Spatial Structural Change[*]

Fabian Eckert[†] Michael Peters[‡]

February 7, 2023

Abstract

We propose a novel theory to study the relationship between local economic development and aggregate structural change. Two forces shape regional variation in wage growth and industrialization: technological catch-up, often associated with spatial convergence, and regional sectoral specialization leading to differences in exposure to aggregate reallocation. We study these forces in the US economy between 1880 and 1920 when its agricultural employment share fell from 50% to 25%, and regional convergence was strong. We show that technological catch-up saved rural America from the adverse consequences of its exposure to the agricultural decline; without catch-up, spatial inequality would have increased.

Keywords: Structural Transformation, Technological Change, Spatial Growth
JEL Codes: O11, O4, R11, R13

[*]We are grateful to Simon Alder, Costas Arkolakis, Timo Boppart, David Lagakos, Claudia Goldin, Doug Gollin, Joe Kaboski, Pete Klenow, 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 Matthew Jaremski for sharing his data on bank branches. We also thank seminar participants at Boston University, Stanford, Penn State, Harvard, 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.

[†]University of California, San Diego; `fpe@ucsd.edu`

[‡]Corresponding Author. Yale University and NBER; `m.peters@yale.edu`

INTRODUCTION

A key economic question is whether poor locations tend to grow faster than rich ones, leading to regional convergence. Popular narratives about the relationship between long-run economic growth and spatial economic development typically highlight one of two forces. On the one hand, the diffusion of technology might allow poor regions far from the technological frontier to benefit temporarily from catch-up growth. On the other hand, structural change systematically shifts aggregate demand across sectors, thereby hurting regions whose sectoral specialization makes them particularly exposed to declining industries. A large macroeconomic literature on regional convergence highlights the catch-up channel but is silent on exposure differences, whereas the structural change literature typically ignores the spatial dimension altogether.

In this paper, we quantify the importance of both channels during a period of rapid sectoral reallocation and unbalanced spatial growth: the first structural transformation of the US economy. Between 1880 and 1920, average incomes grew by 60%, and the agricultural employment share halved from 50% to 25%. At the same time, this period exhibited strong spatial convergence: wages grew substantially faster in regions that started out initially poor, and agricultural employment shares declined faster in more agricultural locations. This episode is a prime example of the value of studying regional productivity convergence and sectoral reallocation in a unified framework: because agricultural regions were poor in 1880, they were the natural beneficiaries of catch-up growth, but also particularly exposed to the declining agricultural sector.

We introduce a quantitative framework of *spatial structural change* tailored to such an analysis. Our theory rests crucially on recent advances in quantitative spatial economics that allow us to combine elements from the macroeconomic literature on structural change and regional convergence in a tractable, yet quantifiable framework. Borrowing from the structural change literature, our economy features two sectors, an agricultural and a non-agricultural, across which workers are not perfectly substitutable, making sectoral reallocation costly. Productivity growth in either sector makes the economy richer, and non-homothetic preferences imply that economic growth reduces the spending share on agricultural goods. We then embed these ingredients into a standard model of economic geography, where locations differ in their sectoral productivity, amenities, and the supply of agricultural land. Workers are spatially mobile subject to moving costs.

A central element of the model is the possibility of spatial productivity convergence. Our modelling follows the macroeconomic literature on cross-country convergence, which posits that a region's productivity growth depends directly on its distance to the technological frontier. The growth process in each location is thus parsimoniously described by two parameters: the growth rate of the technological frontier and the strength of catch-up growth, which regulates the growth premium of locations behind the frontier. As a result, individual locations and sectors can be on different growth trajectories at any point in time. Crucially, our theory links catch-up potential to regional differences in productivity within each sector and *not* to sectoral specialization. Whether agriculturally specialized locations benefit from productivity convergence therefore hinges on whether their productivity is low, relative to other locations.

