

Spatial Structural Change^{*}

Fabian Eckert[†] Michael Peters[‡]

January 2018

Abstract

This paper studies the spatial implications of structural change. By shifting demand towards non-agricultural goods, the structural transformation benefits workers in urban centers and hurts rural locations. At the aggregate level, the economy can respond either through a reallocation of labor from rural to urban areas or through a reduction in agricultural employment within a given location. Using detailed spatial data for the U.S. between 1880 and 2000, we show that spatial reallocation accounts for almost none of the aggregate decline in agricultural employment. We interpret this fact through the lens of a novel quantitative theory of spatial structural change, and show that the absence of the spatial reallocation channel is primarily due to regional productivity shocks, which almost entirely offset the urban bias of the structural transformation. Frictions to labor mobility meanwhile are quantitatively unimportant. The model implies that spatial welfare differences declined substantially during the United States' structural transformation and that the spatial reallocation of factors can account for about 15% of aggregate growth since 1880.

^{*}We thank seminar participants at Boston University, Stanford, Penn State, UNC, the SED, the EEA, the NBER Macroeconomic Across Time and Space Conference, the Edinburgh Workshop on Agricultural Productivity and the Yale International Lunch for very helpful comments. We are particularly thankful to Simon Alder, Timo Boppert, David Lagakos, Doug Gollin, Kjetil Storesletten and Fabrizio Zilibotti for their suggestions. We also thank Costas Arkolakis, Pete Klenow, and Aleh Tsyvinski for their comments. Zara Contractor, Andrés 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 also implies that economic growth is unbalanced across space. In particular, by shifting expenditure away from the agricultural sector, the structural transformation is biased against regions, that have a comparative advantage in the production of agricultural goods. 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 the US economy accommodated this “urban bias” of the structural transformation. Furthermore, we assess the importance of this spatial unbalancedness of the growth process for aggregate productivity and welfare.¹

We start by documenting a striking - and to the best of our knowledge - new empirical fact: the spatial reallocation of people from rural, agriculturally intensive regions to centers of non-agricultural production 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 localities.² While this is seemingly inconsistent with the secular trend in urbanization, whereby the share of urban dwellers among US workers increased from 25% to 75% between 1880 and 2000, we explicitly show that this is not the case. In fact, like the change in agricultural employment, the process of urbanization was also a predominantly within-region phenomenon.

To understand why the structural transformation did not cause more spatial reallocation and to quantify the implications for the spatial distribution of welfare and aggregate productivity, we propose a new quantitative theory of spatial structural change. Our theory combines an otherwise standard, neoclassical model of the structural transformation with an economic geography model featuring intra-national trade and frictional labor mobility. At the aggregate level, non-homothetic consumer demand and unbalanced technological progress reduce the relative demand for agricultural goods as the economy grows. At the spatial level, regions are differentially exposed to such secular demand shifts 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.

There are two candidate explanations for why spatial reallocation accounts for little of the aggregate decline in agricultural employment. It could be that frictions to spatial mobility were prohibitively large for the majority of workers throughout the 20th century. Alternatively, enough people were migrating, but the correlation between regional agricultural employment shares and net spatial outflows was small. Our model squarely points to the second explanation. In particular, US Census data reveals that over the last

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

²This holds true at the state, commuting zone (as defined in Tolbert and Sizer (1996)) or county level. There are roughly 700 commuting zones and 3000 counties. In our quantitative analysis we focus on commuting zones. We discuss this choice in more detail in Section 4.1.

century, 30% of the working-age population lived in states different from their state of birth. According to our model, these are more than enough bodies crossing regional boundaries to accommodate the declining demand in agricultural localities. Moving frictions *per se*, therefore, cannot explain the quantitative insignificance of the spatial reallocation channel.

To explain why such moving flows are only weakly correlated with the initial level of agricultural specialization, our model highlights two forces: the extent to which rural areas tend to have low wages in the future and the migration elasticity, i.e. the sensitivity of migration flows to spatial wage differences. We find that both are quantitatively important. While the structural transformation indeed put downward pressure on rural wages, we show that such shifts are small relative to the level of existing wage differences. And because such wage differences are only imperfectly correlated with the regional agricultural employment share, individuals, in their search for higher earnings, often flock *towards* agricultural areas. Moreover, regional productivities are not constant but subject to stochastic shocks. This further weakens the relationship between agricultural specialization and future earnings. Additionally, we also find that the migration elasticity is relatively small. In particular, we show that spatial population *gross* flows are much larger than population *net* flows. Through the lens of the model, this suggests that idiosyncratic, non-monetary preference shocks are an important determinant of migration decisions.

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

Next, we turn to the evolution of spatial welfare inequality. We find that welfare inequality across US commuting zones declined substantially between 1910 and 2000. In 2000 the interquartile range of the distribution of spatial welfare corresponded to a doubling of lifetime income for the average region in the US. In contrast, in 1910 one would have had to increase regional income by a factor of 2.6. Importantly, and similar to aggregate productivity, the possibility of spatial mobility was an important determinant of 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.

As a theoretical contribution, 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 a result, we do not have to keep track of individuals' expectations about the entire distribution of future wages across locations - the aggregate interest rate is sufficient. Moreover, because our model essentially nests a version of a canonical aggregate model of structural change as the number of regions collapses to one, we view our framework as a natural spatial extension of neoclassical theories of the structural transformation.

Related Literature We combine insights from the macroeconomic literature on the structural transformation with recent advances in quantitative models of economic geography. 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 reallocation of workers across heterogeneous locations.⁴ We show that these two aspects interact in a natural way. The urban bias, which is inherent in the structural transformation, induces changes in demand, which are non-neutral across space and hence affect the spatial equilibrium of the system. Conversely, the spatial topography, in particular the extent to which individuals are spatially mobile, has macroeconomic implications by determining equilibrium factor prices, capital accumulation and aggregate productivity.

Relatively few existing papers explicitly introduce a spatial dimension into an analysis of the structural transformation. An early contribution is [Caselli and 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

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

⁴This 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.

relationship between agricultural specialization and population growth across US counties. Their analysis is more empirically oriented and does not use a calibrated structural model. More recently, [Desmet and Rossi-Hansberg \(2014\)](#) propose a spatial theory of the US transition from manufacturing to services and [Nagy \(2017\)](#) examines the process of city formation in the United States before 1860.

In allowing for a spatial microstructure, this paper also sheds new light on the so-called “agricultural productivity gap”, i.e. the observation that value added per worker is persistently low in the agricultural sector (see e.g. [Gollin et al. \(2014\)](#), [Buera and Kaboski \(2009\)](#) or [Herrendorf and Schoellman \(2015\)](#)). Our model endogenously generates an “agricultural productivity gap”, without resorting to labor market frictions across sectors, as spatial mobility costs keep wages in agricultural areas low.⁵ The existence of such spatial gaps and their implications for aggregate productivity and welfare has been subject to a recent active literature (see e.g. [Young \(2013\)](#), [Bryan et al. \(2014\)](#), [Hsieh and Moretti \(2015\)](#) or [Lagakos et al. \(2017\)](#)). [Bryan and Morten \(2017\)](#) and [Hsieh and Moretti \(2015\)](#) use spatial models related to ours to study the aggregate effects of spatial misallocation. In contrast to us, their models are static and they do not focus on the structural transformation.

On the theoretical side, we build on [Boppart \(2014\)](#) and assume a price independent generalized linear (PIGL) demand structure. This demand structure has more potent income effects than the widely-used Stone-Geary specification, a feature which is required to generate declines in agricultural employment of the magnitude observed in the data. At the same time, we show how it can be introduced in a general equilibrium trade model in a tractable way.⁶

The remainder of the paper is structured as follows. In Section 2, we document the empirical fact that spatial reallocation accounts for essentially none of the aggregate decline in agricultural employment over the last 120 years. Section 3 presents our model. In Section 4, we calibrate the model to time-series and spatial data from the US. In Sections 4.3 and 5, we explain why the spatial reallocation component of the structural transformation is small and we quantify the implications for aggregate productivity and spatial inequality. Section 6 concludes. An Appendix contains the majority of our theoretical proofs and further details on our empirical results.

⁵In an empirical exercise, [Gollin et al. \(2017\)](#) use direct measures of differences in spatial amenities and show that amenities in agricultural, rural localities are much lower than in urban centers. This finding is consistent with our structural estimates of spatial amenities in the US of the 19th century.

⁶Between 1880 and 2000 the aggregate agricultural employment share declines from around 50% to 2%. [Alder et al. \(2018\)](#) show that the PIGL demand system provides a good fit to the data since 1900. A model with Stone-Geary preferences can match the post-war data (see e.g. [Herrendorf et al. \(2013\)](#)), but has difficulties at the longer time horizon as income effects vanish asymptotically. The non-homothetic CES demand system, recently proposed by [Comin et al. \(2017\)](#), has similarly favorable time-series properties. However, it has less tractable aggregation properties making it harder to embed it in a spatial structure.

2 Spatial Reallocation and Structural Change

The structural transformation is not only unbalanced across sectors but also unbalanced across space. In particular, the reallocation of spending across sectors is biased against rural regions, which specialize in agriculture. From an accounting perspective, there are two margins through which the economy can accommodate such changes. First, the structural transformation can induce a process of *spatial reallocation*, whereby labor reallocates from agriculturally intensive regions to places which are more specialized in the production of non-agricultural goods. Second, it can lead to a *regional transformation*, whereby regional agricultural employment shares decline in every locality. Formally, the aggregate decline in the agricultural employment share since 1880 can be decomposed as

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

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

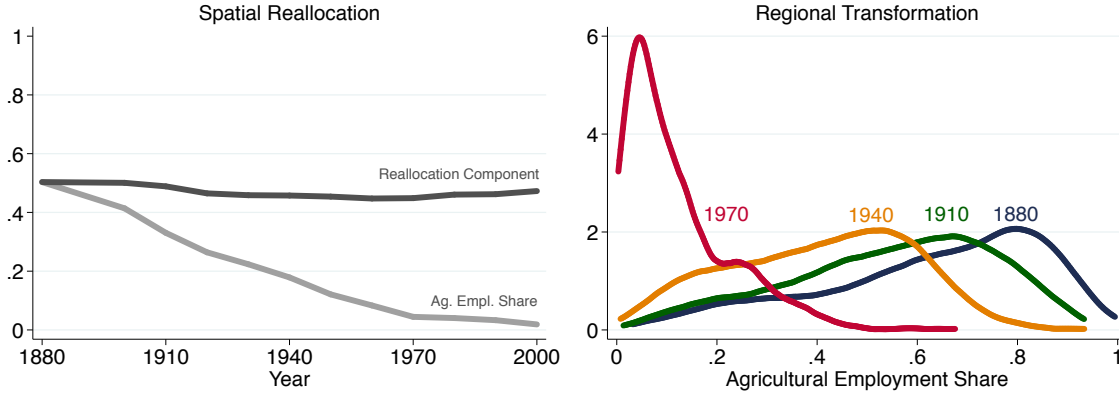
For the case of the U.S., the relative importance of the reallocation and transformation margins is striking: the spatial reallocation of labor accounts for essentially none of the structural transformation observed in the aggregate. To see this, consider Figure 1, where we implement (1) at the commuting zone level.⁷ In particular, we conduct a “shift-share”-analysis and compare the actual employment share, s_{At} , to the implied agricultural employment share, which emerged *solely* from spatial reallocation, $\sum_r s_{rA1880} l_{rt}$. It is clear that - in an accounting sense - spatial reallocation accounts for none of the aggregate decline in agricultural employment. To put it differently, the process of the structural transformation is *not* driven by a reallocation of people from high to low agricultural places. Out of the total decline of about 48%, only 3% is due to the reallocation of workers across commuting zone boundaries.

It follows that most of structural change takes place *within* regions through a transformation of the local structure of employment. This is seen in the right panel of Figure 1, where we display the cross-sectional distribution of regional agricultural shares for different years. First of all, note that there is substantial cross-sectional dispersion. While the majority of commuting zones have agricultural employment shares exceeding 75% in 1880, many regions are already much less agriculturally specialized and have agricultural employment shares below 25%. Secondly, there is a marked leftwards shift, whereby *all* regions see

⁷In the paper, we use the definition of a commuting zone as our definition of a region. There are roughly 700 commuting zones in the US. We describe our data in more detail in Section 4 below. We have replicated Figure 1 on the county level, it looks effectively identical. In addition we have computed the importance of regional transformation versus spatial reallocation for countries around the world using publicly available data from the Integrated Public Use Micro data Series (IPUMS) International data base. We find that regional transformation appears to be an integral part of the structural transformation across countries and the size of the spatial component is always below 5%.

a decline in agricultural employment and undergo a regional transformation as the US economy develops. This shows that structural change is not a process that induces regional specialization, but rather features a fractal property whereby all local economies undergo changes in sectoral structure akin to the aggregate economy.

These patterns are seemingly at odds with the process of urbanization, whereby cities, i.e. centers of non-agricultural production, grow.⁸ This, however, is not the case. While inconsistent with a view of urbanization as a movement of workers towards few large population centers, it is consistent with a world in which workers move towards urban centers *within* labor markets. This is exactly, what happened in the US. As we explicitly show in Section A in the Appendix, the increase in urbanization is also a within-region phenomenon, which is very local in nature. In particular, like for the change in the agricultural employment, the reallocation of individuals from rural to urbanized counties explains almost none of the sharp rise in urbanization. There we also show that this pattern is not exclusive to the US but present in many countries around the world undergoing the structural transformation.



Notes: In the left panel, we show the aggregate agricultural employment share (light grey line) and the predicted agricultural share holding regional agricultural shares at their 1880 level, i.e. $\sum_r s_{rA1880} l_{rt}$, where s_{rA} and l_{rt} are the agricultural employment share and the population share of region r at time t (dark grey line). 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 left panel for expositional purposes only: the agricultural employment share does not change much between 1970-2000. For a detailed description of the construction of the regional data we refer to Section 4.

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

In the remainder of this paper, we develop a theory of spatial structural change, which can speak to these patterns. More specifically, we will use the calibrated model to answer (i) why the spatial reallocation component is small and (ii) whether its absence has important implications for the spatial distribution of welfare and aggregate productivity during the structural transformation.

⁸Over the 1880-2000 period that is the focus of this paper, the fraction of the US population living in cities tripled from about 25% to 75%. See Figure 10 in the Appendix for a time series of the fraction of the US population residing in an urban environment.

3 A Quantitative Theory of Spatial Structural Change

In this section we present a model of spatial structural change. The theory rests on three pillars. We start with what is essentially a neoclassical model of the structural transformation, where the process of structural change is generated from non-homothetic preferences and unbalanced technological progress. We introduce a spatial dimension by embedding this structure into an economic geography model of heterogeneous locations and costly spatial mobility. Finally, we allow for skill-based selection across locations and sectors of production by assuming that individuals differ in their human capital, with skilled workers have both an absolute advantage and a comparative advantage in the non-agricultural sector.

3.1 Environment

We consider an economy consisting of R regions indexed by r . Each region features an agricultural (A) and a nonagricultural (NA) sector, which are indexed by $s = A, NA$ throughout the paper. Since the model is dynamic we additionally index most quantities by t . For expositional simplicity, we sometimes omit the time subscript where there is no risk of confusion.

