

NBER WORKING PAPER SERIES

SPATIAL STRUCTURAL CHANGE

Fabian Eckert  
Michael Peters

Working Paper 30489  
<http://www.nber.org/papers/w30489>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2022, revised July 2025

We are grateful to Simon Alder, Costas Arkolakis, Timo Boppart, David Lagakos, Claudia Goldin, Doug Gollin, Joe Kaboski, Pete Klenow, David Nagy, Pascual Restrepo, Esteban Rossi-Hansberg, Gregor Schubert, Kjetil Storesletten, Aleh Tsyvinski, and Fabrizio Zilibotti for their suggestions. We particularly thank Tasso Adamopoulos for his discussion at the Toronto Trade-Development Workshop and Claudia Steinwender and Andy Foster for their discussions at the AEA and the NBER DEV meetings. We also thank seminar participants at ASU, Berkeley, Boston University, Stanford, LSE, Penn State, Harvard, the NY Fed, the IMF, UNC, the University of Munich, the SED, the EEA, the NBER Macroeconomic Across Time and Space Conference, the NBER Trade and Geography Conference, the NBER Development Meeting, the NBER Inequality and Macroeconomics Meeting, and STEG. Andrés Gvirtz provided outstanding research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Fabian Eckert and Michael Peters. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Spatial Structural Change  
Fabian Eckert and Michael Peters  
NBER Working Paper No. 30489  
September 2022, revised July 2025  
JEL No. O1, R11

## **ABSTRACT**

We develop a quantitative theory of spatial structural change in which economic growth, sectoral reallocation, and regional development are jointly determined in equilibrium. We show that sectoral reallocation inherently disadvantages regions specialized in the declining sector. Yet spatial forces—such as technology diffusion, trade, and migration—can offset or even reverse this bias. These forces are crucial to understand the first structural transformation of the US. Between 1880 and 1920, GDP per capita doubled and the agricultural employment share fell from 50% to 25%. Yet wages rose fastest in agriculture-intensive regions. Calibrated to this period, our model shows that productivity catch-up through technology diffusion enabled rural America to benefit from structural change; in its absence, rural regions would not have stagnated—they would have fallen further behind.

Fabian Eckert  
University of California, San Diego  
f@fpeckert.me

Michael Peters  
Yale University  
Department of Economics  
and NBER  
m.peters@yale.edu

A data appendix is available at <http://www.nber.org/data-appendix/w30489>

## INTRODUCTION

Since Kuznets' seminal work, economists have equated structural change with sectoral reallocation—the shift of labor out of agriculture and into industry and services as economies grow. Yet in his *Reflections on Modern Economic Growth* Kuznets also emphasized a spatial dimension: as part of the sectoral transition, urbanization reshapes where people live and work, creating regional winners and losers. Modern macroeconomic models reproduce the sectoral facts, but because they abstract from geography, they remain silent on those regional outcomes.

This omission is not without consequence since standard economic intuitions risk misidentifying who gains and who loses from structural change. The first US structural transformation is a case in point: from 1880 to 1920, national income doubled while the agricultural employment share halved from 50 to 25 percent. A standard Bartik-type intuition predicts wage declines and out-migration in farm-intensive regions, given their outsized exposure to the contracting sector. Instead, wages in farm-intensive regions rose faster than in the already industrialized cities, and 80 percent of the 25-point fall in the agricultural employment share occurred *within* regions rather than through rural-to-urban migration.

This paper develops a theory of *spatial structural change* in which aggregate growth, sectoral transformation, and spatial reallocation are jointly determined in equilibrium. We embed standard macroeconomic mechanisms of structural change into a dynamic spatial model of growth, building on recent advances in quantitative spatial economics. The economy comprises two sectors—agriculture and manufacturing—and many regions that differ in productivity, amenities, and land endowments, and are connected by trade, migration, and technology diffusion. We show that sectoral reallocation is inherently urban-biased: regions specialized in the shrinking farm sector are most exposed to its decline. Yet this bias can be offset—or even reversed—by spatial forces absent from standard macro models: technology diffusion, trade, and migration. Quantitatively, we find that diffusion accounts for most of the rural–urban wage convergence between 1880 and 1920; without it, rural regions would have fallen further behind, and US industrialization would have been slower.

A theory of spatial structural change requires two key extensions to the canonical static spatial framework. First, to capture the role of income effects, we introduce non-homothetic preferences into a general equilibrium model of trade and migration in a tractable way. Second, since structural change is driven by long-run growth, we explicitly model the evolution of sectoral productivity across regions. Building on the macro literature on technology diffusion, we allow regions to adopt technologies from others. In equilibrium, these endogenous choices yield a parsimonious law of

motion for regional productivity: lagging regions can catch up depending on their distance to the frontier and a single structural “catch-up” parameter that governs the ease of adopting local versus external technologies. When this parameter is positive, convergence occurs; when it is zero, productivity growth is spatially uniform.

The model permits an analytical decomposition of local wage growth and industrialization into contributions from sectoral productivity growth, local demand shifts, and population growth, which congests agricultural land. The decomposition shows that if sectoral productivity grows uniformly across space, trade is free, and population growth is spatially balanced, then structural change necessarily generates urban-biased wage growth—starkly at odds with the rural-biased growth observed in the data. Rural wage convergence requires either faster productivity growth in farm regions through catch-up growth, trade-cost declines that favor rural producers, or rising urban amenities that reduce compensating differentials and pull people out of the countryside.

We structurally estimate the model using a new panel dataset covering 495 commuting zones from 1880 to 1920. To construct it, we link the complete-count Population Censuses of 1880, 1900, and 1920 with the Agricultural and Manufacturing Censuses. The dataset includes commuting-zone-level wages, sectoral employment, land rents, and population, allowing us to trace America’s shift from farm to factory at a spatial scale consistent with modern regional analysis.

A key component of our quantitative analysis is the estimation of sector- and region-specific productivity dynamics. We develop an empirical strategy to separately identify the initial distribution of productivities across space and the strength of catch-up growth, while accounting for regional variation in market access and population growth. Following standard methods in quantitative spatial economics, we invert the model’s static equilibrium conditions using 1880 data to recover initial local productivities, amenities, and land endowments. This reveals that agricultural regions had low productivity in both sectors, positioning them to benefit most from technology diffusion.

We identify the strength of catch-up by targeting the empirical relationships between initial agricultural specialization and subsequent wage growth and sectoral reallocation using indirect inference. This allows us to account for changes in trade costs due to railroad expansion, exogenous variation in population growth from immigration and fertility differences, changing urban amenities, and all general equilibrium interactions in a model-consistent way. The remaining parameters are chosen to match three macro targets (a 22-point fall in the farm share, 2 percent annual GDP-per-capita growth, and a 9 percent drop in the farm-goods relative price) and additional spatial moments.

We use the calibrated model to understand why the first US structural transformation featured rural-biased wage growth. To do so, we decompose local wage growth and industrialization into contributions from sectoral demand shifts, regional productivity

growth, and population change. Technological catch-up in agriculture accounts for the bulk of rural wage convergence, as farm regions are most exposed to—and gain most from—agricultural productivity growth. However, because agricultural catch-up reinforces rural specialization in farming, it would slow industrialization if acting alone. Catch-up in manufacturing is thus essential to replicate the observed decline in rural agricultural employment. While local demand and population growth also play a role, each is quantitatively less important on its own.

Given the central role of technological catch-up, we conduct a counterfactual calibration that shuts down technology adoption from other regions. While this alternative still matches aggregate time-series trends, its spatial predictions diverge sharply: without catch-up, the Bartik-like exposure logic dominates, producing urban-biased wage growth, slower rural industrialization, and excessive worker relocation toward existing industrial centers. This highlights that catch-up growth was essential for rural America to weather the adverse effects of structural change. Without it, initially agricultural regions would not merely have stagnated—they would have fallen further behind.

Although our focus is historical, the lessons are contemporary. Structural change inherently produces uneven spatial outcomes by penalizing regions specialized in declining industries. Whether this bias translates into rising regional inequality depends on the strength of spatial linkages—migration, trade, and especially technology diffusion. In the 19th century, catch-up growth helped offset the urban bias of sectoral reallocation. Today’s shift toward services, by contrast, appears to be amplifying regional inequality—perhaps because technology diffuses more slowly in the service sector than it did in agriculture and manufacturing.

**Related Literature** Our paper contributes to and connects two strands of the literature: macroeconomic theories of structural change and spatial models of trade, migration and development.

First, we build on the macroeconomic literature on structural transformation, which explains the sectoral reallocation of labor through non-homothetic preferences, differential productivity growth, and capital deepening (see, for example, [Kongsamut, Rebelo, and Xie, 2001](#); [Ngai and Pissarides, 2007](#); [Comin, Lashkari, and Mestieri, 2021](#); [Boppart, 2014](#); [Herendorf, Rogerson, and Valentinyi, 2014](#)). This work focuses exclusively on the mechanisms driving changes in sectoral composition over time, but treats the spatial heterogeneity of local labor markets as irrelevant. We show that structural change at the aggregate level has distinct spatial consequences that depend crucially on inherently spatial mechanisms that defy a simple Bartik logic.

Second, our paper contributes to the growing body of work in spatial economics that examines how trade, migration, and productivity shape the distribution of economic

activity across space (see, for example, [Allen and Arkolakis, 2014](#); [Redding and Rossi-Hansberg, 2017](#)). Most existing spatial models are static and abstract from sectoral reallocation. Recent dynamic frameworks allow for evolving local productivity and migration, but do not consider the interaction between long-run structural change and spatial dynamics (see, for example, [Caliendo, Dvorkin, and Parro, 2019](#); [Kleinman, Liu, and Redding, 2023](#); [Desmet, Nagy, and Rossi-Hansberg, 2018](#)). Our contribution is to provide a dynamic quantitative framework that jointly determines aggregate structural change and the spatial evolution of population, wages, and sectoral specialization.

A few notable papers combine geography, sectors, and dynamics, but do not offer a general quantitative theory of spatial structural change. [Caselli and Coleman II \(2001\)](#) develop a stylized two-region model to study the North-South convergence in the US highlighting the role of education. Most closely related to our work, [Nagy \(2023\)](#) studies the formation of US cities before 1860, that is before the wide-spread industrial revolution that made the US the world's foremost industrial power. [Michaels, Rauch, and Redding \(2012\)](#) document the empirical relationship between population density and growth in 1880, without proposing a quantitative framework. [Bohr, Mestieri, and Robert-Nicoud \(2024\)](#) study the relationship between structural change and spatial concentration in a symmetric two-region New Economic Geography model. Recent work also explores aspects of the spatial shift toward services (see, for example [Desmet and Rossi-Hansberg, 2014](#); [Eckert, Ganapati, and Walsh, 2022](#); [Fan, Peters, and Zilibotti, 2023](#)).

**Roadmap** The rest of the paper is organized as follows. Section 1 documents the regional convergence patterns that motivate our analysis. Sections 2 and 3 present the model and its implications. Section 4 describes the calibration, and Section 5 quantifies the role of catch-up growth in rural convergence and structural change. Section 6 reports robustness checks, and Section 7 concludes. Additional details are provided in the Appendix.

## 1. SPATIAL STRUCTURAL CHANGE IN THE US: 1880-1920

The aggregate patterns of structural change in the US are well known: the US economy grew at around 2% a year over the past 150 years and its agricultural employment share declined from 50% in 1880 to about 1% today. In this paper, we focus on a period when this process was particularly fast: between 1880 and 1920 the economy industrialized at break-neck speed and the agricultural employment share dropped by almost 25% in just 40 years.

The *spatial* patterns of this process have received substantially less attention. In this section, we use data from the full-count Decennial Census files and the Census of Man-

ufacturing to measure income growth and sectoral reallocation across US commuting zones and document how the structural transformation played out across space.<sup>1</sup>

We begin by describing the initial conditions of our analysis: the patterns of agricultural specialization and relative wages across US commuting zones in 1880. At that time, disparities in sectoral composition were stark. One fifth of workers lived in highly agricultural commuting zones, where 80% of employment was in agriculture, while another fifth lived in already industrialized zones, where agriculture accounted for just 10% of employment. These differences in specialization were closely tied to income. Wages in the industrialized zones were nearly twice as high as in the agricultural ones, mirroring the large urban-rural wage gaps observed in many developing countries today ([Gollin, Kirchberger, and Lagakos, 2021](#)).

The aggregate decline in agricultural employment masks vast differences in the speed of industrialization across space. As shown in the left panel of Figure 1, some commuting zones saw almost no decline in their agricultural employment shares while others experienced declines of 40 percentage points or more. Industrialization was fastest in commuting zones with intermediate levels of initial agricultural specialization: a typical region with 60% agricultural employment in 1880 experienced a 22-point decline in the subsequent 40 years. In contrast, declines in agricultural employment shares were limited in locations that were already very industrialized and to some extent in the most rural areas, producing a mildly U-shaped relationship between initial agricultural intensity and subsequent industrialization.

A natural Bartik-like intuition suggests that more agricultural regions were hurt by the aggregate shift away from their sector of specialization. After all, a hundred years later in 1980, many previously prosperous manufacturing towns entered periods of decline as manufacturing employment commenced its precipitous downward trend. However, contrary to this intuition, between 1880 and 1920 rural commuting zones prospered. The right panel of Figure 1 shows that the initially most agricultural regions saw faster wage growth: each year, wages in the most agricultural regions grew 1.5 to 2 percentage points faster than those in urban areas and the standard deviation of log wages across commuting zones fell from 0.41 to 0.22 over the period.<sup>2</sup>

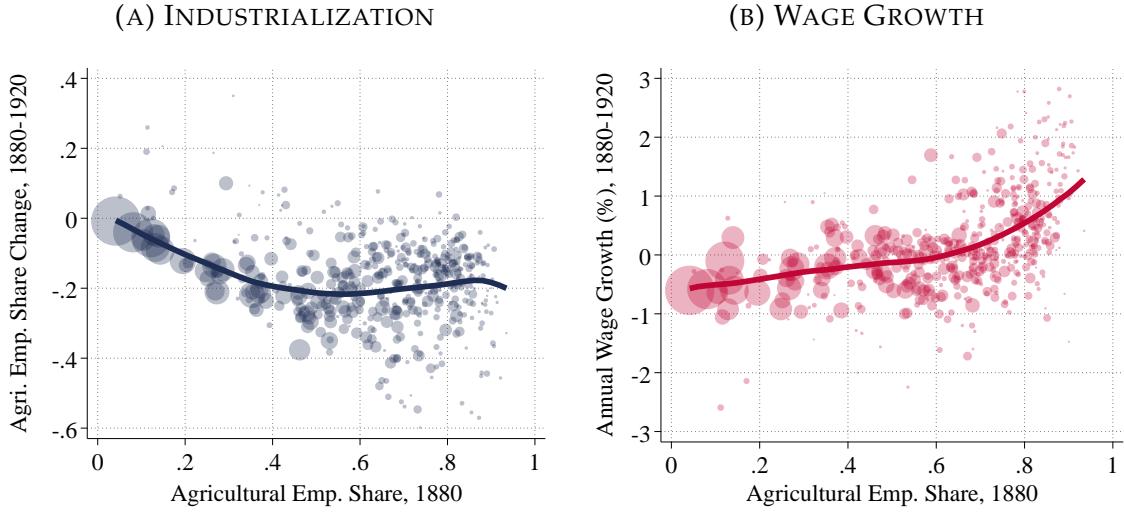
In the next section, we develop a theory of spatial structural change in which aggregate structural change and industrialization and wage growth across regions are equilibrium outcomes. By embedding the canonical drivers of structural transformation—non-homothetic preferences and sector-specific productivity growth—into a spatial economy linked by trade, migration, and technological catch-up, our framework can quantita-

---

<sup>1</sup>We re-aggregating all data to constant-boundary “commuting zones” using the crosswalk by [Eckert, Gvirtz, Liang, and Peters \(2020\)](#). We discuss the data in more detail in Section 5.

<sup>2</sup>Appendix B.2 shows the robustness of the patterns documented in Figure 1 to changes in the spatial unit of observation and the inclusion of various fixed effects.

FIGURE 1: SPATIAL STRUCTURAL CHANGE AND RURAL CONVERGENCE



*Notes:* The left panel plots commuting zones' agricultural employment shares in 1880 against the change in agricultural employment shares from 1880 to 1920. The right panel plots 1880 agricultural shares against relative wage growth over the same period. Wage growth is normalized so that the population-weighted mean is zero. Point sizes reflect total employment in each commuting zone. Each panel includes a nonparametric best-fit line, estimated using an Epanechnikov kernel-weighted polynomial smoother with a bin width of 0.1.

tively decompose the patterns in Figure 1 into the contributions of differential productivity growth, demand shifts, and population movements.

## 2. A THEORY OF SPATIAL STRUCTURAL CHANGE

We consider an economy that consists of a set of discrete locations, indexed by  $r = 1, \dots, R$ , and two sectors, agriculture and manufacturing, indexed by  $s = A, M$ .<sup>3</sup> At each time  $t$ , the economy is populated by a mass  $\bar{L}_t$  of workers who choose where to live subject to migration costs and whether to work as self-employed farmers or as wage earners in manufacturing. Each location  $r$  differs in its endowments of agricultural land, local amenities, and sectoral productivities, which evolve endogenously over time. Goods are traded across locations subject to iceberg trade costs,  $\tau_{rjt}$ , that can change over time. Preferences are non-homothetic, so that the expenditure share on agricultural goods falls as people become richer.

### 2.1 Technology

Each region can produce agricultural and manufacturing goods. A central feature of the model is that labor productivity is endogenous in both sectors since it depends on the technology adoption decisions of individual farmers and firms. These decisions, in turn, interact with the spatial structure of the economy, giving rise to region-specific

<sup>3</sup>We do not explicitly model a non-tradable services sector but allow for amenity congestion to mimic its cost-of-living effects; tradable services were an insignificant part of the 1880 US economy.

paths of productivity growth.

**Agriculture** Input and output markets in the agricultural sector are perfectly competitive. In each location, individuals can work as self-employed farmers by renting agricultural land on a frictionless spot market. The total supply of agricultural land in region  $r$  is fixed and given by  $\bar{T}_r$ . All farmers in location  $r$  produce the same variety, specific to region  $r$ , using land, innate ability, and time.

Each farmer  $i$  has an idiosyncratic farming ability  $z_A^i$  and one unit of time to allocate between production and adopting new technologies. A farmer with ability  $z_A^i$  who rents  $T$  units of land, devotes a fraction of time  $t_P$  to production and  $t_A = 1 - t_P$  to adopting new agricultural techniques, produces

$$(1) \quad y_{rt}^i = A \left( z_A^i t_P \right)^{1-\alpha} T^\alpha \quad \text{where} \quad A = K_{rt} t_A^\gamma.$$

Here  $\gamma$  governs the curvature of the adoption technology and  $K_{rt}$  is a region-specific shifter that determines the returns to technology adoption for farmers in region  $r$ .

A region  $r$  farmer with productivity  $z_A^i$  therefore is the residual claimant to the following profits:

$$(2) \quad \chi_{rt} \left( z_A^i \right) \equiv \max_{(t_P, T)} \left\{ p_{rAt} K_{rt} (1 - t_P)^\gamma \left( z_A^i t_P \right)^{1-\alpha} T^\alpha - R_{rt} T \right\},$$

where  $R_{rt}$  denotes the rental rate of agricultural land and  $p_{rAt}$  is the "farm gate price" of the local agricultural variety; the same variety costs  $p_{rAt} \tau_{rjt}$  in other locations  $j$ .

Equation (2) implies that a farmer's optimal allocation of time is constant at  $t_A = \gamma / (\gamma + 1 - \alpha)$ . As a consequence, all farmers in region  $r$  at time  $t$  choose the same technology  $A$ , which we denote by  $A_{rt}$ . By implication, all heterogeneity in farm output within a location arises from differences in farmer ability  $z_A^i$  alone: farmer profits are proportional to ability  $z_A^i$  and given by  $\chi_{rt} z_A^i$ , where  $\chi_{rt}$  denotes the return per unit of farming ability.

The term  $K_{rt}$  reflects the potential for technology adoption from farmers in other locations and determines the returns to adopting new techniques in location  $r$ . Following the literature on technology diffusion (see, for example, [Acemoglu, Aghion, and Zilibotti, 2006](#); [Desmet et al., 2018](#)), we assume

$$(3) \quad K_{rt} = \mu_A A_{rt-1}^{1-\lambda_A} \mathcal{F}_{rAt}^{\lambda_A}, \quad \text{where} \quad \mathcal{F}_{rAt} = \mathcal{F}_A \left( \{f(d_{rj}) A_{jt-1}\}_j \right),$$

where  $\mu_A$  is an aggregate shifter,  $A_{rt-1}$  is region  $r$ 's lagged agricultural productivity,  $f(d_{rj})$  is a function of distance between locations  $r$  and  $j$ , which parameterizes geographic frictions to spatial technology adoption, and  $\mathcal{F}_A$  is a function that aggre-

gates the distribution of "distance discounted" lagged productivity into a scalar. The parameter  $\lambda_A$  governs the ease of adopting technologies from other locations.

**Manufacturing** Input markets in the manufacturing sector are perfectly competitive, while output markets feature monopolistic competition. Manufacturing entrepreneurs in any region  $r$  can start a manufacturing firm to produce a quantity  $q$  of a differentiated variety  $\omega$  using  $h_{rM}(\omega) = f_E + q(\omega)/Z$  units of manufacturing labor. The technology level  $Z$  is a choice variable of the firm. To produce with productivity  $Z$ , it must hire  $(Z/\psi_{rt})^\theta$  units of labor, where  $\psi_{rt}$  indexes the cost of adopting new technologies in region  $r$ . Because all manufacturing firms in region  $r$  are ex-ante homogeneous, we denote the optimal technology choice in location  $r$  at time  $t$  by  $Z_{rt}$ .

Mirroring the agricultural sector, we model the inverse technology adoption costs as

$$(4) \quad \psi_{rt} = \mu_M Z_{rt-1}^{1-\lambda_M} \mathcal{F}_{rMt}^{\lambda_M} \quad \text{where} \quad \mathcal{F}_{rMt} = \mathcal{F}_M \left( \{f(d_{rj})Z_{jt-1}\}_j \right).$$

The function  $\mathcal{F}_{rMt}$  aggregates the past productivity of neighboring regions  $j$ , weighted by distance via  $f(d_{rj})$ . As in agriculture, the parameter  $\lambda_M$  governs the ease of adopting technologies from other regions. When  $\lambda_M = 0$ , adoption costs depend solely on a region's own productivity. For  $\lambda_M > 0$ , regions with relatively low productivity benefit from proximity to more productive locations, reflecting the potential for productivity catch-up through adoption of more advanced technologies.

Firms operate for a single period and free entry ensures that firms enter until profits net of adoption costs equal the entry cost:

$$(5) \quad \max_Z \left\{ \pi_{rt}(Z) - (Z/\psi_{rt})^\theta w_{rMt} \right\} = w_{rMt} f_E,$$

where  $\pi_{rt}(Z)$  denotes variable profits and  $w_{rMt}$  is the manufacturing wage. We denote the total mass of entering firms—and hence manufacturing varieties—in region  $r$  by  $N_r$  and total manufacturing employment in region  $r$  by  $H_{rMt}$ , which is the sum of labor used for entry, adoption, and production ( $H_{rEt}$ ,  $H_{rZt}$ , and  $H_{rPt}$ ).

## 2.2 Preferences

Consumers derive utility from agricultural and manufacturing consumption bundles. These bundles are CES aggregates of the available varieties in each sector:

$$(6) \quad c_{rA}^i = \left( \sum_{j=1}^R (c_{rjA}^i)^{\frac{\sigma_A-1}{\sigma_A}} \right)^{\frac{\sigma_A}{\sigma_A-1}} \quad \text{and} \quad c_{rM}^i = \left( \sum_{j=1}^R \int_0^{N_j} c_{rj}^i(\omega)^{\frac{\sigma_M-1}{\sigma_M}} d\omega \right)^{\frac{\sigma_M}{\sigma_M-1}},$$

where  $\sigma_s$  denotes the elasticity of substitution across varieties in sector  $s$ . Because goods are subject to iceberg trade costs, the corresponding ideal sectoral price indices  $P_{rMt}$

and  $P_{rAt}$  are location-specific.

To capture income-driven shifts in sectoral expenditure—an important feature of the structural transformation—we assume individuals have non-homothetic preferences over their sectoral consumption bundles,  $c_{rA}^i$  and  $c_{rM}^i$ . Specifically, we follow [Boppart \(2014\)](#) and adopt the Price-Independent Generalized Linear (PIGL) class of indirect utility functions which have useful aggregation properties illustrated below. As a result, the indirect utility function of worker  $i$  with total expenditure  $e^i$  facing local prices  $(P_{rAt}, P_{rMt})$  is:

$$(7) \quad V\left(e^i, P_{rAt}, P_{rMt}\right) = \frac{1}{\eta} \left( \frac{e^i}{P_{rAt}^\phi P_{rMt}^{1-\phi}} \right)^\eta - \nu \ln \left( \frac{P_{rAt}}{P_{rMt}} \right),$$

where  $\eta, \phi \in (0, 1)$  and the term  $e^i / (P_{rAt}^\phi P_{rMt}^{1-\phi})$  is a measure of real income. Applying Roy's identity yields the agricultural expenditure share of consumer  $i$ :

$$(8) \quad \vartheta_A\left(e^i, P_{rAt}, P_{rMt}\right) = \phi + \nu \left( \frac{e^i}{P_{rAt}^\phi P_{rMt}^{1-\phi}} \right)^{-\eta}.$$

As long as  $\nu > 0$ , the agricultural spending share declines with real income and asymptotes to  $\phi$  from above. The "Engel elasticity"  $\eta$  governs the strength of income effects on sectoral demand.<sup>4</sup> For  $\nu = 0$ , preferences collapse to Cobb-Douglas with constant expenditure shares.

## 2.3 Labor Supply to Sectors and Regions

Workers choose in which location to live and in which sector to work. We begin by describing the sectoral choice conditional on having moved to a particular location.

**Sectoral Labor Supply** We model a worker's sectoral choice using a Roy-type framework: each worker  $i$  draws sector-specific ability  $(z_A^i, z_M^i)$  from independent Fréchet distributions with shape parameter  $\zeta$  and scale 1. This setup captures imperfect substitutability of labor across sectors and endogenously generates selection into sectoral employment.

Workers choose their sector of employment to maximize earnings. In manufacturing, a worker with ability  $z_M^i$  earns  $z_M^i w_{rMt}$ . In agriculture, we assume that payments to land are distributed to farmers in proportion to their output within each region. As a result, a farmer with ability  $z_A^i$  earns a share of agricultural revenue proportional to

---

<sup>4</sup>The elasticity of substitution between the value added generated in the two sectors is given by  $EOS = 1 + \eta(\vartheta_A - \phi)^2 / (\vartheta_A(1 - \vartheta_A))$ . Hence, the elasticity is not a structural parameter but varies across space and the income distribution. The elasticity increases in  $\vartheta_A$  (that is, decreases in real income) and satisfies  $\lim_{\vartheta_A \rightarrow \phi} EOS = 1$ .

their contribution to total production. The total earnings of such a farmer are given by:<sup>5</sup>

$$\bar{\chi}_{rt} (z_A^i) = \frac{1}{1-\alpha} \chi_{rt} z_A^i \equiv \bar{\chi}_{rt} z_A^i,$$

where  $\bar{\chi}_{rt}$  captures the total return per efficiency unit of labor in agriculture. As a result, the income (and spending) of worker  $i$  in region  $r$  at time  $t$  is given by  $e_{rt}^i = \max \{z_M^i w_{rMt}, z_A^i \bar{\chi}_{rt}\}$ .