Our theory permits an analytical characterization of the determinants of local wage growth and industrialization. In particular, we derive a concise representation of the two countervailing narratives of unbalanced productivity growth versus regional exposure. First, we show that in the absence of technological catch-up, the structural transformation *necessarily* generates urban-biased growth, that is, faster growth in industrialized, high-wage locations. The intuition is reminiscent of "Bartik"-like instruments: the secular reallocation away from agriculture is hurting regions with a comparative advantage in agricultural production whose specialization makes them particularly exposed to the agricultural decline. Second, our theory illuminates that rural convergence can be driven either by (exogenous) technological catch-up growth or by (endogenous) changes in market access and regional migration that benefit agriculturally specialized labor markets, highlighting the need to jointly analyze the spatial dimension of structural change and technological convergence.

To quantify the strength of catch-up growth and differential spatial exposure, we structurally estimate our model using time series and regional data for the US between 1880 and 1920. Our calibration strategy reflects the theory's two building blocks: a macro model of structural change and a spatial model of regional convergence. We calibrate the key parameters related to structural change to match classic aggregate time-series data: we choose the growth rates of the technological frontier in each sector and the preference parameters governing the non-homotheticity of preferences to match the time series of sectoral prices, GDP growth, and the agricultural employment share.

On the spatial side, we first use our model to infer each region's initial productivity in each sector in an unrestricted way from the joint distribution of local wages and sectoral employment shares in 1880 (while controlling for local employment and the availability of agricultural land in a model-consistent way). We find agriculturally specialized locations were – on average – behind the technological frontier in *both* sectors, providing them with the potential to catch up. We then use indirect inference to estimate the parameters governing the extent of catch-up growth and hence the evolution of productivity between 1880 and 1920. We do so by ensuring the model matches the empirical relationships between initial agricultural specialization and subsequent wage growth, changes in agricultural employment shares, and population flows across space.

Our estimates imply annual productivity growth in both sectors was roughly two percentage points higher in rural locations than in urban regions close to the technological frontier. Moreover, this rural productivity premium played a central role in the spatial convergence of wages. Both in the calibrated model and the data, the wage gap between industrialized labor markets and the rural hinterland shrunk by 40% between 1880 and 1920. By contrast, an alternative "macro-calibration" in which we shut down the possibility of regional catch-up growth shows the rural-urban wage gap would have *increased* by 15%. Hence, catch-up growth saved rural America from the adverse exposure effects of the structural transformation.

We also show technological catch-up in the two sectors played fundamentally different roles. Whereas faster productivity growth in agriculture explains why rural locations experienced faster wage growth, technological catch-up in non-agriculture was the main reason rural labor markets industrialized. The interaction between unbalanced productivity growth and regional differences in sectoral exposure is again central for this finding: precisely because of their agricultural exposure, rural wages are especially sensitive to agricultural productivity growth. But without catch-up growth in non-agricultural technology, rural locations would have increased their agricultural employment share rather than industrialized.

Given the central importance of rural catch-up growth, we also provide direct empirical evidence for possible mechanisms. Empirically, we document that various canonical development indicators, such as educational attainment, capital-deepening, firm size, financial development, and market integration via the expansion of the railroad network, grew substantially faster in rural America between 1880 and 1920. Whereas our theory summarizes these developments in a scalar measure of productivity in each location, these correlations paint a picture of a period in which a multitude of institutional and technological changes came together to systematically benefit remote, agricultural locations.

Although many aspects of our theory are specific to the transition out of agriculture, our framework also provides insights into the spatial incidence of the recent transition toward services, where spatial inequality has increased. Our analysis suggests changes in the potential for catch-up growth could be responsible. Whereas capital, schools, and railroad tracks might have been easy to move into rural regions, the *human* capital required in today's high-skill service production may be unwilling to settle in declining manufacturing towns outside big cities. As a result, today's divergent wage-growth patterns may reflect regional differences in exposure dominating the weakened convergence forces that remain.