Technology We consider an economy with two goods, an agricultural good and a non-agricultural good. Each good is a CES composite of differentiated regional varieties with a constant elasticity of substitution σ . In particular,

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

where Y_{rs} is the amount of goods in sector s produced in region r and σ is the elasticity of substitution. For analytical tractability, we assume that goods are freely traded so that prices are equalized across locations.⁹

Regional production functions are fully neoclassical and given by

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

where K_{rst} and H_{rst} denotes capital and labor (in efficiency units) in region r , sector s and time t . Building on the work of [Herrendorf et al. \(2015\)](#), we assume that capital shares are identical across sectors.¹⁰ It is

⁹This assumption considerably simplifies workers' spatial choice problem, as outlined below.

¹⁰[Herrendorf et al. \(2015\)](#) 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).

conceptually useful to decompose regional productivity A_{rst} as

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

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

Capital 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. We assume that the investment good is a Cobb-Douglas composite of the agricultural and non-agricultural good given in (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 $P_t = P_{At}^\phi P_{NA_t}^{1-\phi}$.¹¹ For the remainder of the paper the investment good will serve as the numeraire of our economy.

Demographics and Labor Supply We assume that individuals are heterogenous in the amount of labor efficiency units they can supply to the market. We follow the macroeconomic literature on structural change and assume that efficiency units are perfectly substitutable across sectors, i.e. there is effectively a single dimension of skill. This implies that the sectoral allocation of labor within locations is only determined from the labor demand side. While this assumption allows us to focus entirely on the novel spatial dimension of our model, in Section 3.4 below we will explicitly introduce skill heterogeneity across sectors before taking the model to the data.

It is analytically attractive to assume that the amount of individuals i 's efficiency units, z^i , is drawn from a Frechet distribution, i.e. $F(z) = e^{-z^{-\zeta}}$. Here, the parameter ζ governs the dispersion of skills across individuals and the average level of efficiency units is given by $E[z] = \Gamma_\zeta$, where $\Gamma_\zeta \equiv \Gamma(1 - \frac{1}{\zeta})$ and $\Gamma(\cdot)$ denotes the Gamma function. Hence, the distribution of earnings of individual i in region r at time, $y_{rt}^i = z^i w_{rt}$, is also Frechet distributed with mean $\Gamma_\zeta w_{rt}$.

In terms of preferences and demographics, ours is an overlapping generations (OLG) economy. Individuals live for two periods, work when they are young and save to be able to consume when they are old. The OLG structure is analytically extremely convenient because it generates a motive for savings (and

¹¹The 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 (see (Herrendorf et al., 2017)).

hence capital accumulation), while still being sufficiently tractable to allow for spatial mobility. It is also empirically attractive in that it captures the importance of cohort effects in accounting for the structural transformation (see [Hobijn et al. \(2018\)](#) and [Porzio and Santangelo \(2017\)](#)).

In our framework individuals make three economic choices: (i) how much to save and consume when they are young, (ii) how to allocate their spending optimally across the two consumption goods in each period of life and (iii) and where to live and work. Letting $V(e_t, P_t)$ be the *indirect* utility of spending an amount e_t at sectoral prices $P_t = (p_{At}, p_{NA_t})$, life-time utility of individual i after having moved to region r , U_{rt}^i , is given by

$$U_{rt}^i = \max_{[e_t, e_{t+1}, s]} \{V(e_t, P_t) + \beta V(e_{t+1}, P_{t+1})\}, \quad (4)$$

subject to

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

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

Preferences To generate the structural transformation at the aggregate level, we follow the existing macroeconomic literature and allow for both non-homothetic preferences, i.e. a form of Engel's Law whereby consumers reduce their relative agricultural spending as they grow richer, and relative price effects, whereby sectoral differences in technological progress induce a reallocation in consumer spending.¹² The most popular choice among non-homothetic preferences is the Stone-Geary specification, which relies on subsistence requirements to generate a less than unitary income elasticity for consumers' spending on agricultural goods (see e.g. [Kongsamut et al. \(2001\)](#)). [Herrendorf et al. \(2013\)](#), for example, show that these preferences do a good job to quantitatively account for the decline in agricultural spending in the US in the post-war period. This is not the case at a longer horizon. In particular, [Alder et al. \(2018\)](#) use aggregate data to show that the Stone-Geary specification is unable to generate the large decline in agricultural employment since 18880 as the effect due to the non-homotheticity vanishes asymptotically. In contrast, they find that *Price-Independent Generalized Linear* ("PIGL") Preferences, introduced in the literature on structural change in [Boppart \(2014\)](#), provide a much better fit to the data.

We therefore follow [Boppart \(2014\)](#) and assume that individual preferences can be represented by the

¹²Both 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. \(2017\)](#).

indirect utility function

$$V(e, P) = \frac{1}{\eta} \left(\frac{e}{p_A^\phi p_{NA}^{1-\phi}} \right)^\eta - \frac{\nu}{\gamma} \left(\frac{p_A}{p_{NA}} \right)^\gamma + \frac{\nu}{\gamma} - \frac{1}{\eta}. \quad (5)$$

This is a slight generalization of the PIGL demand system employed by [Boppart \(2014\)](#).¹³ In particular, Roy's Identity implies that the expenditure share on the agricultural good, $\vartheta_A(e, p)$, is given by

$$\vartheta_A(e, p) \equiv \frac{x_A(e, p) p_A}{e} = \phi + \nu \left(\frac{p_A}{p_{NA}} \right)^\gamma e^{-\eta}, \quad (6)$$

and hence incorporates both income effects (governed by η) and price effects (governed by γ).¹⁴ For $\eta > 0$, the expenditure share on agricultural goods is declining in total expenditure. This captures the income effect of non-homothetic demand, whereby higher spending reduces the relative expenditure share on agricultural goods. Holding real income e constant, the expenditure share is increasing in the relative price of agriculture if $\gamma > 0$. We can also see that this demand system nests important special cases. The case of $\eta = 0$ corresponds to a homothetic demand system, where expenditure shares only depend on relative prices. The case of $\eta = \gamma = 0$ is the Cobb Douglas case where expenditure shares are constant and equal to ϕ .¹⁵

The preference specification in (5) also has the advantageous aggregation properties in our context. In particular, we show below that these preferences (combined with the Frechet distribution of individuals skills) allow for closed-form aggregation of individuals despite the fact that they fall outside the Gorman class.¹⁶

Spatial Mobility A crucial aspect of our theory are agents' endogenous location choices. We assume that individuals have the option to move once, in the beginning of their lives, before they learn the actual realization of their labor efficiency z^i . This structure has two convenient 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.

¹³For $V(e, p)$ to be well-defined, we have to impose additional parametric conditions. In particular, we require that $\eta < 1$, that $\gamma \geq \eta$. These conditions are satisfied in our empirical application. See Section B.7 of the Appendix for a detailed discussion. [Boppart \(2014\)](#) uses this demand system to study the evolution of service sector. In terms of (5) he assumes that p_A is the price of goods and p_{NA} is the price of services and considers the case of $\phi = 0$. In his Appendix, however, [Boppart \(2014\)](#) also discusses the case of (5).

¹⁴Roy's Identity implies that $\vartheta_A(e, p) = x_A^M(p, e) \times \frac{p_A}{e} = -\frac{\partial V(p, e(p, u)) / \partial p_A}{\partial V(p, e(p, u)) / \partial e} \frac{p_A}{e}$, where $x_A^M(p, e)$ denotes the Marshallian demand function. For $V(p, e)$ given in (5), this expression reduces to (6). See Section B.2 in the Appendix for details.

¹⁵Note 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.

¹⁶This is in contrast to the non-homothetic CES demand system, which has recently been analyzed in [Comin et al. \(2017\)](#) and for which no closed form aggregation results exist.

Hence, consider the decision for individual i to move from location j to r . We follow the literature on discrete choice models and assume the value of this bilateral move to agent i can be summarized by

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

where $E[U_{rt}^i]$ is the *expected* utility of living in region r (see (4)), MC_{jr} denotes the cost of moving from j to r , A_r is a location amenity, which summarizes the attractiveness of region r and is common to all individuals and v_r^i is an idiosyncratic error term, which is independent across locations and individuals. Furthermore, 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 v_r^i is drawn from a Gumbel distribution. This implies that the share of people moving from j to r is given by

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

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

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

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

These expressions formalize the three determinants of the spatial reallocation component of the structural transformation. The extent to which individuals leave agricultural areas depends on (1) the size of the moving costs MC_{jr} , (2) the correlation between initial agricultural employment shares s_{rAt-1} and future lifetime utility $E[U_{rt}^i]$ and amenities A_{rt} and (3) the importance of idiosyncratic shocks, parametrized by κ . In particular, the correlation between agricultural employment and population outflows will be large if $E[U_{rt}^i]$ and s_{rAt-1} are negatively correlated and the elasticity of moving flows with respect to fundamental differences is large, i.e. κ is small.

3.2 Competitive Equilibrium

Given the environment above, we can now characterize the equilibrium of the economy. We proceed in three steps. First we characterize the household problem, i.e. the optimal consumption-saving decision

and the spatial choice. We then show that the solution to the household problem together with our distributional assumptions on individuals' skills delivers an analytic solution for the economy's aggregate demand system, despite the fact that our economy does not admit a representative consumer. Finally, we show that the dynamic competitive equilibrium has a structure akin to the neoclassical growth model: given the sequence of interest rates $\{r_t\}_t$, we can solve the entire path of spatial equilibria from static equilibrium conditions. The equilibrium sequence of interest rates can then be calculated from households' savings decisions. The model can then be solved by iteratively computing the sequence of spatial equilibria and finding a fixed point for the path of interest rates.

Individual Behavior

First consider the households' consumption-saving decision given in (4). The two-period OLG structure together with the specification of preferences in (5) 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. Due to the OLG structure, a young and an old generation co-exists at each point in time. We therefore denote the optimal level of expenditure when young (old) of the generation that is born at time t as e_t^Y (e_{t+1}^O).

Proposition 1. *Consider the maximization problem in (4) where $V(e, P)$ is given in (5). The solution to this problem is given by*

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

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

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

where

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

$$\Lambda_{t,t+1} = -\frac{\nu}{\gamma} \left(\left(\frac{p_{At}}{p_{NA,t}} \right)^\gamma + \beta \left(\frac{p_{A,t+1}}{p_{NA,t+1}} \right)^\gamma \right) + (1 + \beta) \left(\frac{\nu}{\gamma} - \frac{1}{\eta} \right).$$

Proof. See Section B.1 in the Appendix. □

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

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

of his income when young and consumes the remainder (and the accrued interest) when old. Because $\eta < 1$, the consumption share $\psi(r_{t+1})$ is decreasing in the interest rate and decreasing in the discount factor β as both increase the value of saving. If $\eta = 0$, i.e. if demand is homothetic, we recover the canonical OLG solution for log utility where the consumption share is simply given by $1/(1 + \beta)$. Importantly, the consumption share $\psi(r_{t+1})$ only depends on the interest rate r_{t+1} and not on relative prices P_t or P_{t+1} . This is due to our assumption that nominal income e is deflated by the same price index as the investment good. This is convenient for tractability and akin to the single-good neoclassical growth model, where the consumption good and the investment good uses all factors in equal proportions. For our purposes, this ensures that an increase in the price of investment good, p_{It} , makes savings more attractive but at the same reduces the marginal utility of spending. Third, lifetime utility U_{rt}^i only depends on the location r via individual income y_{rt}^i . This is due to our assumption that trade is frictionless so that the price indices (which determine $\Lambda_{t,t+1}$) do not vary across space. Finally, lifetime utility is *additively separable* in income y_{rt} and current and future prices P_t and P_{t+1} (which determine $\Lambda_{t,t+1}$).

This latter property is key to make the analysis tractable. In particular, it implies that agents' spatial choice problem reduces to a static decision problem, which depends only on current, not future, equilibrium objects (in particular the distribution of wages). Moreover, we can calculate individuals' lifetime utility $E[U_{rt}^i]$ analytically: because life-time utility is a power function of individual income y_r^i and individual income is Frechet distributed, we get that $E[y^\eta] = \Gamma_{\eta/\zeta} w_{rt}^\eta$. Together with (7), this delivers closed form expressions for individual migration decisions, which we summarize in the following Proposition.

Proposition 2. *Consider the environment above. The expected life-time value of location r at time t is given by*

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

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

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

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

Proposition 2 implies that individuals' migration decisions are fully captured by $\mathcal{W}_{rt}$, which is a summary measure of regional attractiveness. Note that the cross-sectional variation in $\mathcal{W}_{rt}$ stems from differences in average wages, w_{rt} , and in amenities A_{rt} . The former is endogenous and depends on the extent of spatial sorting and aggregate demand conditions. The latter is fully exogenous. Importantly, the expression for individuals' spatial choice probabilities in (13) does *not* feature $\Lambda_{t,t+1}$, which is constant across locations (because of the absence of trade costs) and hence does not determine spatial labor flows. This additive separability of future prices embedded in $\Lambda_{t,t+1}$ is crucial, because it turns individuals' optimal location

choices into a static problem. This structure allows us to calculate the transitional dynamics in the model with a realistic geography, i.e. with about 700 regions.

Equilibrium Aggregation and Aggregate Structural Change

The spatial equilibrium of economic activity is of course driven by aggregate demand and supply conditions. Our economy does not admit a representative consumer, since the PIGL preference specification in (5) falls outside of the Gorman class. In particular, consider a set of individuals $i \in \mathcal{S}$, with spending e_i . The aggregate demand for agricultural products of this set of consumers is given by

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

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

Our model, however, delivers tractable expressions for the economy's aggregate quantities. The distributional assumption on individual skills implies that individual income y^i is Frechet distributed. Combined with Proposition 1, which showed that individuals' expenditure policy functions are linear, this implies that individual spending e is *also* Frechet distributed. This allows us to solve for the aggregate demand system explicitly as a function of equilibrium wages.

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

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

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

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

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

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

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

Proposition 3 is an “almost-aggregation” result. Even though the PIGL preferences fall outside of the Gorman class, (14) shows that the aggregate demand of a given set of consumers resembles that of a representative consumer with mean spending $E_{\mathcal{S}}$ and an adjusted preference parameter $\tilde{v}$. Because the linearity of individuals’ policy functions allows to express aggregate spending directly as a function of equilibrium spatial wages, aggregate sectoral spending PC_t^A is then simply the spatial aggregate over the respective consumer groups and the aggregate value added share of the agricultural sector takes the form in (15). Note that ϑ_t^A can be directly calculated from current and past wages $\{w_{rt}, w_{rt-1}\}$, the spatial allocation of factors $\{L_{rt}, L_{rt-1}\}$. 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 rise. Similarly, changes in relative technological progress (and therefore in sectoral prices) will affect agricultural spending as long as $\gamma \neq 0$.

Equilibrium Conditions

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

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

where s_{rAt} is the agricultural employment share in region r at time t .¹⁷ 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, π_{rAt} , in aggregate spending on agricultural goods $\vartheta_t^APY_t$. The CES structure of consumers’ preferences implies that regional trade shares, π_{rAt} , are given by

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

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

To characterize the equilibrium, we find it useful to express the sectoral market clearing conditions in

¹⁷Because efficiency units are fully substitutable, the equilibrium only determines the allocation of efficiency units. To relate the model to the data, we assume that allocation of efficiency units and people is proportional. Once we explicitly introduce sector-specific human capital in Section 3.4, this indeterminacy will be resolved.

