_{jrt} = \frac{\exp\left(\frac{1}{\kappa}\left(\mathscr{W}_{rt} + \Lambda_{r,t,t+1} - MC_{jr}\right)\right)}{\sum_{l=1}^{R} \exp\left(\frac{1}{\kappa}\left(\mathscr{W}_{lt} + \Lambda_{l,t,t+1} - MC_{jl}\right)\right)}, \tag{23}$$

where $\Lambda_{r,t,t+1} = -\frac{\nu}{\gamma}\left(\left(p_{At}^r / p_{NAt}^r\right)^{\gamma} + \beta\left(p_{At+1}^r / p_{NAt+1}^r\right)^{\gamma}\right) + (1+\beta)(\nu/\gamma - 1/\eta)$ and $p_{st}^r = \left(\sum_j \left(\tau_{jr} p_{jst}\right)^{1-\sigma}\right)^{1/(1-\sigma)}$ and τ_{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 (23). 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 equation (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. $V(e,p) = \frac{1}{\eta}(e/B(p))^{\eta} + C(p) - \frac{1}{\eta}$. The maximization problem is

$$U_r^i = \max_{[e_t, e_{t+1}, s]} \left\{V(e_t, p_t) + \beta V(e_{t+1}, p_{t+1})\right\},$$

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

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

The optimal allocation of spending is determined from the Euler equation

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

From above (36) we get that this equation is

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

This yields

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

so that

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

where $\phi_{t,t+1} \equiv \frac{1}{1+r_{t+1}}\frac{B(p_{t+1})/B(p_t)}{p_{I,t+1}/p_{I,t}}$. Hence, (5) implies that

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

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

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

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

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

and

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

Note also that

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

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 $y^i = \max_s \left\{w_{rs} z_s^i\right\}$. We assumed that individual productivities are Frechet Distributed, i.e. $F_s^h(z) = e^{-\Psi_s^h z^{-\zeta}}$, where Ψ_s^h parametrizes the average level of productivity of individuals of skill h in region r in sector s and ζ the dispersion of skills. Hence, the distribution of sectoral earning $y_{rs}^i \equiv w_{rs} z_s^i$ 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

$$F_{y_r}^h(y) = e^{-\left(\Theta_r^h\right)^{\zeta} y^{-\zeta}} = e^{-\left(\frac{y}{\Theta_r^h}\right)^{-\zeta}} \tag{24}$$

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 (24) we can derive the distribution of $y^{1-\eta}$. As $\eta < 1$, we have that

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

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

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

To derive the aggregate value added share ϑ_t^A, note that

$$\begin{aligned} PY_t^A &= \sum_{r=1}^{R}\left(\vartheta_A\left(E_{\mathscr{S}_r^Y}, p\right) E_{\mathscr{S}_r^Y} L_{rt} + \vartheta_A\left(E_{\mathscr{S}_r^O}, p\right) E_{\mathscr{S}_r^O} L_{rt-1}\right) + \phi I_t \\ &= \phi\left[I_t + \sum_{r=1}^{R}\left(E_{\mathscr{S}_r^Y} L_{rt} + E_{\mathscr{S}_r^O} L_{rt-1}\right)\right] + \tilde{\nu}\left(\frac{p_A}{p_{NA}}\right)^{\gamma} \sum_{r=1}^{R}\left(E_{\mathscr{S}_r^Y}^{1-\eta} L_{rt} + E_{\mathscr{S}_r^O}^{1-\eta} L_{rt-1}\right). \end{aligned}$$

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

$$\sum_{r=1}^{R} E_{\mathscr{S}_r^O} L_{rt-1} = R_t K_t + (1-\delta) K_t = \alpha P Y_t + (1-\delta) K_t.$$

And because the future capital stock is given by the savings of the young generation (see (19)), we get that

$$I_t + \sum_{r=1}^{R} \left(E_{\mathscr{S}_r^Y} L_{rt} + E_{\mathscr{S}_r^O} L_{rt-1} \right) \quad = \quad K_{t+1} + \sum_{r=1}^{R} \left(E_{\mathscr{S}_r^Y} L_{rt} \right) + \alpha P Y_t = P Y_t.$$

This yields (14).