Related Literature We contribute to the literature on structural change by combining elements of a standard macroeconomic model of structural change (e.g., Herrendorf, Rogerson, and Valentinyi (2014)) with recent advances in spatial economics (e.g., Allen and Arkolakis (2014), Redding and Rossi-Hansberg (2017)).[1] Most contributions in the structural change literature seek to explain the process of structural change at the aggregate level.[2] Notable exceptions are Caselli and Coleman II (2001), who use a stylized two-region model to highlight the link between structural change and regional convergence in the US, Nagy (2023), who examines the process of city formation in the US before 1860, and Michaels, Rauch, and Redding (2012), who study the empirical link between population density and population growth in the US in 1880.[3]

We also add to the classic macroeconomic literature on convergence across countries (see, e.g., Acemoglu, Aghion, and Zilibotti (2006) or Desmet, Nagy, and Rossi-Hansberg (2018)) and regions (see, e.g., Barro and Sala-i Martin (1991, 1992) or Blanchard, Katz, Hall, and Eichengreen (1992)). Our analysis highlights the need to explicitly model the spatial and sectoral links between local labor markets to consistently estimate spatial productivity convergence in the presence of secular sectoral reallocation.

Besides our particular question of interest, we also make a distinct theoretical contribution by showing how to tractably integrate a flexible class of non-homothetic preferences, the price-independent generalized linear class (PIGL), recently popularized by Boppart (2014), into a general-equilibrium trade and geography model. Its convenient aggregation properties makes the PIGL class not only a natural choice in our setting, but also potentially useful for other applications.

The paper is structured as follows. Section 1 documents the patterns of regional convergence that motivate our analysis. Section 2 contains our theory. We describe the calibration of our model in Section 3 and quantify the link between catch-up growth and rural convergence in Section 4.

1. Rural Convergence: 1880-1920

In this section, we document the empirical patterns of convergence in agricultural employment shares and averages wages across regions during the first structural transformation of the US between 1880 and 1920.[4] We use data from the full-count Decennial Census files and county-level tabulations of the Census of Manufacturing.

[1] The quantitative spatial literature has studied misallocation (Fajgelbaum, Morales, Suárez Serrato, and Zidar (2019)), trade liberalization (Fajgelbaum and Redding (2022), Caliendo, Dvorkin, and Parro (2019)), and market access (Redding and Sturm (2008)).

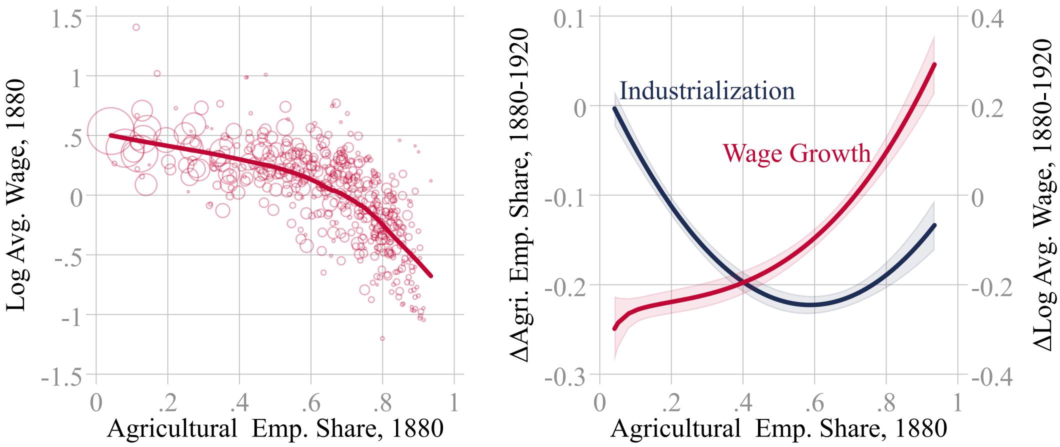
[2] See Kongsamut, Rebelo, and Xie (2001), Comin, Lashkari, and Mestieri (2021), and Boppart (2014) for papers on non-homothetic demand, and Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008) for papers that focus on the supply side.

[3] Recent papers study the spatial dimension of the transition toward services (see, e.g., Desmet and Rossi-Hansberg (2014), Eckert, Ganapati, and Walsh (2020a), or Fan, Peters, and Zilibotti (2022)) and structural change in developing countries (e.g., Pellegrina and Sotelo (2021), Sotelo (2020), Farrokhi and Pellegrina (2020)).

[4] We focus on this period due to data availability and to avoid the Great Depression.