(16), as

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

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

Equation (18) shows that local labor earnings are a *demand-weighted* average of regional sectoral trade shares and hence highlights the urban bias of the structural transformation: a decline of the aggregate agricultural spending share ϑ_t^A tends to reduce regional earnings in locations who have a comparative advantage, in the agricultural sector. Moreover, equation (19) illustrates the spatial comovement of sectoral employment shares. Because regional (scaled) agricultural shares $\frac{s_{rAt}}{1 - s_{rAt}}$ are proportional to the aggregate (scaled) agricultural expenditure share $\frac{\vartheta_t^A}{1 - \vartheta_t^A}$, a decline in the aggregate spending share ϑ_t^A tends to reduce agricultural employment shares in *all* locations. In Section 3.3 we are going to derive yet sharper predictions for these two relationships in a special case of our model.

Additionally, these equations highlight how the spatial distribution of economic activity is fully determined from *static* equilibrium conditions. Note first that GDP is proportional to aggregate labor earnings, i.e. $(1 - \alpha)PY_t = \Gamma_{\zeta} \sum_r L_{rt}w_{rt}$ (see (18)) and that the spatial labor supply function is *only* a function of the the spatial distribution of wages (see Proposition (2)). Likewise the agricultural value added share 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.

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

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

i.e. future capital is simply a fraction $1 - \psi(r_{t+1})$ of aggregate labor earnings. This proportionality between the aggregate capital stock and aggregate GDP is a consequence of the linearity of agents' consumption policy rules. A dynamic equilibrium then requires that the set of interest rates $\{r_t\}_t$ is consistent with the implied evolution of the capital stock. More formally, a competitive equilibrium in our economy is defined in the following way.

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

It is worthwhile pointing out that our model retains many features of the baseline neoclassical growth model. In particular, for given initial conditions $\{K_0, L_{r,-1}, w_{r,-1}\}$ and a path of interest rates $\{r_t\}_t$, the equilibrium evolution of wages and people are solutions to the static equilibrium conditions highlighted above. Given these allocations, the model predicts the evolution of the capital stock according to (20). A full dynamic equilibrium is then a fixed point for the path of equilibrium interest rates. In practice, it is straightforward to calculate this dynamic equilibrium even for a realistic economic geography of hundreds of locations.

The special case of our theory with just a single location reveals that our model is essentially a standard, macroeconomic model of the structural transformation augmented by a spatial layer. With a single region, it is easy to show that GDP is given by

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

where $Z_t \equiv \Gamma_\zeta^{1-\alpha} Z_{At}^\phi Z_{NA_t}^{1-\phi}$. Moreover, capital still accumulates according to (20). Even though consumers' savings rates $1 - \psi_{t+1}$ depend on future interest rates (and hence on the future capital stock K_{t+1}), it can be shown that - given some initial condition K_0 and processes for productivity $\{Z_{At}, Z_{NA_t}\}_t$ - there exists a unique dynamic equilibrium path of capital $\{K_t\}_t$. Section B.4 in the Appendix provides a formal proof. Further, as in the baseline macroeconomic model of the structural transformation, this equilibrium path can be characterized *independently* of the sectoral labor allocation (see e.g. Herrendorf et al. (2014)). To see this, suppose that the economy is on a balanced growth path where aggregate income, capital and wages grow at rate g and the interest rate is constant. In Section B.4 in the Appendix we characterize this path and show that the agricultural share in value added ϑ_t^A is given by

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

where 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, note that in the single region case, sectoral employment shares are equal to sectoral value added shares, i.e. $s_{At} = \vartheta_t^A$ (see (19)). Hence, as in most frictionless models of the structural transformation, value added per worker is equalized across sectors. This implication is inconsistent with the large literature on the “agricultural productivity gap”, which finds that value added per worker in the agricultural sector is relatively low (see e.g. Gollin et al. (2014)). This is no longer the case in our spatial economy with multiple regions. In particular, value added per worker in agriculture *relative* to average value added per worker is given by

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

Hence, there is an agricultural productivity gap, i.e. $\vartheta_t^A < s_{A,t}$, whenever the spatial correlation between

wages and agricultural employment share is *negative*. Differences in regional efficiency combined with frictions to spatial mobility are therefore one potential mechanism to explain why agricultural productivity low. In Section 5 we quantify this spatial agricultural productivity gap for the US.

3.3 The Urban Bias of the Structural Transformation

A key property of our theory is the spatial bias of the structural transformation whereby rural areas are hurt by the secular change in spending patterns. This bias is of course at the heart of the induced spatial reallocation, whereby declining wages push people out of agricultural areas. To see this mechanism more formally, consider a special case of our model where there are no moving costs and individuals have no locational preferences.¹⁸ This parametrization implies that equilibrium wages (and individual welfare) are equalized across space at each point in time. The equilibrium trade shares, π_{rst} are therefore given by $\pi_{rst} = Q_{rst}^{\sigma-1}$, i.e. are fully exogenous and only depend on regional productivity (see (17)). Equations (18) and (19) therefore reduce to

$$\begin{aligned} L_{rt} &= Q_{rAt}^{\sigma-1} \vartheta_t^A + Q_{rNA_t}^{\sigma-1} (1 - \vartheta_t^A) \\ \frac{s_{rAt}}{1 - s_{rAt}} &= \frac{\pi_{rAt} \vartheta_t^A}{\pi_{rNA_t} (1 - \vartheta_t^A)} = \left(\frac{Q_{rAt}}{Q_{rNA_t}} \right)^{\sigma-1} \frac{\vartheta_t^A}{1 - \vartheta_t^A}. \end{aligned}$$

Hence, the size of the local *population* is a demand-weighted average of the exogenous local sectoral productivities and regional agricultural shares are proportional to the aggregate agricultural expenditure share. These equations concisely illustrate the urban bias of the structural transformation and the determinants of the spatial reallocation component. In particular, the structural transformation can again be summarized by a decline in the sectoral spending share, $d\vartheta_t^A < 0$. The induced change in the local population is then given by

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

The first term on the right is what we refer to as the urban bias: A decline in the spending share on agricultural goods reduces population in region r if and only if region r has a *comparative* advantage in

¹⁸Formally, this parametrization is nested in our full model by assuming that $MC_{jr} = 0$ and $A_r = 0$,

the agricultural sector. In fact, it is easy to show that¹⁹

$$\text{sgn}((Q_{rAt}^{\sigma-1} - Q_{rNAAt}^{\sigma-1})d\vartheta_t) = \text{sgn}((s_{rAt} - s_{At})d\vartheta_t).$$

Hence, the structural transformation is biased against regions in which the agricultural employment share *exceeds* the aggregate share of agricultural employment. In particular, if it was not for local productivity shocks, regional population growth and the initial agricultural employment share would be perfectly negatively aligned and spatial reallocation would account for a sizable share of the decline in agricultural employment.

The expressions above also show the extent to which the structural transformation affects the sectoral structure of the local economy. In particular,

$$d \ln \left(\frac{s_{rAt}}{1 - s_{rAt}} \right) = d \ln \left(\frac{\vartheta_t^A}{1 - \vartheta_t^A} \right) + (\sigma - 1) d \ln \left(\frac{Q_{rAt}}{Q_{rNAAt}} \right).$$

In the absence of local productivity shocks, *all* regions' degree of sectoral specialization moves in the same direction, i.e. the process of structural change indeed transforms all regions in the economy. Idiosyncratic shocks to comparative advantage break this tight relation across space as increases in relative agricultural productivity Q_{rA}/Q_{rNA} will increase agricultural employment despite the secular agricultural decline in the aggregate.

These expressions already anticipate that local productivity shocks, dQ_{rst} , are going to be important to explain why spatial reallocation plays a minor role for the decline in the agricultural employment share. In our full model, three additional channels are at work, which also affect the incentives to spatially reallocation. First, costly spatial mobility prevent wages from being equalized and tends to keep individuals in their location. Second, local amenities are an additional determinant of moving flows. Whether amenities contribute to the spatial reallocation component depends on their correlation with regional agricultural employment shares. Finally, idiosyncratic shocks reduce the sensitivity of moving flows to wage differentials and generate gross mobility flows, which exceed the extent of net reallocation. Because the urban bias affects individual moving incentives through changing factor prices, idiosyncratic preference shocks dampen the correlation between agricultural employment shares and population outflows.

3.4 Selection, Human Capital and Labor Supply

So far we assumed that human capital is perfectly substitutable across sectors and that all individuals are ex-ante identical. In preparation for the quantitative exercise, we now extend it to allow for imperfect

¹⁹To see this, note that

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

substitutability of efficiency units across industries and for systematic differences in skill-supply. In particular, we assume that individuals draw a two-dimensional vector of skill-specific efficiency units $z^i = (z_A^i, z_{NA}^i)$ and sort across industries based on their comparative advantage. We also assume that individuals can be of two types - high skilled and low skilled. Their skill type $h \in \{L, H\}$ determines the distribution of z^i . As before, we assume that z^i is drawn from a Frechet distribution

$$F_s^h(z) = e^{-\Psi_s^h z^{-\zeta}}, \quad (22)$$

where Ψ_s^h parametrizes the average level of human capital of individuals of skill type h in sector s and ζ governs the dispersion of skills. Without loss of generality we parameterize Ψ_s^h as

$$\begin{bmatrix} \Psi_A^L & \Psi_{NA}^L \\ \Psi_A^H & \Psi_{NA}^H \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ q & q\mu \end{bmatrix}.$$

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.²⁰ 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.

We are adding these additional ingredients for two reasons. First, the extent of skill substitutability across sectors is an important determinant of the costs of the *regional* transformation. While moving costs are a hurdle for spatial reallocation, an upward sloping relative sectoral supply function within locations makes it costly to reallocate agricultural workers to factories.²¹ Second, systematic differences in skills are important in that they generate spatial sorting. Because skilled workers have a comparative advantage in manufacturing jobs, they are more likely to move to locations which are productive in the manufacturing sector. In this way our model captures the fact that the distribution of skills and local productivity fundamentals is endogenous.

Because of the properties of the Frechet distribution, these additional ingredients leave the rest of the analysis almost unchanged.²² In particular, total earnings of individual i residing in region r are given by $y_r^i = \max \{w_{rA} z_A^i, w_{rNA} z_{NA}^i\}$, where w_{rs} is the prevailing equilibrium wage in region r and sector s (which now is no longer equalized across industries within locations). Average earnings of individuals in skill group h are then given by $E^h[y_r^i] = \Gamma_\zeta \Theta_r^h$, where

²⁰Hence, we abstract from human capital accumulation and simply assume that skills are fully inherited between parents and children. Our timing assumption there implies that individuals know their skill $h \in \{L, H\}$ prior to migrating but not the realization of their z^i .

²¹This is the consequences of worker selection stressed by [Lagakos and Waugh \(2013\)](#): the *average* amount of efficiency units provided to sector s by individuals of skill group h is given by $\frac{H_{rs}^h}{L_r s_{rs}^h} = (s_{rs}^h)^{-\frac{1}{\zeta}} (\Psi_s^h)^{\frac{1}{\zeta}}$, i.e. is decreasing in the sectoral employment share s_{rs}^h as individuals' sorting implies that the marginal worker in sector s is worse than the average worker.

²²The key property is the fact that the Frechet distribution is max-stable.

$$\Theta_{rt}^h = \left(\Psi_A^h w_{rAt}^\zeta + \Psi_{NA}^h w_{rNA t}^\zeta \right)^{1/\zeta}. \quad (23)$$

Note that while Θ_{rt}^h , which we also refer to as *regional income*, differs across skill-types, it is equalized across sectors and can be directly calculated from regional wages (w_{rA}, w_{rNA}) . In particular, high skilled individuals put a higher relative weight on non-agricultural wages $w_{rNA t}$ and hence find locations with a strong manufacturing sector particularly attractive. This spatial sorting captures a distinct role for the *supply* of agricultural labor. A shift of people towards urban areas will, for example, reduce the relative supply of agricultural goods at given prices. Spatial mobility costs are therefore one example of cohort-specific “sectoral moving costs” highlighted by [Hobijn et al. \(2018\)](#) and [Porzio and Santangelo \(2017\)](#).

As we show explicitly in Section B.5 of the Appendix, the characterization of the equilibrium is essentially the same as in our baseline economy. The main differences are that (i) regional incomes Θ_{rt}^h now take the role of the equilibrium wage w_{rt} , (ii) the equilibrium conditions in 18 and 19 are weighted averages of sectoral labor supplies by skill group, (iii) the relative supply of sectoral efficiency units is given by $s_{rs}^h = \Psi_s^h (w_{rs}/\Theta_{rt}^h)^\zeta$ and hence depends on the relative wage with an elasticity ζ and (iv) spatial mobility choices differ by skill-group as the regional attractiveness in Proposition 2 is skill-specific and given by $\mathcal{W}_{rt}^h = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_{rt}^h)^\eta + A_{rt}$. Hence, the law of motion of the population in 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.

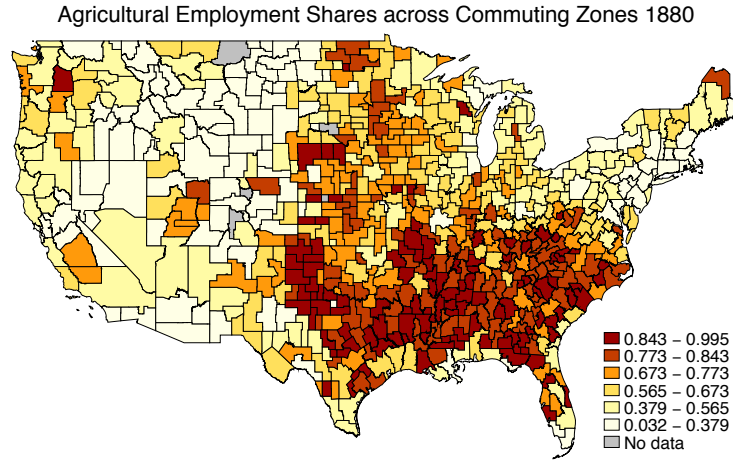
4 Spatial Structural Change in the US: 1880 - 2000

We now calibrate the model to key moments of the spatial growth experience of the United States over the last 120 years. We do so by exploiting a novel panel data set on the regional development of the United States between 1880 and 2000.

4.1 Data

To compose our panel we draw on various data sets published by the US Census Bureau. In particular, we use information from the Census of Manufacturing for 1880 and 1910, the Population Census for 1880-2000 and the County and City Data Books for 1940-2000. From these sources we construct a panel data set of total workers $\{L_{rt}\}_{rt}$, average manufacturing wages $\{w_{rt}\}_{rt}$, and sectoral employment shares $\{s_{rAt}\}_{rt}$ for all US counties at 30 year intervals between 1880 and 2000.²³ We define the agricultural sector to comprise agriculture, fishing and mining industries following the 1950 Census Bureau industrial

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



Notes: The figure shows the agricultural employment share across US commuting zones in 1880.