B.4 The Quasi-Spaceless Economy

If wages are equalized across space, $w_{rt} = w_t$, the final good prices are given by

$$p_{st} = \left[\sum_{r=1}^{R} \left(\frac{1}{Q_{rst} Z_{st}} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha} \left(\frac{R_t}{\alpha} \right)^{\alpha} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = \frac{1}{Z_{st}} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha} \left(\frac{R_t}{\alpha} \right)^{\alpha},$$

as $\sum_{r=1}^{R} Q_{rst}^{\sigma-1}$. Hence, the choice of numeraire implies that

$$1 = p_{At}^{\phi} p_{NAt}^{1-\phi} = \left(\frac{1}{Z_{At}} \right)^{\phi} \left(\frac{1}{Z_{NAt}} \right)^{1-\phi} \left(\frac{w_t}{1-\alpha} \right)^{1-\alpha} \left(\frac{R_t}{\alpha} \right)^{\alpha} = \left(\frac{1}{Z_{At}} \right)^{\phi} \left(\frac{1}{Z_{NAt}} \right)^{1-\phi} \left(\frac{P Y_t}{\sum_{r=1}^{R} L_{rt} \Gamma_{\zeta}} \right)^{1-\alpha} \left(\frac{P Y_t}{K_t} \right)^{\alpha}.$$

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

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

Note that the LHS is increasing and continuous in k_{t+1} and satisfies $\lim_{k_{t+1} \to \infty} LHS = \infty$ and $\lim_{k_{t+1} \to 0} LHS = 0$. Hence, there is a unique mapping $k_{t+1} = m(k_t, Z_t, Z_{t+1})$. Furthermore, $m(.)$ 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 $\{w_{rt}, L_{rt}\}_{rt}$, is determined from labor market clearing condition (17)

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

and the spatial labor supply condition $L_{rt} = \sum_{j=1}^{R} \rho_{jrt} L_{jt-1}$, where ρ_{jrt}, ϑ_t^A and π_{rst} are given in (12), (14) and (16) and also only depend on $\{w_{rt}, L_{rt}\}_r$. Given $\{w_{rt}, L_{rt}\}_r$, the future capital stock is then determined from (19) and the regional agricultural employment shares from expression (18).

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

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

where Θ_r^h, given in equation (21), denotes regional income for skill group h and λ_{rt} is the share of skilled individuals.[36] 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 ρ_{jrt}^h is still given in (12) with $\mathscr{W}_{rt}^h = \left(\Gamma_{\eta/\zeta}/\eta\right)\psi(r_{t+1})^{\eta-1}\left(\Theta_r^h\right)^{\eta} + A_{rt}$. Finally, the within-region allocation of factors across sectors, i.e. the counterpart to (18), is given by

$$\frac{\pi_{rAt}}{\pi_{rNAt}}\frac{\vartheta_t^A}{1-\vartheta_t^A} = \frac{w_{rA}\left(H_{rA}^L + H_{rA}^H\right)}{w_{rNA}\left(H_{rNA}^L + H_{rNA}^H\right)} = \left(\frac{w_{rAt}}{w_{rNAt}}\right)^{\zeta}\frac{(1-\lambda_r)\Psi_A^L\left(\Theta_r^h\right)^{1-\zeta} + \lambda_r\Psi_A^H\left(\Theta_r^H\right)^{1-\zeta}}{(1-\lambda_r)\Psi_{NA}^L\left(\Theta_r^h\right)^{1-\zeta} + \lambda_r\Psi_{NA}^H\left(\Theta_r^H\right)^{1-\zeta}}. \tag{25}$$

These equations determine $\left\{w_{rAt}, w_{rNAt}, L_{rt}^L, L_{rt}^H\right\}_r$. The skill specific employment shares can then be calculated as $s_{rst}^h = \Psi_s^h\left(w_{rs}/\Theta_r^h\right)^{\zeta}$. The capital accumulation is still given by (19), i.e.

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

The main difference to the case with substitutable skills is the within-region sectoral supply equation (25). In the baseline model, s_{rAt} is determined residually from equation (18). 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 $R_t = \alpha PY_t/K_t$. 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 equation (10) is constant and given

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

by

$$\psi = \left(1 + \beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}}\right)^{-1}. \tag{26}$$

To solve for the interest rate, note that capital grows at rate g, so that expression (19) 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,

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

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 equation (27) yields $(PI_t/PY_t) = (g+\delta)(\alpha/R)$ and $PC_t/PY_t = 1-(g+\delta)(\alpha/R)$. From equation (27) we also get that $\psi = (\alpha/(1-\alpha))(1+g/R)$.