FIGURE 1: SPATIAL STRUCTURAL CHANGE AND RURAL CATCH-UP

(A) RURAL POVERTY (B) RURAL CONVERGENCE



Notes: The left panel shows a scatter plot between commuting zones' agricultural employment shares and average earnings in 1880 and a Lowess fit line. The size of the points is proportional to the total workforce in each commuting zone. The right panel shows two fitted fractional polynomial curves along with 95% confidence intervals. They show the relationship between commuting zones' agricultural employment share in 1880 and (1) the change in the agricultural employment share between 1880 and 1920 (left axis) and (2) their average earnings growth between 1880 and 1920 (right axis) relative to the nationwide average. In fitting the polynomials, we weight by commuting zones' total employment in 1880.

We focus on states that had joined the Union by 1860, and aggregate county-level observations to constant-boundary "commuting zones" using the crosswalk by Eckert, Gvirtz, Liang, and Peters (2020b).[5]

The left panel of Figure 1 shows the correlation of agricultural specialization and average wages across US regions in 1880. The relationship is tight and strongly negative: agricultural specialization and poverty were almost synonymous in the 1880 US economy. In the right panel, we document the striking importance of rural convergence. The red line shows wages converged dramatically between 1880 and 1920: more agricultural regions showed substantially faster wage growth than less agricultural ones. Quantitatively, the urban-rural wage gap declined by 0.4 log points between 1880 and 1920.

The blue line shows rural locations also caught up in their employment structure: on average, they saw much faster declines in their agricultural employment share. However, this convergence in agricultural employment shares was not monotone but exhibited a distinct *U*-shape. The regions that industrialized the most were regions in an intermediate range of agricultural specialization. The typical commuting zone with an agricultural employment share of 60% in 1880 experienced a 20-percentage-point decline.

The patterns of rural convergence shown in Figure 1 are robust to changes in the spatial unit of observation and the inclusion of various fixed effects. This is seen in

[5]We discuss the data in more detail in Section 4.

TABLE 1: SPATIAL STRUCTURAL CHANGE AND RURAL CATCH-UP

	PANEL A: LOG AVERAGE WAGES IN 1880				
s_{rAt}	-0.820***	-0.794***	-0.930***	-0.778***	-0.753***
	(0.0268)	(0.0352)	(0.0533)	(0.0168)	(0.0218)
R^2	0.546	0.762	0.700	0.736	0.757
	PANEL B: WAGE GROWTH				
s_{rAt}	0.251***	0.357***	0.426***	0.331***	0.346***
	(0.0220)	(0.0381)	(0.0603)	(0.0192)	(0.0263)
R^2	0.839	0.855	0.732	0.718	0.712
	PANEL C: CHANGE IN AGRI. EMP. SHARE				
s_{rAt}	-0.484***	-0.466***	-0.372***	-0.380***	-0.383***
	(0.0279)	(0.0342)	(0.0693)	(0.0161)	(0.0208)
s^2_{rAt}	0.451***	0.427***	0.241***	0.346***	0.357***
	(0.0317)	(0.0386)	(0.0634)	(0.0188)	(0.0227)
R^2	0.309	0.362	0.170	0.234	0.316
Observations	990	990	990	3910	3910
Geography	CZ	CZ	CZ	County	County
FEs		State	State	State	CZ
Weighted	Yes	Yes		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 in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively.

Table 1, where we report the results of three regressions: (1) log wages on agricultural employment shares in 1880 (Panel A), (2) local wage growth on the initial agricultural employment share (Panel B), and (3) changes in agricultural employment shares on initial agricultural employment shares and their square (Panel C).

For all regressions, we report our baseline results, corresponding to Figure 1, in column 1. In columns 2 and 3, we document that these results are not driven by weighting commuting zones by their size and are robust to the inclusion of state fixed effects. In columns 4 and 5, we perform the same analysis at the county level with state and commuting-zone fixed effects, respectively. Across all of these specifications, we see the same pattern of spatial convergence shown in Figure 1: the period of US industrialization between 1880 and 1920 was a time of distinct regional integration. Initially poor agricultural regions caught up to more industrialized regions in terms of wages, and industrialization exhibited a U-shape as a function of initial agricultural specialization.