Figure 2: Agricultural Employment Share across US Commuting Zones in 1880

classification system (outlined in [Ruggles et al. \(2015\)](#)). All remaining employed workers are assigned to the non-agricultural sector. We construct average manufacturing wages from county level total manufacturing payroll data and manufacturing head counts obtained from the same source. Table (16) in Appendix (C) 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 as constructed by [Tolbert and Sizer \(1996\)](#).²⁴ We do so for three reasons. First, we need stable regional boundaries over time. Second we want the union of all regions to cover the entire continental United States. Lastly, regions should be large enough such that both an agricultural and a non-agricultural sector is present in each region. Commuting zones partition the territory of the United States into 741 polygons, all of which exhibited non-zero non-agricultural employment shares in 1880.

This procedure leaves us with a panel data set, which features sectoral employment shares, total employment and average manufacturing earnings for all continental commuting zones from 1880 to 2000. As a visual example, consider Figure 2, where we depict the distribution of agricultural employment shares across commuting zones. While some regions (in particular, commuting zones in Northeastern states like Massachusetts or New York) already have agricultural employment shares of less than 10%, many commuting zones in the South have more than 75% of their population employed in the agricultural sector. For the main calibration of the model we employ the cross-sections 1880, 1910, 1940, 1970 and 2000 only and normalize the size of the total US workforce to unity in each period.²⁵

In addition, we rely heavily on the 1940 edition of the decennial Micro Census by the US Census Bureau

²⁴To do so, we construct a crosswalk between counties and 1990 commuting zones for every decade between 1880 and 2000. Using this we re-aggregate the county level data for the various years to commuting zones, employing area weights to allocate workers wherever counties are split.

²⁵A 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. \(2018\)](#)).

(see [Ruggles et al. \(2015\)](#)). This is the most recent Census which contains individual identifiers for all US counties and it is the first Census for which earnings and education variables are available.²⁶ We use this information to calibrate the spatial distribution of skilled workers.

To estimate the moving cost parameters we exploit the information in the Population Census. In the model, individuals move once in their lifetime to access their preferred labor market. In the data we therefore look at patterns of lifetime mobility. All censuses between 1880 and 2000 contain information on the state of residence and the state of birth. Focusing on workers between 26 and 50 years of age, we can therefore calculate state-by-state life-time mobility matrices for the periods 1880-1910, 1910-1940, 1940-1970 and 1970-2000.

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

4.2 Calibration

Calibration Strategy

In this section, we outline our calibration strategy. Section B in the Appendix provides considerably more detail.

Aggregate time series 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.²⁷ We calibrate the time series of aggregate sectoral productivities, $\{Z_{At}, Z_{Mt}\}_t$, to match the evolution of relative prices and a GDP growth rate of 2%.

Skill Supply 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 a high school education in 1940 and hold the aggregate share of skilled workers fixed.²⁸ This choice yields an

²⁶We use the Public-use Micro Census data compiled and maintained [Ruggles et al. \(2015\)](#). In the publicly available samples counties are censored and only become available after 70 years. As a result we cannot identify all counties in Census cross-sections beyond 1940, which is the cross-section most recently de-censored (2010).

²⁷In the counterfactual analysis, interest rates will of course be free to vary over time.

²⁸Because 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.

aggregate skilled employment share of about 0.3. We then calibrate $\{\lambda_{r1880}\}_r$ for the model to exactly replicate the the spatial skill distribution in 1940, which is the first year for which we can calculate educational attainment at the commuting zone level. We calibrate the parameter ζ , i.e. the dispersion of individual productivity, to match the dispersion of earnings in the 1940 Census data. In particular, the model implies that the variance of log earnings within region-skill cells is given by $\frac{\pi^2}{6}\zeta^{-2}$. We therefore identify ζ from $\zeta = \frac{\pi}{6^{1/2}} \text{var}(\hat{u}_{rsh}^i)^{-1/2}$, where $\text{var}(\hat{u}_{rsh}^i)$ is the variance of the estimated residuals from a regression of log earnings on region, sector and skill-group fixed effects in the 1940 Census Data.

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

Moving Costs and Idiosyncratic Location Preferences We specify the costs of moving, denominated in utils, as a quadratic function of distance, d_{jr} , i.e.

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

whenever $j \neq r$ and zero otherwise. Because in the theory workers move only once (at the beginning of their working life), for the remainder of the paper all mention of migration refers to lifetime migration.²⁹ Here, $\tau > 0$ corresponds to a fixed cost of moving and δ_1 and δ_2 parameterize how mobility costs vary with distance. We normalize d_{jr} so that the maximum distance in the US is 1.

In the Census data we observe the state of birth and the current state of residence of every worker in every decade. In the model we aggregate commuting zone flows to the state level and choose $(\kappa, \tau, \delta_1, \delta_2)$ so as to make the model fit the observed state level flows. Because according to our model migration flows are skill-specific, we calibrate our model to the data on migration flows between 1910 and 1940, as 1940 is the first year where the population census contains information on educational attainment. Note also that 1940 is exactly in the middle of our sample period. We chose τ so as to match the aggregate lifetime interstate migration rate between 1910 and 1940 exactly.³⁰ We then chose $(\kappa, \delta_1, \delta_2)$ to minimize the distance between spatial mobility rates in the data and the model.³¹

²⁹The aggregate lifetime migration rate refers to the fraction of people who live in a different location than where they were born (see Molloy et al. (2011)).

³⁰As shown in Section B.6 in the Appendix, the number of stayers in a commuting zone is a monotone function in τ given $(\kappa, \delta_1, \delta_2)$.

³¹More 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 .

Productivities and Amenities We calibrate local productivities and amenities $\{Q_{rst}, A_{rt}\}_{rst}$ as structural residuals, i.e. we force the model to match the spatial data on agricultural employment shares, populations and average manufacturing wages perfectly with $\{Q_{rst}, A_{rt}\}_{rst}$ absorbing any residual variation. In Appendix (B.1) we 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 $\{Q_{rst}\}_{rst}$. Intuitively, the local employment shares contain information on $\frac{Q_{rAt}}{Q_{rNA_t}}$ while the level of wages along with the total number of workers informs the level of Q_{rNA_t} . The vector of amenities $\{A_{rt}\}_{rt}$ can then be inferred from the observed population flows together with evolution of skill shares.

Preference parameters We use the micro data from the Consumer Expenditure Survey in 1936 (“Study of Consumer Purchases in the United States, 1935-1936”) to estimate the extent of non-homotheticities, i.e. η from the cross-sectional relationship between sectoral spending shares and the level of expenditure. We then use the time-series of the aggregate agricultural employment share to identify the remaining parameters (ϕ, v, γ) . As γ determines the price elasticity of demand, we discipline γ with the elasticity of substitution. Comin et al. (2017) estimate this elasticity to be around 0.7 in post-war data for the US. Given our calibration strategy for the underlying distribution of productivity and amenities, internal consistency requires us to match the time series of agricultural employment shares exactly. As we discuss in detail in Section A.5 in the Appendix, the income effects as implied from the cross-sectional spending-food relationship are not strong enough to explain the entire decline in agricultural employment in the time-series.³² We therefore allow the parameter ϕ to be time-specific to fully account for the residual decline in agricultural employment and choose v to minimize the required time-variation in ϕ_t . Intuitively, v is chosen for the model to explain as much of the aggregate process of structural change as possible, given the income and price elasticities η and γ . Recall that ϕ does not enter the household’s decision problem directly. Finally, we chose the rate of time preference β to be consistent with the aggregate rate of investment.

Other parameters Finally, we need values of the capital share α , the rate of depreciation δ and the preference parameter σ . We do not estimate these parameters within our model, but rely on central values from the literature.

Calibration Results and Model Fit

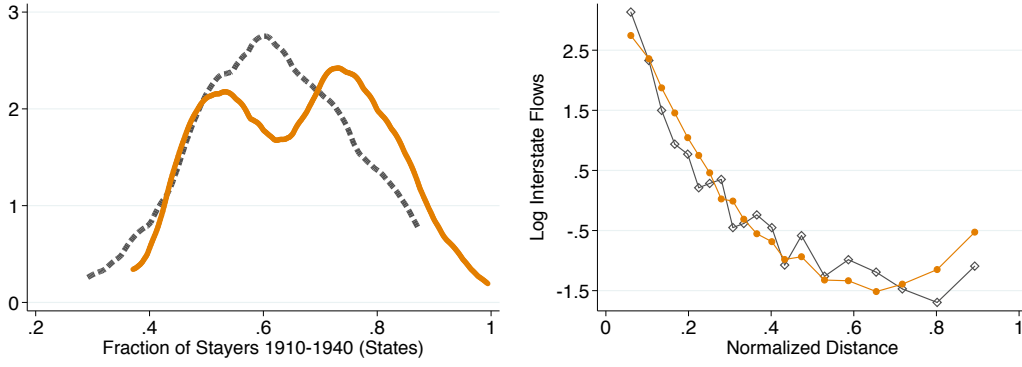
In Table 1 we report the calibrated parameters and the main targeted moment, both in the data and the model. Naturally, the parameters are calibrated jointly.

³²This 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.

Parameter	Target	Value	Data	Moments
				Model
Skill Supply				
ζ	Skill heterogeneity	Residual Earnings variance in 1940	1.62	0.62
ξ	Share of skilled individuals	Share with more than high school degree in 1940	0.3	0.3
μ	Comparative advantage	Rel. non-ag. share of skilled workers in 1940	3.41	1.21
q	Absolute advantage	Skill premium in 1940	0.68	1.61
$[\lambda_r, 1880]$	Initial distribution of skilled individuals	Skill distribution in 1940	$\{\lambda_r, 1940\}$	match data perfectly
Regional Fundamentals				
$[Q_{rAt}]_{rst}$	Agricultural productivity	Regional empl. shares and earnings	See Appendix	$\{e_{rt}^{Man}, s_{A,rt}\}$
$[Q_{rNAr}]_{rst}$	Non-agricultural productivity			
$[A_{rt}]_{rt}$	Amenities	Regional population		$\{L_{rt}^h\}$
Time Series Implications				
$[Z_{NAr}]$	Non-agricultural productivity	Aggregate growth rate of GDP pc	See Appendix	2%
$[Z_{Ar}]$	Agricultural productivity	Relative price of ag. goods	See Appendix	$\{P_{A,t}/P_{NA,t}\}$
Preference Parameters				
β	Discount rate	Investment rate along the BGP	0.29	0.15
ϕ	Ag. share in price index	Time series of ag. empl. share	See Appendix	$\{s_{A,t}\}$
ν	PGL Preference parameter	Time series of ag. empl. share	0.017	0.20
η	Non-homotheticity	Ag.share - expenditure relationship	0.35	0.04
γ	Price sensitivity	Elasticity of substitution in 2000	0.35	0.7
Moving Costs and Mobility				
τ	Fixed costs of moving	Lifetime interstate migration rate	1.63	0.32
κ	Dispersion of idiosyncratic spatial tastes	Observed state-to-state flows	0.42	Estimated with NLS
$[\delta_1, \delta_2]$	Distance (miles) 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 annually)	.
α	Capital share in production function	set exogenously	1/3	.

Notes: The table contains the calibrated parameters. See Section 4.2 for the calibration strategy.

Table 1: Calibrated structural parameters



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

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

where ρ_{jr} denote 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 3: Lifetime state-to-state migration: Model vs Data

The presence of *spatial mobility costs* is an important component of our theory. To get a sense of how well the model matches important features of the lifetime migration data, we report the distribution of stayers across states and the relationship between moving flows and distance in Figure 3. In the left panel we depict the cross-sectional distribution of the share of “stayers”, i.e. 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 distance to more attractive places, there is sizable heterogeneity in the extent to which regions are able to retain workers. Figure 3 shows that the model matches this cross-sectional heterogeneity, even though it is only calibrated to match the average rate of mobility. 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

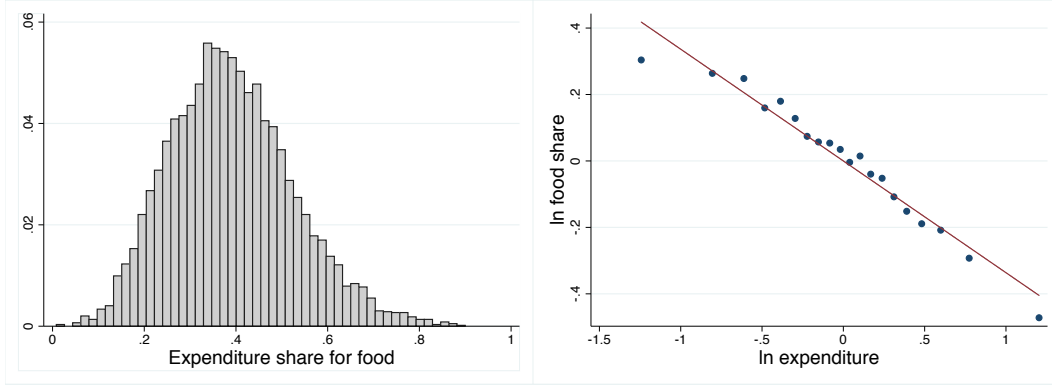
$$\log \frac{\rho_{jr}}{1 - \rho_{jj}} = \alpha_j + \beta_r + u_{jr},$$

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 actual data and the data stemming from the model. The right panel of Figure 3 shows that the model captures this systematic pattern of lifetime migration flows well. In particular, spatial mobility is very local, i.e. spatial flows are steeply decreasing in distance.³³

³³To get a sense of the *economic* magnitude of our estimates of moving costs, consider an individual of skill h born in region j . The expected value of moving (to her preferred location) is given by

$$\mathcal{U}_{jt}^{h,Mov} = E \left[\max_{r \neq j} \left\{ \mathcal{U}_{jr}^{ih} \right\} \right] = \tilde{\gamma} - \tau + \ln \left\{ \left[\sum_{r \neq j} \exp \left(\mathcal{W}_r^h - f(d_{jr}) \right)^{1/\kappa} \right]^\kappa \right\},$$

where τ denotes the fixed cost of moving, $f(d_{jr}) = \delta_1 d_{jr} + \delta_2 d_{jr}^2$ is the component of moving costs, which is distance-specific, and $\tilde{\gamma}$ is a constant, which depends on the Euler constant and $\Lambda_{t,t+1}$. Similarly, the expected value of staying in



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

Figure 4: Expenditure Shares on Food in 1936

In terms of *consumers' preferences*, note that the demand system implies that expenditure shares at the individual level are given by $\vartheta_A(e, p) = \phi + v \left(\frac{p_A}{p_M} \right)^\gamma e^{-\eta}$. For $\phi \approx 0$, this implies that there is a log-linear relationship between expenditure shares and total expenditure. In Figure 4 we depict the cross-sectional distribution of the expenditure share for food (left panel) and 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 . From this plot, it is clear that there is substantial heterogeneity in the expenditure share on food in the cross-section of households. Moreover, the expenditure share is systematically declining in the level of expenditure and the cross-sectional relationship is essentially log-linear, as predicted by the theory. The slope coefficient implies that $\eta = 0.32$.³⁴

An important part of our calibration strategy is that we calibrate the cross-sectional distribution of sectoral productivities $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$ and amenities $\{A_{rt}\}_{rt}$ as structural residuals of the model in line with the new quantitative spatial economics literature (Redding and Rossi-Hansberg, 2017). Hence, given the remaining parameters reported in Table 1, the model matches the population distribution, the regional agricultural employment shares and average manufacturing earnings in each cross-section exactly. 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 0.5. Similarly, the correlation of productivity and amenities is also positive. Hence, productive places are *also* relatively pleasant places to live, in line with recent direct evidence for developing countries by Gollin et al. (2017). In Section B.9 in the Appendix, we provide additional details about the estimates of spatial fundamentals.

region j , $\mathcal{W}_{jt}^{h, Stay}$, is given by $\tilde{\gamma} + \mathcal{W}_{jt}^h$ as $\tau = f(d_{jj}) = 0$. In our calibrated model, we find that for the median commuting zone, the expected value of moving, $\mathcal{W}_{jt}^{h, Mov} - \mathcal{W}_{jt}^{h, Stay}$ is slightly positive and amounts to roughly 5% of the estimated fixed cost of moving. See also Section B.9 in the Appendix, where we discuss this in more detail.