B.7 Regularity conditions for PIGL preferences

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

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

For these expenditure shares to be positive, we need that

$$\vartheta_A(e,p) \geq 0 \Rightarrow e^{\eta} \geq -\frac{\nu}{\phi} \left(\frac{p_A}{p_{NA}}\right)^{\gamma}, \tag{28}$$

and $\vartheta_{NA}(e,p) \geq 0 \Rightarrow e^{\eta} \geq (\nu/(1-\phi))(p_A/p_{NA})^{\gamma}$. Note first that (28) is trivially satisfied as long as $\nu > 0$. 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

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

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

$$e^{\eta} \geq \frac{\nu}{1-\phi}\left(\frac{p_A}{p_{NA}}\right)^{\gamma} \tag{29}$$

$$(1-2\phi-\gamma)e^{\eta}+\frac{(1-\phi)\phi}{\nu}\left(\frac{p_A}{p_{NA}}\right)^{-\gamma}e^{2\eta} \geq \nu(1-\eta)\left(\frac{p_A}{p_{NA}}\right)^{\gamma}. \tag{30}$$

Lemma 5. *A sufficient condition for (29) to be satisfied is that (30) holds and that*

$$\gamma>(1-\phi)\eta. \tag{31}$$

Proof. Equation (30) can be written as

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

Letting $x=\frac{e^{\eta}}{\frac{\nu}{1-\phi}\left(\frac{p_A}{p_{NA}}\right)^{\gamma}}$, 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 (30) 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 condition (30) to be satisfied, it has to be the case that $x>x^*=1-((1-\gamma)/2\phi)$ as $h(x^*)<h(0)$. Hence, condition (30) implies inequality (29) if

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

Rearranging terms yields $(1-\phi)\eta < \gamma$, which is inequality (31). □

Hence, the preferences are well defined as long (30) is satisfied and (31) holds. Because the RHS of (30) is increasing in e in the relevant range, i.e. as long as (30) 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 (24) we know that

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

where

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

Hence, as long as aggregate productivity in 1880 (and hence w_{rs1880}) is high enough, we can make $P(e_{it} \leq \kappa)$ arbitrarily small.

B.8 Additional Properties of the Calibrated Model

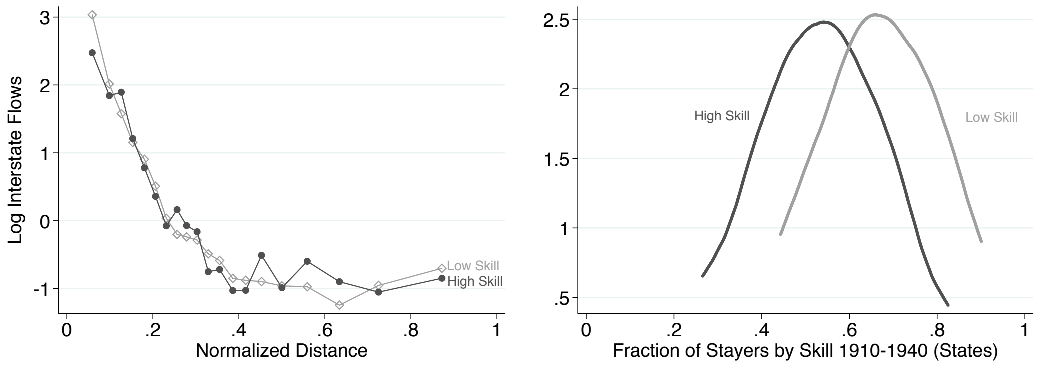
In this section we provide additional results for the fit of the calibrated model.

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 τ to differ by skill type, since as outlined in the calibration section, the total number of stayers is monotone in τ. 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

$$\Delta_{jt}^h \equiv \frac{\mathscr{U}_{jt}^{h,Mov} - \mathscr{U}_{jt}^{h,Stay}}{\tau}, \tag{32}$$

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 Δ_{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.



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

	Expected value of moving relative to fixed costs Δ_{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 Δ_{jt}^h calculated as (32).