We focus much of the analysis in the rest of the paper on the relationship between initial agricultural employment shares and subsequent wage growth and industrialization. However, in our quantitative analysis, we also incorporate data on employment growth. Michaels et al. (2012) document a U-shaped relationship between initial agricultural specialization and subsequent population growth — a finding we confirm below. Any attempt to connect a quantitative theory with data on employment growth in the historic US has to grapple with the fact that much of it occurred for reasons outside most theoretical models. In our case, the vast inflows of foreign immigrants (who predominantly settled in cities), the large discrepancy in local fertility rates (that were much higher in rural areas), and the fact that the US territory was still expanding are of particular importance. Below, we provide an explicit methodology to account for regional employment growth due to such exogenous factors without explaining their determinants within our theory.

Next, we present a theory of spatial structural change that can speak to the patterns of convergence documented in Figure 1.

2. Theory

Our theory of spatial structural change uses the workhorse model of economic geography (see, e.g., Redding and Rossi-Hansberg (2017)) to combine a macroeconomic theory of structural change (see, e.g., Herrendorf et al. (2014)) with insights from the literature on cross-country convergence (see, e.g., Acemoglu et al. (2006) or Desmet et al. (2018)). We provide detailed derivations in Section 2 of the Appendix.

2.1 Preferences, Technology, and Labor Supply

The economy consists of a set of discrete locations, indexed by $r = 1,...,R$, and two sectors, agriculture and non-agriculture, indexed by $s = A,M$, respectively. At time t, the economy is inhabited by a mass $\bar{L}_t$ of workers. We suppress time subscripts when describing the static elements of our model.

Preferences Individuals value the consumption of agricultural and non-agricultural goods. Preferences for these sectoral outputs are non-homothetic to generate the shifts in sectoral demand associated with the structural transformation. Following Boppart (2014), we assume preferences fall in the non-homothetic PIGL (Price-Independent Generalized Linear) class. As we show in detail in Section 2.3, these preferences have convenient aggregation properties that make them a natural choice for models of trade and economic geography.

PIGL preferences do not have an explicit utility representation but are defined implicitly via the indirect utility function. We parametrize the indirect utility of an agent with expenditure y facing prices (P_{rA}, P_{rM}) as

$$V\left(y, P_{rA}, P_{rM}\right) = \frac{1}{\eta}\left(\frac{y}{P_{rA}^{\phi} P_{rM}^{1-\phi}}\right)^{\eta} - \nu \ln\left(\frac{P_{rA}}{P_{rM}}\right), \tag{1}$$

where $\eta, \phi \in (0,1)$.

For now, we assume trade costs are zero for the agricultural good. Doing so allows us to treat the agricultural good as the numeraire, that is, $P_{rA} = P_A = 1$, and to simplify the notation. Our quantitative exercise below features trade costs in both sectors.

Applying Roy's identity yields the following expression for an individual's expenditure share on the agricultural good:

$$\vartheta_A(y, P_M) = \phi + \nu \left(y / P_{rM}^{1-\phi} \right)^{-\eta}. \quad (2)$$

Equation (2) shows the demand system is akin to a Cobb-Douglas specification with a non-homothetic adjustment. Conveniently, the term $y/P_{rM}^{1-\phi}$, which we also sometimes refer to as "real income," emerges as a summary statistic for such non-homotheticities. Since $\eta > 0$, consumers reduce their relative agricultural spending as they grow richer as long as $\nu > 0$. Moreover, the expenditure share asymptotes to ϕ as incomes grow large. If $\nu = 0$ and $\eta = 1$, equation (1) reduces to a Cobb Douglas utility function with constant expenditure shares.[6] We refer to the elasticity parameter η as the "Engel elasticity" because it determines the shape of consumers' Engel curves. The larger the Engel elasticity, the stronger the effect of real income on consumer demand.[7]

Technology Each region can produce agricultural and non-agricultural goods. A representative local firm produces the agricultural good using the following technology:

$$Y_{rA} = Z_{rA} H_{rA}^{1-\alpha} T_r^{\alpha},$$

where Z_{rA} is the local productivity in agriculture, H_{rAt} is agricultural labor (measured in efficiency units), and T_r denotes agricultural land. We assume agricultural land is in fixed supply in each region. As result, the land share, α, indexes the strength of decreasing returns to scale.