³⁴For 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.

Finally, we also related our measured regional productivities Q_{rst} to direct measures of local productivity. In particular, we show, using their data, that the regional growth in market access documented by [Donaldson and Hornbeck \(2016\)](#) is correlated with an increase in $\{Q_{rA}, Q_{rNA}\}_r$ by our model. See Section B.11 in the Appendix for details.

4.3 Why is the reallocation component of the structural transformation small?

In Section 2, and especially Figure 1 we showed that the structural transformation in the US is entirely driven by a within-region transformation - the reallocation of individuals from agricultural regions to non-agricultural regions played only a minor role. This section explains why this is the case.

In our model, individuals spatially relocate to arbitrage away wage differences and to move towards high amenity regions. Hence, spatial mobility contributes to the decline in aggregate agricultural employment, if (i) moving frictions are not prohibitively large, (ii) regional employment shares are negatively correlated with *future* earning and *future* amenities and (iii) idiosyncratic taste shocks are of limited importance for moving decisions.

Observed gross and net moving flows are informative about (i) and (iii). The high level of gross life-time interstate migration observed throughout the 20th century hints at the fact that moving frictions play a minor role in explaining the absence of the reallocation component in the structural transformation. Similarly, the high level of bi-directional flows in the data, leads the model to assign an important role to idiosyncratic preference shocks, i.e. there is a low elasticity of moving flows with respect to spatial wage differences. Finally, turning to (ii), our theory highlights three forces, which shape the correlation between initial regional employment shares and future earnings. First of all, to the extent that wages show some persistence and population mobility is slow, individuals will leave rural areas if wages in such regions are low. Secondly, as discussed above, the urban bias of the structural transformation endogenously induces a negative correlation between agricultural employment and future earnings as consumer demand systematically shifts towards non-agricultural goods. Thirdly, the effect of changes in regional productivity depends on the correlation between local productivity growth with agricultural employment shares.

In Table 2 we provide direct reduced form evidence for each of these margins. In the first two columns, we report the results from a regression of log earnings on the (log of the) agricultural employment share. There is a significant negative correlation, especially for high-skilled workers. In columns 3 and 4, we test for evidence of the urban bias. In particular, we consider the specification

$$\ln \Theta_{rt}^h = \delta_t + \alpha \ln s_{rAt-1} + \beta \ln s_{rA,t-1} \Delta s_{At} + \gamma \ln \Theta_{rt-1}^h + u_{rt}, \quad (24)$$

where $\ln \Theta_{rt}^h$ denotes regional income, $\ln s_{rAt-1}$ is the lagged agricultural employment share and $\Delta s_{A,t}$ denotes the change in the aggregate agricultural employment share between t and $t - 1$. The coefficient α captures the direct effect of agricultural specialization. The coefficient β captures the urban bias and

	<i>Current Earnings</i>		<i>The Urban Bias</i>		<i>Local Productivity Growth</i>		<i>Future Amenities</i>
	$\ln \Theta_{rt}^L$	$\ln \Theta_{rt}^H$	$\ln \Theta_{rt+1}^L$	$\ln \Theta_{rt+1}^H$	$d \ln Q_{rA}$	$d \ln Q_{rNA}$	A_{rt+1}
$\ln s_{rAt}$	-0.059*** (0.004)	-0.149*** (0.005)					
$\ln s_{rAt-1}$			-0.032*** (0.005)	-0.036*** (0.005)	0.175*** (0.036)	-0.365*** (0.055)	-0.223*** (0.018)
$\ln s_{rAt-1} \times \Delta s_{At}$			0.516*** (0.066)	0.475*** (0.079)			
Year FE	✓	✓	✓	✓	✓	✓	✓
Lagged Earnings			✓	✓			
Lagged Fundamentals					✓	✓	✓
N	2580	2580	1935	1935	2580	2580	2580
R^2	0.894	0.918	0.908	0.924	0.428	0.409	0.088

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

Table 2: Income and Agricultural Specialization

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. Columns 3 and 4 show direct evidence for the urban bias: a decline in agricultural employment share particularly hurts regions with a large initial agricultural share. Moreover, this urban bias is present for both low and high skilled workers. In columns 5 and 6 we regress the growth rate of local productivity, $d \ln Q_{rst}$, on the initial agricultural employment share and the respective current level of productivity, $\ln Q_{rst}$. Regions with a larger agricultural employment share tend to have higher (lower) productivity growth in the agricultural (non-agricultural) sector. However, even after controlling for the level of productivity, productivity shocks are quite noisy - the R^2 is only 0.4. Finally, the last column shows that rural areas tend to also have low future amenities. This is consistent with [Gollin et al. \(2017\)](#), who provide direct evidence that measured amenities in rural areas are much lower than in cities.

The results in Table 2 show that there are multiple forces at play that determine the relationship between agricultural shares and future earnings. This suggests that the relationship between initial agricultural employment and future earnings might be weak. This is exactly what we find. If, for example, we run a regression of future earnings or amenities on a full set of 100 fixed effects for the percentiles of initial agricultural employment shares, we explain at most 1/3 of the cross-sectional variation.³⁵

To quantify the importance of these different margins, let us go back to the “frictionless” parametrization of our model introduced in Section 3.3 and also assume that there are *no* changes in regional productivities.³⁶ This implies that the urban bias is the *only* driver of spatial mobility (see (21)). When we calibrate this model to the same aggregate time-series data, we find that this model has strikingly different implications for the link between initial agricultural employment shares and subsequent population flows. Consider for example Figure 5, where we depict the share of commuting zones experiencing population

³⁵See Section B.10 in the Appendix for more details.

³⁶This parametrization therefore assumes that there are (i) no moving costs ($MC_{jr} = 0$), (ii) no *changes* in regional fundamentals ($Q_{rst} = Q_{rs,1880}$), (iii) no amenities ($A_{rt} = 0$) and (iv) no idiosyncratic preferences for particular locations ($\kappa \rightarrow 0$).



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

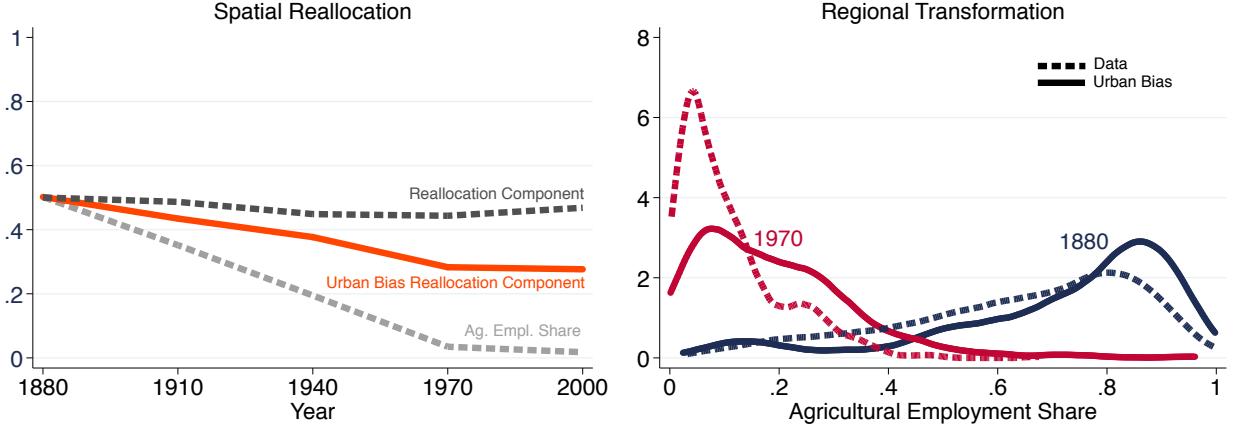
Figure 5: Inflows and Outflows

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 very high agricultural employment shares in 1910, about 20% see net population inflows. Similarly, 60% of the most non-agricultural areas see their population decline.³⁷ This is very different for the “urban bias” economy, where agricultural specialization and population outflows are perfectly aligned. In particular, this model implies that the *only* regions experiencing population inflows are the 15% of commuting zones with the lowest agricultural share in 1910. Hence, if only the urban bias mechanism was at play, the structural transformation would have induced much more population growth in non-agricultural, urban localities.³⁸

This model therefore predicts a much bigger role for population mobility in the process of the structural transformation. This is seen in Figure 6, which has the exact same structure as Figure 1. The spatial reallocation of individuals now explains about one third of the aggregate decline in agricultural employment. Interestingly, as the model *overestimates* the role of spatial reallocation, it *underestimates* the extent of the structural transformation at the local level. In particular, there are more regions which remain dominated by the agricultural sector throughout the 20th century. Consider, for example, the two red densities on the far left, which corresponds to 1970. In the data (the dashed line), the vast majority of commuting zones have an agricultural employment share below 20%. In the model, there is a substantial number of regions, where 30%-40% of the workforce is still employed in the agricultural sector. The reason for

³⁷Note that Figure 5 reports the share of regions with net outflows, i.e. the extensive margin of population flows. In the aggregate, net population flows are by construction zero.

³⁸Note that this strong form of sorting is implied by (21): only counties with $s_{rAt} < s_{At}$ see their population increase. And because population size and agricultural employment shares are strongly negatively correlated, far less than half of the regions are predicted to experience population inflows.



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

Figure 6: Spatial Reallocation without frictions

this is that worker mobility and 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.

Finally, in Table 3 we provide a decomposition into the different channels. In the first two rows, we again report the results for the data and the “urban bias” economy. While the data hardly shows any decline in the agricultural employment share, the urban bias model implies that the agricultural employment share would have declined from 50% to 35% only through the spatial reallocation of workers. In the next three rows, we shut off different components of our theory. Spatial productivity shocks itself accounts for a sizable part of the “lack of spatial reallocation”. In particular, if productivity Q_{rst} was fixed at its level in 1880, the extent of spatial reallocation would have doubled from 3.2% to 6.5%. The reason is that productivity shocks introduce noise in the relationship between earnings growth and initial agricultural specialization.

The next two rows show that neither moving costs nor local amenities can explain the absence of reallocation. A reduction of moving costs would in fact *reduce* the role of spatial reallocation: the agricultural employment share would only fall by 0.6%, which is even less than the 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 are in fact a positive contributor to the spatial reallocation component. This is due to the fact that rural regions have (on average) low future amenities (see Table 2). Spatial amenities are therefore a force driving people out of agricultural areas.

The final difference between the baseline and the “urban bias” model is the importance of idiosyncratic

	Spatial growth	Moving costs	Amenities	Idio. shocks	Spatial Reallocation 1880-2000	Migration patterns Net migration rate	Turnover (gross/net)
Urban bias model					-15.2%	14.0%	1*
Baseline model	✓	✓	✓	✓	-3.2%	12.9%	2.8
Convergence		✓	✓	✓	-6.5%	10.7%	3.2
Moving Costs	✓		✓	✓	-0.6%	41.1%	2.4
Amenities	✓	✓		✓	+2.6%	6.0%	6.2

Notes: The table decomposes the spatial reallocation component into different margins. The first two rows contain the results for the frictionless and the baseline model. The third row abstracts from productivity convergence ($Q_{rst} = Q_{rs1880}$). The fourth row abstracts from moving costs ($MC_{jr} = 0$). The fifth row abstracts from regional amenities ($A_{rt} = 0$). The frictionless model has no idiosyncratic taste shocks (i.e. $\kappa \rightarrow 0$).

Table 3: Why is there no reallocation?

shocks. The reason why our model infers that idiosyncratic shocks are empirically important is prevalence of *bi-directional flows*. If idiosyncratic shocks were absent, all individuals would agree on the ranking of potential destinations and we would not see people moving in both directions. Hence, the gross flows between any pair of locations would coincide with the net flows. The presence of idiosyncratic shock breaks this implication and gross flows exceed the amount of net flows. In the last two columns of Table 3 we therefore report the net migration rate and the rate of “turnover”, i.e. the ratio of the gross and net migration rate. The latter can be interpreted as the number of bodies that have to be moved across space to reallocate one person “on net” between regions.

In the data and in our baseline model, this number is around 3, i.e. if two people move from region j to region r , one individual moves from r to j . Moreover, this number does not decrease dramatically across our different exercises and, if anything, increases. In the absence of moving costs for example, the net migration rate increases by 4, but so does the gross migration rate, keeping the turnover margin stable. If there were no regional amenities, the net migration rate would shrink to 6%, but the gross migration rate would stay roughly constant.³⁹ This is very different in the “urban bias” economy, where gross flows and net flows coincide. Importantly, the net migration rate between the baseline model and the frictionless economy is roughly the same. Hence, the extent of spatial reallocation is quite similar - but the correlation of net-reallocation and agricultural specialization is much higher in the frictionless economy.⁴⁰

³⁹Intuitively, amenities induce correlated moving flows across agents. This increases the net migration rate relative to the gross migration rate.

⁴⁰To 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.

5 The Aggregate Implications of Spatial Structural Change

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. We are particularly interested in quantifying the role of spatial mobility in shaping aggregate outcomes. Both the urban bias of the structural transformation and the stochastic evolution of regional productivities induce changes in the marginal product of labor across locations over time. The ease of labor mobility therefore crucially determines the extent to which such marginal product differences can be kept in check. Labor mobility therefore tends to increase allocative efficiency and reduces spatial inequality.

To quantify the importance of spatial mobility, we compare our baseline model with two alternative parameterizations. First, we consider a model without any labor mobility by making the costs of moving prohibitively high. Secondly, we consider a model, where spatial sorting is costless and based only on local productivity and aggregate demand conditions. More specifically, we assume that both moving costs are absent ($MC_{jr} = 0$) and agents do not derive utility from non-pecuniary features of particular locations ($A_r = 0$ and $\kappa = 0$), so that spatial sorting is only based on expected earnings.

We report the outcomes of these three different models in Table 4. 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. In the second row, we report the spatial reallocation component. Without spatial mobility, the reallocation component is naturally equal to zero. In the absence of moving costs and spatial tastes, the reallocation component is positive but still only slightly larger than the one in the data. This again shows the quantitative importance of regional productivity shocks: in Table 3, we showed that the reallocation component was 15% if regional technologies were held constant at their 1880 level. Conversely, even without spatial amenities and moving costs, the model generates a reallocation component, which is close to the one observed in the data.