Conveniently, the distribution of income inherits the shape of the distribution of individual ability so that:

$$(9) \quad F_{rt}(e) = \exp \left\{ - (e/\bar{w}_{rt})^{-\zeta} \right\} \quad \text{where} \quad \bar{w}_{rt} = \left( \bar{\chi}_{rt}^{\zeta} + w_{rMt}^{\zeta} \right)^{1/\zeta}.$$

The term  $\bar{w}_r$  is proportional to average income in region  $r$ , with a constant of proportionality of  $\Gamma_\zeta \equiv \Gamma(1 - 1/\zeta) > 0$  where  $\Gamma(\cdot)$  denotes the Gamma function. Since  $\bar{w}_r$  captures all regional and intertemporal variation in average income, we refer to it simply as average income throughout the remainder of the paper.

The Fréchet assumption also leads to analytical expressions for the agricultural employment share,  $s_{rAt}$ , and the aggregate supply of farmers' efficiency units,  $H_{rAt}$ :

$$(10) \quad s_{rAt} = (\bar{\chi}_{rt}/\bar{w}_{rt})^\zeta \quad \text{and} \quad H_{rAt} = \Gamma_\zeta L_{rt} s_{rAt}^{\frac{\zeta-1}{\zeta}},$$

where analogous equations hold for manufacturing.

Equation (10) shows that  $\zeta$  governs the sectoral labor supply elasticity: higher  $\zeta$  implies more elastic labor supply. As in [Lagakos and Waugh \(2013\)](#), this reflects selection: the more workers choose a sector, the lower their average ability. In the limit, as  $\zeta \rightarrow \infty$  so that idiosyncratic differences in ability vanish, labor becomes fully elastic across sectors as in most macro models of structural change.

**Spatial Mobility** At the beginning of each period, worker  $i$  from location  $j$  chooses their *expected* utility-maximizing location without knowing the realization of their sectoral ability:

$$r^i = \arg \max \{ \mathcal{V}_{rt} \mu_{jr} \mathcal{B}_{rt} u_{rt}^i \},$$

where  $\mathcal{V}_{rt}$  denotes expected consumption utility which depends on wages and prices in the destination  $r$ , the term  $\mu_{jr} \in (0,1]$  captures migration frictions,  $\mathcal{B}_{rt}$  denotes

---

<sup>5</sup>With non-homothetic preferences, the income distribution is consequential for aggregate sectoral spending. Distributing land payments to local workers instead of immobile local land owners (see, for example, [Redding and Rossi-Hansberg, 2017](#)) implies that we must keep track of only one income distribution.

destination amenities, and  $u_{rt}^i$  is a worker- and destination-specific preference shifter. We present an expression for  $\mathcal{V}_{rt}$  in the next section.

We model amenities as  $\mathcal{B}_{rt} = B_{rt} L_{rt}^{-\rho}$ , where  $B_{rt}$  is an exogenous amenity shifter and  $\rho > 0$  governs the strength of congestion forces, such as the increasing cost of non-tradables like housing.

For aggregation purposes, we assume that idiosyncratic preferences over destinations,  $u_{rt}^i$ , are drawn from an i.i.d. Fréchet distribution with shape parameter  $\varepsilon$  and scale parameter 1. This assumption yields an analytic expression for the fraction of workers born in origin  $j$  that choose location  $r$  and allows us to express the aggregate law of motion for local employment as follows:

$$(11) \quad L_{rt} = \sum_j m_{jrt} n_{jt-1} L_{jt-1} \quad \text{where} \quad m_{jrt} = \frac{(\mu_{jr} \mathcal{V}_{rt} \mathcal{B}_{rt})^\varepsilon}{\sum_k (\mu_{jk} \mathcal{V}_{kt} \mathcal{B}_{kt})^\varepsilon}.$$

Here,  $m_{jrt}$  denotes the probability of moving from  $j$  to  $r$  and  $n_{jt-1}$  is the exogenous rate of employment growth in location  $j$  between periods  $t-1$  and  $t$ . We follow [Cruz and Rossi-Hansberg \(2024\)](#) in introducing  $n_{jt-1}$  to account for sources of regional employment growth that are not modeled explicitly, such as international migration or differences in local fertility rates. A region's employment at time  $t$  is thus determined by its employment in the previous period, the local rate of employment growth, and endogenous migration decisions of workers.

## 2.4 Aggregation and Equilibrium

This section derives closed-form expressions for the aggregate demand system, characterizes the laws of motion for local productivity, and formally defines the equilibrium of the economy.

**Aggregate Demand and Spatial Welfare** To compute the equilibrium, we need to characterize both the expected consumption utility of consumers  $\mathcal{V}_{rt}$  and the aggregate demand system. The combination of PIGL preferences and the Fréchet distribution of sectoral ability allows us to derive closed-form expressions for both these objects. In particular, we show that both aggregate spending patterns and expected utility admit representative-agent formulations that depend only on local average income and sectoral price indices:

**Proposition 1.** *Let  $\vartheta_{rAt}$  denote the aggregate expenditure share on agricultural goods in region*

$r$  and  $\mathcal{V}_{rt}$  the expected consumption utility of workers in region  $r$ . Then

$$(12) \quad \vartheta_{rAt} = \frac{\int \vartheta_A(e, P_{rAt}, P_{rMt}) e dF_{rt}(e)}{\int e dF_{rt}(e)} = \phi + \nu^{RC} \left( \bar{w}_{rt} / (P_{rAt}^\phi P_{rMt}^{1-\phi}) \right)^{-\eta},$$

$$(13) \quad \mathcal{V}_{rt} = \int V(e, P_{rAt}, P_{rMt}) dF_{rt}(e) = \frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \left( \bar{w}_{rt} / (P_{rAt}^\phi P_{rMt}^{1-\phi}) \right)^\eta - \nu \ln(P_{rAt} / P_{rMt}),$$

where  $\nu^{RC} = \nu \frac{\Gamma_{\zeta/(1-\eta)}}{\Gamma_\zeta} > 0$ .

*Proof.* See Section A.2.2 in the Appendix.  $\square$

Proposition 1 contains closed-form expressions of aggregate demand and expected consumption utility in terms of average earnings and local price indices. In particular, it allows us to treat demand *as if* arising from a representative household with non-homothetic PIGL preferences and a composite preference parameter  $\nu^{RC}$  that depends on  $\nu$  and the underlying income distribution parameterized by  $\zeta$ .

An important implication of Proposition 1 is that it leads to an explicit characterization of aggregate demand and sectoral revenue. Specifically, we can write aggregate revenue in sector  $s$  and region  $r$  as

$$(14) \quad \mathcal{R}_{rMt} = N_{rt} p_{rMt}^{1-\sigma_M} \mathcal{D}_{rMt} \quad \text{and} \quad \mathcal{R}_{rAt} = p_{rAt}^{1-\sigma_A} \mathcal{D}_{rAt},$$

where the terms  $\mathcal{D}_{rst}$  capture the *market access* of a location (see [Donaldson and Hornbeck, 2016](#)):

$$(15) \quad \mathcal{D}_{rst} = \sum_j \tau_{rjt}^{1-\sigma_s} P_{jst}^{\sigma_s-1} \vartheta_{jst} \Gamma_\zeta L_{jt} \bar{w}_{jt} \quad \text{for } s = A, M,$$

and  $\vartheta_{jst} \Gamma_\zeta L_{jt} \bar{w}_{jt}$  denotes aggregate spending on sector  $s$  goods in region  $j$ . The market access terms summarize the sector-specific demand for producers in region  $r$ , incorporating trade costs, price competition, and sectoral spending. Crucially, economic growth raises manufacturing market access relative to agricultural market access because the agricultural expenditure share is falling as real income rises.

**Adopting Technologies from Other Regions and Productivity Catch-Up** A central feature of our theory is the possibility to benefit from the productivity of neighboring regions. This mechanism is embedded in our specification of adoption costs, which depend on both local productivity and a distance-weighted average of productivities in other locations. Despite the fact that technology adoption is endogenous in our theory, the resulting productivity processes admit simple, closed-form expressions in *equilibrium*.

**Proposition 2.** Let  $A_{rt}$  and  $Z_{rt}$  denote the optimal technology choices of farmers and manufacturing firms in location  $r$  at time  $t$ . Sectoral productivity growth in region  $r$  is then given by

$$(16) \quad g_{Art} \equiv \ln(A_{rt}/A_{rt-1}) = g_A + \lambda_A \ln(\mathcal{F}_{rAt}/A_{rt-1})$$

$$(17) \quad g_{Zrt} \equiv \ln(Z_{rt}/Z_{rt-1}) = g_M + \lambda_M \ln(\mathcal{F}_{rZt}/Z_{rt-1}),$$

where  $g_A \equiv \ln\left(\left(\frac{\gamma}{\gamma+1-\alpha}\right)^\gamma \mu_A\right)$  and  $g_M \equiv \ln\left(\left(\frac{\sigma_M-1}{\theta-(\sigma_M-1)} f_E\right)^{1/\theta} \mu_M\right)$ .

*Proof.* See Section A.2 in the Appendix.  $\square$

Proposition 2 contains a key result of the paper: productivity growth in each sector consists of a nationwide term  $g_s$  shared by all regions, and a location-specific adjustment that depends on  $\lambda_s$ , that is, the extent to which adoption from other locations is possible, and location  $r$ 's productivity *relative* to the sectoral productivity distribution as summarized by  $\mathcal{F}_{rAt}$  and  $\mathcal{F}_{rZt}$ . This second term introduces a “benefit of backwardness”: when  $\lambda_s > 0$ , less productive regions grow faster, leading to convergence in spatial productivity levels. In the absence of technology adoption from other locations ( $\lambda_s = 0$ ), growth is uniform across space and the productivity distribution remains stationary. Spatial technology adoption is therefore intrinsically connected to regional productivity catch-up growth. We thus refer to  $\{\lambda_s\}_s$  as the *strength of productivity catch-up growth* or simply as productivity catch-up parameters.

The result in Proposition 2 for the agricultural sector follows directly from farmers' constant time allocation across adoption and production activities, which implies that  $A_{rt} \propto K_{rt}$ . Combined with the specification of the adoption technology in equation (3), equation (16) follows directly.

The law of motion for manufacturing technology emerges only in general equilibrium and exploits the free entry condition. To derive equation (17), consider first firms' optimal technology choice in partial equilibrium. Firms choose their technology to maximize variable profits, which, using the expression for sectoral revenue in equation (14), are given by

$$\pi_{rt}(Z) = \frac{1}{\sigma_M} \left( \frac{\sigma_M}{\sigma_M - 1} \frac{w_{rMt}}{Z_{rt}} \right)^{1-\sigma_M} \mathcal{D}_{rMt}.$$

As a result, equation (5) implies that the optimal technology choice is given by

$$(18) \quad Z_{rt} = \left( \frac{1}{\theta} \left( \frac{\sigma_M - 1}{\sigma_M} \right)^{\sigma_M} \frac{\psi_{rt}^\theta \mathcal{D}_{rMt}}{w_{rMt}^{\sigma_M}} \right)^{\frac{1}{\theta - (\sigma_M - 1)}},$$

that is, firms' desired productivity is increasing in market access and decreasing in local

manufacturing wages and adoption costs.<sup>6</sup> However, imposing the free entry condition implies that the optimal technology level depends *only* on the cost of adoption. This, in turn, yields Proposition 2. The intuition for this result is reminiscent of Young (1998): when demand increases in a location, profitability and the adoption incentives rise. However, local wages increase too, because free entry requires that profits net of adoption costs are pinned down by the entry cost. In Appendix Section A.2, we show formally that, in equilibrium,  $Z_{rt}$  is proportional to  $\psi_{rt}$ , so that equation (17) follows directly from the functional form of  $\psi_{rt}$  in equation (4).

Importantly, equations (16) and (17) do not mechanically imply a link between sectoral specialization and future productivity growth. A region's productivity growth depends on its *level* of physical productivity, while its sectoral specialization reflects *relative* physical productivity as well as other sources of comparative advantages, such as agricultural land supply, employment density, and trade linkages.

**Equilibrium** Given initial conditions  $\{L_{r0}, A_{r0}, Z_{r0}\}_r$  for the distribution of workers and sectoral productivity across space, an equilibrium consists of a sequence of prices, allocations, and technology levels that satisfy individual optimization, firm behavior, market clearing, and dynamic consistency. Formally:

**Definition.** Given  $\{L_{r0}, A_{r0}, Z_{r0}\}_r$ , an equilibrium is a sequence of prices  $\{P_{rAt}, P_{rMt}\}_{rt}$ , wages  $\{w_{rMt}\}_{rt}$ , rental rates for agricultural land  $\{R_{rt}\}_{rt}$ , manufacturing varieties  $\{N_{rt}\}_{rt}$ , employment allocations  $\{H_{rAt}, H_{rEt}, H_{rPt}, H_{rZt}\}_{rt}$ , local employment  $\{L_{rt}\}_{rt}$ , individual consumption  $\{c_{rjAt}^i, [c_{rjMt}^i(\omega)]_{\omega}\}_{rjt}^i$ , and productivity processes  $\{A_{rt}, Z_{rt}\}_{rt}$ , such that (i) consumers' consumption and location choices maximize utility, (ii) the creation of local varieties is consistent with free entry, (iii) firms maximize profits, (iv) all markets clear, and (v) productivity evolves according to the laws of motion in Proposition 2.

Given the initial distribution of productivity  $\{A_{r0}, Z_{r0}\}_r$ , the evolution of productivity follows equations (16) and (17). The market-clearing conditions for both sectors, together with the spatial law of motion for employment (equation 11) and the aggregate labor supply across sectors (equation 10), jointly determine the endogenous equilibrium objects: skill prices in each sector ( $w_{rMt}, \bar{\chi}_{rt}$ ), agricultural employment shares  $s_{rAt}$ , and the mass of manufacturing varieties  $N_{rt}$ . Appendix A.2 presents the full equilibrium system.

### 3. THE ECONOMICS OF SPATIAL STRUCTURAL CHANGE

In this section, we derive the implications of the model for the patterns of spatial structural change documented in Figure 1. Specifically, we characterize how local wage

---

<sup>6</sup>We require  $\theta > \sigma_M - 1$  to make profits concave in  $Z$ .

growth and industrialization respond to local productivity growth, local population growth, and changes in sectoral demand. The central result of this section is the Bartik-style logic mentioned above: aggregate structural change *per se* systematically disadvantages regions specialized in the agricultural sector. Explaining why rural location flourished therefore requires additional forces that counteract this inherent urban bias of structural change.

To build toward our main proposition, we show in Appendix A.3 that sectoral revenue productivity can be written as:

$$\begin{aligned}\mathcal{R}_{rMt}/H_{rMt} &= w_{rMt} = \bar{f}_E Z_{rt}^{\frac{\sigma_M-1}{\sigma_M}} \mathcal{D}_{rMt}^{\frac{1}{\sigma_M}}, \\ \mathcal{R}_{rAt}/H_{rAt} &= \bar{\chi}_{rt} = \bar{\alpha} A_{rt}^{\frac{\sigma_A-1}{\sigma_A}} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A}} \bar{T}_r^{\alpha \frac{\sigma_A-1}{\sigma_A}} H_{rAt}^{-\left(1-(1-\alpha)\frac{\sigma_A-1}{\sigma_A}\right)},\end{aligned}$$

where  $\bar{f}_E$  and  $\bar{\alpha}$  are inconsequential constants. These expressions show that sectoral revenue productivity, depends on a Cobb–Douglas composite of physical productivity and market access. In addition, the expressions highlight that revenue productivity is constant in manufacturing but declining in agriculture. In manufacturing this is the direct result of free entry which implies that additional labor translates into new varieties instead of more output per existing variety. The decreasing returns in agriculture reflect diminishing returns to land ( $\alpha$ ) and deteriorating terms of trade ( $\sigma_A$ ).

Motivated by these expressions, we define two *revenue productivity shifters* that summarize how physical productivity, market access, and the size of local population impact local revenue in each sector:

$$(19) \quad \mathcal{Z}_{rt} = \tilde{f}_E Z_{rt}^{\frac{\sigma_M-1}{\sigma_M}} \mathcal{D}_{rMt}^{\frac{1}{\sigma_M}},$$

$$(20) \quad \mathcal{A}_{rt} = \tilde{\alpha} A_{rt}^{\frac{\sigma_A-1}{\sigma_A}} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A}} (L_{rt}/T_r)^{-\alpha \frac{\sigma_A-1}{\sigma_A}} L_{rt}^{-\frac{1}{\sigma_A}}.$$

These shifters are sufficient to determine equilibrium wages and employment shares in each region.

**Proposition 3.** *Average income  $\bar{w}_{rt}$  and agricultural employment shares  $s_{rAt}$  satisfy:*

$$(21) \quad 1 = \left(\frac{\mathcal{Z}_{rt}}{\bar{w}_{rt}}\right)^\zeta + \left(\frac{\mathcal{A}_{rt}}{\bar{w}_{rt}}\right)^{\frac{\zeta}{\kappa}} \quad \text{and} \quad \frac{s_{rAt}^\kappa}{1-s_{rAt}} = \left(\frac{\mathcal{A}_{rt}}{\mathcal{Z}_{rt}}\right)^\zeta,$$

where  $\kappa \equiv \zeta - (\zeta - 1)(1 - \alpha) \frac{\sigma_A-1}{\sigma_A} \geq 1$ .

*Proof.* See Section A.3.2 in the Appendix. □

Proposition 3 shows that local income and sectoral specialization are fully determined

by  $\mathcal{Z}_r$  and  $\mathcal{A}_r$ . Physical productivity, market access terms, and local population are all endogenous local outcomes that depend on spatial linkages and equilibrium prices. The sectoral revenue shifters summarize all these factors and show that they are isomorphic in as far as wages and employment shares of an individual region are concerned.

Local income is increasing in both revenue productivities, while sectoral specialization depends only on their relative magnitude which summarizes the sectoral comparative advantage of a region. The mapping from these endogenous statistics to equilibrium allocations depends on the parameter  $\kappa$ , which reflects the elasticity of labor supply, the importance of agricultural land, and the price elasticity of agricultural goods.

The result in Proposition 3 is quite general: it follows solely from static equilibrium conditions and does not rely on assumptions about consumer preferences, the structure of spatial labor supply, or the dynamics of local productivity. As a result, it applies to a broad class of models in which sectoral and spatial reallocation jointly determine local economic outcomes.

Importantly, Proposition 3 links the drivers of structural change – productivity growth, changes in sectoral demand as captured by the market access terms, and population dynamics – to local wage growth and industrialization in equilibrium. We formalize this insight in the following Proposition:

**Proposition 4.** *Changes in local wages and agricultural employment shares are given by:*

$$(22) \quad d\ln \bar{w}_{rt} = \phi(s_{rAt}) d\ln \mathcal{Z}_{rt} + (1 - \phi(s_{rA})) d\ln \mathcal{A}_{rt}$$

$$(23) \quad ds_{rAt} = \psi(s_{rAt}) (d\ln \mathcal{Z}_{rt} - d\ln \mathcal{A}_{rt}),$$

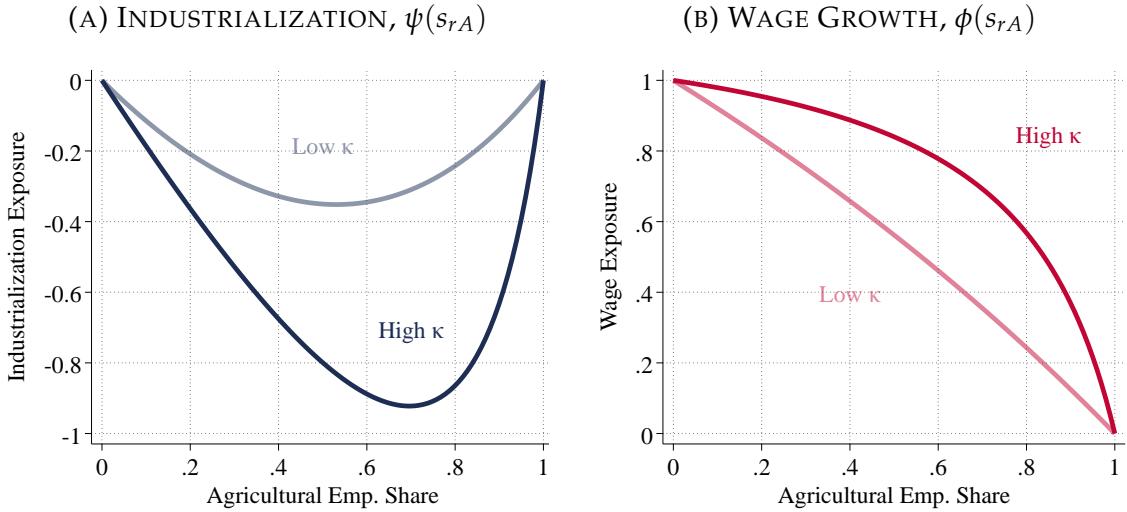
where  $\phi(s_{rAt}) = \frac{1-s_{rAt}}{1-s_{rAt}+s_{rAt}\frac{1}{\kappa}} \in (0, 1)$ ,  $\psi(s_{rAt}) = -\frac{\zeta s_{rAt}}{\kappa} \phi(s_{rAt}) < 0$ , and  $\kappa$  is defined in Proposition 3.

*Proof.* See Section A.3.3 in the Appendix. □

Proposition 4 delivers a complete analytical characterization of local wage growth and industrialization. It shows that regional heterogeneity in outcomes arises from two distinct sources: variation in the *incidence* of shocks to revenue productivity ( $d\ln \mathcal{Z}_{rt}$  and  $d\ln \mathcal{A}_{rt}$ ), and variation in *exposure* to those shocks, as captured by the elasticities  $\phi(s_{rAt})$  and  $\psi(s_{rAt})$ .

Proposition 4 establishes that sectoral specialization,  $s_{rAt}$ , is a sufficient statistic for regional exposure to structural change reflecting a Bartik-like intuition. As Figure 2 shows, the wage elasticity  $\phi(s_{rAt})$  declines monotonically with agricultural employment: rural areas are disproportionately affected by agricultural shocks and relatively insulated from manufacturing shocks. By contrast, the industrialization elasticity  $\psi(s_{rAt})$  is U-shaped, implying that gains in relative manufacturing productivity induce the largest

FIGURE 2: SPATIAL HETEROGENEITY IN EXPOSURE



*Notes:* This figure plots the exposure elasticities  $\psi(s_{rA})$  and  $\phi(s_{rA})$  from Proposition 4 as functions of the agricultural employment share. The dark line corresponds to a higher value of  $\kappa$ , and the lighter line to a lower value. The parameter  $\kappa \equiv \zeta - (\zeta - 1)(1 - \alpha) \frac{\sigma_A - 1}{\sigma_A} \geq 1$ , where  $\zeta$  denotes the elasticity of labor supply across sectors,  $\sigma_A$  is the demand elasticity for agricultural goods, and  $\alpha$  is the land share in agricultural production.

reallocation of employment in regions with intermediate specialization. Both elasticities are increasing in (the absolute value of)  $\kappa$ : when sectoral labor supply is highly elastic (high  $\zeta$ ) and agriculture exhibits limited decreasing returns (low  $(1 - \alpha) \frac{\sigma_A - 1}{\sigma_A}$ ), exposure to structural change is amplified.

Proposition 4 is central to understanding the economics of spatial structural change. It shows that aggregate structural change *always* has unbalanced spatial effects. Even when revenue productivity growth does not differ across locations, differences in agricultural employment shares across regions invariably imply heterogeneity in wage growth and industrialization across space.

In fact, the shape of the exposure elasticities in Figure 2 implies that if the aggregate agricultural employment share is decreasing, wage growth is necessarily urban-biased. To see this, consider a "macro-spatial" economy in which growth is balanced across space, but not sectors so that  $d \ln \mathcal{A}_{rt} = d \ln \mathcal{A}_t$  and  $d \ln \mathcal{Z}_{rt} = d \ln \mathcal{Z}_t$ .<sup>7</sup> In such an economy, the agricultural employment declines if and only if  $d \ln \mathcal{Z}_t > d \ln \mathcal{A}_t$ , since  $\psi(s_{rAt}) < 0$ . Equation (22) then implies that wage growth is *decreasing* in the agricultural employment share because rural locations are relatively more exposed to  $d \ln \mathcal{A}_t$ . By contrast, rural-biased wage growth in a macro-spatial economy requires that  $d \ln \mathcal{A}_t > d \ln \mathcal{Z}_t$ , which, however, would lead to *increasing* agricultural employment everywhere.<sup>8</sup>

<sup>7</sup>This arises naturally in our model when there is no productivity catch-up ( $\lambda_s = 0 \forall s$ ), trade is frictionless ( $\tau_{rjt} = 1 \forall r, j, t$ ), and population growth is uniform.

<sup>8</sup>The knife-edge case of spatially balanced wage growth emerges for  $d \ln \mathcal{A}_t = d \ln \mathcal{Z}_t$ , in which case no sectoral reallocation of employment would take place.

Note that the urban bias of structural change does *not* require *physical* productivity to grow faster in manufacturing than in agriculture. Even if physical agricultural productivity rises more quickly—that is,  $d \ln A_t > d \ln Z_t$ —agricultural *revenue* productivity may still grow more slowly than in manufacturing, that is,  $d \ln A_t < d \ln Z_t$ . This disconnect arises because revenue productivity depends on relative prices and relative demand. For instance, non-homothetic demand or an elasticity of substitution below one imply that a falling relative price of agricultural goods can go hand in hand with a falling expenditure share and hence faster revenue growth in the manufacturing sector.

The urban-biased wage growth predicted by aggregate structural change stands in stark contrast to the empirical patterns in Figure 1: between 1880 and 1920, sectoral reallocation toward manufacturing went hand-in-hand with rural-biased wage growth. Proposition 4 implies that, within our framework, this reversal can only be explained by spatially uneven revenue productivity growth—specifically, growth tilted toward agricultural regions. Our model identifies three mechanisms that could account for this pattern: faster physical productivity growth in rural areas via adoption of technologies from more advanced regions, faster increases in market access in rural areas, or reductions in rural populations that alleviate decreasing returns in agriculture. To assess the relative contribution of each, we now structurally estimate the model using detailed spatial data from this period.

## 4. STRUCTURAL ESTIMATION

We now describe how we estimate the structural parameters of our model. After summarizing the data in Section 4.1, we outline our estimation strategy in Section 4.2. Section 4.3 shows that our model matches a variety of targeted and non-targeted moments and successfully replicates the most salient features of the structural transformation in the US between 1880 and 1920 across time and space, most importantly the patterns in Figure 1.

### 4.1 Data Description

We treat each period in the model as spanning 20 years and use commuting zones (CZs) as our geographic units (Tolbert and Sizer, 1996). All workers are assigned to either agriculture or a residual non-agricultural sector combining manufacturing and services that we refer to simply as “manufacturing” throughout.

We obtain total employment by sector and county from the US Census Bureau’s Decennial Full Count Census files for 1860, 1880, 1900, and 1920 (via IPUMS; see Ruggles, Genadek, Goeken, Grover, and Sobek, 2015). These data also contain information on children and immigrants, which we use to estimate the exogenous component of local employment growth,  $n_{rt}$ .

We obtain county-level data on average wages from the Census of Manufacturing and agricultural land values from the Census of Agriculture (both via NHDGS; see [Manson, Schroeder, Van Riper, and Ruggles, 2017](#)).<sup>9</sup> The Census of Agriculture also provides data on the distribution of farm sizes (in acres) within each county, which we use to estimate the sectoral labor supply elasticity. To estimate spatial labor supply elasticities, we use inter-CZ migration flows constructed from linked Census microdata ([Ruggles et al., 2015](#)).