We model the non-agricultural sector in the standard "CES-monopolistic-competition" way. Individual firms pay a fixed cost of entry, f_E, denoted in units of non-agricultural labor. Upon entering, each firm produces a differentiated variety, indexed by ω, using the same constant-returns-to-scale, labor-only production technology with productivity Z_{rM}. Firms operate for a single period, which we define as 20 years in our empirical application. We assume free entry, so new firms enter until their profits equal their fixed costs. Total demand for non-agricultural labor in region r, H_{rM}, is therefore the sum of entry and production labor, H_{rE} and H_{rP}, respectively. The market for non-agricultural varieties is monopolistically competitive.

In each location, a representative firm assembles the differentiated non-agricultural

[6] In our quantitative application, we choose the level of regional productivity to ensure expenditure shares are between 0 and 1. This amounts to assuming consumers are sufficiently rich to be willing to consume non-agricultural goods in positive quantities.

[7] The elasticity of substitution between the value added generated in the two sectors is given by $\varrho = 1 + \eta(\vartheta_A - \phi)^2 / (\vartheta_A(1 - \vartheta_A))$. Hence, it is not a structural parameter but varies across space and the income distribution. Note ϱ is increasing in ϑ_A (i.e., decreasing in real income) and satisfies $\lim_{\vartheta_A \to \phi} \varrho = 1$.

varieties into a final consumption good:

$$Y_{rM} = \left(\int_0^N y(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} = \left(\sum_{j=1}^{R} \int_0^{N_j} y(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}.$$

Here, N is the total number of varieties available and N_r denotes the number of varieties produced in region r. Non-agricultural varieties are subject to the usual iceberg trade costs. The presence of such trade costs implies that the composition and price of the final non-agricultural good differs across locations.

Sectoral Labor Supply Structural change exerts pressure on local economies to reallocate labor across industries. Workers' ability to move out of agriculture depends on the extent to which their skills are substitutable across sectors. To capture this reallocation margin, we model sectoral labor supply using the typical Roy-type machinery.

An individual worker i in region r can supply z_s^i efficiency units to sector s that are drawn from a sector-specific Fréchet distribution, $P\left(z_s^i \leq z\right) = F_s(z) = e^{-z^{-\zeta}}$. The parameter ζ captures the dispersion of efficiency units across workers in sector s.

We denote total payments per efficiency unit of labor in region r and sector s by w_{rs} and assume the payments to agricultural land in a location are distributed to local agricultural workers and included in w_{rA}. Each worker i chooses a sector of employment to maximize their income, y_r^i, so that $y_r^i = \max_s \left\{ z_s^i w_{rs} \right\}$. As a result, the income distribution in each location inherits the Fréchet distribution of the underlying efficiency units, that is,

$$F_r(y) = e^{-(y/\overline{w}_r)^{-\zeta}} \quad \text{where} \quad \overline{w}_r = \left(w_{rA}^{\zeta} + w_{rM}^{\zeta} \right)^{1/\zeta},$$

where the term $\overline{w}_r$ denotes average earnings in region r. Similarly, sectoral employment shares and aggregate labor supply are given by:

$$s_{rs} = \left(w_{rs}/\overline{w}_r\right)^{\zeta} \quad \text{and} \quad H_{rs} = \Gamma_{\zeta} L_r \left(w_{rs}/\overline{w}_r\right)^{\zeta-1}, \tag{3}$$

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

Equation (3) highlights that ζ governs the sectoral-labor-supply elasticity: the higher ζ, the higher the elasticity of labor supply. As $\zeta \to \infty$, the heterogeneity in efficiency units disappears and labor is fully elastic across industries. This limiting case is the benchmark of most macroeconomic models of the structural transformation. We show below that the parameter ζ is a crucial determinant of the spatial exposure to sectoral reallocation.

Spatial Mobility At the beginning of each period, workers can move to another location. We denote the distribution of workers across regions at the beginning and end of a period by $\{L^Y_{rt}\}_r$ and $\{L_{rt}\}_r$, respectively.