In the lower panel, we report the implications for GDP growth and aggregate TFP.⁴¹ Spatial mobility contributed heavily to aggregate productivity growth in the US during the 20th century. Without labor mobility, income per capita would have been lower by almost 17.5% and this decline is entirely accounted for by lower TFP, i.e. more misallocation.^{42 43} If, on the other hand, spatial mobility was only based on regional earnings and mobility was free, aggregate productivity and income per capita would have been higher by about 2.2% and 3% respectively. Compared to the losses from the absence of spatial mobility, these numbers are relatively small. Also note that the frictionless net migration rate is only 22%. This suggests that the 13%, which we see in the data, was sufficient to capture most of the productivity gains through spatial reallocation.⁴⁴

⁴¹We calculate TFP from the perspective of an aggregate production function, i.e. $TFP_t = Y_t / (K_t^\alpha L_t^{1-\alpha})$.

⁴²Income per capita is not proportional to TFP because of capital accumulation. As the capital-output ratio is essentially constant in all parametrizations, we have that $Y \propto TFP^{\frac{1}{1-\alpha}}$.

⁴³

⁴⁴To put these results in perspective, [Bryan and Morten \(2017\)](#), using a static spatial equilibrium model, estimate higher productivity losses from limited labor mobility for Indonesia. In their specifications without agglomeration and endogenous

	<i>Baseline</i>	<i>Counterfactuals</i>	
		No Spatial Mobility	No Spatial Frictions
Decline in Agricultural Employment Share	-48.1%	-47.8%	-48.2%
Spatial Reallocation Component	-3.2%	0	-3.7%
Net migration rate	13%	0	22%
Increase in GDP pc (rel. to 1880)	10.78	8.89	11.09
<i>Difference to baseline</i>		-17.57%	2.87%
Increase in TFP (rel. to 1880)	4.88	4.27	4.99
<i>Difference to baseline</i>		-12.59%	2.22%
Agricultural productivity Gap (2000)	0.83	0.73	0.83
Spatial Income Inequality (2000)	0.19 / 0.21	0.58 / 0.58	0

Notes: The model with no spatial mobility corresponds to the baseline model except that individuals are not allowed to move (i.e. $MC_{jr} \rightarrow \infty$). The model without spatial frictions corresponds to the baseline model except that moving costs are zero (i.e. $MC_{jr} = 0$), that there are no amenities ($A_{rt} = 0$) and no idiosyncratic tastes (i.e. $\kappa = 0$). The net migration rate is the average migration rate between 1910 and 2000. TFP is measured as $TFP_t = Y_t / (K_t^\alpha L_t^{1-\alpha})$. The “Agricultural Productivity Gap” is measured as aggregate value added per worker in agriculture relative the rest of the economy, i.e. $(VA_t^A / L_t^A) / (VA_t / L_t)$. Spatial income inequality is measured as the spatial IQR in log average earnings ($\ln \Theta_{rt}^h$). We report the results for low and high skilled individual.

Table 4: Reallocation and Aggregate Productivity: Sectors versus Space

These *spatial* productivity gains are in fact reflected in the aggregate *sectoral* productivity gaps. Both the baseline model and the model without spatial frictions, for example, imply an agricultural productivity gap of around 17% (Gollin et al., 2014). Hence, value added per worker in agriculture is about 17% lower than for the average worker even though there are no sectoral frictions. These productivity differences are due to skill-based sorting, whereby high-skilled individuals both work in the non-agricultural sector within regions and move towards urban locations spatially. Without spatial mobility this gap would increase to 27%. Hence, even though the spatial reallocation component is seemingly small, it has important aggregate effects for sectoral productivity differences by reducing the spatial dispersion in the marginal product of factors.

One implication of these spatial differences is the regional dispersion in average earnings. In the last line of Table 4, we report the interquartile range of log earnings for low and high skilled individuals across US commuting zones. In the baseline model, this difference amounts to roughly 0.2 log points for both skill groups. Without spatial mobility, labor is inelastically supplied at the region level and spatial inequality would increase by a factor of three and amount to almost 0.6 log points. If, on the other hand, spatial mobility were costless and there were no amenities, utility equalization implies that regional earnings would be equalized within skill groups at each point in time.

These differences in spatial earnings, translate into spatial welfare differences. We focus on $\mathcal{W}_{rt}^h = E^h[U_r] + A_{rt}$ (see (12)) as our measure of regional welfare for skill group h . We showed above that

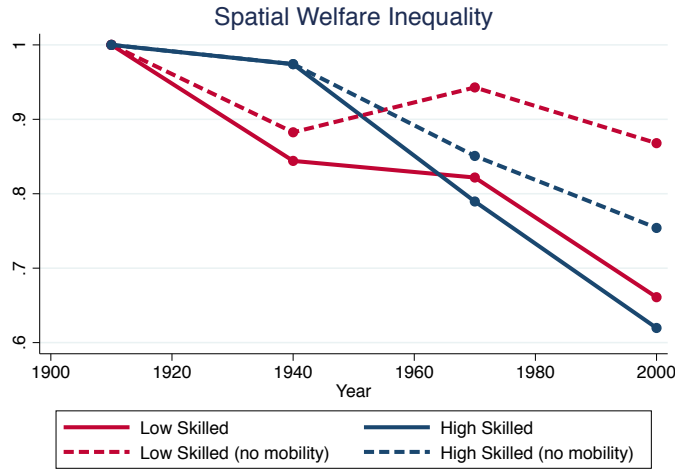
amenities, which we for simplicity abstract from in our model, they find that productivity would increase by about 17% if moving costs were zero and there were no amenity differences. There are two main reasons, why our results are small. First of all, by explicitly considering capital as a factor of 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.

regional income Θ_r^h (which, recall, is the sole determinant of $E^h[U_r]$) and regional amenities are negatively correlated with the agricultural employment share. Hence, welfare is systematically lower in agriculturally intensive places. It is also the case that $\mathcal{W}_{rt}^h$ is affected by the urban bias of the structural transformation: similar to average income (as reported in Table 2), a given decline of the aggregate agricultural share reduces welfare more the more agriculturally specialized the respective location.

To quantify to the evolution of spatial welfare inequality during the 20th century, we convert utility differences in “life-time-income” equivalents. Specifically, let $T_{rt}^h(\Delta)$ be the increase of expected lifetime income, an individual with skill h in region r requires to increase the average utility of living in region r by Δ , i.e.

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

Letting $\Delta_t^{h,IQR}$ denote the interquartile range of $\mathcal{W}_{rt}^h$ at time t , our measure of spatial inequality for individuals of skill h is then given as the regional average of $T_{rt}^h(\Delta_t^{h,IQR})$, i.e. $T_t^h(\Delta_t^{h,IQR}) = \frac{1}{R} \sum T_{rt}^h(\Delta_t^{h,IQR})$. In Figure 7, we depict the evolution spatial welfare inequality, relative to the level in 1910. The solid lines refer to our baseline calibration. Spatial welfare inequality decreased substantially during the 20th century. For both high and low skill workers, the required increase in lifetime income to raise utility from the 25% quantile to the 75% quantile of the spatial distribution of welfare decreased by about 35% since 1910. In terms of levels, in 1910 the average region would have required an increase in average lifetime income by about 160% to increase utility by the interquartile range of regional welfare differences. In 2000, lifetime income would only have to be doubled.⁴⁵



Notes: The figure shows average increase in life-time income required to increase utility from the 25% quantile to the 75% quantile of the cross-sectional welfare distribution at time t . See Section B.8 in the Appendix for details.

Figure 7: The Evolution of Spatial Welfare Inequality

Spatial mobility is an important factor for this decline in spatial welfare inequality. Without labor mobility

⁴⁵The decline in inequality is not driven by changes in the distribution of regional amenities. The dispersion in amenities is roughly constant between 1910 and 2000 and spatial welfare differences had declined even in the absence of amenities.

to prevent the regional distribution of income from fanning out during the structural transformation, spatial welfare would have declined much less. To see this, Figure 7 also depicts the evolution of spatial inequality in the model without labor mobility. Two features are noteworthy. First of all, the decline in spatial inequality is much lower compared to the model with spatial mobility. Because amenities evolve exogenously, the only difference between these two models is solely the endogenous distribution in spatial earnings. Hence, spatial arbitrage through labor mobility goes a long way to compressing the spatial distribution of welfare.

Secondly, labor mobility is a particularly important adjustment mechanism for low-skilled workers. This discrepancy between workers of different skills is the consequence of the urban bias of the structural transformation. The secular shift away from agriculture hurts agricultural regions and, in particular, unskilled workers who have a comparative advantage in the agricultural sector. Moving out of rural locations as a response to this decrease in labor demand is an important mechanism to prevent regional wages, and in particular, agricultural wages, from falling. 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 Conclusion

The structural transformation, i.e. the systematic reallocation of employment out of the agricultural sector, is 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 regions and hurts rural locations. In this paper, we use a novel theory of spatial structural change and detailed regional data to analyze the spatial nature of the structural transformation of the US between 1880 and 2000.

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

To understand why this is the case and whether this mode of adjustment has important aggregate implications, we construct a new quantitative model of the structural transformation that explicitly incorporates a spatial layer. The model combines basic features from 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 very tractable and can be applied to a realistic geographical structure.

Our analysis yields three main results. First of all, we show that the insignificance of the spatial reallocation channel is mostly due to the evolution of regional productivities. Because local productivities are 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 spatial mobility was an important contributor of aggregate productivity growth in the US. Without labor mobility, aggregate income would have been 17% smaller. Finally, spatial welfare inequality declined substantially during the structural transformation and spatial mobility played an important role in that decline by keeping the regional income distribution compressed.

References

- Acemoglu, Daron and Veronica Guerrieri**, “Capital Deepening and Nonbalanced Economic Growth,” *Journal of Political Economy*, 2008, 116 (3), 467–498.
- Adao, Rodrigo, Costas Arkolakis, and Federico Esposito**, “Trade, Agglomeration Effects, and Labor Markets: Theory and Evidence,” 2018.
- Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf**, “The Economics of Density: Evidence From the Berlin Wall,” *Econometrica*, 2015, 83 (6), 2127–2189.
- Alder, Simon, Timo Boppart, and Andreas Muller**, “A Theory of Structural Change that Can Fit the Data,” 2018.
- Allen, Treb and Costas Arkolakis**, “Trade and the Topography of the Spatial Economy,” *The Quarterly Journal of Economics*, 2014, 129 (3), 1085–1140.
- Alvarez-Cuadrado, Francisco and Markus Poschke**, “Structural Change out of Agriculture: Labor Push versus Labor Pull,” *American Economic Journal: Macroeconomics*, 2011, 3 (3), 127–158.
- , **Ngo Van Long, and Markus Poschke**, “Capital Labor Substitution, Structural Change, and Growth,” *Theoretical Economics*, 2017, 12 (3), 1229–1266.
- Baumol, William J**, “Macroeconomics of Unbalanced Growth: the Anatomy of Urban Crisis,” *American Economic Review*, 1967, 57 (3), 415–426.
- Boppart, Timo**, “Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non-Gorman Preferences,” *Econometrica*, 2014, 82 (6), 2167–2196.
- Bryan, Gharad and Melanie Morten**, “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia,” 2017.

- , **Shyamal Chowdhury, and Ahmed Mushfiq Mobarak**, “Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh,” *Econometrica*, 2014, 82 (5), 1671–1748.
- Buera, Francisco J and Joseph P Kaboski**, “Can Traditional Theories of Structural Change Fit the Data?,” *Journal of the European Economic Association*, 2009, 7 (2-3), 469–477.
- Caselli, Francesco and Wilbur John Coleman II**, “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation,” *Journal of Political Economy*, 2001, 109 (3), 584–616.
- Chenery, Hollis B**, “Patterns of Industrial Growth,” *American Economic Review*, 1960, 50 (4), 624–654.
- Comin, Diego, Danial Lashkari, and Marti Mestieri**, “Structural Change with Long-Run Income and Price Effects,” 2017.
- Desmet, Klaus and Esteban Rossi-Hansberg**, “Spatial Development,” *American Economic Review*, 2014, 104 (4), 1211–43.
- Diamond, Rebecca**, “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000,” *American Economic Review*, March 2016, 106 (3), 479–524.
- Donaldson, Dave and Richard Hornbeck**, “Railroads and American Economic Growth: A Market Access Approach,” *The Quarterly Journal of Economics*, 2016, 131 (2), 799–858.
- Eckert, Fabian, Andres Gvirtz, and Michael Peters**, “Constructing a Population Panel for US Commuting Zones 1790-2000,” 2018.
- Fajgelbaum, Pablo and Stephen J. Redding**, “External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914,” Working Paper 20217, National Bureau of Economic Research June 2014.
- Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suarez Serrato, and Owen M. Zidar**, “State Taxes and Spatial Misallocation,” Working Paper 21760, National Bureau of Economic Research November 2015.
- Gollin, Douglas, David Lagakos, and Michael E. Waugh**, “The Agricultural Productivity Gap,” *The Quarterly Journal of Economics*, 2014, 129 (2), 939–993.
- , **Martina Kirchberger, and David Lagakos**, “In Search of a Spatial Equilibrium in the Developing World,” Working Paper 23916, National Bureau of Economic Research October 2017.
- , **Stephen Parente, and Richard Rogerson**, “The Role of Agriculture in Development,” *American Economic Review*, May 2002, 92 (2), 160–164.
- Herrendorf, Berthold and Todd Schoellman**, “Why is Measured Productivity so Low in Agriculture?,” *Review of Economic Dynamics*, 2015, 18 (4), 1003–1022.

- , **Christopher Herrington, and Akos Valentinyi**, “Sectoral Technology and Structural Transformation,” *American Economic Journal: Macroeconomics*, October 2015, 7 (4), 104–33.
- , **Richard Rogerson, and Akos Valentinyi**, “Two Perspectives on Preferences and Structural Transformation,” *American Economic Review*, December 2013, 103 (7), 2752–89.
- , —, and —, “Chapter 6 - Growth and Structural Transformation,” in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 2 of *Handbook of Economic Growth*, Elsevier, 2014, pp. 855 – 941.
- , —, and —, “Structural Change in Investment and Consumption: A Unified Approach,” 2017. Working Paper.
- Hobijn, Bart, Todd Schoellman, and Alberto J. Vindas Q.**, “Structural Transformation: A Race between Demographics and Technology,” 2018. Mimeo, Arizona State University.
- Hsieh, Chang-Tai and Enrico Moretti**, “Housing Constraints and Spatial Misallocation,” Working Paper 21154, National Bureau of Economic Research May 2015.
- Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie**, “Beyond Balanced Growth,” *The Review of Economic Studies*, 2001, 68 (4), 869–882.
- Kuznets, Simon**, “Quantitative Aspects of the Economic Growth of Nations: II. Industrial Distribution of National Product and Labor Force,” *Economic Development and Cultural Change*, 1957, pp. 1–111.
- Lagakos, David and Michael E. Waugh**, “Selection, Agriculture, and Cross-Country Productivity Differences,” *American Economic Review*, April 2013, 103 (2), 948–80.
- , **Mushfiq Mobarak, and Michael E Waugh**, “The Welfare Effects of Encouraging Rural-Urban Migration,” 2017.
- Lindbeck, Assar, ed.**, *Nobel Lectures, Economics 1969-1980*, World Scientific Publishing Co., Singapore, 1992.
- Michaels, Guy, Ferdinand Rauch, and Stephen J Redding**, “Urbanization and Structural Transformation,” *The Quarterly Journal of Economics*, 2012, 127 (2), 535–586.
- Molloy, Raven, Christopher L. Smith, and Abigail Wozniak**, “Internal Migration in the United States,” *Journal of Economic Perspectives*, September 2011, 25 (3), 173–96.
- 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.