Table 9: The Economic Magnitude of Moving Costs

Regional Fundamentals In this section, we describe additional details for the estimated spatial productivities $\{Q_{rAt}, Q_{rNAt}\}_{rt}$ and amenities $\{A_{rt}\}_{rt}$. First we 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. In particular, we report the results from the specification

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

where δ_t is a year fixed effect. In columns 1 to 3 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

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

	Dep. variable: $\ln s_{Art}$			
Non Ag. Productivity ($\ln Q_{rNA}$)	-0.331***			-0.344***
	(0.020)			(0.007)
Ag. Productivity ($\ln Q_{rA}$)		0.404***		0.408***
		(0.022)		(0.006)
Amenities (A_r)			-0.430***	-0.055***
			(0.075)	(0.013)
Skill share in 1880 (λ_{r1880})	0.139	-1.594***	-0.466**	-0.117***
	(0.123)	(0.150)	(0.196)	(0.032)
ln 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

	$\ln Q_{rNAt}$		$\ln Q_{rAt}$		A_{rt}	
$\ln Q_{rNAt-1}$	0.832***	0.416***				
	(0.011)	(0.039)				
$\ln Q_{rAt-1}$			0.725***	0.045		
			(0.050)	(0.043)		
A_{rt-1}					0.897***	0.175
					(0.063)	(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

where x denotes either log sectoral productivity, $\ln Q_{rst}$, or the level of amenities A_{rt} and δ_t and δ_r are 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 $\{Q_{rAt}, Q_{rNAt}\}_{rt}$ 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, $\{Q_{rAt}, Q_{rNAt}\}_{rt}$, 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

$$\ln Q_{rst} = \delta_r + \beta \times MAC_{rt} + u_{rst}, \tag{33}$$

where δ_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 33 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 33 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 $\{Q_{rAt}, Q_{rNAt}\}_{rt}$.

	$\ln Q_{rNAt}$	$\ln Q_{rAt}$	$\ln Q_{rNAt}$	$\ln Q_{rAt}$	$\ln Q_{rNAt}$	$\ln Q_{rAt}$	$\Delta \ln Q_{rNAt}$	$\Delta \ln Q_{rNAt}$
Market Access Costs	-1.897***	-2.077***	-5.205***	-4.291***	-3.937***	-3.760***		
	(0.224)	(0.125)	(0.290)	(0.195)	(0.353)	(0.309)		
Change in Market Access Costs							-4.005***	-2.395***
							(0.388)	(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 Table 5 in the main text we discuss the evolution of spatial inequality. We construct this the numbers in this table in the following way. From Proposition 2 we have that expected utility in region r is given by

$$\mathscr{W}_{rt}^h\left(\Theta_r^h\right)=\frac{\Gamma_{\eta/\zeta}}{\eta}\psi\left(r_{t+1}\right)^{\eta-1}\left(\Theta_r^h\right)^{\eta}+\Lambda_{t,t+1}+A_{rt}. \tag{34}$$

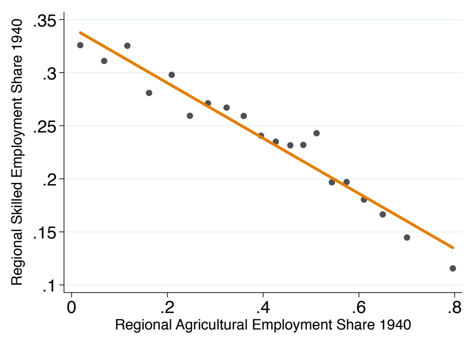
Let $T_{rt}^h(\Delta)$ be the increase in average income Θ_r^h required to increase utility by Δ, i.e. $\mathscr{W}_{rt}^h\left(\Theta_r^h T_{rt}^h(\Delta)\right)=\mathscr{W}_{rt}^h\left(\Theta_r^h\right)+\Delta$. (34) implies that

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

Let $\Delta_t^{h,IQR}$ be the interquartile range in regional welfare $\mathscr{W}_{rt}^h$ at time t, i.e. $\Delta_t^{h,IQR}=\mathscr{W}_t^{h,75}-\mathscr{W}_t^{h,25}$, where $\mathscr{W}_t^{h,x}$ is the x-quantile of the distribution of $\mathscr{W}_{rt}^h$.

B.11 Agricultural Employment Shares and Skilled Employment

The 1940 cross-section of the US Micro Census published by the US Census Bureau is the first to contain an educational variable and the last to contain the full set of county identifiers. We construct skilled employment shares on the county level as the fraction of workers employed in this county that have at least a completed high school degree. In Figure 13 we depict the spatial correlation between agricultural employment shares and skill shares. The calibration strategy outlined in the main body of the paper allows us to match the distribution of skilled employment shares in 1940 exactly.



Notes: The figure shows the binscatter plot of the agricultural employment shares in US counties and their skilled employment shares in 1940.

Figure 13: Agricultural Employment Shares and Skilled Employment Shares in 1940