We aggregate all data to constant-boundary CZs using the crosswalk in [Eckert et al. \(2020\)](#), yielding a balanced panel of 495 CZs from 1860 to 1920 (see Section B.1 in the Appendix for a map). We supplement these regional data with aggregate time series data on GDP and prices from the Historical Statistics of the United States ([Carter, Gartner, Haines, Olmstead, Sutch, Wright et al., 2006](#)). Appendix B.1 provides full details on our data sources and processing.

## 4.2 Estimation Strategy

We estimate the model using a hybrid approach. First, we use cross-sectional data from 1880 to invert the static equilibrium conditions and recover region-specific fundamentals, that is productivities, amenities, and land endowments, for that year. This part of our calibration strategy only uses static equilibrium conditions; it does not depend on the parameters or specification of the dynamic productivity process. Second, we use dynamic moments from 1880–1920 to identify the structural parameters governing productivity catch-up growth via indirect inference. This two-step structure allows us to identify the initial productivity distribution independently from any assumptions on the process of productivity growth. Our estimation strategy does not require the economy to be in steady state.

Table 1 summarizes the parameter values and the moments used for estimation.

**Regional Fundamentals in 1880:**  $[Z_{r1880}, A_{r1880}, B_{r1880}, T_{r1880}]$  Proposition 3 shows that local wages and sectoral employment shares exactly identify the two sectoral revenue productivity terms. We then use the structure of the model and observed land rents to decompose these revenue productivities into physical productivity in each sector, sectoral market access, and land supply consistent with labor, land, and goods market clearing.<sup>10</sup> This implies that the model matches the observed wages, rental rates, and sectoral employment shares in 1880 exactly, given local employment. We then

---

<sup>9</sup>We compute average wages in each region,  $\bar{w}_{rt}$  as total manufacturing payroll divided by total manufacturing employment in the Census of Manufacturing. In our model, average wages are equal across sectors, reflecting selection on unobserved ability. As a result, our calibration does not require agricultural wage data, which are unavailable at the county level.

<sup>10</sup>In the data, we observe land values rather than rental rates. We assume that the land values in the data are proportional to the land rents in the model.

choose regional amenities in 1880 so that the migration probabilities implied by the 1880 equilibrium generate the observed distribution of employment across commuting zones in 1880, given the distribution of employment in 1860.<sup>11</sup>

The resulting estimates from this inversion procedure are important for the subsequent calibration steps. In particular, as shown in the left panel of Figure 3, we find that in 1880, agricultural regions lagged behind in both agricultural and manufacturing productivity; their comparative advantage in agriculture stemmed primarily from abundant land rather than from high agricultural productivity. This implies that initially agricultural regions stand to benefit from catch-up productivity growth. We exploit this fact in calibrating the productivity catch-up parameters  $\lambda_A$  and  $\lambda_M$ .

In Appendix B.3.3, we report the relationship between all inferred regional fundamentals and agricultural employment shares in 1880. Interestingly, we find that agricultural regions also exhibited slightly higher amenity levels—consistent with historical accounts of poor sanitation and safety in early industrial cities.<sup>12</sup>

**Technology Adoption Across Regions:**  $[\lambda_s, \mathcal{F}_s]_s$  We assume the following functional form for the index capturing the potential for technology adoption from other regions for farmers in region  $r$ :

$$(24) \quad \mathcal{F}_{rAt} = \mathcal{F}_A \left( \{f(d_{rj}) A_{jt-1}\}_j \right) = \left( \frac{1}{R} \sum_{j=1}^R d_{rj}^{-\iota} A_{jt-1}^\zeta \right)^{1/\zeta},$$

where  $d_{rj}$  is the distance between region centroids. In our baseline calibration, we set  $\zeta \rightarrow \infty$  and  $\iota = 0$ , so that  $\mathcal{F}_{rAt} = \max_j \{A_{jt-1}\}$  and hence also refer to  $\mathcal{F}_{rAt}$  as the *agricultural productivity frontier*.<sup>13</sup> We parametrize the manufacturing frontier,  $\mathcal{F}_{rZt}$ , in the same way.

We estimate the productivity catch-up parameters  $\lambda_A$  and  $\lambda_M$  by indirect inference targeting the two core patterns of rural convergence in Figure 1. Specifically, we simulate the model and choose  $\lambda_s$  to match the coefficients  $\beta_w$  and  $\beta_{s_A}$  of the following empirical regressions:

$$(25) \quad d \ln \bar{w}_{rt} = \delta_t + \beta_w s_{rAt} + \nu_{rt}$$

$$(26) \quad ds_{rAt} = \delta_t + \beta_{s_A} s_{rAt} + \gamma_{s_A} s_{rAt}^2 + \nu_{rt},$$

---

<sup>11</sup>Note that the calibrated amenity values implicitly control for differences in the geographic size of commuting zones: holding wages constant, more populous commuting zones are associated with higher inferred amenities.

<sup>12</sup>Many large cities lacked functioning sewer systems in 1880. For example, New York City’s sewer system was only completed in 1902. Before 1900, industrializing cities offered higher wages than rural areas, but were often dirtier and more dangerous to live in (see, e.g., [Boustan, Bunten, and Hearey, 2018](#)).

<sup>13</sup>In practice, to guard against measurement error, we define  $\mathcal{F}_{rAt}$  as the average sectoral productivity among the 5% of regions with the highest productivity in that sector.

with time differences taken over 20-year intervals. The quadratic specification in the second equation captures the U-shaped industrialization pattern in Figure 1. In the data, consistent with Figure 1, we find that  $\beta_w = 0.27$  and  $\beta_{s_A} = -0.49$ ; see Column 1 of Table A.1.

Proposition 4 provides a direct identification argument for why the two moments  $\beta_w$  and  $\beta_{s_A}$  are informative about the productivity catch-up parameters. The proposition shows that increases in agricultural productivity raise both the average wage and the agricultural employment share of a region, while increases in manufacturing productivity raise wages but shift employment toward manufacturing. As shown in the left panel of Figure 3, agricultural regions in 1880 lagged behind in both sectors. Productivity catch-up in either sector therefore raises wages, but the employment response differs: agricultural catch-up reinforces specialization in agriculture, while manufacturing catch-up pushes toward industrialization. This asymmetry underpins the separate identification of  $\lambda_A$  and  $\lambda_M$ .

**Changes in Amenities**  $B_{rt}$  The existing literature documents a systematically faster growth of amenities in urban areas during our study period.<sup>14</sup> The corresponding fall of compensating differentials to reside in cities could have contributed to rural-biased wage growth. To allow for the possibility of urban-biased amenity growth, we model the *growth* in local amenities as  $d \ln B_{rt} = \xi s_{rAt}$  and treat  $\xi$  as a structural parameter. We estimate  $\xi$  using indirect inference by matching the empirical relationship between agricultural specialization and population *growth*. Specifically, we target the coefficient  $\beta_L$  from the regression

$$(27) \quad d \ln L_{rt} = \delta_t + \beta_L s_{rAt} + \nu_{rt},$$

with time differences taken over 20-year intervals. In the data, we find that  $\beta_L = -0.1$ , implying essentially no correlation between agricultural specialization and subsequent population growth. Given faster rural wage growth, amenities must have grown more quickly in urban CZs, as suggested by the findings of [Boustan et al. \(2018\)](#).

**Aggregate Productivity Growth ( $g_A$  and  $g_M$ ) and Consumer Preferences ( $\eta$ ,  $\phi$ , and  $\nu$ )** We estimate aggregate sectoral productivity growth ( $g_A$ ,  $g_M$ ) and the Engel elasticity  $\eta$  to match three aggregate moments computed over the 1880-1920 period: GDP per capita growth, the decline in the agricultural employment share, and the change in relative sectoral prices. The parameter  $\phi$  corresponds to the asymptotic agricultural

---

<sup>14</sup>See, for example, [Boustan et al. \(2018\)](#) who write “The second episode, from the late nineteenth century to the mid-twentieth century, featured steadily rising wages coupled with a declining urban wage premium. From 1880 to 1940, wages in each sector more than tripled while the wage premium fell modestly from 41 percent to 36 percent. Because workers continued to move to the cities during this period, the slumping wage premium suggests an increase in urban consumption amenities like public health.”

employment share in the model and we set  $\phi = 0.01$ , the approximate US agricultural employment share in 2020. We normalize  $\nu = 1$  without loss of generality.<sup>15</sup>

**Skill Heterogeneity  $\zeta$**  We estimate the Fréchet dispersion parameter  $\zeta$  using data on the farm-size distribution in each CZ from the Agricultural Census. The data report the number of farms with less than 3, 10, 20, 50, 100, 500, and 1000 acres of farmland; we index these seven size bins by  $j$ . In the model, the optimal farm size is proportional to farmer ability and is hence also Fréchet distributed. This implies that  $\zeta$  can be estimated from the regression

$$(28) \quad \ln(-\ln F_{rt}^T(x^j)) = \delta_{rt} - \zeta \ln x^j + u_{rt}^j,$$

where  $x^j$  denotes the size cutoff of bin  $j$  (e.g., 3, 10, ... 1000),  $F_{rt}^T(x^j)$  is the share of farms in region  $r$  at time  $t$  smaller than  $x^j$ , and  $u_{rt}^j$  is an error term, which we assume to reflect measurement error. If  $\zeta$  is large, the share of farms in the upper tail of the size distribution declines quickly in size. Note that the region-time fixed effects  $\delta_{rt}$  control for local productivity  $A_{rt}$  and local rental rates  $R_{rt}$ , which are common across farmers within region  $r$ : our estimation of  $\zeta$  uses variation across farm size bins  $j$  *within* each CZ and year. As we show in Appendix Section B.3.4, this yields an estimate of  $\zeta = 1.6$ , which is very robust across different specifications.<sup>16</sup>

**Spatial Labor Supply** The dynamics of spatial labor supply are governed by bilateral migration costs  $\mu_{rj}$ , the dispersion of location preference shocks,  $\varepsilon$ , and the rate of exogenous population growth in each location,  $n_{rt}$ .

We parametrize migration costs as a function of distance,  $\mu_{rj} = d_{rj}^{-\omega}$  and estimate  $\varepsilon$  and  $\omega$  in two steps. First, we take the logarithm of the migration share expression in equation (11) to derive the following estimating equation:

$$(29) \quad \ln m_{rjt} = \delta_{rt}^o + \delta_{jt}^d - \omega \varepsilon \ln d_{rj},$$

where  $\delta_{rt}^o$  and  $\delta_{jt}^d$  are origin and destination fixed effects. As we show in Appendix B.3.2, estimating equation (29) using linked Census data, yields  $\omega \varepsilon = 2.8$ .<sup>17</sup>

Second, we estimate  $\varepsilon$  using a method-of-moments estimator to match an average elasticity of labor supply across regions of 2, consistent with Peters (2022). Note that due to non-homothetic preferences, the labor supply elasticity is not a structural parameter

<sup>15</sup>Note that  $\nu$  is not separately identified from “real income”  $e^i / (P_{rA}^\phi P_{rM}^{1-\phi})$  — see equation (8). Hence,  $\nu$  can be normalized by choice of units as long as  $\nu > 0$ , that is, agricultural goods are necessities.

<sup>16</sup>Using data on individuals earnings in the US around 2000, Lagakos and Waugh (2013) estimate a Fréchet shape parameter of around 5 in agriculture. In Section 6, we therefore explore the robustness of our results with respect to  $\zeta$ .

<sup>17</sup>Our estimate of this composite parameter is qualitatively consistent with Allen and Donaldson (2020), who find a value of 2.16 for cross-county migration in the US during the same time period.

but, in addition to  $\varepsilon$ , also depends on the Engel elasticity  $\eta$ , the taste parameter  $\nu$ , and a set of endogenous variables.

To estimate the exogenous region-specific employment growth shifters  $n_{rt}$ , we use historical data on international migration, local birth rates, and the age structure. As we describe in detail in Appendix B.3.1, we choose  $n_{rt}$  to match the net employment growth from international migration and new births for each region; we scale total employment growth across all regions to ensure our model also matches the aggregate rate of employment growth between 1880 and 1920.

**Trade Costs** We take trade costs from [Donaldson and Hornbeck \(2016\)](#), who reconstruct the entire historical transportation network of the US economy and provide direct estimates of iceberg trade costs between all county pairs for all years in our study. Since the railroad expanded massively from 1870 to 1890, their trade cost estimates decrease over time. In Section 6 below, we quantify the contribution of these changes to the observed patterns of structural change in Figure 1.

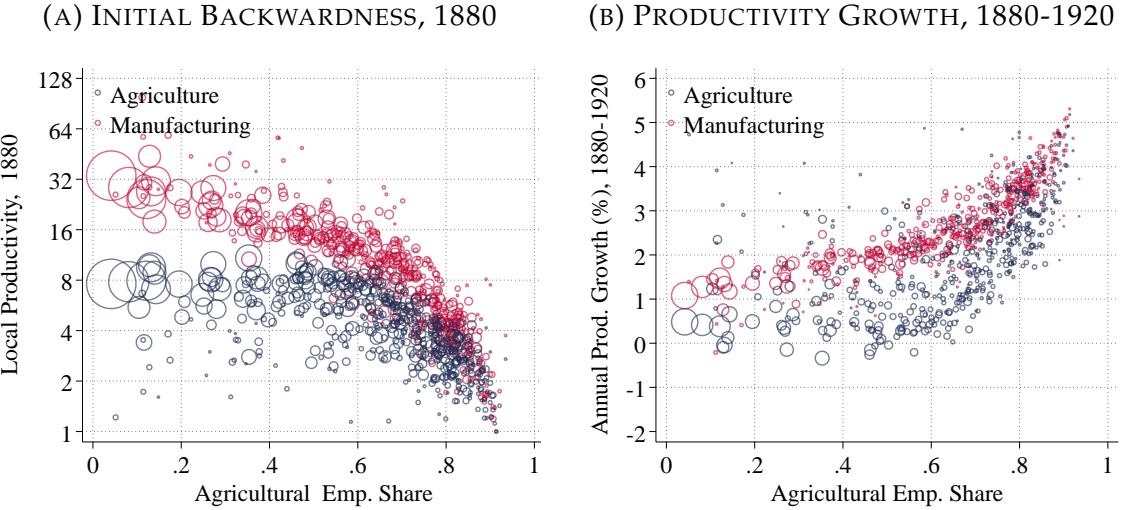
**Other Parameters** The remaining parameters are taken from established sources in the literature. Following [Donaldson and Hornbeck \(2016\)](#), we use a trade elasticity of  $\sigma_A = 8.22$  for the agricultural sector and, for symmetry in our baseline calibration, set  $\sigma_M = 8.22$  for manufacturing as well. We obtain the share of agricultural land in production,  $\alpha$ , from [Valentinyi and Herendorf \(2008\)](#) who find  $\alpha = 0.4$ . Lastly, we take the congestion elasticity of  $\rho = 0.15$  from [Allen and Donaldson \(2020\)](#), which is estimated using the same time period and Census data used in our study.

### 4.3 Estimates and Model Fit

Table 1 presents our parameter estimates alongside the moments used for identification. We distinguish between parameters estimated by solving the equilibrium of the model and targeting the respective empirical moments ("in-model," Panel A), parameters estimated outside the model using model-derived estimating equations ("out-of-model," Panel B), and parameters borrowed from the literature ("exogenously-set" Panel C). All structural residuals are inferred within the equilibrium loop since they depend on the other structural parameters.

Panel A shows that our model replicates the key empirical features of the structural transformation in the US between 1880 and 1920 both across sectors and regions. First, it matches three central time-series patterns: (i) the 22-point decline in agricultural employment, (ii) the average annual growth rate of GDP per capita of 2%, and (iii) the modest decline of the relative price of agricultural goods of 9 percentage points. To match these moments, we estimate 20-year frontier growth rates of 1% in agriculture and 17% in manufacturing, and an Engel elasticity of 0.8, implying a strong role for demand-side forces in driving structural transformation.

FIGURE 3: RURAL PRODUCTIVITY CATCH-UP GROWTH



*Notes:* The left panel displays the initial productivity distribution,  $\ln Z_{r1880}$  and  $\ln A_{r1880}$ , with the agricultural employment share in 1880. The right panel displays the correlation of the estimated rate of annual local productivity growth between 1880 and 1920, that is,  $\frac{1}{40} \ln (Z_{r1920} / Z_{r1880})$  and  $\frac{1}{40} \ln (A_{r1920} / A_{r1880})$ , with the agricultural employment share in 1880.

Second, the calibrated model successfully replicates the rural convergence patterns documented in Figure 1. It exactly matches the empirical coefficients from the linear regressions of wage growth and industrialization on initial agricultural employment shares (equations 25 and 26). As shown in Figure 4, the model-implied relationships closely align with the data—even though we target only the linear coefficients. The model captures both the strong rural bias in wage growth (right panel) and the non-monotonic, U-shaped pattern of industrialization across space (left panel).

Notably, although the quadratic term in regression (26) is not targeted, the model-generated estimate,  $\gamma_{sA} = 0.42$ , closely matches the empirical value of 0.45. This success stems from the shape of the model-implied exposure to industrialization shown in Figure 2: the U-shape in exposure is inherited by the pattern of industrialization, despite substantial variation in the incidence of shocks across regions.

To match these convergence patterns, our estimates imply an important role for catch-up growth. We estimate  $\lambda_A = 0.73$  and  $\lambda_M = 0.28$ , suggesting that backward regions experienced faster productivity growth in both sectors. The right panel of Figure 3 illustrates the magnitude of differential productivity growth implied by these parameters, given the initial productivity differences shown in the left panel. Quantitatively, rural labor markets experienced a growth premium of around three percentage points, with a slightly larger premium in the agricultural sector.

The model also replicates key patterns of spatial labor mobility. The left panel of Figure 5 shows local population growth by initial share of agricultural employment in both the model and data. Both display a weak, non-monotonic relationship: population

TABLE 1: MODEL PARAMETERS AND FIT

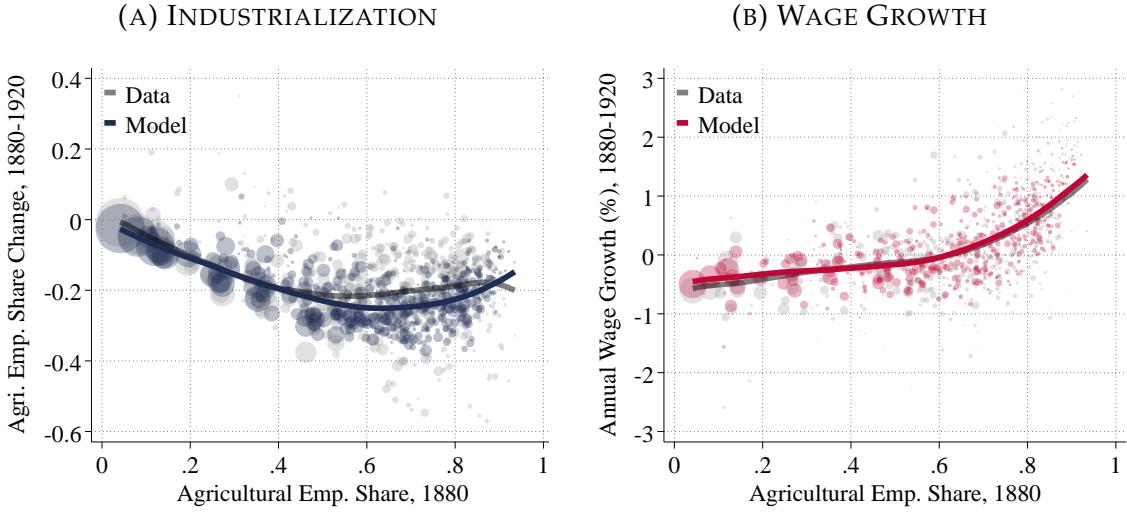
STRUCTURAL PARAMETERS		ESTIMATION METHOD			
DESCRIPTION	VALUE	PANEL A: IN-MODEL (MOMENT, DATA, MODEL)			
$g_A$	Agg. prod. growth: agriculture	0.01	Annual GDP p.c. growth, 1880-1920	0.02	0.02
$g_M$	Agg. prod. growth: manufacturing	0.17	Change in $P_M/P_A$ , 1880-1920 (ppt.)	-8.95	-8.95
$\lambda_A$	Catch-up parameter: agriculture	0.73	$\beta_{sA}$ in regression (26)	-0.49	-0.49
$\lambda_M$	Catch-up parameter: manufacturing	0.28	$\beta_w$ in regression (25)	0.27	0.27
$\epsilon$	Location taste heterogeneity	2.47	Avg. Migration Elasticity	2.00	2.00
$\eta$	Engel elasticity	0.80	$\Delta$ agri. emp. share, 1880-1920 (ppt.)	-21.37	-21.38
$\xi$	Spatial bias in amenity growth	-1.10	$\beta_L$ in regression (27)	-0.10	-0.10
PANEL B: OUT-OF-MODEL (STRATEGY)					
$\zeta$	Labor Supply Elasticity	1.60	Farm size distribution, 1880		
$\omega$	Migration Cost Distance Elasticity	1.13	Gravity relationship of migration flows		
$\phi$	Asymptotic exp. share on agri. goods	0.01	Agri. emp share 2020		
PANEL C: EXOGENOUSLY-SET (SOURCE)					
$\sigma_A$	Trade elasticity agriculture	8.22	<a href="#">Donaldson and Hornbeck (2016)</a>		
$\sigma_M$	Trade elasticity manufacturing	8.22	<a href="#">Donaldson and Hornbeck (2016)</a>		
$\rho$	Amenity congestion elasticity	0.15	<a href="#">Allen and Donaldson (2020)</a>		
$\alpha$	Land Share in Production Function	0.40	<a href="#">Valentini and Herrendorf (2008)</a>		
$\nu$	PIGL preference parameter	1.00	Normalization		
$f_E$	Fixed cost of entry	1.00	Normalization		
STRUCTURAL RESIDUALS			MOMENT		
$A_{r1880}$	Local productivity: agriculture		CZ agricultural employment shares, 1880		
$Z_{r1880}$	Local productivity: manufacturing		CZ average wages, 1880		
$T_{r1880}$	Local agricultural land supply		CZ agricultural land rents, 1880		
$B_{r1880}$	Local amenity		CZ net employment growth, 1880-1900		
TRADE COSTS AND DEMOGRAPHICS			SOURCE/TARGET		
$\tau_{rj}^A$	Trade cost: agriculture		<a href="#">Donaldson and Hornbeck (2016)</a>		
$\tau_{rj}^{NA}$	Trade cost: manufacturing		<a href="#">Donaldson and Hornbeck (2016)</a>		
$n_{rt}$	Exogenous population growth		Local birth and immigration rates		

*Notes:* The table contains the values for all structural parameters and targeted moments of our model. "In-Model" refers to parameters calibrated by solving the model repeatedly and adjusting parameters to match empirical moments. "Out-of-Model" refers to parameters estimated using structural estimating equations derived from the model. All structural residuals are inferred as part of the equilibrium loop.

grew slightly faster in the most rural and most urban regions. This is surprising, given that rural areas experienced faster wage growth, and implies that urban amenities must have improved relative to rural areas. Indeed, our estimate of  $\xi = -1.1$  implies that a 10-point higher agricultural employment share at time  $t$  was associated with 11% slower amenity growth between  $t$  and  $t + 1$ , consistent with independent evidence on rising urban amenities in this period (see, [Boustan et al., 2018](#)).

The fact that agricultural specialization and subsequent population growth are uncorrelated suggests that spatial reallocation from more to less agricultural regions did not play an important role in the aggregate decline in agricultural employment ([Eckert, Juneau, and Peters, 2023](#)). The right panel of Figure 5 decomposes the 22 percentage point decline of the national agricultural employment share into declines that occur within locations (holding constant regional populations in 1880), across locations (hold-

FIGURE 4: RURAL CONVERGENCE – MODEL AND DATA



Notes: The figure displays the correlation of industrialization (left panel) and wage growth (right panel) with the agricultural employment share. We show the data in lighter-shaded colors and model output in darker shades.

ing constant regional agricultural employment shares in 1880), and a covariance term.<sup>18</sup> We find that nearly all of the aggregate decline occurred *within* regions while spatial reallocation only played a minor role. Figure 5 also shows that our model replicates these patterns remarkably well. Even though all workers are spatially mobile, the aggregate shift out of agriculture is overwhelmingly a local phenomenon.

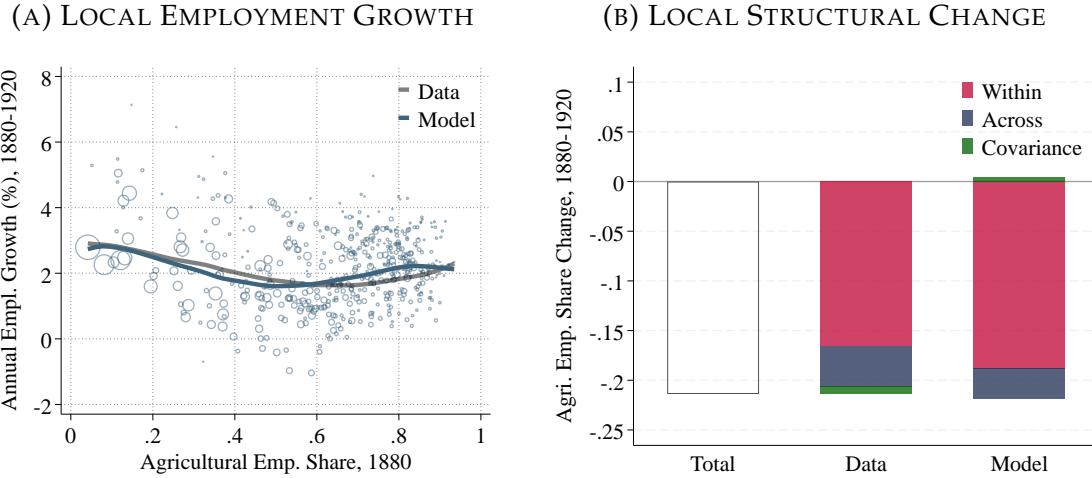
#### 4.4 Rural Productivity Catch-up Growth: Additional Evidence

Our structural estimates imply that rural labor markets experienced significantly faster productivity growth between 1880 and 1920. This section provides three additional pieces of evidence in support of this finding. First, we estimate productivity growth using a fully flexible model inversion to show that our findings are not an artifact of our functional form assumptions on the dynamic productivity process. Second, we show that the rural bias in productivity growth does not reflect spatial differences in human capital accumulation. Finally, we provide direct evidence of rural-biased productivity growth using data from the Agricultural Census.

**Model-Inversion** Our theory proposes a specific and parsimonious model of technological convergence. A natural concern is whether our finding of rural productivity catch-up is driven by this modeling assumption. To address this, we estimate sectoral productivity— $A_{rt}$  and  $Z_{rt}$ —using direct model inversion. As in our baseline inversion for 1880, we recover  $A_{rt}$ ,  $Z_{rt}$ , and  $T_{rt}$  separately for 1880, 1900, and 1920, using observed wages, employment shares, land rents, and population. This procedure rationalizes the

<sup>18</sup>Formally, we implement the following decomposition,  $\Delta s_{At} = \sum_r l_{rt} \Delta s_{rAt} + \sum_r \Delta l_{rt} s_{rAt} + \sum_r \Delta l_{rt} \Delta s_{rAt}$ , where  $\Delta$  indicates time differences over 40 years and  $l_{rt}$  is the share of national employment at time  $t$  attributable to commuting zone  $r$ .