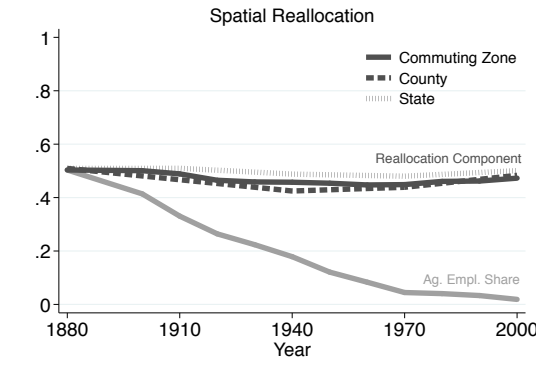
- Porzio, Tommaso and Gabriella V. Santangelo**, “Structural Change and the Supply of Agricultural Workers,” 2017.
- Redding, Stephen J. and Daniel M. Sturm**, “The Costs of Remoteness: Evidence from German Division and Reunification,” *American Economic Review*, December 2008, 98 (5), 1766–97.
- **and 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 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.

First we consider other definitions of regions other than that of a commuting zone used throughout the main body of the paper. Figure 8 replicates Figure 1 on the county and state 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 the three region definitions is chosen. 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 aggregation we consider) and smallest for state (the largest aggregation we consider). All in all Figure 8 reinforces our result of spatial reallocation as a highly local phenomenon, which operates predominantly at the intra-county level.



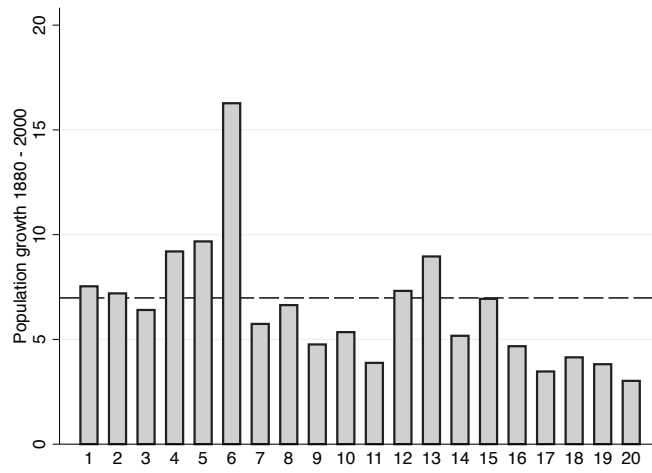
Notes: We show the aggregate agricultural employment share (light grey line) and the predicted agricultural share holding regional agricultural shares at their 1880 level, i.e. $\sum_r s_{rA1880} l_{rt}$, where s_{rA} and l_{rt} are the agricultural employment share and the population share of region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

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

Next we provide additional evidence for this (non-)relationship between past agricultural specialization and subsequent population growth. 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 5. 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



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

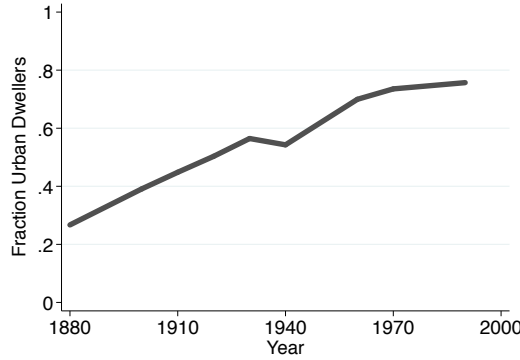
4 we include a set of twenty fixed effects of the initial agricultural share quantiles. While these fixed effects are jointly statistically significant, their explanatory power is still very small. Figure 9 shows this relationship graphically. More specifically, we report average population growth between 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.

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

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

Table 5: Agricultural Specialization in 1880 and Population growth 1880-2000

The absence of the spatial reallocation channel seems to be inconsistent with sharp increase in the rate of urbanization shown in Figure 10. We now show that this is not the case. In particular, we show that both the secular decrease in agricultural employment and the increase in urbanization are very local phenomena. In the left panel of Figure 11, we show the increase in urbanization since 1880. The share



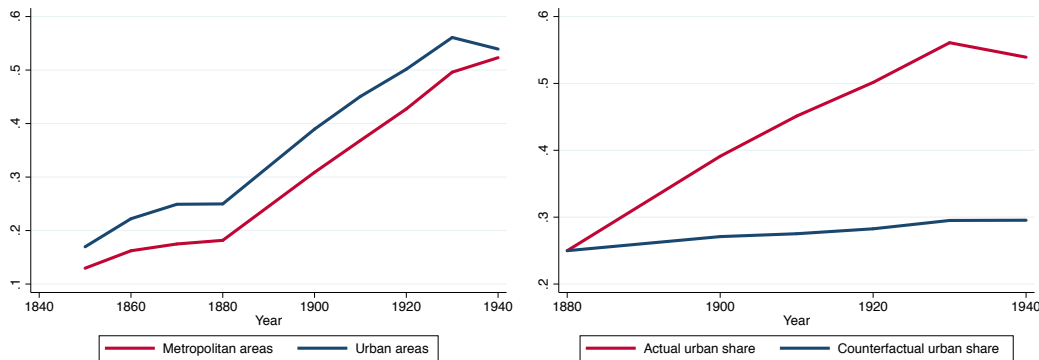
Notes: The graph is constructed from the decennial US micro census available on IPUMS. The line graphs the fraction of Americans who live in an urban environment rather than a rural one over time. An urban environment is defined as a locality with more than 2500 inhabitants.

Figure 10: Fraction of US population living in urban environment

of people living in urban areas (i.e. cities with more than 2500 inhabitants) increases from just shy of 20% in 1850 to more than 50% of the population in 1940. A different measure, the share of people living in metropolitan areas, shows a similar pattern. In the right panel we again decompose this time-series evolution into a within and across commuting zone component. In particular, we calculate the increase in the rate of urbanization due to spatial reallocation as

$$u_t^{CF} = \sum_r u_{r,1880} \frac{L_{rt}}{\sum_r L_{rt}},$$

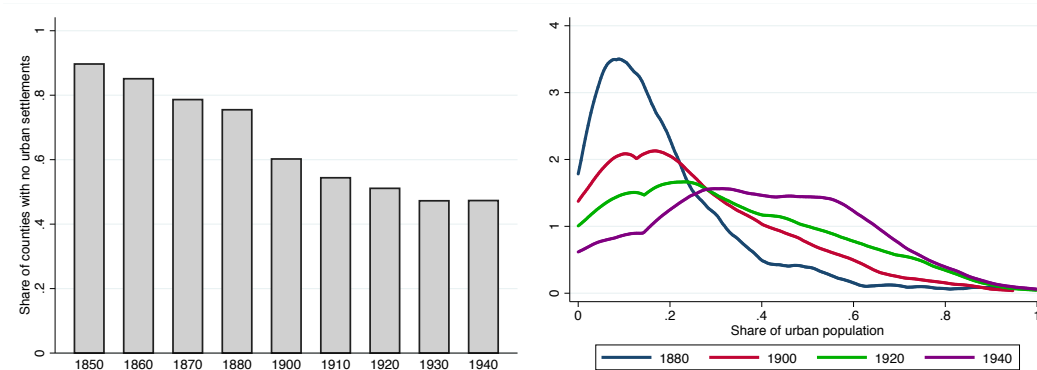
where $u_{r,1880}$ is the urbanization rate in commuting zone r in 1880. If the increase in urbanization stemmed from individuals migrating into high-urbanization commuting zone, this counterfactual urbanization rate would be close to the actual time series.



Notes: In the left panel we show the aggregate share of people living in urban areas and metropolitan areas between 1850 and 1940. In the right panel, we again show the share of people living in urban areas and counterfactual urban share calculated according to $u_t^{CF} = \sum_r u_{r,1880} \times \frac{L_{rt}}{\sum_r L_{rt}}$, where $u_{r,1880}$ is the urban share in 1880.

Figure 11: Urbanization within and across commuting zones

Figure 11 shows that this is not the case: as for the agricultural employment share, the cross-commuting zone population flows explain a minor share of actual increase observed in the data. Figure 12 provides an alternative way to see this within-commuting zone pattern of urbanization. In the left panel we display the “extensive” margin of urbanization, i.e. the share of counties, which have no urban centers. Expectedly, this number is declining. Similarly, the right panel 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.



Notes: In the left panel we depict the share of counties without any urban areas. In the right panel we show the cross-sectional distribution of the share of the urban population across counties.

Figure 12: Urbanization within commuting zones

Finally we ask whether the insignificance of the reallocation component of structural change is a more general feature of the structural transformation or particular to the United States. 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, consider Table 6 which uses 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 and finds it to be generally very small.

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	-0.076	-0.10	-0.09	-0.12	-0.29	-0.12	-0.02
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 6: The Reallocation Component of the Structural Transformation: International Evidence

The regions available for these countries in IPUMS are generally larger than commuting zones in the United States, but smaller than US states. Given Figure 8 for the United States, we take this quasi-absence

of the reallocation component at the sub-state resolution as indicative of the local nature of urbanization and reallocation being an important feature of the process of structural change more generally.

B Theory: Proofs and Derivations

B.1 Proof of Proposition 1

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

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

subject to

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

Substituting for e_{t+1} yields

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

The optimal allocation of spending is determined from the Euler equation

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

From above (26) we get that this equation is

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

This yields

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

so that

$$e_t = \frac{1}{1 + \left(\phi_{t,t+1}^\eta \frac{1}{\beta}\right)^{\frac{1}{\eta-1}}} y_{rt}^i$$

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

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, (4) implies that

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

This can be written as

$$U_{jr} = A_{jr} w_{rt}^\varepsilon \Phi_{t,t+1} + \Lambda_{t,t+1},$$

where

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

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

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

and

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

Note also that

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

This proves the results for Proposition 1.

B.2 Consumption expenditure shares

We can derive the expenditure shares from the indirect utility function $V_{jr}(e, p)$ in (5) from Roy's identity. The indirect utility function is defined by

$$V(p, e(p, u)) = u.$$

Hence,

$$\frac{\partial V(p, e(p, u))}{\partial p_j} + \frac{\partial V(p, e(p, u))}{\partial e} \frac{\partial e(p, u)}{\partial p_j} = 0.$$

The expenditure function is given by

$$e(p, u) = \min_x \sum p_j x \text{ s.t. } u(x) \geq u.$$

Hence,

$$\frac{\partial e(p, u)}{\partial p_j} = x_j(p, u),$$

where $x_j(p, u)$ is the Hicksian demand function. We also have that the Hicksian and the Marshallian demands are linked via

$$x^H(p, u) = x^M(p, e(p, u)).$$

Hence,

$$\frac{\partial V(p, e(p, u))}{\partial p_j} + \frac{\partial V(p, e(p, u))}{\partial e} x_j^M(p, e(p, u)) = 0$$

Rearranging terms yields

$$x_j^M(p, e(p, u)) = - \frac{\frac{\partial V(p, e(p, u))}{\partial p_j}}{\frac{\partial V(p, e(p, u))}{\partial e}}.$$

The expenditure share on good j is therefore given by

$$\vartheta_j(e, p) = x_j^M(p, e) \frac{p_j}{e} = - \frac{\frac{\partial V(p, e(p, u))}{\partial p_j} p_j}{\frac{\partial V(p, e(p, u))}{\partial e} e}. \quad (25)$$

We consider an indirect utility function from the PIGL class

$$V(e, P) = \frac{1}{\varepsilon} \left(\frac{e}{B(p)} \right)^\varepsilon + C(P) - \frac{1}{\varepsilon}.$$

We get that

$$\begin{aligned}\frac{\partial V_{rj}}{\partial e} &= e^{\varepsilon-1} B(p)^{-\varepsilon} \\ \frac{\partial V_{rj}}{\partial p_j} &= -e^{\varepsilon} B(p)^{-\varepsilon-1} \frac{\partial B(p)}{\partial p_j} + \frac{\partial C(p)}{\partial p_j}.\end{aligned}\tag{26}$$

(25) therefore implies that

$$\begin{aligned}\vartheta_j(e, p) &= \frac{e^{\varepsilon} B(p)^{-\varepsilon-1} \frac{\partial B(p)}{\partial p_j} p_j - \frac{\partial C(p)}{\partial p_j} p_j}{e^{\varepsilon-1} B(p)^{-\varepsilon} e} \\ &= \frac{\partial B(p)}{\partial p_j B(p)} p_j - \frac{\partial C(p)}{\partial p_j C(p)} p_j C(p) \left(\frac{e}{B(p)} \right)^{-\varepsilon} \\ &= \eta_j^B - \eta_j^C \left(\frac{e}{B(p)} \right)^{-\varepsilon} C(p),\end{aligned}$$

where η_j^B and η_j^C are the price elasticities of the B and C function respectively.

The specification we consider is

$$V(e, P) = \frac{1}{\eta} \left(\frac{e}{p_A^\phi p_M^{1-\phi}} \right)^\eta - \frac{\nu}{\gamma} \left(\frac{p_A}{p_M} \right)^\gamma - \frac{1}{\eta} + \frac{\nu}{\gamma}.$$

Hence, we have that

$$\begin{aligned}\frac{\partial V}{\partial e} &= \left(\frac{e}{p_A^\phi p_M^{1-\phi}} \right)^\eta \frac{1}{e} \\ \frac{\partial V}{\partial p_A} &= -\phi \left(\frac{e}{p_A^\phi p_M^{1-\phi}} \right)^\eta \frac{1}{p_A} - \nu \left(\frac{p_A}{p_M} \right)^\gamma \frac{1}{p_A}.\end{aligned}$$

Hence,

$$x_A(e, p) = \phi \frac{e}{p_A} + \nu \left(\frac{p_A}{p_M} \right)^\gamma \frac{1}{p_A} \left(p_A^\phi p_M^{1-\phi} \right)^\eta e^{1-\eta}.\tag{27}$$

The expenditure share is

$$\vartheta_A(e, p) = \frac{x_A(e, p) p_A}{e} = \phi + \nu \left(\frac{p_A}{p_M} \right)^\gamma \left(\frac{e}{p_A^\phi p_M^{1-\phi}} \right)^{-\eta}.$$

B.3 Earning, Expected Earnings and Aggregate Demand

Consider individual i in region r . Given her optimal occupational choice, the earnings of individual i are given by

$$y^i = \max_s \{w_{rs} z_s^i\}.$$

We assumed that individual productivities are Frechet Distributed, i.e.

$$F_s^h(z) = e^{-\Psi_s^h z^{-\zeta}}, \quad (28)$$

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 $y_{sr}^i \equiv w_{s,r} z_s^i$ is also Frechet and given by

$$F_{y_{rs}}(y) = P\left(z \leq \frac{y}{w_{rs}}\right) = e^{-\Psi_s^h w_{rs}^\zeta y^{-\zeta}}.$$

Using standard arguments about the max stability of the Frechet, the distribution of total earnings y^i is also Frechet and given by

$$F_{y_r}^h(y) = e^{-(\Theta_r^h)^\zeta y^{-\zeta}} = e^{-\left(\frac{y}{\Theta_r^h}\right)^{-\zeta}} \quad (29)$$

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

$$E[y_{r,h}^i] = \Gamma\left(1 - \frac{1}{\zeta}\right) \Theta_r^h.$$

From (29) we can derive the distribution of $y^{1-\eta}$. As $\eta < 1$, we have that

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

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

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

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

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

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

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

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

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

This yields (15).