FIGURE 5: NON-TARGETED MOMENTS



*Notes:* The left panel shows the relationship between employment growth between 1880 and 1920 and the agricultural employment share in 1880. We show the data in grey and the model-generated data in blue. The size of the markers reflects the relative size of different commuting zones. The fit lines are Epanechnikov kernel-weighted polynomial fit lines with a bin width of 0.1. The right panel decomposes the aggregate decline in the agricultural employment share between 1880 and 1920,  $\Delta s_{At}$ , as follows:  $\Delta s_{At} = \sum_r l_{rt} \Delta s_{rAt} + \sum_r \Delta l_{rt} s_{rAt} + \sum_r \Delta l_{rt} \Delta s_{rAt}$ , where  $\Delta$  indicates time differences over 40 years and  $l_{rt}$  is the share of national employment at time  $t$  attributable to commuting zone  $r$ . We refer to the first components as “within,” the second as “across,” and the third as “covariance.”

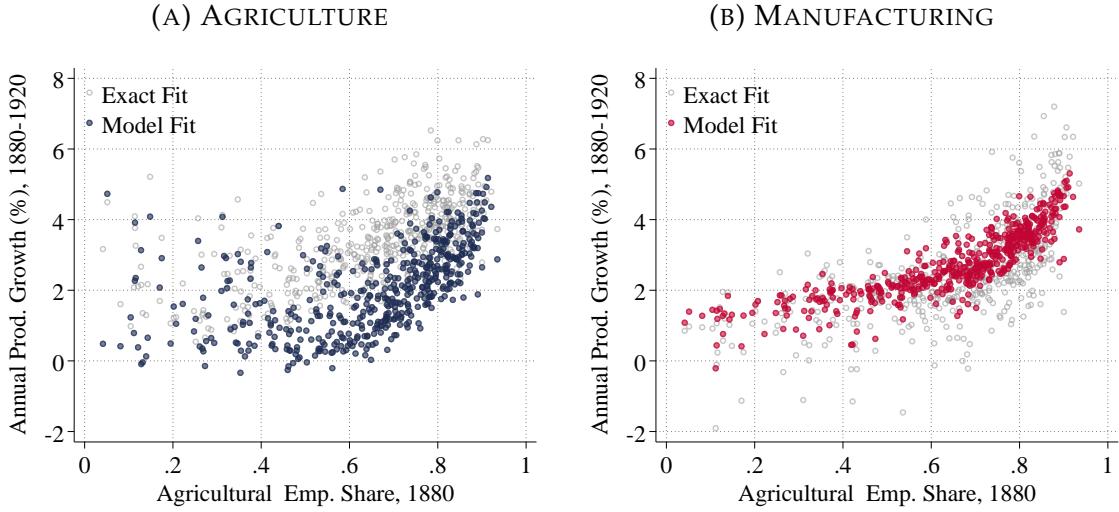
data exactly at each point in time, without imposing any assumptions on the dynamics of productivity or land supply.

Figure 6 compares the resulting productivity growth estimates (in gray) to those from our baseline calibration (in red and blue). In both agriculture (left panel) and manufacturing (right panel), we again find a rural-biased pattern: initially more agricultural locations experienced faster productivity growth. The close alignment of the estimated productivity terms between the two approaches confirms that rural productivity catch-up is not a mechanical outcome of our assumptions on the productivity process.

**Adjusting for Human Capital** Our baseline model abstracts from human capital. Yet differential human capital accumulation across regions could, in principle, account for rural-biased productivity growth (e.g., [Caselli and Coleman II, 2001](#)). To address this possibility, we extend the model to allow for arbitrary human capital differences across sectors and space (see Appendix B.5). Specifically, we assume that individuals draw sector-specific ability from a Fréchet distribution with mean  $h_{rt}^s \Gamma_\zeta$ , where  $h_{rt}^s$  denotes the average human capital in sector  $s$  and region  $r$ .

This extension preserves the model’s structure and is isomorphic to the baseline formulation, with the sole difference that physical productivity is scaled by average human capital in each region and sector. For instance, agricultural productivity becomes  $A_{rt} = (h_{rt}^A)^v \bar{A}_{rt}$  for a known constant  $v$ , where  $\bar{A}_{rt}$  denotes physical productivity net of human capital. Given estimates of  $h_{rt}^s$ , we can therefore recover “purged” productivity

FIGURE 6: RURAL CATCH-UP: MODEL INVERSION VS. STRUCTURAL ESTIMATES



Notes: The figure plots annualized productivity growth as estimated from our model and from full model inversion against the initial agricultural employment share.

measures  $\bar{A}_{rt}$  from our original estimates of  $A_{rt}$ .

We model human capital in a Mincerian way as  $h_{rt}^s = \exp(\gamma_s \text{sch}_{rt})$ , where  $\gamma_s$  is the sector-specific return to schooling and  $\text{sch}_{rt}$  denotes average years of schooling among workers in location  $r$  at time  $t$ . When we combine Census data and model-implied estimating equations, we find substantial returns to schooling in manufacturing but near-zero returns in agriculture, consistent with prior work (e.g., [Porzio, Rossi, and Santangelo, 2022](#)). Using these estimates, we construct estimates of  $h_{rt}^s$  for every location and sector.<sup>19</sup>

When we recover the purged productivity terms,  $\bar{A}_{rt}$ , from our original estimates, we find that their cross-sectional and dynamic patterns closely mirror those in our original estimates. While rural regions had lower educational attainment in 1880, changes in education over time were not systematically related to initial agricultural specialization. We conclude that human capital growth cannot explain rural-biased productivity convergence, though it likely contributed to overall productivity growth—particularly in manufacturing where returns to education are important.

**Direct Productivity Measurements** Finally, we present model-free evidence on agricultural productivity growth using data from the Census of Agriculture. Specifically, we use agricultural yields—output per acre—as a direct proxy for local agricultural productivity. As shown in Appendix B.3.6, rural regions were less productive in 1880 but experienced faster yield growth between 1880 and 1920. These patterns reinforce two key findings of our structural estimation: initially agricultural regions had low agricultural productivity, but experienced productivity catch-up, even in the agricultural

<sup>19</sup>Since educational attainment is not observed in the US Census prior to 1940, we infer educational attainment in 1900 and 1920 using retrospective data on older cohorts in the 1940 Census.

sector.

## 5. THE ECONOMICS OF SPATIAL STRUCTURAL CHANGE: QUANTITATIVE RESULTS

We now use our calibrated model to decompose the patterns of spatial structural change shown in Figure 1 into their underlying driving forces. This analysis delivers one of the central results of the paper: quantitatively, explaining the joint patterns of rural wage convergence and rural industrialization observed between 1880 and 1920 requires strong productivity catch-up in rural locations in both sectors.

We document this finding in two complementary ways, both of which build on our theoretical decomposition contained in Proposition 4. In Section 5.1, we directly leverage Proposition 4 to decompose wage growth and industrialization across regions into the impact of productivity growth, changes in demand, and local employment growth. We find that spatial differences in productivity growth account for the bulk of rural wage convergence and industrialization, while changes in demand and population growth played a secondary role.

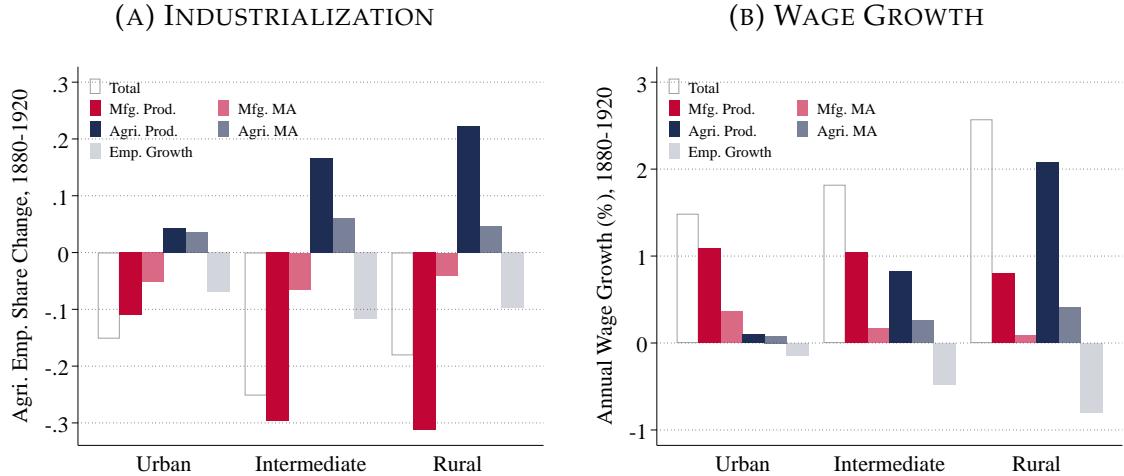
In Section 5.2, we supplement this decomposition with a complimentary exercise where we re-estimate our entire model under the assumption of no catch-up growth in productivity. Under this calibration, the model continues to match the time series aspects of the structural transformation and incorporates other spatially-biased demand shifts such as falling transport costs or the arrival of international migrants. We show that this parametrization of our theory fails to account for the spatial patterns of structural change: it predicts faster wage growth in urban areas and substantially slower industrialization in all locations. These findings confirm the quantitative importance of the urban-bias of *aggregate* structural change, and highlight the pivotal role of productivity catch-up in producing the rural-biased wage growth in Figure 1.

### 5.1 Sources of Spatial Structural Change

Proposition 4 contains a model-based decomposition of local wage growth and local industrialization into changes in sectoral revenue productivity and region-specific exposure. Changes in revenue productivity can be further decomposed into five economic forces: (i) local productivity growth in agriculture and manufacturing ( $d \ln A_{rt}, d \ln Z_{rt}$ ), (ii) changes in market access ( $d \ln D_{rAt}, d \ln D_{rMt}$ ), and (iii) local employment growth ( $d \ln L_{rt}$ ). The local impact of each channel varies with a region's exposure as summarized by its sectoral employment share.

To implement the decomposition, for each commuting zone, we compute the impact of each of these five factors multiplied by their respective exposure elasticities separately

FIGURE 7: THE MECHANISMS OF SPATIAL STRUCTURAL CHANGE



Notes: The figure reports the decomposition of local industrialization,  $ds_{rA}$ , and local wage growth,  $d\ln \bar{w}_{rt}$ , into changes in market access ( $d\ln \mathcal{D}_{rst}$ , "MA"), sectoral productivity ( $d\ln Z_{rt}$  and  $d\ln A_{rt}$ ), and local employment ( $d\ln L_{rt}$ ), all pre-multiplied by the relevant exposure elasticities, see Proposition 4. We define urban (rural) locations as regions in the lower (upper) quartile of the distribution of agricultural employment shares in 1880. We refer to all commuting zones within the interquartile range as "intermediate."

for local wage growth and local industrialization for 1880-1900 and 1900-1920. We then aggregate these results across time periods and among urban, intermediate and rural locations, which we define as all regions below, within, and above the interquartile range of agricultural employment shares in 1880.<sup>20</sup>

Figure 7 shows the results. We depict the impact of physical productivity growth in dark blue and red, the impact of changes in market access in light blue and red and the impact of population growth in gray. The aggregate effect, shown in white, is the sum of these five forces.

Our first quantitative finding is that catch-up productivity growth in agriculture was essential to explain rural wage convergence. The right panel shows that wage growth in rural areas is overwhelmingly driven by *agricultural* productivity growth, while urban wage growth is primarily driven by productivity growth in *manufacturing*. Specifically, holding all else constant, agricultural TFP growth alone would have raised rural wages by nearly 2% annually.

Our second quantitative finding is that rural industrialization was driven by productivity catch-up in manufacturing as shown in the left panel of Figure 7. In fact, if only manufacturing productivity had increased, the overall decline in agricultural employment would have been larger than what is observed empirically. This reflects that agricultural productivity growth acts in the opposite direction—it strengthens comparative advantage in agriculture and works against reallocation.

<sup>20</sup>See Appendix B.2 for details. There we also show that the decomposition of Proposition 4, which relies on a first-order approximation, provides a good fit to the full non-linear solution in our model.

A corollary of our findings on the role of sectoral productivity growth is that the contributions of demand-side forces and population growth are secondary. While falling trade costs improves market access in many regions and structural change shifts demand toward manufacturing due to non-homothetic preferences, quantitatively they play a less important role, both to explain local wage growth and local industrialization. The effect of population growth is similarly nuanced. In the presence of decreasing returns in agriculture, rising population density contributed to rural industrialization, but at the same time put downward pressure on rural wages.

The decompositions in Figure 7 highlight how the interaction between exposure and incidence shapes spatial economic outcomes: incidence determines who is affected, while exposure determines how strongly. For any force to have a large impact on a region, both its incidence and the region’s exposure to it must be high. For example, manufacturing productivity growth was a key driver of rural industrialization but had limited effects on rural wage growth, since wages respond little to rising productivity in sectors that employ a small share of the local workforce. Conversely, even though physical manufacturing productivity grew less in urban areas than in rural ones, its wage impact was larger in cities due to their greater exposure to the manufacturing sector. A similar logic applies to population growth: although average growth was spatially balanced, it depressed rural wages more because rural areas were more exposed to the sector—agriculture—where congestion effects are strongest.

Together, these findings deliver a sharp conclusion. First, productivity catch-up growth played the key role in explaining the empirical patterns of spatial structural change during the first structural transformation of the US. Second, while faster productivity growth in agriculture was at the heart of rural wage growth, rural industrialization was driven by productivity convergence in manufacturing. Third, neither changes in market access nor population growth can, on their own, account for the pro-rural nature of US economic growth between 1880 and 1920.

## 5.2 The Inherent Urban-Bias of Structural Change

A key implication of Proposition 4 is the inherent urban-bias of the structural transformation: spatially balanced productivity growth leads to faster wage growth in industrialized locations if and only if agricultural employment declines. While this result is exactly true with common population growth across space and free trade, we now use our quantitative model to show that this pattern also arises in the presence of trade costs and endogenous migration decisions.

To do so, we turn to a counterfactual analysis to quantify the full, system-wide impact of spatial productivity catch-up growth taking all general equilibrium interactions into account. In particular, we compare our baseline model to an alternative calibration in

which the spatial distribution of productivity remains stationary, but the model still generates the same *aggregate* facts in the time-series. Specifically, we set  $\lambda_A = \lambda_M = 0$  and re-estimate the aggregate productivity shifters  $g_A$  and  $g_M$  to match the observed aggregate growth rate and the change in relative prices. Because sectoral productivity grows at the common rates  $g_s$  in all locations, there are no benefits of backwardness and we refer to this parametrization as the “no productivity catch-up” calibration.

The results are stark. Without productivity catch-up growth, structural change becomes strongly urban-biased. Rural wage convergence disappears, rural industrialization slows markedly, and employment spatially reallocates towards urban centers. Aggregate structural change also takes a different form: although the total decline in agricultural employment remains similar, it is now driven much more by the spatial reallocation of labor toward industrialized labor markets than by falling agricultural employment within local labor markets.

TABLE 2: THE “NO PRODUCTIVITY CATCH-UP” CALIBRATION

DESCRIPTION	VALUE	MOMENT	DATA	MODEL	TARGETTED
<b>Panel A: Re-calibrated</b>					
$g_A$ Agg. prod. growth: agriculture	0.29	Annual GDP p.c. growth, 1880-1920	0.02	0.02	Yes
$g_M$ Agg. prod. growth: manufacturing	0.30	Change in $P_M/P_A$ , 1880-1920 (ppt.)	-8.95	-8.95	Yes
<b>Panel B: Exogenously Set</b>					
$\lambda_A$ Catch-up parameter: agriculture	0.00	$\beta_{sA}$ in regression (26)	-0.49	-0.15	No
$\lambda_M$ Catch-up parameter: manufacturing	0.00	$\beta_w$ in regression (25)	0.27	-0.11	No
<b>Panel C: Not-Recalibrated</b>					
$\epsilon$ Location taste heterogeneity	2.47	Avg. Migration Elasticity	2.00	2.01	No
$\eta$ Engel elasticity	0.80	Agri. emp. share change, 1880-1920 (ppt.)	-21.37	-17.51	No
$\xi$ Spatial bias amenity growth	-1.10	$\beta_L$ in regression (27)	-0.10	-0.65	No

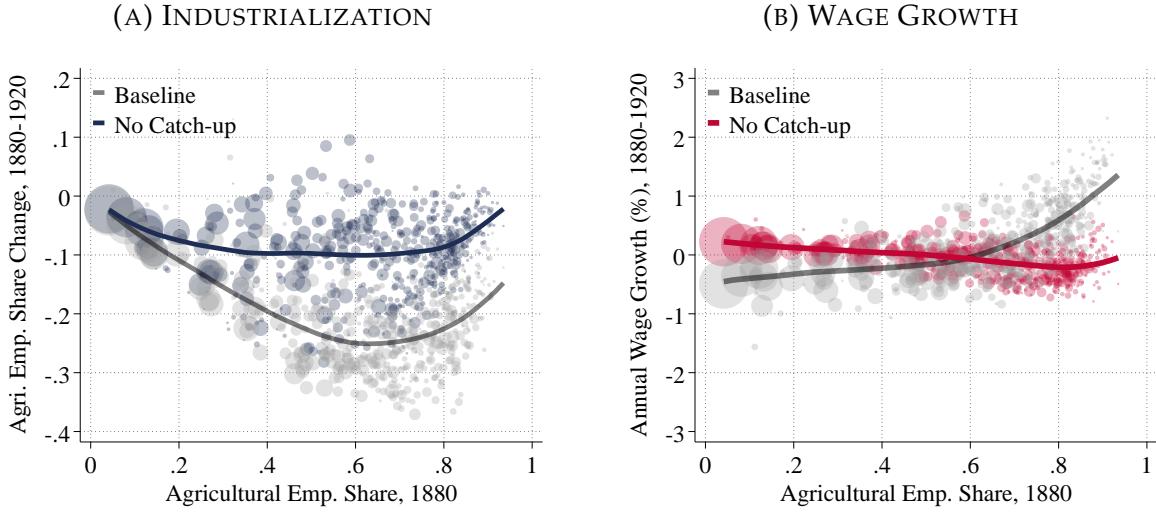
*Notes:* The table reports the structural parameters estimated by simulating the model and their corresponding moments for the “no productivity catch-up” calibration. We set  $\lambda_A$  and  $\lambda_M$  to zero; the table reports the moments we use to estimate  $\lambda_A$  and  $\lambda_M$  in the baseline calibration for comparison. We recalibrate  $g_A$  and  $g_M$  to match the same targets as before. We also report the remaining structural parameters calibrated within the model and their corresponding moments, however, we are not re-calibrating these moments. All other parameters are the same in both calibrations and reported in Table 1.

These findings are summarized in Table 2 and Figure 8. Table 2 reports the parameter values and key moments for the “no productivity catch-up” calibration. Note first that the model still matches the aggregate patterns of structural change. Overall growth and the change in relative prices are, by construction, consistent with the data. In addition, the model still generates a large decline in agricultural employment of 17.5 percentage points.<sup>21</sup>

Crucially, Figure 8 shows that the model has drastically counterfactual predictions for the spatial patterns of structural change. Without productivity catch-up, the model

<sup>21</sup>The location fundamentals in 1880, that is, initial productivity, land endowments and local amenities, are also the same in both calibrations, because they are estimated from static equilibrium conditions and therefore independent of  $g_s$  and  $\lambda_s$ .

FIGURE 8: THE ROLE OF RURAL PRODUCTIVITY CATCH-UP



Notes: In the left (right) panel, we show local industrialization (local wage growth) as a function of the initial agricultural employment share. We depict the baseline calibration in gray and the calibration without productivity catch-up in red and blue, respectively. The size of the markers reflects the relative size of different commuting zones.

neither replicates the observed rural wage convergence, nor the U-shaped relationship between initial agricultural employment and industrialization. In the baseline model (gray), rural locations experienced both faster wage growth and stronger industrialization. In the no-catch-up model (red and blue), urban locations pull ahead: annual wage growth is about 0.5 percentage points faster in cities, and industrialization in rural America is far more modest. Regions with a 60% agricultural employment share in 1880 saw a 25-point decline in agricultural employment in the baseline, but less than a 10-point decline in the no-catch-up model. These patterns are also reflected in the implied coefficients  $\beta_w$  and  $\beta_{s_A}$ , which become smaller in absolute value and, in the case of  $\beta_w$ , even reverse sign (see Table 2).

Figure 8 also highlights how productivity catch-up growth shaped the composition of aggregate structural change. The decline of aggregate agricultural employment by 17.5 percentage points exceeds the decline within each local labor market. The reason is that in the absence of productivity catch-up growth, faster wage growth in urban centers pulls people into cities and hence out of agriculture.<sup>22</sup> In terms of the decomposition shown in Figure 5, the “within” component only accounts for about 50% of aggregate structural change, with 50% of it now accounted for by the spatial reallocation of workers from high to low agricultural employment share labor markets. In contrast, the model with productivity catch-up growth matches the empirical finding that most of the agricultural decline was a within-region transformation, not a product of migration.

These findings summarize the central result of the paper. Aggregate structural change is

<sup>22</sup>Note that the correlation between population growth and agricultural employment,  $\beta_L$ , falls to  $-0.64$ , suggesting that population growth is strongly biased toward already-industrialized places-see Table 2.

inherently biased against rural locations. Moreover, this Bartik-like channel is not only a theoretical possibility but quantitatively important: differences in spatial exposure generate large differences in local wage growth. What made the US experience between 1880 and 1920 exceptional is that this urban bias was counteracted—and ultimately overturned—by strong productivity convergence. Faster productivity growth did not just help rural America to catch up to more industrialized areas, but rather prevented it from falling further behind.

## 6. ROBUSTNESS

We now assess the robustness of our central finding to alternative parameter values and modeling assumptions: throughout a set of alternative calibrations, catch-up productivity growth plays the central role for understanding the empirical patterns of spatial structural change. Appendix Section B.4 provides additional details.

**Alternative Parameter Values** We begin by exploring the sensitivity of our results to key structural parameters. As highlighted in Propositions 3 and 4, the elasticity of sectoral labor supply ( $\zeta$ ), the agricultural land share ( $\alpha$ ), and the elasticity of substitution in manufacturing ( $\sigma_M$ ) play important roles in shaping the transmission of sectoral productivity growth into wages and employment. We re-estimate the model with a higher labor supply elasticity ( $\zeta = 4$ ), a lower agricultural land share ( $\alpha = 0.2$ ), and a lower elasticity of substitution in manufacturing ( $\sigma_M = 4$ ), respectively.

We also probe our assumption of fixed land supply. As robustness, we consider a version of the model in which  $T_{rt} = \bar{T}_r R_{rt}^\varrho$ , so that agricultural land supply responds to the local rental rate. We set  $\varrho = 1.2$  based on estimates using historical data on land supply and rental rates from the Census of Agriculture, where we instrument for local demand shifts using inflows of international migrants (see Appendix B.4 for details).

Across all these parameter permutations, the rural bias in productivity growth remains the central force driving the patterns of spatial structural change. Table 3 summarizes the estimated strength of sectoral productivity growth in rural and urban locations, and its role for wage growth and industrialization. For comparison, we report the results of our baseline calibration in the first row.<sup>23</sup> Physical productivity growth in both sectors remains substantially higher in rural relative to urban locations (Columns 1 and 2). Industrialization remains overwhelmingly driven by manufacturing productivity growth across all locations (Columns 3 and 4). Similarly, wage growth in rural areas continues to be driven almost entirely by agricultural productivity catch-up, while wage growth in urban areas is accounted for by growth in the manufacturing sector (Columns 5 and 6). These results reinforce our finding that productivity catch-up growth in

---

<sup>23</sup> Appendix B.4 presents the full decomposition—mirroring the baseline results shown in Figure 7—for each of the alternative calibrations.

TABLE 3: ROBUSTNESS

Annual Growth (%) in...		Total Change in $s_{rAt}$ due to...		Annual Growth (%) in $w_{rt}$ due to...	
$A_{rt}$	$Z_{rt}$	$A_{rt}$	$Z_{rt}$	$A_{rt}$	$Z_{rt}$
<b>Panel A: Urban Commuting Zones</b>					
Baseline	0.57	1.54	0.04	-0.11	0.10
$\zeta = 4$	0.51	1.37	0.07	-0.14	0.07
$\alpha = 0.2$	1.07	1.53	0.07	-0.13	0.15
$\varrho = 1.2$	0.19	1.55	0.02	-0.12	0.05
$\sigma_M = 4$	0.59	0.97	0.04	-0.07	0.10
$\tau_{rjt} = \tau_{rj1880}$	0.68	1.60	0.05	-0.12	0.12
$n_{rt} = n_t$	0.63	1.55	0.05	-0.12	0.11
$\zeta = 5, \varrho = 1$	0.48	1.54	0.03	-0.11	0.07
<b>Panel B: Rural Commuting Zones</b>					
Baseline	2.97	3.86	0.22	-0.31	2.07
$\zeta = 4$	3.27	2.81	0.51	-0.45	1.92
$\alpha = 0.2$	2.25	4.12	0.18	-0.35	1.57
$\varrho = 1.2$	2.08	3.85	0.20	-0.32	1.78
$\sigma_M = 4$	2.96	3.72	0.22	-0.25	2.05
$\tau_{rjt} = \tau_{rj1880}$	3.09	4.01	0.23	-0.32	2.15
$n_{rt} = n_t$	2.95	3.90	0.22	-0.32	2.04
$\zeta = 5, \varrho = 1$	3.00	3.88	0.22	-0.31	2.08
					0.79

*Notes:* The table reports the contribution of agricultural and manufacturing productivity growth to spatial wage growth and industrialization across several model specifications. Columns 1 and 2 display estimated productivity growth in each sector for urban and rural areas, defined as the bottom and top quartiles of 1880 agricultural employment shares, respectively. Columns 3 and 4 present the corresponding contributions to changes in the agricultural employment share over the same period. Columns 5 and 6 show the portion of wage growth from 1880 to 1920 attributable to sectoral productivity growth in urban and rural regions.

agriculture was essential for rural wage growth, whereas productivity convergence toward the manufacturing frontier was the primary driver of rural industrialization.

While they do not change our quantitative conclusions, the direction in which these alternative parameter values influence our findings is instructive. A higher sectoral labor supply elasticity raises the sensitivity of wages and employment shares to sector-specific shocks by making it easier for workers to switch sectors. As a result, the size of the estimated productivity growth is lower to explain the same variation in wages and employment shares. Introducing an elastic land supply reduces the amount of agricultural productivity growth required to rationalize observed wage gains, since land expansion itself boosts labor productivity in agriculture by alleviating the decreasing returns to labor. Finally, lowering the elasticity of substitution in manufacturing reduces the estimated rate of productivity growth in that sector, because the decline in trade costs generates more economic growth if regional varieties are more complementary.

**Structural Modeling Assumptions** We next assess the role of three structural modeling assumptions: the evolution of trade costs, the specification of local population growth, and the presence of geographic frictions to technology adoption. Specifically, we re-estimate our model while holding trade costs fixed at their 1880 levels, removing the effects of the expanding railroad network and other advances in transportation technologies implicit in the annual trade cost matrices provided by [Donaldson and Hornbeck \(2016\)](#); we replace the differential exogenous population growth rates across locations with a spatially balanced growth rate of population  $n_{rt} = n_t$  equal to aggregate population growth; and we allow the technological frontier to vary across space by setting  $\varsigma = 5$  and  $\iota = 1$  in equation (24), so that access to the technologies of other regions declines with geographic distance.

The results of these exercises are contained in the last three rows of the respective panels of Table 3. Falling trade costs have a rural-biased wage growth effect, but their contribution pales compared to the importance of productivity catch-up. If trade costs were held constant at their 1880 level, our model would “need” slightly faster productivity growth in both sectors and locations. Similarly, the fact that birth rates and immigration inflows were unbalanced across space or that our baseline calibration abstracts from geographic frictions to technology adoption does not interfere much with our inference of rural productivity catch-up growth and the decomposition of local wage growth and industrialization.

## 7. CONCLUSION

Aggregate structural change reallocates resources out of agriculture and into other sectors. Because agricultural activity is unevenly distributed across space, some locations are more exposed to this reallocation than others. As a result, a natural conjecture is that structural change creates regional winners and losers with urban areas benefiting from industrial expansion and rural labor markets contracting as their sector of specialization loses relevance. Yet the historical US data reveal the opposite pattern—between 1880 and 1920, rural areas experienced both faster wage growth and faster industrialization.

This paper reconciles these facts by developing a theory of spatial structural change in which sectoral reallocation interacts with region-specific productivity dynamics, trade across regions, and internal migration. Our theory stresses that structural change itself indeed inherently favors urban regions, but that rural areas can benefit if they are positioned behind the technological frontier and productivity catch-up is strong. Quantitatively, we find that rural productivity catch-up in both agriculture and manufacturing was essential to explain the observed pattern of wage growth and industrialization. Agricultural productivity catch-up accounts for rural wage convergence, while productivity catch-up in manufacturing drove rural industrialization. Absent these forces,

the model predicts urban-biased wage growth, limited industrialization of rural labor markets, and a larger role for migration in the decline of agricultural employment.

Although our analysis focuses on the US during its first structural transformation, our framework offers a general approach to study structural change in a unified way across sectors and space. It opens several avenues for future work, two of which we think are particularly interesting. First, applying the model to the transition from manufacturing to services may shed light on the growing regional divergence in advanced economies (see, for example, [Austin, Glaeser, and Summers, 2018](#); [Chatterjee and Giannone, 2021](#)). Through the lens of our model, divergence suggests that in this second sectoral transition, catch-up growth is weak and the urban bias in aggregate structural change dominates. Second, applying the model to other countries currently undergoing their first structural transformation can help assess whether productivity catch-up is a general feature of spatial structural change or specific to US experience of the late 19th century.

## REFERENCES

ABRAMITZKY, R., L. BOUSTAN, K. ERIKSSON, J. FEIGENBAUM, AND S. PÉREZ (2021): "Automated Linking of Historical Data," *Journal of Economic Literature*, 59, 865–918.

ABRAMITZKY, R., L. BOUSTAN, K. ERIKSSON, M. RASHID, AND S. PÉREZ (2022): "Census Linking Project: 1850-1860 Crosswalk," Tech. rep.

ACEMOGLU, D., P. AGHION, AND F. ZILIBOTTI (2006): "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association*, 4, 37–74.

ALLEN, T. AND C. ARKOLAKIS (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129, 1085–1140.

ALLEN, T. AND D. DONALDSON (2020): "Persistence and Path Dependence in the Spatial Economy," Tech. rep., National Bureau of Economic Research.

AUSTIN, B., E. GLAESER, AND L. SUMMERS (2018): "Jobs for the Heartland: Place-Based Policies in 21st-Century America," *Brookings Papers on Economic Activity*, 2018, 151–255.

BOHR, C., M. MESTIERI, AND F. ROBERT-NICOUD (2024): "Heterothetic cobb douglas: Theory and applications," Tech. rep., Centre for Economic Policy Research.

BOPPART, T. (2014): "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and non-Gorman Preferences," *Econometrica*, 82, 2167–2196.

BOUSTAN, L., D. BUNTEN, AND O. HEAREY (2018): "Urbanization in American Economic History, 1800–2000," in *The Oxford Handbook of American Economic History Volume 2*, Oxford University Press.

CALIENDO, L., M. DVORKIN, AND F. PARRO (2019): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 87, 741–835.

CARTER, S. B., S. S. GARTNER, M. R. HAINES, A. L. OLMSTEAD, R. SUTCH, G. WRIGHT, ET AL. (2006): *Historical Statistics of the United States: Millennial Edition*, vol. 3, Cambridge: Cambridge University Press.

CASELLI, F. AND W. J. COLEMAN II (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 109, 584–616.

CHATTERJEE, S. AND E. GIANNONE (2021): "Unequal Global Convergence," Tech. rep., Working Paper.

COMIN, D., D. LASHKARI, AND M. MESTIERI (2021): "Structural Change with Long-run Income and Price Effects," *Econometrica*, 89, 311–374.

CRUZ, J.-L. AND E. ROSSI-HANSBERG (2024): "The Economic Geography of Global Warming," *Review of Economic Studies*, 91, 899–939.

DESMET, K., D. K. NAGY, AND E. ROSSI-HANSBERG (2018): "The Geography of Development," *Journal of Political Economy*, 126, 903–983.

DESMET, K. AND E. ROSSI-HANSBERG (2014): "Spatial Development," *American Economic Review*, 104, 1211–43.

DONALDSON, D. AND R. HORNBECK (2016): "Railroads and American Economic Growth: A "Market Access" Approach," *The Quarterly Journal of Economics*, 131, 799–858.

ECKERT, F., S. GANAPATI, AND C. WALSH (2022): "Urban-Biased Growth: A Macroeconomic Analysis," Tech. rep., National Bureau of Economic Research.

ECKERT, F., A. GVIRTZ, J. LIANG, AND M. PETERS (2020): "A Method to Construct Geographical Crosswalks with an Application to US Counties since 1790," Tech. rep., National Bureau of Economic Research.

ECKERT, F., J. JUNEAU, AND M. PETERS (2023): "Sprouting Cities: How Rural America Industrialized," in *AEA Papers and Proceedings*, vol. 113, 87–92.

FAN, T., M. PETERS, AND F. ZILIBOTTI (2023): "Growing Like India: The Unequal Effects of Service-led Growth," *Econometrica*, 91, 1457–1494.

GOLLIN, D., M. KIRCHBERGER, AND D. LAGAKOS (2021): "Do Urban Wage Premia Reflect Lower Amenities? Evidence from Africa," *Journal of Urban Economics*, 121, 103301.

HERRENDORF, B., R. ROGERSON, AND A. VALENTINYI (2013): "Two Perspectives on Preferences and Structural Transformation," *American Economic Review*, 103, 2752–89.

— (2014): "Growth and Structural Transformation," *Handbook of Economic Growth*, 2, 855–941.

HORNBECK, R. AND M. ROTEMBERG (2024): "Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies," *Journal of Political Economy*, 132, 3547–3602.

KLEINMAN, B., E. LIU, AND S. J. REDDING (2023): "Dynamic Spatial General Equilibrium," *Econometrica*, 91, 385–424.

KONGSAMUT, P., S. REBELO, AND D. XIE (2001): "Beyond Balanced Growth," *The Review of Economic Studies*, 68, 869–882.

LAGAKOS, D. AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-country Productivity Differences," *American Economic Review*, 103, 948–980.

MANSON, S., J. SCHROEDER, D. VAN RIPER, AND S. RUGGLES (2017): "IPUMS National Historical Geographic Information System: Version 12.0," *Minneapolis: University of Minnesota*. 2017.

MICHAELS, G., F. RAUCH, AND S. J. REDDING (2012): "Urbanization and Structural Transformation," *The Quarterly Journal of Economics*, 127, 535–586.

NAGY, D. K. (2023): "Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion," *Review of Economic Studies*, 90, 3238–3281.

NGAI, L. R. AND C. A. PISSARIDES (2007): "Structural Change in a Multisector Model of Growth," *American Economic Review*, 97, 429–443.

PETERS, M. (2022): "Market Size and Spatial Growth—Evidence From Germany's Post-War Population Expulsions," *Econometrica*, 90, 2357–2396.

PORZIO, T., F. ROSSI, AND G. SANTANGELO (2022): "The Human Side of Structural Transformation," *American Economic Review*, 112, 2774–2814.

REDDING, S. J. AND E. ROSSI-HANSBERG (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

RUGGLES, S., K. GENADEK, R. GOEKEN, J. GROVER, AND M. SOBEK (2015): "Integrated Public Use Microdata Series: Version 6.0," *Minneapolis: University of Minnesota*.

——— (2017): "Integrated Public Use Microdata Series: Version 7.0," *Minneapolis: University of Minnesota*.

SILVA, J. S. AND S. TENREYRO (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658.

TOLBERT, C. M. AND M. SIZER (1996): "US Commuting Zones and Labor Market Areas: A 1990 Update," Tech. rep.

VALENTINYI, Á. AND B. HERRENDORF (2008): "Measuring Factor Income Shares at the Sectoral Level," *Review of Economic Dynamics*, 11, 820–835.

YOUNG, A. (1998): "Growth Without Scale Effects," *Journal of Political Economy*, 106, 41–63.

# FOR ONLINE PUBLICATION:

## APPENDIX

### A. ADDITIONAL THEORETICAL RESULTS AND DERIVATIONS

The theory appendix contains additional proofs and derivations omitted in the body of the paper.

#### A.1 Results in Sections 2.2 and 2.3

**The PIGL Demand Function** Consider the indirect utility function given in equation (7). Roy's Identity implies that sectoral expenditure shares are given by:

$$(A.1) \quad \vartheta_s \equiv \vartheta_s(y, P_{rA}, P_{rM}) = -\frac{\frac{\partial V(y, P_{rA}, P_{rM})}{\partial P_{rs}} P_{rs}}{\frac{\partial V(y, P_{rA}, P_{rM})}{\partial y} y}.$$

Using equation (7), we have

$$(A.2) \quad \vartheta_A = \phi + \nu \left( \frac{y}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^{-\eta} \quad \text{and} \quad \vartheta_M = (1 - \phi) - \nu \left( \frac{e}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^{-\eta}.$$

**Sector Labor Supply** We denote total payments per efficiency unit of labor in region  $r$  by  $w_{rM}$  and  $\bar{\chi}_{rt}$ . We denote the efficiency units individual  $i$  can supply to each sector by  $z_A^i$  and  $z_M^i$ . Individuals draw their efficiency units from a sector-specific Fréchet distribution,  $P(z_s^i \leq z) = F_s(z) = e^{-z^{-\zeta}}$ . Workers choose their sector so as to maximize their labor income. A standard set of arguments implies the following analytical expressions for the key objects of our theory.

1. The manufacturing employment share is given by

$$(A.3) \quad s_{rM} = \left( \frac{w_{rM}}{\bar{w}_r} \right)^\zeta \quad \text{where} \quad \bar{w}_r = \left( w_{rM}^\zeta + \bar{\chi}_{rt}^\zeta \right)^{1/\zeta},$$

and  $s_{rA} = 1 - s_{rM} = (\bar{\chi}_{rt}/\bar{w}_r)^\zeta$ .

2. The aggregate amounts of sectoral human capital are

$$H_{rs} = \Gamma_\zeta L_r \left( \frac{w_{rs}}{\bar{w}_r} \right)^{\zeta-1} = \Gamma_\zeta L_r s_{rs}^{\frac{\zeta-1}{\zeta}},$$

where  $\Gamma_x \equiv \Gamma(1 - 1/x)$  and  $\Gamma$  denotes the Gamma function.

3. Total sectoral earnings are

$$w_{rM} H_{rM} = w_{rM} \Gamma_\zeta L_r \left( \frac{w_{rM}}{\bar{w}_r} \right)^{\zeta-1} = \bar{w}_r \Gamma_\zeta L_r \left( \frac{w_{rM}}{\bar{w}_r} \right)^\zeta = \bar{w}_r \Gamma_\zeta L_r s_{rM}.$$

and  $\bar{\chi}_{rt} H_{rA} = \bar{w}_r \Gamma_\zeta L_{rA}$ .

4. The distribution of realized labor income,  $y_r^i$ , inherits the Fréchet distribution of the underlying efficiency units of labor and is given by

$$(A.4) \quad F_r(y) \equiv P(y_r^i \leq y) = e^{-(w_{rM}^\zeta + \bar{\chi}_{rt}^\zeta)y^{-\zeta}} = e^{-(y/\bar{w}_r)^{-\zeta}}.$$

Hence, a worker's expected income in region  $r$  prior to moving there is given by  $E[y_r^i] = \Gamma_\zeta \bar{w}_r$ . Due to the law of large numbers this also corresponds to the ex-post average income in location  $r$ , so that,  $Y_r = \bar{w}_r \Gamma_\zeta L_r$ .

## A.2 Results in Section 2.4

In this section we include detailed derivations for the results reported in Section 2.4, including Propositions 1 and 2.

### A.2.1 Proof of Proposition 1 (PIGL Aggregation)

**Aggregate Demand** Let  $F_r(y)$  be the distribution of income derived in equation (A.4). Integrating over the sectoral expenditure shares of individual workers in region  $r$  in equation (A.2) yields an expression for a region's aggregate expenditure share:

$$\vartheta_{rs} \equiv \frac{\int \vartheta_A(y, P_{rA}, P_{rM}) y dF_r(y)}{\int y dF_r(y)} = \phi + \nu \left( \frac{1}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^{-\eta} \frac{\int y^{1-\eta} dF_r(y)}{\int y dF_r(y)}.$$

Given that  $F_r(y) = e^{-(y/\bar{w}_r)^{-\zeta}}$ , we have that  $P(y^{1-\eta} < m) = e^{-\left(\frac{m}{\bar{w}_r}\right)^{-\frac{\zeta}{1-\eta}}}$ . Hence,

$$\frac{\int y^{1-\eta} dF_r(y)}{\int y dF_r(y)} = \frac{\Gamma_{\frac{\zeta}{1-\eta}} \bar{w}_r^{1-\eta}}{\Gamma_\zeta \bar{w}_r} = \frac{\Gamma_{\frac{\zeta}{1-\eta}}}{\Gamma_\zeta} \bar{w}_r^{-\eta},$$

so that

$$(A.5) \quad \vartheta_{rA} = \phi + \nu \frac{\Gamma_{\frac{\zeta}{1-\eta}}}{\Gamma_\zeta} \left( \frac{\bar{w}_r}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^{-\eta} = \phi + \nu^{RC} \left( \frac{\bar{w}_r}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^{-\eta},$$

where we defined the composite parameter  $\nu^{RC} \equiv \nu^{\frac{\Gamma_\zeta}{1-\eta}}$ .

**Indirect Utility** Using the indirect utility function in equation (7), we derive the following expression for the expected utility in region  $r$ :

$$E[V(y, P_{rA}, P_{rM})] = \frac{1}{\eta} \left( \frac{1}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^\eta \int y^\eta dF_r(y) - \nu \ln \left( \frac{P_{rA}}{P_{rM}} \right).$$

Workers effectively draw their income from the Fréchet distribution in equation (A.4) upon arriving in their region of choice. As a result,  $y^\eta$  itself is drawn from a Fréchet

distribution with a shape parameter  $\zeta/\eta$  and scale  $\bar{w}_r^\eta$ :  $P(y^\eta < m) = e^{-\left(\frac{m}{\bar{w}_r^\eta}\right)^{-\frac{\zeta}{\eta}}}$ . By implication,  $\int y^\eta dF_r(y) = \Gamma\left(1 - \frac{1}{\zeta/\eta}\right) \bar{w}_r^\eta = \Gamma_{\frac{\zeta}{\eta}} \bar{w}_r^\eta$ , so that

$$(A.6) \quad E[V(y, P_{rA}, P_{rM})] = \frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \left( \frac{\bar{w}_{rt}}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^\eta - \nu \ln \left( \frac{P_{rA}}{P_{rM}} \right).$$

This is the expression in equation (13).

### A.2.2 Proof of Proposition 2 (Catch-Up Growth)

**Agricultural Sector** Consider first the agricultural sector. We consider the general case where the overall land supply elastic, i.e.  $T_{rt} = \bar{T}_r R_{rt}^\varrho$ , where  $R_{rt}$  denotes the equilibrium rental rate of agricultural land and  $\varrho$  is the supply elasticity: if  $\varrho = 0$ , agricultural land is in fixed supply; if  $\varrho > 0$  agricultural land can be expanded through cultivation. Our baseline model corresponds to the case of  $\varrho = 0$ . The case of  $\varrho > 0$  is explored in Section 6.

Agricultural profits are given by (see equation (2))

$$\chi_{rt} (z_A^i) \equiv \max_{(t_P, T)} \left\{ p_{rAt} \mu_A A_{rt-1}^{1-\lambda_A} (\mathcal{F}_{rt-1}^A)^{\lambda_A} (1 - t_P)^\gamma (z_A^i t_P)^{1-\alpha} T^\alpha - R_{rt} T \right\}$$

The optimal allocation of farmer's time is thus given by  $t_P = \frac{1-\alpha}{\gamma+1-\alpha}$ . This implies that

$$(A.7) \quad A_{rt} = \mu_A A_{rt-1}^{1-\lambda_A} (\mathcal{F}_{rt-1})^{\lambda_A} \left( \frac{\gamma}{\gamma+1-\alpha} \right)^\gamma,$$

which is the expression in equation (16).

**Manufacturing Sector** Given the CES demand system, firms in region  $r$  that chose a technology  $Z$  set a constant markup and hence a price  $p_{rMt} = \frac{\sigma_M}{\sigma_M-1} \frac{w_{rMt}}{Z_t}$ . This implies that profits are given by

$$(A.8) \quad \pi_{rt}(Z) = \frac{1}{\sigma_M} \left( \frac{\sigma_M}{\sigma_M-1} \right)^{1-\sigma_M} Z^{\sigma_M-1} w_{rMt}^{1-\sigma_M} \mathcal{D}_{rMt}$$

where  $\mathcal{D}_{rMt} = \sum_j \tau_{rjt}^{1-\sigma_M} P_{jMt}^{\sigma_M-1} \vartheta_{jMt} \Gamma_\zeta L_{jt} \bar{w}_{jt}$  denotes manufacturing market access (see equation (15)). The optimal technology choice is thus given by

$$Z_{rt} = \arg \max_Z \left\{ \pi_{rt}(Z) - \left( \frac{Z}{\psi_{rt}} \right)^\theta w_{rMt} \right\} = \left( \frac{1}{\theta} \left( \frac{\sigma_M-1}{\sigma_M} \right)^{\sigma_M} \frac{\psi_{rt}^\theta \mathcal{D}_{rMt}}{w_{rMt}^{\sigma_M}} \right)^{\frac{1}{\theta-(\sigma-1)}},$$

which is equation (18) in the main text. Note that we can also express the first order condition of the optimal technology choice as

$$(A.9) \quad \frac{\sigma_M - 1}{\theta} \pi_{rt}(Z_{rt}) = \left( \frac{Z_{rt}}{\psi_{rt}} \right)^\theta w_{rMt}.$$

Similarly,

$$\max_Z \left\{ \pi_{rt}(Z) - \left( \frac{Z}{\psi_{rt}} \right)^\theta w_{rMt} \right\} = \pi_{rt}(Z_{rt}) \left( 1 - \frac{\sigma_M - 1}{\theta} \right),$$

that is the net profits after paying workers for the adoption of the production technology are proportional to the gross profits  $\pi_{rt}(Z_{rt})$ . Finally, free entry requires that profits net of entry costs are zero, that is  $\pi_{rt}(Z_{rt}) \left( 1 - \frac{\sigma_M - 1}{\theta} \right) = f_E w_{rMt}$ . Combining this with equation (A.9) yields  $Z_{rt} = \left( \frac{\sigma_M - 1}{\theta - (\sigma - 1)} f_E \right)^{1/\theta} \psi_{rt}$ . Using  $\psi_{rt} = \mu_M \mathcal{F}_{rZt}^{\lambda_M} Z_{rt-1}^{1-\lambda_M}$ , we arrive at

$$Z_{rt} = \left( \frac{\sigma_M - 1}{\theta - (\sigma - 1)} f_E \right)^{1/\theta} \mu_M \mathcal{F}_{rZt}^{\lambda_M} Z_{rt-1}^{1-\lambda_M}.$$

This is expression equation (17) in Proposition 2.

### A.2.3 The Equilibrium System

In this section we derive the characterization of the equilibrium system. We first derive total revenue and the return per agricultural efficiency unit,  $\bar{\chi}_{rt}$ , in the agricultural sector. The optimal amount of land a farmer with productivity  $z_A^i$  hires is given by

$$T_r(z_A^i) = \left( \frac{\alpha}{R_{rt}} p_{rAt} A_{rt} \right)^{\frac{1}{1-\alpha}} \frac{1-\alpha}{\gamma+1-\alpha} z_A^i.$$

Total profits after paying for agricultural land are given by

$$(A.10) \quad \chi_{rt}(z_A^i) = (1-\alpha) \left( \frac{1-\alpha}{\gamma+1-\alpha} \right) (p_{rAt} A_{rt})^{\frac{1}{1-\alpha}} \left( \frac{\alpha}{R_{rt}} \right)^{\frac{\alpha}{1-\alpha}} z_A^i.$$

Land market clearing implies that

$$(A.11) \quad T_{rt} = \bar{T}_r R_{rt}^o = \int_{z \in A_g} T(z) L_{rt} dF_{rt}^A(z) = \left( \frac{\alpha}{R_{rt}} \right)^{\frac{1}{1-\alpha}} (p_{rAt} A_{rt})^{\frac{1}{1-\alpha}} \frac{1-\alpha}{\gamma+1-\alpha} H_{rAt}.$$

The equilibrium rental rate for land is therefore given by

$$R_{rt} = \alpha^{\frac{1}{\varrho + \frac{1}{1-\alpha}}} (p_{rAt} A_{rt})^{\frac{1}{\varrho + \frac{1}{1-\alpha}}} \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{\frac{1}{\varrho + \frac{1}{1-\alpha}}} \left( \frac{H_{rAt}}{\bar{T}_r} \right)^{\frac{1}{\varrho + \frac{1}{1-\alpha}}}.$$

Substituting this equation into equation (A.10) yields

$$\begin{aligned}\chi_{rt}(z_A^i) &= (1-\alpha) \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{\frac{(\varrho+1)(1-\alpha)}{\varrho(1-\alpha)+1}} (p_{rAt} A_{rt})^{\frac{\varrho+1}{\varrho(1-\alpha)+1}} \alpha^{\frac{\varrho\alpha}{\varrho(1-\alpha)+1}} \left( \frac{\bar{T}_r}{H_{rAt}} \right)^{\frac{\alpha}{\varrho(1-\alpha)+1}} z_A^i \\ &\equiv (1-\alpha) \bar{\chi}_{rt} z_A^i,\end{aligned}$$

where

$$(A.12) \quad \bar{\chi}_{rt} = \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{\frac{(\varrho+1)(1-\alpha)}{\varrho(1-\alpha)+1}} (p_{rAt} A_{rt})^{\frac{\varrho+1}{\varrho(1-\alpha)+1}} \alpha^{\frac{\varrho\alpha}{\varrho(1-\alpha)+1}} \left( \frac{\bar{T}_r}{H_{rAt}} \right)^{\frac{\alpha}{\varrho(1-\alpha)+1}}.$$

Note also that equation (A.11) implies that  $\bar{\chi}_{rt} = \frac{1}{\alpha} \frac{R_{rt} T_r}{H_{rAt}}$ .

To solve for the trade equilibrium in the manufacturing sector, we need to solve for the employment allocations. Letting  $\mathcal{R}_{rMt}$  denote total manufacturing revenue in region  $r$  at time  $t$ , monopolistic competition requires that production workers receive a share  $\frac{\sigma_M - 1}{\sigma_M}$  of total revenue and aggregate profits account for the residual share  $\frac{1}{\sigma_M}$ , that is

$$w_{rt} H_{Prt} = \frac{\sigma_M - 1}{\sigma_M} \mathcal{R}_{rMt} \quad \text{and} \quad N_{rt} \pi_{rt}(Z_{rt}) = \frac{1}{\sigma_M} \mathcal{R}_{rMt}.$$

Hence,

$$H_{Prt} = (\sigma - 1) \frac{N_{rt} \pi_{rt}(Z_{rt})}{w_{rt}} = \frac{(\sigma - 1) \theta}{\theta - (\sigma - 1)} N_{rt} f_E = \frac{(\sigma - 1) \theta}{\theta - (\sigma - 1)} H_{Ert},$$

where the second equality uses the free entry condition, and the third equality exploits that  $N_{rt} = \frac{1}{f_E} H_{Ert}$  as each firm requires  $f_E$  workers. This implies that the allocation of manufacturing employment across activities is constant and given by

$$(A.13) \quad \frac{H_{Prt}}{H_{Mrt}} = \frac{\sigma_M - 1}{\sigma_M} \quad \text{and} \quad \frac{H_{Zrt}}{H_{Mrt}} = \frac{1}{\theta} \frac{\sigma_M - 1}{\sigma_M} \quad \text{and} \quad \frac{H_{Ert}}{H_{Mrt}} = \frac{\theta - (\sigma - 1)}{\sigma \theta}.$$

Armed with equations (A.12) and (A.13) we can now state the final equilibrium conditions. The equilibrium is characterized by the following system of equations:

1. **Spatial labor supply:** The spatial labor supply function is given in the law of motion for the local population in equation (A.14) given by

$$(A.14) \quad L_{jt} = \sum_r m_{rjt} n_{rt-1} L_{rt-1}.$$

Together with the expression for expected utility  $\mathcal{V}_{rt}$  given in equation (13), equation (A.14) determines the spatial supply function as a function of local wages  $\bar{w}_{rt}$  and local prices  $\{P_{rAt}, P_{rMt}\}_r$ .

2. **Market clearing in agriculture:** The supply of agricultural efficiency units is given by  $H_{rAt} = \Gamma_\zeta L_{rt} (\bar{\chi}_{rt} / \bar{w}_{rt})^{\zeta-1}$  (see equation 10). Using the expression for  $\bar{\chi}_{rt}$

in equation (A.12), we can write  $\bar{\chi}_{rt}$  as

$$(A.15) \quad \bar{\chi}_{rt}^{\zeta-1+\frac{\varrho(1-\alpha)+1}{\alpha}} = \bar{w}_{rt}^{\zeta-1} \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{\frac{(\varrho+1)(1-\alpha)}{\alpha}} (p_{rAt} A_{rt})^{\frac{\varrho+1}{\alpha}} \alpha^\varrho \left( \frac{\bar{T}_r}{\Gamma_\zeta L_{rt}} \right).$$

Given agricultural prices,  $p_{rAt}$ , and the fact that aggregate income,  $p_{rAt} Y_{rAt}$ , is equal to agricultural earnings (including payments for land),  $\bar{\chi}_{rt} H_{rAt}$ , the CES demand system implies that

$$(A.16) \quad \bar{\chi}_{rt} H_{rAt} = \sum_j \frac{(\tau_{rj} p_{rAt})^{1-\sigma_A}}{\sum_{k=1}^R (\tau_{kj} p_{kAt})^{1-\sigma_A}} \vartheta_{jAt} \Gamma_\zeta L_j \bar{w}_j.$$

Given equation (A.15) and the expression for  $H_{rAt}$ , we can compute  $\bar{\chi}_{rt}$  and  $H_{rAt}$  as a function of  $P_{rAt}$ ,  $\bar{w}_{rt}$ , and  $L_{rt}$ .

**Market clearing in manufacturing:** For manufacturing products, sales of firm  $\omega$  located in region  $r$  are given by

$$p_{rMt}(\omega) y_{rt}(\omega) = \sum_j \left( \frac{\tau_{rjM} p_{rMt}(\omega)}{P_{jMt}} \right)^{1-\sigma_M} \vartheta_{jMt} \Gamma_\zeta L_{jt} \bar{w}_{jt}.$$

The mass of manufacturing firms that enter a location,  $N_{rt}$ , is also equal to the number of varieties produced in region  $r$ . Aggregating over the measure of varieties,  $N_{rt}$  yields:

$$\mathcal{R}_{rMt} = w_{rMt} H_{rMt} = N_{rt} \left( \frac{\sigma_M}{\sigma_M - 1} \right)^{1-\sigma_M} w_{rMt}^{1-\sigma_M} Z_{rt}^{\sigma_M-1} \sum_j \left( \frac{\tau_{rjM}}{P_{jMt}} \right)^{1-\sigma_M} \vartheta_{jMt} \Gamma_\zeta L_{jt} \bar{w}_{jt}.$$

Using that  $N_{rt} = \frac{1}{f_E} H_{Ert} = \frac{1}{f_E} \frac{\theta - (\sigma - 1)}{(\sigma - 1)\theta} H_P$ , we can solve for skill prices in the manufacturing sector as

$$(A.17) \quad w_{rMt}^{\sigma_M} = \frac{1}{f_E \sigma_M} \frac{\theta - (\sigma - 1)}{\theta} \left( \frac{\sigma_M - 1}{\sigma_M} \right)^{\sigma_M-1} Z_{rt}^{\sigma_M-1} \sum_j \tau_{rjM}^{1-\sigma_M} P_{jMt}^{\sigma_M-1} \vartheta_{jMt} \Gamma_\zeta L_{jt} \bar{w}_{jt},$$

where the manufacturing spending share  $\vartheta_{rMt}$  is given in equation (A.5).