B.4 Uniqueness and Existence of the Dynamic Equilibrium in the Economy without Space

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

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

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

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

Note that the LHS is increasing and continuous in k_{t+1} and satisfies $\lim_{k_{t+1} \rightarrow \infty} LHS = \infty$ and $\lim_{k_{t+1} \rightarrow 0} LHS = 0$. Hence, there is a unique mapping

$$k_{t+1} = m(k_t, Z_t, Z_{t+1}).$$

Furthermore, $m(\cdot)$ is increasing in all arguments.

B.5 The Equilibrium with Heterogeneous Skills

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

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

and the spatial labor supply condition (8)

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

where ρ_{jrt} , ϑ_t^A and π_{rst} are given in (13), (15) and (17) and also only depend on $\{w_{rt}, L_{rt}\}_r$. Given $\{w_{rt}, L_{rt}\}_r$, the future capital stock is then determined from (20) and the regional agricultural employment shares from (19).

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

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

where Θ_r^h , given in (23), denotes regional income for skill group h and λ_{rt} is the share of skilled individuals.⁴⁶ 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 (13) with $\mathcal{W}_{rt}^h = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_r^h)^{\eta} + A_{rt}$. Finally, the within-region allocation of factors across sectors, i.e. the counterpart to (19), is given by

$$\frac{\pi_{rAt}}{\pi_{rNA_t}} \frac{\vartheta_t^A}{1 - \vartheta_t^A} = \frac{w_{rA} (H_{rA}^L + H_{rA}^H)}{w_{rNA} (H_{rNA}^L + H_{rNA}^H)} = \left(\frac{w_{rAt}}{w_{rNA_t}} \right)^{\zeta} \frac{(1 - \lambda_r) \Psi_A^L (\Theta_r^h)^{1-\zeta} + \lambda_r \Psi_A^H (\Theta_r^H)^{1-\zeta}}{(1 - \lambda_r) \Psi_{NA}^L (\Theta_r^h)^{1-\zeta} + \lambda_r \Psi_{NA}^H (\Theta_r^H)^{1-\zeta}}. \quad (30)$$

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

$$s_{rst}^h = \Psi_s^h \left(w_{rs} / \Theta_r^h \right)^{\zeta}.$$

⁴⁶To 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.

The capital accumulation is still given by (20), i.e.

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

The main difference to the case with substitutable skills is the within-region sectoral supply equation (30). In the baseline model, s_{rAt} is determined residually from (19). With an upward sloping supply curve, the relative wages have to be consistent with sectoral labor supplies.

B.6 Balanced Growth Path Relationships

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

$$\psi = \left(1 + \beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}} \right)^{-1}. \quad (31)$$

To solve for the interest rate, note that (20) implies that

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

where g is the growth rate of the economy. Hence,

$$1 + g = (1 - \psi) \frac{(1 - \alpha)}{\alpha} (r + \delta) = \left(\frac{\beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}}}{1 + \beta^{\frac{1}{1-\eta}} (1+r)^{\frac{\eta}{1-\eta}}} \right) \frac{(1 - \alpha)}{\alpha} (r + \delta). \quad (32)$$

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 (32) yields

$$\frac{PI_t}{PY_t} = (g + \delta) \frac{\alpha}{R} \quad \text{and} \quad \frac{PC_t}{PY_t} = 1 - (g + \delta) \frac{\alpha}{R}.$$

From (32) we also get that

$$\psi = 1 - \frac{\alpha}{1 - \alpha} \frac{1 + g}{R}. \quad (33)$$

B.7 Regularity conditions for PIGL preferences

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

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

For these expenditure shares to be positive, we need that

$$\vartheta_A(e, p) \geq 0 \Rightarrow e^\eta \geq -\frac{v}{\phi} \left(\frac{p_A}{p_M} \right)^\gamma, \quad (34)$$

and

$$\vartheta_{NA}(e, p) \geq 0 \Rightarrow e^\eta \geq \frac{v}{1-\phi} \left(\frac{p_A}{p_M} \right)^\gamma.$$

Note first that (34) is trivially satisfied as long as $v > 0$. Also note that satisfying both of these restrictions automatically implies that $\vartheta_s(e, p) \leq 1$. In addition, as we show in Section A.2 in the Online Appendix, for the Slutsky matrix to be negative semi definite, we need that

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

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

$$e^\eta \geq \frac{v}{1-\phi} \left(\frac{p_A}{p_M} \right)^\gamma \quad (35)$$

$$(1-2\phi-\gamma)e^\eta + \frac{(1-\phi)\phi}{v} \left(\frac{p_A}{p_M} \right)^{-\gamma} e^{2\eta} \geq v(1-\eta) \left(\frac{p_A}{p_M} \right)^\gamma. \quad (36)$$

Lemma 5. *A sufficient condition for (35) to be satisfied is that (36) holds and that*

$$\gamma > (1-\phi)\eta. \quad (37)$$

Proof. Equation (36) can be written as

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

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

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

Now let $h(x) = (x-1) + (\phi x - (2\phi + \gamma))x$. For (36) 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 (36) to be satisfied, it has to be the case that $x > x^* = 1 - \frac{1 - \gamma}{2\phi}$ as $h(x^*) < h(0)$. Hence, condition (36) implies (35) if

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

Rearranging terms yields $(1 - \phi)\eta < \gamma$, which is (37). □

Hence, the preferences are well defined as long (36) is satisfied and (37) holds. Because the RHS of (36) is increasing in e in the relevant range, i.e. as long as (36) is satisfied, this implies that the preferences are well defined as long as e is high enough.

Now note that $e_{it} = \psi_{t+1} \times y_{it}$, where y_{it} denotes total earnings of individual i . From (29) we know that

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

where

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

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.

	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 (39).

Table 7: The Economic Magnitude of Moving Costs

B.8 Spatial Welfare Inequality

In Figure 7 in the main text we depict the evolution of spatial inequality. We construct this figure in the following way. From (52) we have that expected utility in region r is given by

$$\mathcal{W}_{rt}^h(\Theta_r^h) = \frac{\Gamma_{\eta/\zeta}}{\eta} \psi(r_{t+1})^{\eta-1} (\Theta_r^h)^\eta + \Lambda_{t,t+1} + A_{rt}. \quad (38)$$

Let $T_{rt}^h(\Delta)$ be the increase in average income Θ_r^h required to increase utility by Δ , i.e. $\mathcal{W}_{rt}^h(\Theta_r^h T_{rt}^h(\Delta)) = \mathcal{W}_{rt}^h(\Theta_r^h) + \Delta$. (38) implies that

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

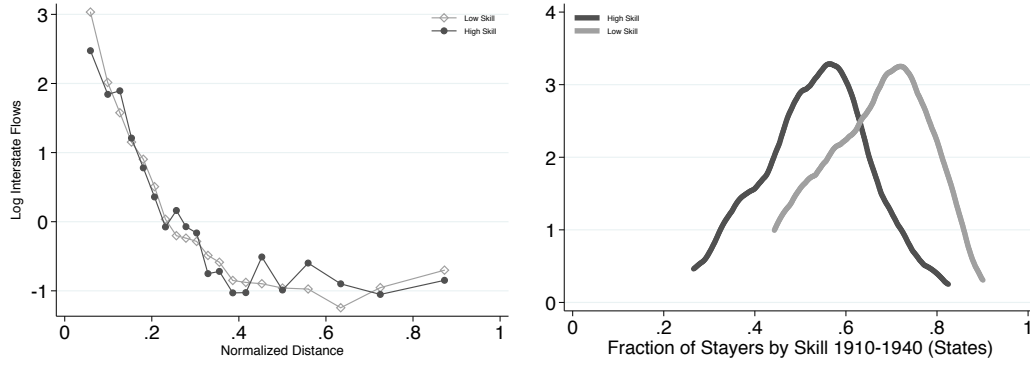
Let $\Delta_t^{h,IQR}$ be the interquartile range in regional welfare $\mathcal{W}_{rt}^h$ at time t , i.e. $\Delta_t^{h,IQR} = \mathcal{W}_t^{h,75} - \mathcal{W}_t^{h,25}$, where $\mathcal{W}_t^{h,x}$ is the x -quantile of the distribution of $\mathcal{W}_{rt}^h$. Figure 7 shows the time series evolution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

B.9 Additional Properties of the Calibrated Model

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

$$\Delta_{jt}^h \equiv \frac{\mathcal{U}_{jt}^{h,Mov} - \mathcal{U}_{jt}^{h,Stay}}{\tau}, \quad (39)$$

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 7 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 7 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:

$$\log \frac{\rho_{jr}}{1 - \rho_{jj}} = \alpha_j^h + \alpha_r^h + \hat{u}_{jr}^h$$

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 people.

Figure 13: Lifetime interstate Migration by Skill in 1940 in the Data

Migration by skill type In the model low and high skill workers are subject to the same distance costs, which yields an elasticity of moving flows to distance that is very similar for both groups. As the left panel of the below figure shows this is in line with the data. In the model high skill types are slightly more likely to move but not a lot. In the data low skilled people 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.

Regional Fundamentals In this section, we describe additional details for the estimated spatial productivities $\{Q_{rAt}, Q_{rNat}\}_{r,t}$ and amenities $\{A_{rt}\}_{r,t}$. This allows us to analyze the 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 8. We start by analyzing the cross-sectional relationship between income and agricultural specialization. In the first column, we report the relationship between the agricultural employment share and average manufacturing earnings. Both of these variables are directly observed in the data and strongly negatively correlated. Filtered through the lens of the structural model, this implies that average regional earnings for low skilled workers, Θ_{rt}^L , are also lower in rural areas. In column three we show that this gradient is even stronger for high-skilled individuals: holding unskilled income fixed, regions with a higher agricultural employment share offer particularly bad earnings opportunities for high skilled workers.

In the remaining columns we directly focus on the relationship between agricultural employment and the

	e_{rt}^{Man}	$\ln \Theta_{rt}^L$	$\ln \Theta_{rt}^H$	log Ag. empl. share ($\ln s_{rA}$)			
Ag. empl share ($\ln s_{rA}$)	-0.098*** (0.012)	-0.050*** (0.011)	-0.045*** (0.005)				
Low-skilled earnings ($\ln \Theta_{rt}^L$)			1.051*** (0.022)				
Non Ag. Productivity ($\ln Q_{rNA}$)				-0.331*** (0.020)			-0.344*** (0.007)
Ag. Productivity ($\ln Q_{rA}$)					0.404*** (0.022)		0.408*** (0.006)
Amenities (A_r)						-0.430*** (0.075)	-0.055*** (0.013)
Skill share in 1880 (λ_{r1880})		0.096** (0.038)	0.013 (0.009)	0.139 (0.123)	-1.594*** (0.150)	-0.466** (0.196)	-0.117*** (0.032)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ln pop 1880	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	3585	2580	2580	3225	3225	2580	2580
R^2	0.989	0.941	0.991	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.

Table 8: Fundamental Determinants of Agricultural Specialization

underlying regional fundamentals. In particular, we report the results from the specification

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

where δ_t is a year fixed effect. In columns four to six 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 9 we focus directly on the dynamics of spatial productivity and amenities. In particular, we consider a simple autoregressive specification

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

where y denotes either log sectoral productivity, $\ln Q_{rst}$, or the level of amenities A_{rt} and δ_r is a region fixed effect. 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, there seems to be more mean-reversion in the non-agricultural sector and the regional fixed effects explain less of the cross-sectional variation. The last two columns show the same result for regional amenities. In contrast to productivity, amenities show more persistence.

	$\ln Q_{NA,r,t}$		$\ln Q_{A,r,t}$		$\ln A_{r,t}$	
Lag	0.700*** (0.016)	0.204*** (0.017)	0.438*** (0.020)	0.109*** (0.023)	0.877*** (0.022)	-0.080** (0.038)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Commuting Zone FE	No	Yes	No	Yes	No	Yes
N	2580	2580	2580	2580	1935	1935
R^2	0.689	0.879	0.277	0.580	0.630	0.856

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

Table 9: The Process of Spatial Fundamentals

B.10 The Correlation of Agricultural Employment and Future Earnings

	Earnings		Amenities
	Low-skilled	High-skilled	
Share of Variance Explained	32.7%	48.6%	27%

Notes: The table reports the average R^2 of a regression of regional log earning (by skill groups) and regional amenities on 100 fixed effects for the respective percentiles of the agricultural employment shares. Hence, the regression we run is $y_{rt} = \sum_{q=1}^{100} \theta_{rt-1}^q + u_{rt}$, where θ_{rt-1}^q is an indicator variable for the q percentile of the distribution of agricultural employment shares at time $t - 1$. We run the regression for each period 1910, 1940, 1970 and 2000 and report the time average of the R^2 in the table. The individual R^2 for the four years are 24.9% 41.6% 35.2% and 29.0% for low skilled earnings, 48.2%, 61.5%, 48.0%, 36.8% for high-skilled earnings and 21.6% 34.2% 26.3% and 25.8% for amenities.

Table 10: Agricultural Employment and Future Local Attractiveness

B.11 Local Productivities and Market Access

As stated in the main part of the paper an important part of our calibration approach is that we calibrate the cross-sectional distribution of sectoral productivities $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$ and amenities $\{A_{rt}\}_{rt}$ as structural residuals of the model as is commonly done in the new quantitative spatial economics literature (see (Redding and Rossi-Hansberg, 2017) for a recent and excellent review of this literature).

The local sectoral productivities calibrated this way, $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$, are residuals necessary to fit the data on wages and sectoral employment conditional on the calibrated parameters of the model. While they reflect 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.

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

$$\ln Q_{rst} = \alpha_r + \beta \times MAC_{rt} + u_{rst}, \quad (40)$$

where α_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 40 separately for the agricultural and the non-agricultural sector. The results are contained in Table 11.

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 40 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 $\{Q_{rAt}, Q_{rNA_t}\}_{rt}$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Q_NA	Q_A	Q_NA	Q_A	Q_NA	Q_A	Q_NA_growth	Q_A_growth
log_cost	-1.897*** (0.224)	-2.077*** (0.125)	-5.205*** (0.290)	-4.291*** (0.195)	-3.937*** (0.353)	-3.760*** (0.309)		
cost_dh_growth							-4.005*** (0.388)	-2.395*** (0.328)
_cons	-5.508*** (0.539)	-2.470*** (0.290)	2.753*** (0.729)	3.059*** (0.489)	-0.412 (0.884)	1.734* (0.773)	0.330** (0.114)	0.0951 (0.0968)
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

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: We use data on the cost of reaching any other county from a given county provided by [Donaldson and Hornbeck \(2016\)](#), take the destination population weighted sum and aggregate them to the commuting zone level. This gives an index of market access for 1880 and 1910 for every commuting zone. Here we relate changes in agricultural productivity implied by the model to changes in market access costs. The commuting zones that saw there market access cost decrease the most are the ones that saw there agricultural productivity increase the most.

Table 11: Agricultural Productivity and Market Access Cost