These equations fully determine the equilibrium. In particular, upon substituting for  $\vartheta_{rMt}$  and  $\mathcal{V}_{rt}$  (which are fully determined from  $\bar{w}_{rt}$  and  $\{P_{rMt}, P_{rAt}\}$  - see Proposition 1), equations (A.14), (A.16), and (A.17) are  $3 \times R$  equations in the  $3 \times R$  unknowns  $\{P_{rAt}, P_{rMt}, L_{rt}\}$ .

### A.3 Results in Section 3

In this section we derive all results stated in Section 3.

### A.3.1 Market Access and Revenue Productivity

We follow [Donaldson and Hornbeck \(2016\)](#) to define market access for manufacturing firms in region  $r$  as  $\mathcal{D}_{rMt} = \sum_j \tau_{rjt}^{1-\sigma_M} P_{jMt} \vartheta_{jMt} \Gamma_\zeta L_{jt} \bar{w}_{jt}$ . Equation (A.17) then implies that wages in the manufacturing sector are given by

$$(A.18) \quad w_{rMt} = \left( \frac{1}{f_E \sigma_M} \frac{\theta - (\sigma - 1)}{\theta} \left( \frac{\sigma_M - 1}{\sigma_M} \right)^{\sigma_M - 1} \right)^{1/\sigma_M} Z_{rMt}^{\frac{\sigma_M - 1}{\sigma_M}} \mathcal{D}_{rMt}^{\frac{1}{\sigma_M}}.$$

Defining  $\tilde{f}_E = f_E \frac{\sigma \theta}{\theta - (\sigma - 1)} \left( \frac{\sigma_M}{\sigma_M - 1} \right)^{\sigma_M - 1}$ , manufacturing revenue is given by

$$(A.19) \quad \mathcal{R}_{rMt} = w_{rMt} H_{rMt} = \left( \frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} Z_{rt}^{(\sigma-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} H_{rMt}.$$

Similarly, market access of agricultural producers in region  $r$  is given by

$$\mathcal{D}_{rAt} = \sum_j \tau_{rjt}^{1-\sigma_A} P_{jAt} (1 - \vartheta_{jAt}) \Gamma_\zeta L_{jt} \bar{w}_{jt}.$$

Equation (A.16) then implies that  $\bar{\chi}_{rt} H_{rAt} = p_{rAt}^{1-\sigma_A} \mathcal{D}_{rAt}$ . Substituting  $p_{rAt} = \left( \frac{\bar{\chi}_{rt} H_{rAt}}{\mathcal{D}_{rAt}} \right)^{\frac{1}{1-\sigma_A}}$  into the expression for  $\bar{\chi}_{rt}$  in equation (A.12) yields

$$(A.20) \quad \bar{\chi}_{rt} = \left( \frac{1 - \alpha}{\gamma + 1 - \alpha} \right)^{(\varrho+1)(1-\alpha)\delta} (\mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} H_{rAt}^{-\frac{1}{\sigma_A-1}} A_{rt})^{(\varrho+1)\delta} \alpha^{\varrho\alpha\delta} \left( \frac{\bar{T}_r}{H_{rAt}} \right)^{\alpha\delta},$$

where

$$(A.21) \quad \delta \equiv \frac{(\sigma_A - 1)}{(\varrho(1 - \alpha) + 1)(\sigma_A - 1) + (\varrho + 1)}.$$

Agricultural revenue is thus given by

$$\mathcal{R}_{rAt} = \bar{\chi}_{rt} H_{rAt} = \left( \frac{1 - \alpha}{\gamma + 1 - \alpha} \right)^{(\varrho+1)(1-\alpha)\delta} \left( \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt} \right)^{(\varrho+1)\delta} \alpha^{\varrho\alpha\delta} \bar{T}_r^{\alpha\delta} H_{rAt}^{1-\alpha\delta - \frac{\varrho+1}{\sigma_A-1}\delta}.$$

For  $\varrho = 0$ , equation (A.21) implies that  $\delta = \frac{\sigma_A - 1}{\sigma_A}$  so that

$$\mathcal{R}_{rAt} = \left( \frac{1 - \alpha}{\gamma + 1 - \alpha} \right)^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}} \left( \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt} \right)^{\frac{\sigma_A-1}{\sigma_A}} \bar{T}_r^{\alpha\frac{\sigma_A-1}{\sigma_A}} H_{rAt}^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}}.$$

### A.3.2 Proof of Proposition 3

Equation (A.18) implies that

$$(A.22) \quad w_{rMt} = \left( \frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} Z_{rt}^{(\sigma-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} \equiv \mathcal{Z}_{rt}.$$

Similarly, equation (A.20) implies that

$$\bar{\chi}_{rt} = \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{(\varrho+1)(1-\alpha)\delta} (\mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt})^{(\varrho+1)\delta} \alpha^{\varrho\alpha\delta} \bar{T}_r^{\alpha\delta} H_{rAt}^{-\left(\alpha+\frac{\varrho+1}{\sigma_A-1}\right)\delta}.$$

$H_{rAt}$  is given by  $H_{rAt} = \Gamma_\zeta L_r (\bar{\chi}_{rt}/\bar{w}_r)^{\zeta-1}$ . Hence,

$$(A.23) \quad \bar{\chi}_{rt} = \mathcal{A}_{rt} \left( \frac{\bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-(\zeta-1)\left(\frac{\varrho+1}{\sigma_A-1}+\alpha\right)\delta},$$

where

$$(A.24) \quad \mathcal{A}_{rt} \equiv \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{(\varrho+1)(1-\alpha)\delta} (\mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt})^{(\varrho+1)\delta} \alpha^{\varrho\alpha\delta} (\bar{T}_r)^{\alpha\delta} (\Gamma_\zeta L_{rt})^{-\left(\frac{\varrho+1}{\sigma_A-1}+\alpha\right)\delta}.$$

This implies that

$$(A.25) \quad \frac{\bar{\chi}_{rt}}{\bar{w}_{rt}} = \left( \frac{\mathcal{A}_{rt}}{\bar{w}_{rt}} \right)^{\frac{1}{\kappa(\varrho)}},$$

where

$$(A.26) \quad \kappa(\varrho) = 1 + (\zeta - 1) \left( \left( \frac{\varrho+1}{\sigma_A-1} \right) + \alpha \right) \delta,$$

and  $\delta$  is defined in equation (A.21). Regional wages  $\bar{w}_{rt}$  are thus implicitly defined by

$$1 = \left( \frac{w_{rMt}}{\bar{w}_{rt}} \right)^\zeta + \left( \frac{\bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^\zeta = \left( \frac{\mathcal{Z}_{rt}}{\bar{w}_{rt}} \right)^\zeta + \left( \frac{\mathcal{A}_{rt}}{\bar{w}_{rt}} \right)^{\frac{\zeta}{\kappa(\varrho)}}.$$

For the case of  $\varrho = 0$ ,  $\delta = \frac{\sigma_A-1}{\sigma_A}$ , so that

$$(A.27) \quad \mathcal{A}_{rt} = \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}} (\mathcal{D}_{rAt}^{\frac{1}{\sigma_A}} A_{rt}^{\frac{\sigma_A-1}{\sigma_A}}) \left( \frac{\bar{T}_r}{\Gamma_\zeta L_{rt}} \right)^{\alpha\frac{\sigma_A-1}{\sigma_A}} (\Gamma_\zeta L_{rt})^{-\frac{1}{\sigma_A}}$$

and  $\bar{w}_{rt}$  is given by

$$1 = \left( \frac{w_{rMt}}{\bar{w}_{rt}} \right)^\zeta + \left( \frac{\bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^\zeta = \left( \frac{\mathcal{Z}_{rt}}{\bar{w}_{rt}} \right)^\zeta + \left( \frac{\mathcal{A}_{rt}}{\bar{w}_{rt}} \right)^{\frac{\zeta}{\kappa}},$$

where  $\kappa = \kappa(0) = \frac{1}{\zeta - (\zeta-1)(1-\alpha)\frac{\sigma_A-1}{\sigma_A}}$ .

To derive the second result in Proposition 3, recall that  $\bar{w}_{rt} = \bar{\chi}_{rt} s_{rAt}^{-1/\zeta}$ . Hence,

$$\bar{\chi}_{rt} = \mathcal{A}_{rt} \left( \frac{\bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-(\zeta-1)\left(\left(\frac{\varrho+1}{\sigma_A-1}\right)+\alpha\right)\delta} = \mathcal{A}_{rt} s_{rAt}^{-\left(\frac{\zeta-1}{\zeta}\right)\left(\left(\frac{\varrho+1}{\sigma_A-1}\right)+\alpha\right)\delta}.$$

Also note that  $\frac{s_{rAt}}{1-s_{rAt}} = \left(\frac{\bar{\chi}_{rt}}{w_{rMt}}\right)^\zeta = \left(\frac{\bar{\chi}_{rt}}{\mathcal{Z}_{rt}}\right)^\zeta$ . Substituting for  $\bar{\chi}_{rt}$  yields  $\frac{s_{rAt}^{\kappa(\varrho)}}{1-s_{rAt}} = \left(\frac{\mathcal{A}_{rt}}{\mathcal{Z}_{rt}}\right)^\zeta$ . For  $\varrho = 0$ , we have

$$\kappa(0) = 1 + (\zeta - 1) \left( \left( \frac{\varrho + 1}{\sigma_A - 1} \right) + \alpha \right) \delta = \zeta - (\zeta - 1)(1 - \alpha) \frac{\sigma_A - 1}{\sigma_A} = \kappa$$

so that  $\frac{s_{rAt}^\kappa}{1-s_{rAt}} = \left(\frac{\mathcal{A}_{rt}}{\mathcal{Z}_{rt}}\right)^\zeta$ .

### A.3.3 Proof of Proposition 4

**The wage exposure elasticity  $\phi(s_{rA})$**  The first result in Proposition 3, directly implies that  $d \ln w_{rMt} = d \ln \mathcal{Z}_{rMt}$ . Equation (A.25) implies that

$$\bar{\chi}_{rt} = \bar{w}_{rt}^{\frac{\kappa(\varrho)-1}{\kappa(\varrho)}} \mathcal{A}_{rt}^{\frac{1}{\kappa(\varrho)}}.$$

Note that  $\kappa(0) = \kappa$  as given in the main text.

Then we can take the total derivative for the expression for the average income in a location in equation (A.3) to obtain:

$$\begin{aligned} d \ln \bar{w}_{rt} &= s_{rAt} d \ln \bar{\chi}_{rt} + (1 - s_{rAt}) d \ln w_{rMt} \\ &= \frac{s_{rAt} \frac{1}{\kappa(\varrho)}}{(1 - s_{rAt}) + s_{rAt} \frac{1}{\kappa(\varrho)}} d \ln \mathcal{A}_{rt} + \frac{1 - s_{rAt}}{(1 - s_{rAt}) + s_{rAt} \frac{1}{\kappa(\varrho)}} d \ln \mathcal{Z}_{rt} \\ &\equiv \phi(s_{rAt}) d \ln \mathcal{Z}_{rt} + (1 - \phi(s_{rAt})) d \ln \mathcal{A}_{rt}, \end{aligned}$$

where

$$\phi(s_{rAt}) = \frac{1 - s_{rAt}}{(1 - s_{rAt}) + s_{rAt} \frac{1}{\kappa(\varrho)}}$$

**The industrialization elasticity  $\psi(s_{rA})$**  To derive the change in  $s_{rAt}$ , we take the total derivative of equation (A.3) and combine it with the expression for  $d \ln \bar{w}_{rt}$  above to obtain:

$$\begin{aligned} d \ln s_{rAt} &= \zeta (d \ln \bar{\chi}_{rt} - d \ln \bar{w}_{rt}) = -\frac{\zeta}{\kappa(\varrho)} (d \ln \bar{w}_{rt} - d \ln \mathcal{A}_{rt}) \\ &= -\frac{\zeta}{\kappa(\varrho)} \phi(s_{rAt}) (d \ln \mathcal{Z}_{rt} - d \ln \mathcal{A}_{rt}) \end{aligned}$$

Finally, using that  $ds_{rAt} = s_{rAt} d \ln s_{rAt}$ , yields  $ds_{rAt} = \psi(s_{rAt}) (d \ln \mathcal{A}_{rt} - d \ln \mathcal{Z}_{rt})$ , where  $\psi(s_{rAt}) = -\frac{\zeta s_{rAt}}{\kappa(\varrho)} \phi(s_{rAt})$ , which is the expression in Proposition 4.

## B. ADDITIONAL DATA DETAILS AND EXHIBITS

The material presented in this section complements the quantification section of the main paper. It contains a detailed description of the data, additional figures and tables, and details of our estimation procedure.

## B.1 Description of Data Sources and Data Construction

The spatial unit of observation throughout the paper is the commuting zone, as defined by [Tolbert and Sizer \(1996\)](#). Because county boundaries changed substantially over time, we use the crosswalk developed by [Eckert et al. \(2020\)](#) to map historical counties to time-invariant commuting zones.

We restrict the sample to states that were part of the Union by 1870, excluding later entrants to ensure consistent and complete Census coverage.<sup>24</sup> Accordingly, we drop Colorado (admitted 1876), 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). The map in Figure [A.1](#) shows agricultural employment shares across commuting zones in 1880 for our final sample.

### B.1.1 Full Count Decennial Census, 1860-1920

**Source and Description** We obtained the full count decennial census micro-data files for the years 1860, 1880, 1900, and 1920 from the IPUMS database (see [Ruggles, Genadek, Goeken, Grover, and Sobek, 2017](#)).

We selected the following variables: state, county, age, years since immigration (“yrimmig”), state of birth (if applicable), and industry of employment using 1950 Census codes (ind1950). We use the county and state identifiers included in the data to assign each observation to a commuting zone. The full-count Census is the only data set from which we use the 1860 cross-section, since our calibration strategy for regional amenities in 1880 requires the 1860 employment counts by region.

**Sample Selection, Processing, and Use** In the data, we define different groups of observations used in various parts of the paper. We define “workers” as observations with an industry identifier and age between 20 and 60 years. We define “agricultural workers” as workers who work in Agriculture, Forestry, and Fishing, corresponding to ind1950 codes 105, 116, and 126. For each commuting zone, dividing the total agricultural worker count by the total number of workers yields the agricultural employment share we use throughout the paper. The map in Figure [A.1](#) depicts a map of the agricultural employment share in 1880.

“Immigrant workers” are workers who immigrated within the last 20 years. “Old workers” are workers between the ages of 40 and 60. “Young workers” are workers between the ages of 20 and 40. We use these groups of observations to inform the location- and decade-specific labor force growth rate  $n_{rt}$ .

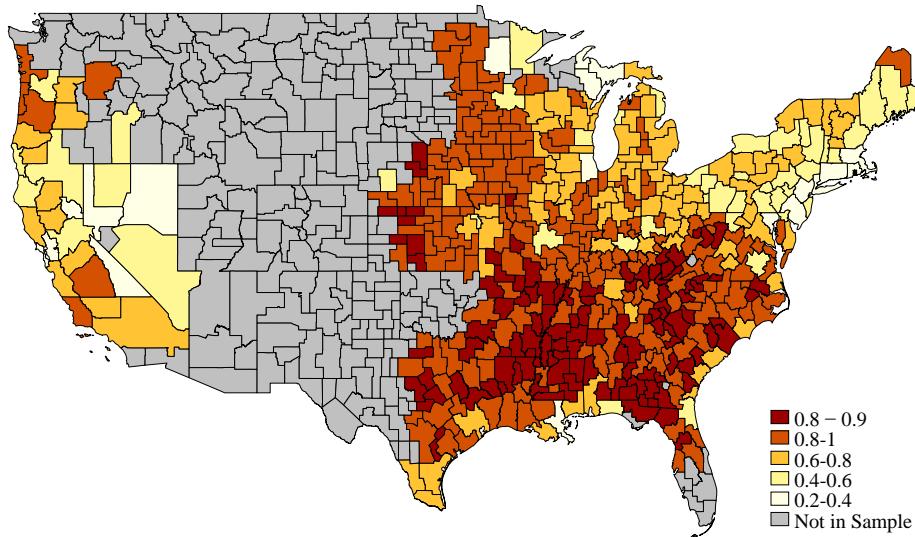
For each state, we also compute the number of workers born in any state. We use the resulting “lifetime state-to-state migration matrix” to estimate the elasticity of migration flows to distance.

In the data, we define several groups of observations used throughout the paper. “Workers” are individuals with an industry identifier and an age between 20 and 60. “Agricultural workers” are workers employed in Agriculture, Forestry, and Fishing, corresponding to ind1950 codes 105, 116, and 126. “Immigrant workers” are defined

---

<sup>24</sup>We adopt 1870 as the cutoff because data collection for the 1880 Census extended over several years, and states admitted after 1870 often had incomplete records.

FIGURE A.1: AGRICULTURAL EMPLOYMENT SHARES ACROSS COMMUTING ZONES, 1880



*Notes:* The map displays all commuting zones in the United States, shaded by their agricultural employment share in 1880. Darker shades indicate higher agricultural specialization. Commuting zones shown in grey are excluded from our sample because their states had not yet joined the Union by 1870.

as workers who immigrated within the past 20 years. We classify “young workers” as those aged 20 to 40 and “old workers” as those aged 40 to 60.<sup>25</sup>

For each commuting zone, we compute the agricultural employment share as the ratio of agricultural workers to all workers. Figure A.1 shows the agricultural employment share in 1880 across commuting zones.

Finally, for each state, we compute the number of workers born in each other state. The resulting lifetime state-to-state migration matrix allows us to estimate the elasticity of migration flows with respect to distance.

### B.1.2 Census of Manufacturing

**Source and Description** We obtain county-level tabulations of the Census of Manufacturing for 1880, 1900, and 1920 from the NHGIS database (see [Manson et al., 2017](#)). We retain only two variables: total manufacturing payroll and total manufacturing employment.

**Sample Selection, Processing, and Use** We drop all counties with zero or missing manufacturing payroll or employment. Using the crosswalk from [Eckert et al. \(2020\)](#), we aggregate the county-level data to the commuting zone level. We compute commuting zone-level average wages by dividing total manufacturing payroll by total manufacturing employment. Throughout the paper, we refer to this measure simply as “average wage” or “earnings.” In our model, average earnings are identical across sectors, so that the average manufacturing wage in the data corresponds to the average commuting zone wage in the model,  $\bar{w}_{rt}$ .

<sup>25</sup>These classifications are used to estimate the location- and decade-specific labor force growth rates,  $n_{rt}$ .

### B.1.3 Census of Agriculture

**Source and Description** We obtain county-level tabulations of the Census of Agriculture for 1880, 1900, and 1920 from the NHGIS database (see [Manson et al., 2017](#)). We retain the following variables: average land value per acre, acres of agricultural land, total number of farms, number of farms by farm-size bin, and total value of agricultural output.

**Sample Selection, Processing, and Use** We drop all counties with zero or missing values for average land value per acre. Using the crosswalk from [Eckert et al. \(2020\)](#), we aggregate county-level data to the commuting zone level by computing a land-area-weighted average of land values across all counties within each commuting zone. As a result, the agricultural land value reflects the value of an average acre of agricultural land within a commuting zone. We interpret these data in 1880 as proportional to land rents in the model and use them to identify the supply of agricultural land,  $T_r$ . Finally, we compute average farm size at the county level by dividing the total number of improved acres by the total number of farms.

### B.1.4 Linked Census Files

**Source and Description** Economists have developed algorithms to match individuals across sequential Decennial Census waves based on names and other observable characteristics. IPUMS provides a linked file identifying individuals appearing in both the 1880 and 1900 Censuses (see [Ruggles et al., 2017](#)). We supplement this with linked files constructed by [Abramitzky, Boustan, Eriksson, Rashid, and Pérez \(2022\)](#), which cover various pairs of Census years.

**Sample Selection, Processing, and Use** Both samples are restricted to men, as surname changes make it difficult to match women over time. We retain only observations who qualify as workers according to our definition applied to the full-count Census files. Using the linked data, we compute the share of workers who move from commuting zone  $r$  to commuting zone  $r'$  between 1880 and 1900. The resulting “commuting-zone-to-commuting-zone migration matrix” is used to estimate the elasticity of migration flows with respect to distance.

### B.1.5 Historical Statistics of the United States

**Source and Description** For aggregate time series data, we use the canonical Historical Statistics of the United States (see [Carter et al., 2006](#)). We rely on the series for real GDP, the price of farm goods, and the price of all commodities excluding farm goods.

**Sample Selection, Processing, and Use** Both the GDP and price series provide target moments in our estimation. We interpret the farm goods price series as agricultural prices in the model, and the non-farm commodity price series as manufacturing prices. Our estimation targets the growth rate of real GDP and the change in relative prices between 1880 and 1920.

## B.2 Robustness of Empirical Results in Figure 1

In this section, we show that the patterns documented in Figure 1 are robust to changes in the spatial unit of observation and the inclusion of fixed effects. Table A.1 reports results from three key regressions: (1) log wages on agricultural employment shares in 1880 (Panel A), capturing the urban-rural wage gap; (2) local wage growth on initial agricultural employment shares (Panel B); and (3) changes in agricultural employment shares on initial agricultural employment shares and their square (Panel C).

Column 1 reproduces our baseline results corresponding to Figure 1. Columns 2 and 3 show that the results are robust to including state fixed effects and to unweighted regressions that do not account for employment size. Columns 4 and 5 repeat the analysis at the county level, with state and commuting zone fixed effects, respectively. Across all specifications, we find consistent evidence of spatial convergence: initially poor, agricultural regions experienced faster wage growth than more industrialized areas, and the process of industrialization displayed the same non-monotonic pattern observed in Figure 1.

## B.3 Details on Estimation Moments and Methods

### B.3.1 Local Employment Growth $\{n_{rt}\}$

Our theory accounts for local labor force growth solely through interregional migration. Empirically, other factors affecting the size of the local labor force are births, immigration, and deaths, all of which likely differ across commuting zones. These determinants of labor force entry and exit also generate aggregate employment growth. In this section, we show which determinants of local labor force growth vary substantially across commuting zones and how we account for them in our analysis.

Figure A.2 shows proxies for the three most important sources of local employment growth besides internal migration: births, immigration from outside the US, and deaths. The rightmost panel shows the number of children per adult (“birth rates”). We measure local “birth rates” as the fraction of children between 0 and 20 relative to the number of working adults aged 20-60. Rural locations have substantially higher birth rates which drive part of local employment growth.

The middle panel graphs the share of immigrants in the local workforce against a commuting zone’s initial agricultural employment shares. Immigrants are workers that immigrated within the last 20 years from outside the US. Immigrants predominantly settled in urban locations, and thus raised the employment of such manufacturing locations.

The rightmost panel of Figure A.2 provides evidence that - compared to births and immigration - labor force exit rates do not vary systematically across space. If death and retirement rates varied substantially across regions, the fraction of young workers (20-40 years old) in the total workforce should vary a lot, too. However, the figure shows that the fraction of young workers is essentially uncorrelated with the local agricultural employment share. We thus assume that the rate of labor force exit is constant across locations.

Our theory accounts for local labor force growth solely through migration internal to the US. Empirically, however, other forces also affect the local labor force: births,

TABLE A.1: SPATIAL STRUCTURAL CHANGE AND RURAL CATCH-UP

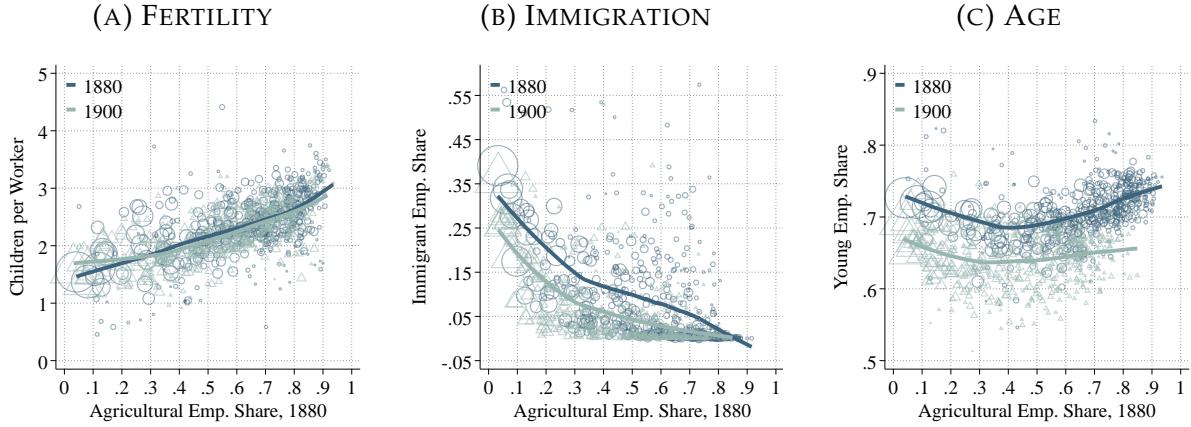
	(1)	(2)	(3)	(4)	(5)
<b>Panel A: Log Average Wages, 1880</b>					
$s_{rAt}$	-0.969*** (0.042)	-1.062*** (0.052)	-1.382*** (0.072)	-1.002*** (0.027)	-1.020*** (0.035)
R-squared	.516	.804	.427	.717	.796
Observations	495	494	495	1,956	1,899
<b>Panel B: Annual Wage Growth (%), 1880-1920</b>					
$s_{rAt}$	1.364*** (0.106)	1.720*** (0.179)	2.095*** (0.196)	1.521*** (0.094)	1.591*** (0.129)
R-squared	.821	.85	.676	.697	.725
Observations	990	990	990	3,912	3,912
<b>Panel C: Changes in Agricultural Employment Shares, 1880-1920</b>					
$s_{rAt}$	-0.485*** (0.028)	-0.468*** (0.034)	-0.366*** (0.062)	-0.381*** (0.016)	-0.384*** (0.021)
$s_{rAt}^2$	0.452*** (0.032)	0.430*** (0.039)	0.269*** (0.057)	0.348*** (0.019)	0.359*** (0.023)
R-squared	.311	.386	.104	.241	.399
Observations	990	990	990	3,912	3,912
Geography	CZ	CZ	CZ	County	County
Year FEs (B,C only)	Yes	Yes	Yes	Yes	Yes
Geo. FEs		State		State	CZ
Emp. Weights	Yes	Yes	No	Yes	Yes

*Notes:* All regressions in panels B and C are pooled for the two periods 1880-1900 and 1900-1920 and include a fixed effect for each period. Data on wages are from the Census of Manufacturing; all other data are from the full-count US Decennial Census files. Robust standard errors are in parentheses. The variation in sample size across regressions reflects the fact that the fixed-effect regressions drop observations without variation within the fixed-effect category. Connecticut consists of a single commuting zone and 57 commuting zones consist of a single county. The symbols \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5% and 1% level, respectively.

immigration from outside the US, and deaths, all vary across commuting zones and contribute to aggregate employment growth. In this section, we document the importance of these forces and describe how we incorporate them into our analysis.

Figure A.2 shows proxies for the three primary drivers of local employment growth beyond internal migration: births, immigration from abroad, and deaths. The left panel plots local birth rates, measured as the ratio of children (ages 0–20) to working adults (ages 20–60). Birth rates were substantially higher in rural areas, contributing to faster local labor force growth in these regions. The middle panel displays the share of recent immigrants—defined as workers aged 20–60 who arrived within the past 20 years—in the local workforce. Immigrants settled predominantly in urban areas, boosting employment growth in more industrialized labor markets. Finally, the right panel shows that labor force exit rates do not vary systematically across space, by plotting the share of young workers among all workers across commuting zones. If mortality or retirement patterns differed across locations, the fraction of young workers (ages 20–40) would vary accordingly. Because the share of young workers is essentially

FIGURE A.2: IMMIGRATION, FERTILITY, AND AGE STRUCTURE ACROSS SPACE



Notes: The left panel plots a proxy for the local birth rate, measured as the share of children (ages 0–20) relative to working adults (ages 20–60), against the initial agricultural employment share. The middle panel plots the share of immigrants—defined as workers aged 20–60 who immigrated within the past 20 years—in the local workforce against the initial agricultural employment share. The right panel plots the share of young workers (ages 20–40) among all workers (ages 20–60) in each commuting zone. Workers are defined as individuals with a reported industry identifier. The underlying data source for all panels is the US Decennial Census files for 1880 and 1900. Marker size is proportional to total employment in each region. Each panel displays the fit from a local polynomial regression.

uncorrelated with initial agricultural employment shares, we conclude that labor force exit rates were approximately constant across locations.

We now describe how we estimate the exogenous component of local employment growth  $n_{rt}$  in each region. We denote by  $L_{rt}^Y$  the number of workers in a location at the beginning of period  $t$ , that is before making their moving decisions.  $L_{rt}$  is the number of workers working in region  $r$  during period  $t$ , i.e., the mass of workers that chose to move to (or remain in) location  $r$  during period  $t$ . The local rate of exogenous labor force growth,  $n_{rt}$ , is thus defined by  $L_{rt+1}^Y = n_{rt}L_{rt}$ . To calibrate  $n_{rt}$ , note that the following accounting identity describes the law of motion of the total labor force in region  $r$  at the beginning of period  $t$ :

$$L_{rt}^Y = L_{rt-1} - Exit_{rt-1,t} + Entry_{rt-1,t} = L_{rt-1} \left( 1 - \frac{Exit_{rt-1,t}}{L_{rt-1}} + \frac{Entry_{rt-1,t}}{L_{rt-1}} \right),$$

where  $Exit_{rt-1,t}$  is the number of workers that exit the labor force between periods  $t-1$  and  $t$  but do not leave the location to work elsewhere. Similarly,  $Entry_{rt-1,t}$  is the number of workers entering the labor force between periods  $t-1$  and  $t$  that did not immigrate from another domestic region between  $t-1$  and  $t$ .

Given our assumption of a constant labor force exit rate across regions, we set the exit rate equal to a common constant,  $\delta$ , so that  $\frac{Exit_{rt-1,t}}{L_{rt-1}} = \delta$ . The gross rate of local labor force growth prior to workers making their migration decisions is thus given by

$$n_{rt-1} = \frac{L_{rt}^Y}{L_{rt-1}} = 1 - \delta + \frac{Entry_{rt-1,t}}{L_{rt-1}}.$$

Let  $C_{rt}$  denote the number of children in  $r$  at time  $t-1$  and  $I_{rt}$  denote the number of working immigrants in location  $r$  that arrived between  $t-1$  and  $t$ . Since we assume

differences in entry rates to be due to differences in fertility rates and immigration only, we relate  $C_{rt}$  and  $I_{rt}$  to  $Entry_{rt-1,t}$  according to

$$\frac{Entry_{rt-1,t}}{L_{rt-1}} = x \times \frac{C_{rt} + I_{rt}}{L_{rt-1}},$$

where  $x$  is a scalar that reflects measurement error, e.g., some children die, time is not discrete (i.e., the 16 year old children enter the labor market earlier than the 5 yr old children), or immigrants might move across locations within the US in-between Census years. Then

$$n_{rt-1} = 1 - \delta + x \times \frac{C_{rt} + I_{rt}}{L_{rt-1}}.$$

Note that  $C_{rt}$ ,  $I_{rt}$  and  $L_{rt-1}$  are observed in the data. Hence, this equation determines  $n_{rt-1}$  as a function of the scalars  $\delta$  and  $x$ .

We determine these scalars in the following way. First, we choose the scalar  $x$  to ensure that this accounting equation satisfies the *aggregate* rate of employment growth in the Census, that is we ensure that the following equation holds in the data:

$$\text{Total employment at } t = \sum_r \left( 1 - \delta + x \frac{C_{rt} + I_{rt}}{L_{rt-1}} \right) L_{rt-1}.$$

Rearranging terms implies that

$$x = \frac{\text{Total employment in } t - (1 - \delta) \text{ Total employment in } t-1}{\sum_r (C_{rt} + I_{rt})}.$$

Hence, for a given exit rate  $\delta$ , we pick the scale  $x$  for the aggregate birth and immigration inflow to account for all employment growth. And then we use this  $x$  to calculate - in the model - the number of workers in region  $r$  prior to their migration choices as

$$L_{rt}^Y = n_{rt-1} L_{rt-1} = L_{rt-1} (1 - \delta) + x (C_{rt} + I_{rt}).$$

Hence, local labor force growth prior to worker's migration choices depends on the observable  $(C_{rt} + I_{rt})$  and it has the correct slope for our model to be consistent with aggregate employment growth.

To pick the exit rate  $\delta$ , note that the fraction of old workers at time  $t$  is given by

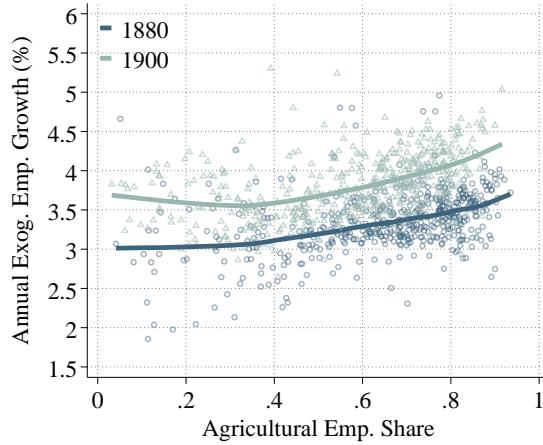
$$(A.28) \quad \text{Share of old workers}_t = \frac{(1 - \delta) \sum_r L_{rt-1}}{\sum_r L_{rt}^Y} = (1 - \delta) \frac{\sum_r L_{rt-1}}{\sum_r L_{rt}}.$$

Because  $\frac{\sum_r L_{rt-1}}{\sum_r L_{rt}}$  is simply the ratio of the total labor force at  $t - 1$  divided by the total labor force at  $t$ , which are both observed, we can calculate  $\delta$  for any target of the share of old workers. A generation in our model corresponds to 20 years in the data. In calibrating  $\delta$ , we think of 0-20 year olds as not working, of 20-40 year olds as "young" workers, and of "40-60" year olds as "old workers." The share of old workers in our data is 0.34, 0.35 and 0.37 in 1880, 1900, and 1920, respectively. Because, empirically, some people above 60 are still in the workforce, we take a number of 0.45. Together

with a rate of employment growth of about 35% observed in the data (at the 20 year horizon), equation (A.28) implies that  $\delta$  is given by  $1 - 0.45 \times 1.35 = 0.4$ .

In Figure A.3, we show the calibrated exogenous rate of employment growth,  $n_{rt}$ , for the two time periods 1880-1900 and 1900-1920. The figure shows that, on net, exogenous employment growth was slightly higher in agricultural regions. The relationship between agricultural specialization and subsequent exogenous employment growth weakens somewhat over these periods suggesting employment growth became somewhat more balanced as fertility rates in more rural regions started to decline.

FIGURE A.3: EXOGENOUS EMPLOYMENT GROWTH



*Notes:* The figure plots the calibrated annualized rate of exogenous employment growth across commuting zones between 1880–1900 and 1900–1920. For clarity, we omit one region in each period with a growth rate exceeding 3 to better display the variation among the remaining observations.

### B.3.2 Migration Gravity Equations

In this section, we describe our estimation of the distance elasticity of migration costs,  $\omega$ . In the model, the mass of workers migrating from region  $r$  to region  $j$  between two periods is given by:

$$M_{rjt} = m_{rjt} L_{rt}^Y = \frac{(\mu_{rj} \mathcal{V}_{jt} \mathcal{B}_{jt})^\varepsilon}{\sum_j (\mu_{rk} \mathcal{V}_{kt} \mathcal{B}_{kt})^\varepsilon} L_{rt-1} n_{rt-1}.$$

We project the moving cost between two regions on the physical distance between them, that is we set  $\mu_{rj} = d_{rj}^{-\omega}$ , where the parameter  $\omega$  parameterizes the distance cost of migration. The larger  $\omega$ , the more the destination utility of areas further away is discounted. In our empirical estimation, we set  $d_{rr}$  for all  $r$  to the average distance between county centroids within commuting zone  $r$ , and  $d_{rj}$  for all regions  $r \neq j$  to the distance between commuting zone centroids.

Taking logs on both sides and grouping terms then yields:

$$(A.29) \quad \ln M_{rjt} = \alpha_{jt} + \beta_{rt} - \omega \varepsilon \ln d_{rj}$$

where  $\alpha_{jt} = \varepsilon \ln(\mathcal{V}_{jt} \mathcal{B}_{jt})$  and  $\beta_{rt} = \ln(L_{rt-1} n_{rt-1}) - \ln \left( \sum_k (d_{rk}^{-\omega} \mathcal{V}_{kt} \mathcal{B}_{kt})^\varepsilon \right)$ .

Equation (A.29) suggests a fixed effect regression of commuting zone migration flows to recover the elasticity of migration flows to distance,  $\omega\epsilon$ , relevant in our model.

The Decennial Census files do not contain information on workers' migration history at the commuting zone or county level. Hence, it is impossible to directly construct cross-commuting zone migration flows. We therefore rely on information from the linked Census files described in our data section above. We estimate equation (A.29) using Poisson Pseudo Maximum Likelihood (PPML), as proposed by [Silva and Tenreyro \(2006\)](#) because the migration matrix across commuting zones contains many zeros. More specifically, we estimate the following equation using PPML:

$$(A.30) \quad M_{rjt} = \exp(\alpha_{jt} + \beta_{rt} - \omega\epsilon \ln d_{rj}) + \epsilon_{rjt}.$$

Columns (1) and (2) in Table A.2 report the estimates based on two different linked-Census files by [Ruggles et al. \(2015\)](#) ("IPUMS") and [Abramitzky, Boustan, Eriksson, Feigenbaum, and Pérez \(2021\)](#) ("ABE"). These files differ slightly in their technique to link individuals across census years. Reassuringly, both produce similar estimates: we estimate an elasticity of migration flows with respect to geographic distance ( $\omega\epsilon$ ) of around 2.75, which we use in our baseline calibration.

TABLE A.2: MIGRATION GRAVITY REGRESSIONS

	(1)	(2)	(3)	(4)	(5)
Log Distance	-2.922*** (0.033)	-2.632*** (0.011)	-3.925*** (0.057)	-2.262*** (0.031)	-2.291*** (0.031)
Geography	CZ	CZ	State	State	State
Fixed Effects	O-D	O-D	O-D, Year	O-D, Year	O-D, Year
Pseudo R-squared	.814	.899	.93		
R-squared				.81	.802
Observations	254,762	349,230	3,983	3,983	3,935

Notes: All regressions contain origin and destination fixed effects. (1) PPML with census data from IPUMS linked by IPUMS. (2) PPML with census data from IPUMS linked by Abramitzky Boustan Eriksson. (3) PPML in state flow data from IPUMS, pooled across all years. (4) OLS regression in state flow data from IPUMS adding a 1 to all flows, pooled across all years. (5) OLS regression in state flow data from IPUMS dropping zero flow observations, pooled across all years. Note the linked data is only available for one cross-section: 1880-1900. For the regressions using state data we pool data on lifetime migration between 1880-1900 and 1900-1920 and add year fixed effects into the regressions.

Linking data across census years requires a set of assumptions and large amounts of data processing. For robustness, we therefore repeat the estimation on a different data set that we can directly compute from the cross-sectional Census data but that only contains state-to-state flows. In particular, as discussed in the data section above, we use the information on the state of birth of each worker contained in the Decennial Census files to construct a matrix of lifetime state-to-state migration flows for all workers between 20 and 40. Column 3 of Table A.2 presents the PPML estimates of the distance elasticity in the state-to-state data. The coefficient on log distance is more negative than when using the commuting zone data highlighting that there are, by construction, less flows across states than across commuting zones making distance appear as a larger impediment of migration.

The state-to-state migration matrix contains relatively few pairs of states with zero flows between them. Across the two cross-sections (1880–1900 and 1900–1920), only about 50 state pairs exhibit zero flows. As a result, we can estimate the gravity regression using simple OLS instead of PPML. Columns 4 and 5 of Table A.2 report results from two alternative specifications: one that replaces zero flows with ones (Column 4) and one that omits state pairs with no migration flows (Column 5).

It is important to note that our theory only approximately delivers a gravity equation at the state level. Because of Jensen’s inequality, the distance elasticities estimated from aggregated state-level flows do not map exactly into the structural parameter  $\omega\epsilon$ . Nevertheless, it is reassuring that the estimates from state-level regressions are similar to those based on commuting zone flows. Moreover, since a fraction of moves in the model occur across commuting zones within the same state, we would expect the state-level distance elasticities to be larger than their commuting zone counterparts in model-generated data as well.

### B.3.3 Local fundamentals in 1880: $[A_{r1880}, Z_{r1880}, B_{r1880}, T_{r1880}]$

Table A.3 shows the relationship between the inferred regional fundamentals and agricultural employment shares in 1880. Specifically, we run a set of cross-sectional regressions of  $srA1880$  against the estimated fundamentals and population density in 1880. In 1880, agricultural regions had both low agricultural productivity,  $A_{r1880}$ , and low manufacturing productivity,  $Z_{r1880}$ . Agricultural specialization was therefore a reflection of a comparative advantage in agriculture but not of an absolute advantage. Columns 3 and 4 show that more agricultural areas had a higher relative productivity in the agricultural sector and fewer workers per unit of agricultural land  $\ell_r = L_r/T_r$ . This pattern is implied by the fact that, empirically, agricultural land rents in rural regions were relatively low in 1880. Table A.3 also shows that more agricultural regions had slightly higher amenities in 1880.

### B.3.4 Estimation of $\zeta$

In this section, we report the results from estimating regression (28), using data on the distribution of farm sizes within each commuting zone from the Census of Agriculture. The data report the share of farms with less than 3, 10, 20, 50, 100, 500, and 1000 hectares. Table A.4 summarizes the estimation results.

Columns 1 and 2 of Table A.4 pool data across commuting zones and decades (1880, 1900, and 1920), including a full set of region-year fixed effects. For the full sample of farms, we estimate a value of  $\zeta$  around 1.4. In Column 2, we restrict the sample to farms with at least 10 hectares in order to focus on full-time farmers, excluding smaller part-time operations. This restriction yields an estimate of  $\zeta$  around 1.6.

Columns 3 through 5 report results separately by decade. The decade-specific estimates are very similar to the pooled estimate in Column 2, suggesting that the distribution of farm sizes was stable over time. Based on these results, we set  $\zeta = 1.6$  in our baseline calibration.

TABLE A.3: DETERMINANTS OF AGRICULTURAL SPECIALIZATION

	(1)	(2)	(3)	(4)	(5)	(6)
$\log A_{r1880}$	-0.197*** (0.016)					
$\log Z_{r1880}$		-0.208*** (0.007)			-0.032*** (0.006)	
$\log A_{r1880}/Z_{r1880}$			0.255*** (0.010)		0.208*** (0.009)	
$\log \ell_{r1880}$				-0.131*** (0.006)	-0.116*** (0.004)	
$\log B_{r1880}$					0.060*** (0.013)	-0.007*** (0.002)
R-squared	.275	.724	.53	.491	.0555	.972
Observations	495	495	495	495	495	495

Notes: The table reports the results of a set of bivariate regressions  $s_{rA1880} = \alpha + \beta x_r + u_r$ , where  $x_r = \ln Z_{rA1880}$  (column 1),  $x_r = \ln Z_{rM1880}$  (Column 2),  $x_r = \ln(Z_{rA1880}/Z_{rM1880})$  (Column 3),  $x_r = \ln \ell_{r1880}$  (Column 4) and  $x_r = \ln B_{r1880}$  (Column 5). Robust standard errors are in parentheses. The symbols \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5% and 1% level, respectively.

TABLE A.4: ESTIMATING  $\zeta$ 

	(1)	(2)	(3)	(4)	(5)
Log Relative Farmsize	-1.405*** (0.012)	-1.650*** (0.015)	-1.581*** (0.022)	-1.649*** (0.023)	-1.688*** (0.029)
Fixed Effects	CZ-Year	CZ-Year	CZ	CZ	CZ
Minimum Farm Size	-	10	10	10	10
Emp. weights	Yes	Yes	Yes	Yes	Yes
Sample	1880-1920	1880-1920	1880	1900	1920
R-squared	.866	.896	.894	.897	.897
Observations	10,259	8,860	2,929	2,969	2,962

Notes: The table reports the results from estimating equation (28) using data on the farm size distribution from the Census of Agriculture. Standard errors are clustered at the region-year level, and all regressions are weighted by local population. Columns 1 and 2 pool data across 1880, 1900, and 1920; Columns 3 to 5 report separate estimates by decade. Columns 2 to 5 restrict the sample to farms with at least 10 hectares.

### B.3.5 Aggregating Trade Costs to CZs

To construct trade cost matrices between commuting zones, we draw on the data from [Hornbeck and Rotemberg \(2024\)](#), who digitize information on the US transportation network by decade from 1830 to 1920. Their dataset, “NSFtranspCost.dta” in their replication package, contains county-to-county transportation cost estimates based on the existing infrastructure in each decade. All county definitions are held fixed at 1890 boundaries, allowing for consistent comparison over time.

Transportation costs are reported in dollars per ton of goods shipped and are calculated using centroid-to-centroid distances, where intra-county costs are set to zero. We follow [Hornbeck and Rotemberg \(2024\)](#), in converting these figures to unit-less ad-valorem cost parameters by assuming an average value of \$35 per ton of agricultural goods in 1890. Specifically, for each origin-destination pair, we compute trade costs as

$$(A.31) \quad \tau_{od} = 1 + \frac{t_{od}}{35},$$

where  $t_{od}$  is the reported cost per ton from county  $o$  to county  $d$ . This yields an origin-destination matrix of bilateral trade cost shifters at the county level.

To construct trade costs between commuting zones (CZs), we use the crosswalk provided by [Eckert et al. \(2020\)](#) from 1890 county boundaries to CZs. The resulting CZ-by-CZ trade cost matrices allow us to track the evolution of bilateral market access over time in our spatial model.

### B.3.6 Agricultural Productivity Growth: Direct Evidence from the Agricultural Census

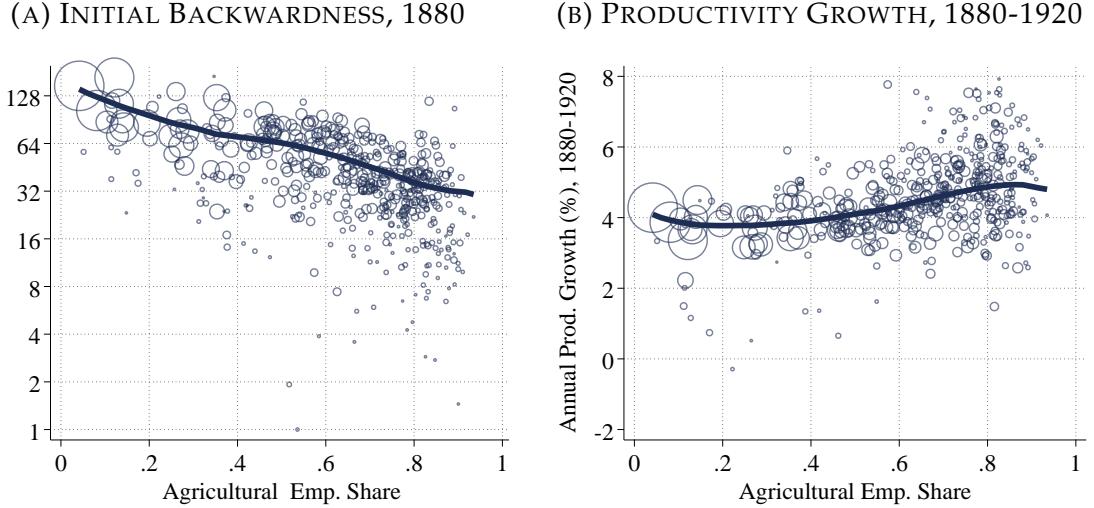
To provide direct evidence on the relationship between agricultural productivity and initial agricultural specialization, we use data from the Census of Agriculture to compute agricultural productivity growth based on observed land yields. We define average agricultural productivity in region  $r$  at time  $t$  as the total value of agricultural output per acre of agricultural land. The left panel of Figure A.4 shows that agricultural productivity was substantially lower in more agricultural regions in 1880, consistent with the findings from our model-inversion exercise in the body of the paper. The right panel shows that these initially less productive regions subsequently experienced faster agricultural productivity growth, providing direct evidence of catch-up dynamics.

### B.3.7 Human Capital or Productivity?

Our baseline model abstracts from regional differences in human capital. In this section, we extend the model to allow for variation in human capital across regions and sectors, and develop a strategy to separately identify the effects of human capital and physical productivity on wage growth and industrialization.

**Extending the Model** Suppose the sector-specific human capital distribution in region  $r$  at time  $t$  is given by  $F_{rst}(z) = \exp\left(-\left(\frac{z}{h_{rt}^s}\right)^{-\zeta}\right)$ , where  $h_{rt}^s$  parametrizes the level of human capital for sector  $s \in A, M$ . Our baseline model corresponds to the special case

FIGURE A.4: DIRECT MEASURES OF AGRICULTURAL PRODUCTIVITY



Notes: In the left (right) panel we show the correlation between log agricultural output per acre (the growth rate of agricultural output per acre) against the agricultural employment share.

where  $h_{rt}^s = 1$ .

In this generalized setting, average earnings in region  $r$  satisfy:

$$(A.32) \quad \bar{w}_{rt} = \left( \left( h_{rt}^A \bar{\chi}_{rt} \right)^\zeta + \left( h_{rt}^M w_{rMt} \right)^\zeta \right)^{1/\zeta}.$$

The agricultural employment share and the human-capital-adjusted agricultural labor supply are:

$$(A.33) \quad s_{rA} = \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^\zeta \quad \text{and} \quad H_{rAt} = \Gamma_\zeta L_r h_{rt}^A s_{rA}^{\frac{\zeta-1}{\zeta}} = \Gamma_\zeta L_r h_{rt}^A \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{\zeta-1}.$$

Manufacturing wages are still given by  $w_{rMt} = \left( \frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} Z_{rt}^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} \equiv \mathcal{Z}_{rt}$ , so that

$$(A.34) \quad h_{rt}^M w_{rMt} = h_{rt}^M \mathcal{Z}_{rt} \equiv \mathcal{Z}_{rt}^{HC}.$$

Similarly, agricultural profitability  $\bar{\chi}_{rt}$  satisfies

$$\bar{\chi}_{rt} = \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}} \left( \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt} \right)^{\frac{\sigma_A-1}{\sigma_A}} T_r^{\alpha \frac{\sigma_A-1}{\sigma_A}} H_{rAt}^{-\left( \alpha + \frac{1}{\sigma_A-1} \right) \frac{\sigma_A-1}{\sigma_A}}.$$

Using the expression in equation (OA.11) yields

$$\begin{aligned} h_{rt}^A \bar{\chi}_{rt} &= h_{rt}^A \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}} \left( \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt} \right)^{\frac{\sigma_A-1}{\sigma_A}} T_r^{\alpha \frac{\sigma_A-1}{\sigma_A}} \left( \Gamma_\zeta L_r h_{rt}^A \right)^{-\left( \alpha + \frac{1}{\sigma_A-1} \right) \frac{\sigma_A-1}{\sigma_A}} \\ &\quad \times \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-\left( \alpha + \frac{1}{\sigma_A-1} \right) \frac{\sigma_A-1}{\sigma_A} (\zeta-1)}. \end{aligned}$$

Defining

$$(A.35) \quad \mathcal{A}_{rt}^{HC} \equiv \left[ \left( h_{rt}^A \right)^{1-\alpha} \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{1-\alpha} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt} \bar{T}_r^\alpha (\Gamma_\zeta L_r)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)} \right]^{\frac{\sigma_A-1}{\sigma_A}}$$

yields

$$(A.36) \quad \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} = \left( \frac{\mathcal{A}_{rt}^{HC}}{\bar{w}_{rt}} \right)^{\frac{1}{\kappa}}.$$

where  $\kappa = \zeta - (\zeta - 1)(1 - \alpha) \frac{\sigma_A - 1}{\sigma_A}$ . Equations (OA.10), (OA.12), and (OA.14) imply that average earnings  $\bar{w}_{rt}$  are given by

$$(A.37) \quad 1 = \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^\zeta + \left( \frac{h_{rt}^M w_{rMt}}{\bar{w}_{rt}} \right)^\zeta = \left( \left( \frac{\mathcal{A}_{rt}^{HC}}{\bar{w}_{rt}} \right)^{\frac{1}{\kappa}} \right)^\zeta + \left( \frac{\mathcal{Z}_{rt}^{HC}}{\bar{w}_{rt}} \right)^\zeta.$$

Similarly, note that

$$h_{rt}^A \bar{\chi}_{rt} = \mathcal{A}_{rt}^{HC} \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right) \frac{\sigma_A-1}{\sigma_A} (\zeta-1)} = \mathcal{A}_{rt}^{HC} \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-(\kappa-1)} = \mathcal{A}_{rt}^{HC} s_{rAt}^{-\frac{\kappa-1}{\zeta}},$$

where the last equality uses (OA.11). Hence,  $\frac{s_{rA}}{1-s_{rA}} = \left( \frac{h_{rt}^A \bar{\chi}_{rt}}{h_{rt}^M w_{rMt}} \right)^\zeta = \left( \frac{\mathcal{A}_{rt}^{HC}}{\mathcal{Z}_{rt}^{HC}} \right)^\zeta s_{rA}^{-(\kappa-1)}$ , so that

$$(A.38) \quad \frac{s_{rA}^\kappa}{1-s_{rA}} = \left( \frac{\mathcal{A}_{rt}^{HC}}{\mathcal{Z}_{rt}^{HC}} \right)^\zeta.$$

Given the human capital adjusted revenue productivity terms  $\mathcal{A}_{rt}^{HC}$  and  $\mathcal{Z}_{rt}^{HC}$ , equations (OA.15) and (OA.16) are the same equations as in Proposition 3.

**Implications** We establish the following result in our extended model with human capital differences across regions and sectors:

**Proposition 5.** Consider the model with human capital and let  $\bar{Z}_{rt}$  and  $\bar{A}_{rt}$  denote productivity in both sectors. The equilibrium allocations in this model are the same as the ones in our baseline model without human capital if we redefine, in our baseline model,

$$Z_{rt}^{Base} \equiv (h_{rt}^M)^{\frac{\sigma_M}{\sigma_M-1}} \bar{Z}_{rt} \quad \text{and} \quad A_{rt}^{Base} = \left( h_{rt}^A \right)^{1-\alpha} \bar{A}_{rt}.$$

*Proof.* Given data on  $s_{rA}$  and  $\bar{w}_r$ , we can uniquely solve for  $\mathcal{A}_{rt}^{HC}$  and  $\mathcal{Z}_{rt}^{HC}$ . We now show that given these data and the other structural parameters there is a one-to-one mapping between  $\mathcal{A}_{rt}^{HC}$  and  $\mathcal{Z}_{rt}^{HC}$  and  $Z_{rt}^{Base}$  and  $A_{rt}^{Base}$ . The fact that  $Z_{rt}^{Base} \equiv h_{rt}^M \bar{Z}_{rt}$

follows immediately from (OA.12), which implies that

$$\begin{aligned}\mathcal{Z}_{rt}^{HC} = h_{rt}^M \mathcal{Z}_{rt} &= h_{rt}^M \left( \frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} (\bar{Z}_{rt})^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} \\ &= \left( \frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} (Z_{rt}^{Base})^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M}.\end{aligned}$$

To derive the expression for  $A_{rt}^{Base}$  note that (see equation OA.13)

$$\begin{aligned}\mathcal{A}_{rt}^{HC} &= \left[ \left( h_{rt}^A \right)^{1-\alpha} \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{1-\alpha} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} \bar{A}_{rt} \bar{T}_r^\alpha (\Gamma_\zeta L_r)^{-\left( \alpha + \frac{1}{\sigma_A-1} \right)} \right]^{\frac{\sigma_A-1}{\sigma_A}} \\ &= \left[ \left( \frac{1-\alpha}{\gamma+1-\alpha} \right)^{1-\alpha} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt}^{Base} \bar{T}_r^\alpha (\Gamma_\zeta L_r)^{-\left( \alpha + \frac{1}{\sigma_A-1} \right)} \right]^{\frac{\sigma_A-1}{\sigma_A}}.\end{aligned}$$

Furthermore, it can be shown that  $\mathcal{D}_{rAt}$  and  $\mathcal{D}_{rMt}$  also only depend on  $A_{rt}^{Base}$  and  $Z_{rt}^{Base}$  conditional on the data  $s_{rA}$  and  $\bar{w}_r$  and the other structural parameters.  $\square$

Proposition 5 highlights that differences in human capital, across sectors and regions, are isomorphic to differences in productivity.<sup>26</sup> This result highlights two concerns for our empirical analysis. First, the initial rural productivity disadvantage we observe in 1880 could, in principle, reflect lower human capital rather than lower physical productivity. Second, the faster growth we estimate for rural regions could be driven by improvements in education rather than by the diffusion of technology.

While both interpretations are consistent with rural convergence, they differ importantly in mechanism and policy implications. If convergence stems from human capital accumulation, it would suggest a central role for public investment in education. If, instead, it reflects physical productivity catch-up, it points to technology diffusion and access to innovation as primary drivers.

**Linking Human Capital to Observables** To quantify the potential role of human capital, we link human capital levels to observed measures of schooling. Specifically, we model

$$(A.39) \quad h_{rt}^M = e^{\gamma_M sch_{rt}} \quad \text{and} \quad h_{rt}^A = e^{\gamma_A sch_{rt}},$$

where  $sch_{rt}$  denotes the average years of schooling in region  $r$  at time  $t$ .

A key challenge in our setting is that the US Census does not report educational attainment before 1940. We therefore exploit the fact that education is decided when people are young and look at different cohorts in 1940 to learn about the educational distribution in the past. We measure the average years of schooling in region  $r$  in 1900

---

<sup>26</sup>The fact that  $h_{rt}^A$  enters with an elasticity  $1 - \alpha$  reflects the fact that human capital only augments labor, not agricultural land. The fact that  $h_{rt}^M$  enters with an elasticity  $\frac{\sigma_M}{\sigma_M-1} = 1 + \frac{1}{\sigma_M-1}$  reflects the fact that higher human capital both increases overall output akin to productivity  $Z_{rt}$  and it also increases the number of local varieties  $N_{rt}$ , which increase productivity with an elasticity  $\frac{1}{\sigma_M-1}$ .

and 1920 as

$$sch_{r1920} = \frac{1}{N_{r1940}^{40-60}} \sum_{a=40}^{60} sch_{r1940}(a) \quad \text{and} \quad sch_{r1900} = \frac{1}{N_{r1940}^{60-80}} \sum_{a=60}^{80} sch_{r1940}(a),$$

where  $N_{r1940}^{40-60}$  ( $N_{r1940}^{60-80}$ ) denotes the number of people in region  $r$  in the 1940 Census between 40 and 60 years old (60 and 80 years old).<sup>27</sup> Our approach assumes that old-age migration is limited, and thus that elderly residents reflect the human capital of the locations where they spent most of their working lives.

Note that the case of  $\gamma_A = \gamma_M = \gamma$  is the case of sector neutral human capital: each year of schooling raises human capital in both sectors by  $\gamma$ . The case of  $\gamma_A = 0$  is the case where schooling does not raise the efficiency for agricultural work, that is as in [Porzio et al. \(2022\)](#) educational growth shifts labor supply toward manufacturing. For given  $\gamma_M$  and  $\gamma_A$  and measures of  $sch_r$ , we can directly compute  $h_{rt}^s$  and hence identify  $\bar{A}_{rt}$  and  $\bar{Z}_{rt}$  given our accounting solution  $Z_{rt}^{Base}$  and  $A_{rt}^{Base}$ . We can then directly measure physical productivity and compute its growth rate once human capital differences have been taken into account. We now explain how we measure  $sch_r$  and how we estimate  $\gamma_s$ .

**Estimating the Returns to Education** To recover the returns  $\gamma_M$  and  $\gamma_A$ , we use cross-sectional variation across educational groups within regions.

The model implies that the agricultural employment share among workers with education level  $j$  satisfies:

$$(A.40) \quad \ln \left( \frac{s_{rA1940}^j}{1 - s_{rA1940}^j} \right) = \delta_r - \zeta (\gamma_M - \gamma_A) sch_{r1940}^j + u_r^j.$$

where  $\delta_r$  captures region fixed effects that control for unobserved skill prices  $w_{rM1940}$  and  $\bar{\zeta}_{r1940}$ . A steeper negative slope indicates that returns to education are larger in manufacturing than in agriculture.

According to equation (A.40), we infer that educational returns are low in agriculture, that is,  $\gamma_M > \gamma_A$ , if people with higher human capital are systematically less likely to work in agriculture.

The model implies a second estimating equation to isolate  $\gamma_M$ :

$$(A.41) \quad \ln \bar{w}_{r1940}^j = \delta_r + \gamma_M sch_{r1940}^j - \frac{1}{\zeta} \ln s_{rM1940}^j + u_r^j.$$

Equation (A.41) looks exactly like a Mincer regression with a human capital return  $\gamma_M$ , except that we need to control for the manufacturing share to control for unobserved selection. To estimate  $\gamma_M$ , we exploit data on earnings observed in the 1940 Census.

**Results** In Table A.5, we report the results of estimating equations (A.40) and (A.41). Column 1 presents the estimates from equation (A.40). The estimated coefficient of -0.15 corresponds to  $\zeta(\gamma_M - \gamma_A)$ . Given our prior estimate of  $\zeta = 1.6$ , this implies

<sup>27</sup>We abstain from exploiting the same procedure for 1880, because we would need to focus on individuals between 80 and 100 years old.

$\gamma_M - \gamma_A = 0.15/1.6 \approx 0.094$ , suggesting that returns to education are substantially higher in the manufacturing sector than in agriculture.

Column 2 reports the estimate from equation (A.41), yielding  $\gamma_M = 0.1$ . Taken together, these results imply that  $\gamma_A = 0.1 - 0.094 \approx 0$ , indicating that human capital has essentially no effect on earnings in the agricultural sector.

Columns 4 and 5 replicate the analysis, weighting each observation by total employment in the commuting zone. The point estimates are very similar, confirming the robustness of our results to alternative weighting schemes.

TABLE A.5: ESTIMATING THE RETURN TO EDUCATION

	(1)	(2)	(3)	(4)	(5)	(6)
Years of Education	-0.150*** (0.001)	0.102*** (0.001)	0.016*** (0.001)	-0.132*** (0.004)	0.092*** (0.002)	0.003 (0.003)
Cty FEs	Yes	Yes	Yes	Yes	Yes	Yes
Cty Emp. Weights	No	No	No	Yes	Yes	Yes
Std. Err Cluster	Cty	Cty	Cty	Cty	Cty	Cty
R-squared	.702	.557	.406	.874	.746	.738
Observations	42,969	42,969	42,969	42,969	42,969	42,969

*Notes:* The table shows the estimated coefficient on the average years of education variable in a number of different regressions. Column 1 shows estimates from equation (A.40). Column 2 shows estimates from equation (A.41). Column 3 shows estimates from equation (A.42). Column 4 shows estimates from equation (A.40) with employment weights. Column 5 shows estimates from equation (A.41) with employment weights. Column 6 shows estimates from equation (A.42) with employment weights. The symbols \*, \*\*, and \*\*\* denote statistical significance at the 10%, 5% and 1% level, respectively.

In columns 3 and 6, we estimate  $\gamma_A$  directly. In particular, we estimate equation (A.41) by directly focusing on the agricultural sector. Because equation (A.41) should hold for either sector, we can estimate  $\gamma_A$  from

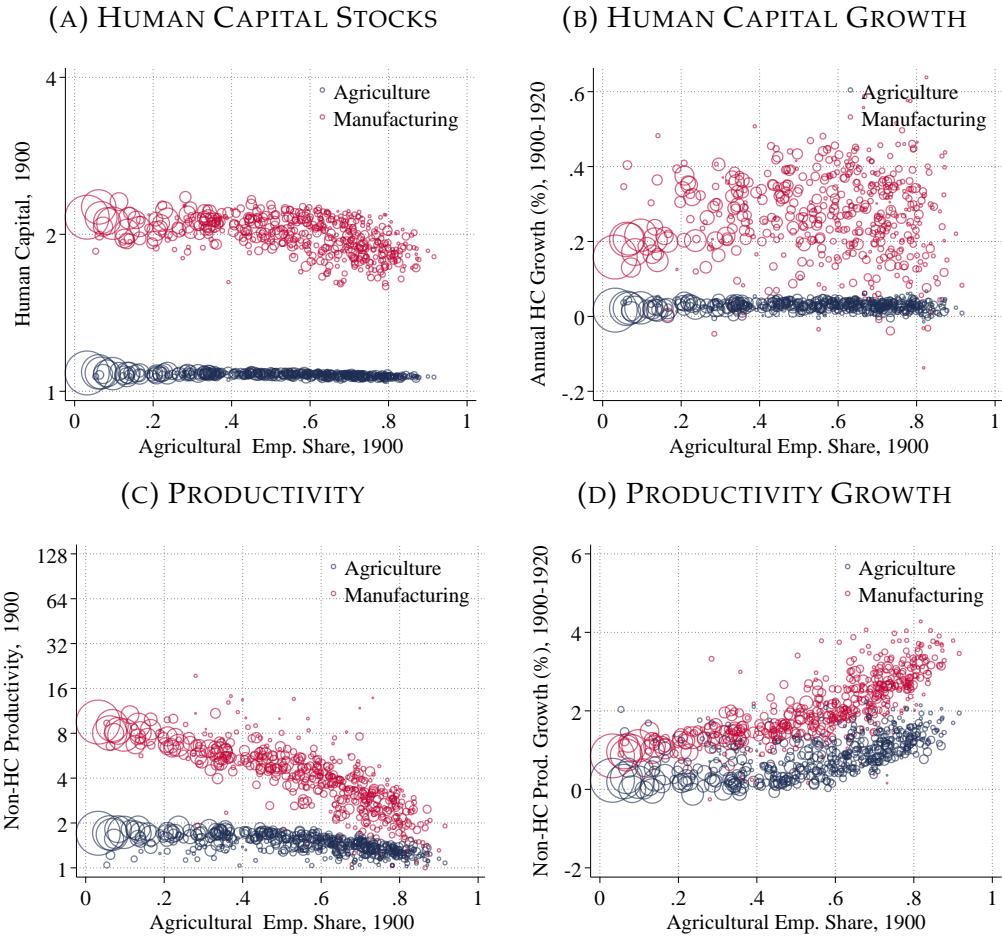
$$(A.42) \quad \ln \bar{w}_{r1940}^j = \delta_r + \gamma_A s_{rA1940}^j - \frac{1}{\zeta} \ln s_{rA1940}^j + u_r^j,$$

where the dependent variable now denotes average earnings among agricultural workers. This indeed leads to estimates for  $\gamma_A$ , which are close to 0, indicating that schooling does not lead to higher earnings in the agricultural sector.

Based on Table A.5, we thus take  $\gamma_A = 0.01$  and  $\gamma_M = 0.095$  to compute regional human stocks for manufacturing and agriculture for all commuting zones in 1900 and 1920 according to equation (A.39).

**Discussion** Figure A.5 contrasts patterns in human capital and physical productivity across space. Panels (A) and (B) show the level and growth of human capital in agriculture and manufacturing, while Panels (C) and (D) report the corresponding measures of physical productivity. Human capital in manufacturing is strongly negatively correlated with initial agricultural employment shares in 1900, whereas agricultural human capital varies little across space. Because we estimate  $\gamma_A \approx 0$ , the agricultural human

FIGURE A.5: CATCH-UP GROWTH: PRODUCTIVITY VERSUS HUMAN CAPITAL



*Notes:* The figure compares human capital and productivity across commuting zones by initial agricultural employment share. Panels (A) and (B) report sector-specific human capital levels in 1900 and their average annual growth between 1900 and 1920, respectively. Human capital stocks are based on local educational attainment and sector-specific returns to schooling. Panels (C) and (D) plot the corresponding patterns for physical productivity. Physical productivity is measured as the productivity residuals from our calibration cleansed of the contribution of human capital.

capital is uncorrelated with the initial agricultural employment share. At the same time, human capital growth is not very spatially-biased. In manufacturing, human capital growth is somewhat faster in rural areas, reflecting educational catch-up. In contrast, initial productivity in both sectors is lower in more agricultural regions (Panel C), but rural areas exhibit substantially faster productivity growth over time (Panel D). Together, the figure shows that while human capital and productivity share some spatial patterns, rural productivity convergence remains quantitatively much more pronounced.

#### B.4 Details of Robustness Exercises

This appendix provides additional details on the robustness exercises discussed in Section 6 of the main text.

**Full Robustness Results** Table 3 reports the robustness of our results for wage growth and industrialization, corresponding to the patterns in Figure 7. For brevity, the main text focused only on the contributions of agricultural and manufacturing productivity

TABLE A.6: ESTIMATING THE LAND SUPPLY ELASTICITY  $\varrho$ 

	(1)	(2)	(3)	(4)	(5)	(6)
$\ln R_{rt}$	0.520*** (0.040)	1.588*** (0.390)	1.940*** (0.303)	0.372*** (0.058)	1.049* (0.602)	0.701 (2.411)
Fixed Effects	Year, CZ	Year, CZ	Year, CZ	Year, CZ	Year, CZ	Year, CZ
Emp. Weights	No	No	No	Yes	Yes	Yes
IV	No	Immigrants	Population	No	Immigrants	Population
R-squared	.901	-.598	-1.2	.936	-.552	.0525
Observations	1,485	1,485	1,485	1,485	1,485	1,485

Notes: The table reports estimates of the land supply elasticity  $\varrho$  from equation (A.43). All specifications include commuting zone and year fixed effects. Columns 1 and 4 report OLS estimates. Columns 2 and 5 instrument  $\ln V_{rt}$  with the log number of immigrants in region  $r$  at time  $t$ , while columns 3 and 6 use the log of total population as an instrument. Columns 4–6 re-estimate the corresponding specifications using population-weighted regressions. Standard errors are clustered at the commuting zone-year level.

growth ( $d \ln A_{rt}$  and  $d \ln Z_{rt}$ ). Figures A.7 and A.6 present the full decompositions. Figure A.7 reports results for wage growth (left panel of Figure 7) while Figure A.6 reports results for industrialization (right panel).

**Estimating the Elasticity of Land Supply  $\varrho$**  In our second robustness exercise, we allow for an elastic supply of agricultural land, modeled as  $T_{rt} = \bar{T} r R_{rt}^\varrho$ . Our baseline calibration assumes  $\varrho = 0$ ; here, we estimate  $\varrho$  using data on land values.

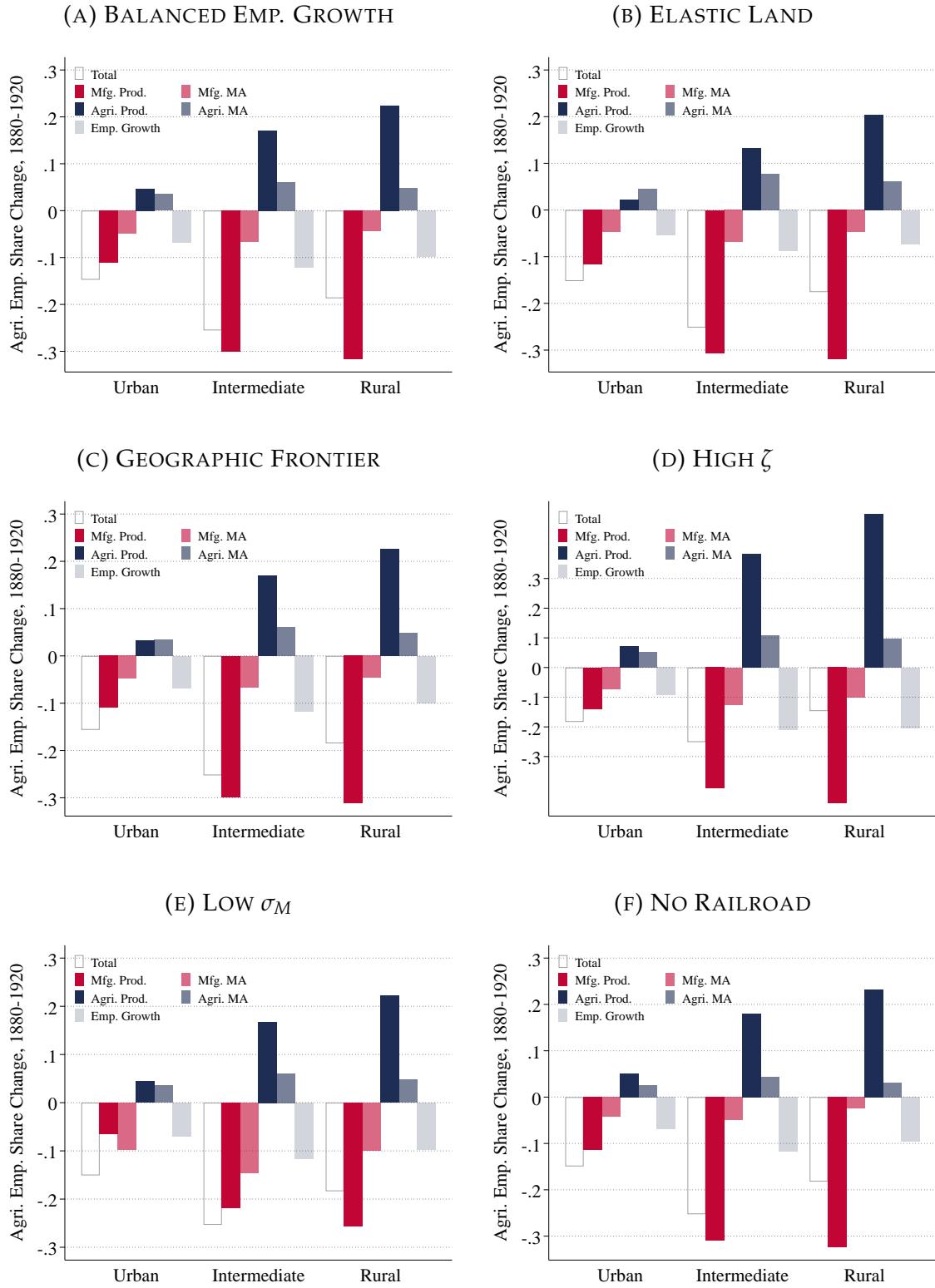
Let  $V_{rt}$  denote the value of agricultural land in region  $r$  at time  $t$ . We assume that land values are proportional to rental rates according to  $R_{rt} = V_{rt} m_r k_t u_{rt}$ , where  $m_r$  and  $k_t$  are region and time-specific wedges, and  $u_{rt}$  is an idiosyncratic error term capturing transitory deviations. Substituting this into the land supply equation yields

$$(A.43) \quad \ln T_{rt} = \delta_r + \delta_t + \varrho \ln V_{rt} + e_{rt},$$

where  $\delta_r$  and  $\delta_t$  are region and time fixed effects, and  $e_{rt} = \varrho \ln u_{rt}$ . If  $u_{rt}$  is uncorrelated with land values  $V_{rt}$  conditional on fixed effects, we can estimate  $\varrho$  using OLS. Otherwise, an instrumental variables strategy is required to isolate exogenous variation in  $V_{rt}$ .

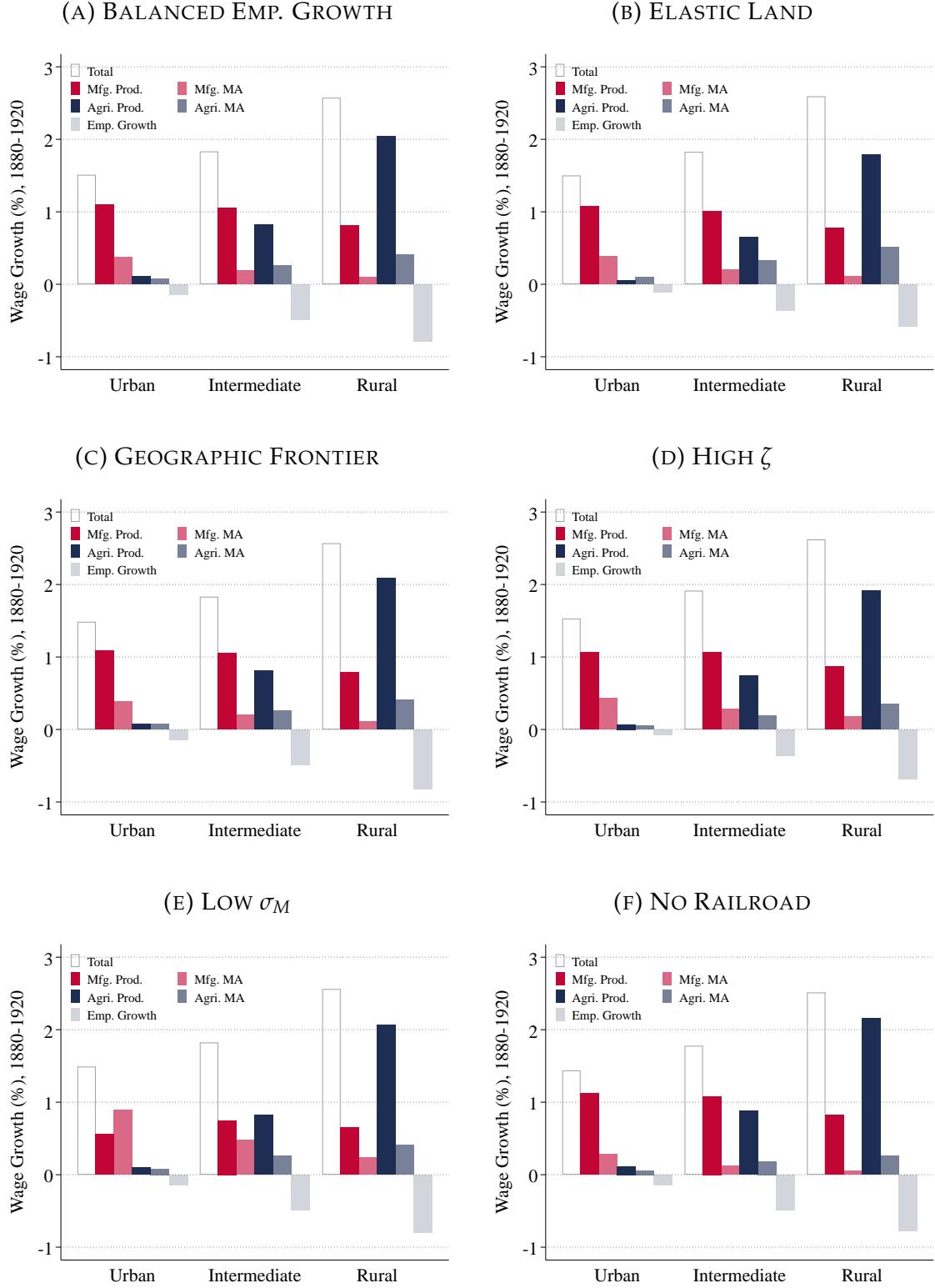
Table A.6 reports our estimates. Column 1 presents the OLS estimate, yielding  $\varrho \approx 0.5$ . Columns 2 and 3 implement two IV strategies, instrumenting  $\ln V_{rt}$  with either (i) the log number of immigrants or (ii) the log of total population in region  $r$  at time  $t$ . Both instruments plausibly shift the demand for land without directly affecting supply conditions captured by  $u_{rt}$ . As expected, the IV estimates are larger than the OLS estimate, with  $\varrho \approx 1.2$ . Columns 4–6 report population-weighted regressions. While the point estimates are similar, standard errors increase substantially, reflecting the noisier nature of the weighted specification. For our robustness analysis, we adopt a value of  $\varrho = 1.2$ , corresponding to the upper range of the IV estimates and providing a useful contrast to the baseline case of  $\varrho = 0$ .

FIGURE A.6: DECOMPOSITION ROBUSTNESS: INDUSTRIALIZATION



*Notes:* The figure decomposes local industrialization,  $ds_{rA}$ , into contributions from changes in market access ( $d \ln D_{rt}$ , "MA"), sectoral productivity growth ( $d \ln Z_{rt}$  and  $d \ln A_{rt}$ ), and local employment growth ( $d \ln L_{rt}$ ), each pre-multiplied by the corresponding exposure elasticities from Proposition 4. Urban (rural) locations are defined as commuting zones in the bottom (top) quartile of the 1880 agricultural employment share distribution; intermediate locations fall between these quartiles. Each panel shows a different model calibration relative to the baseline: balanced employment growth ( $n_t$ ), elastic land supply ( $\varrho = 1.2$ ), geographic frontier ( $\zeta = 5, \iota = 1$ ), high labor supply elasticity ( $\zeta = 4$ ), low substitution elasticity in manufacturing ( $\sigma_M = 4$ ), and fixed trade costs at 1880 levels ("no railroad,"  $\tau_{rrt} = \tau_{rr1880}$ ).

FIGURE A.7: DECOMPOSITION ROBUSTNESS: WAGE GROWTH



*Notes:* The figure decomposes local wage growth,  $d \ln \bar{w}_{rt}$ , into contributions from changes in market access ( $d \ln D_{rst}$ , "MA"), sectoral productivity growth ( $d \ln Z_{rt}$  and  $d \ln A_{rt}$ ), and local employment growth ( $d \ln L_{rt}$ ), each pre-multiplied by the corresponding exposure elasticities from Proposition 4. Urban (rural) locations are defined as commuting zones in the bottom (top) quartile of the 1880 agricultural employment share distribution; intermediate locations fall between these quartiles. Each panel shows a different model calibration relative to the baseline: balanced employment growth ( $n_t$ ), elastic land supply ( $\varrho = 1.2$ ), geographic frontier ( $\zeta = 5, \iota = 1$ ), high labor supply elasticity ( $\zeta = 4$ ), low substitution elasticity in manufacturing ( $\sigma_M = 4$ ), and fixed trade costs at 1880 levels ("no railroad,"  $\tau_{rrt} = \tau_{rr1880}$ ).