We assume workers learn their labor productivity in each sector only after arriving at a destination. The indirect utility of worker i from location r in location r' at time t is thus given by

$$\mathcal{U}^i_{rr't} \equiv \mathcal{V}_{rt}\mathcal{B}_{rt}\mu_{rr'}u^i_{rt}, \text{ s.t. } \mathcal{V}_{rt} \equiv \int V(y, p_{rt})\, dF_{rt}(y) \text{ and } \mathcal{B}_{rt} = B_r L_{rt}^{-\rho}.$$

The term $\mathcal{V}_{rt}$ denotes expected consumption utility reflecting a worker's uncertainty about the efficiency units of labor drawn upon arrival in region r. In Section 2.3 below, we derive a closed-form expression for $\mathcal{V}_{rt}$. The term $\mathcal{B}_{rt}$ is an amenity term, which comprises an exogenous and endogenous part. The parameter $\rho > 0$ indexes the strength of congestion forces such as the scarcity of local housing or other non-traded goods. The matrix $\mu_{rr'} \in (0, 1]$ reflects the cost of moving: destination utility is discounted depending on a worker's region of origin. We assume workers who stay put enjoy the full local utility, that is, $\mu_{rr} = 1$. Finally, u^i_{rt} reflects a worker-location-specific preference shifter, which is drawn prior to choosing a region, i.i.d. from a Fréchet distribution with shape parameter ε.

Using standard properties of the Fréchet distribution, the share of workers moving from location r to r' can be written as

$$m_{rr't} = \frac{\left(\mu_{rr'}\mathcal{V}_{r't}\mathcal{B}_{r't}\right)^{\varepsilon}}{\sum_j \left(\mu_{rj}\mathcal{V}_{jt}\mathcal{B}_{jt}\right)^{\varepsilon}}. \tag{4}$$

In addition to internal migration, we also allow for changes in the local labor force that are not modelled explicitly. In particular, international immigration was substantial during the time period of our study, and local birth rates varied considerably. To capture these factors, we follow Cruz and Rossi-Hansberg (2021) and allow for an exogenous component of employment growth, n_{rt}, that links the beginning-of-period distribution of workers, $\{L^Y_{rt}\}_r$, to the end-of-period workforce of the previous period, $\{L_{rt-1}\}_r$, according to $L^Y_{rt} = n_{rt-1}L_{rt-1}$. As a result, the law of motion for local employment takes the form

$$L_{rt} = \sum_{r'} m_{r'rt}L^Y_{r't} = \sum_{r'} m_{r'rt}n_{r't-1}L_{r't-1},$$

where $m_{r'rt}$ is given in equation (4). The size of region r is thus determined by its relative attractiveness $(m_{r'rt})$, its size in the past (L_{rt-1}), and exogenous employment growth (n_{rt-1}).

2.2 Catch-Up Growth and Productivity Convergence

A key aspect of our theory is the possibility of *catch-up growth*. To model the potential of a location to improve its productivity (relative to others), we follow a large macroeconomic literature on cross-country convergence, which posits that productivity growth depends on the level of productivity relative to the productivity frontier; see, for example Acemoglu et al. (2006), Akcigit, Alp, and Peters (2021), and Desmet et al. (2018).

Specifically, we adopt a parsimonious parametrization of the region- and sector-specific productivity terms Z_{rA} and Z_{rM}. Let $\overline{Z}_{st}$ denote a common sector-specific productivity shifter that grows at the constant rate g_s. We assume $Z_{rst} \leq \overline{Z}_{st}$, and hence also refer to $\overline{Z}_{st}$ as the *sectoral frontier*. In our application, we take $\overline{Z}_{st}$ to be the highest productivity in sector s in the US.

We thus model the evolution of region r's productivity in sector s as

$$d \ln Z_{rst} = g_s + \lambda_s \ln \left(\frac{\overline{Z}_{st}}{Z_{rst}} \right) \quad \text{for } s = A, M. \tag{5}$$

The productivity process in equation (5) exhibits different spatial biases depending on the value of a single parameter, λ_s. If $\lambda_s = 0$, sectoral productivity grows at the same rate in all regions and the spatial productivity distribution in sector s is stationary. If $\lambda_s > 0$, less productive regions benefit from their backwardness and grow at a faster rate. If $\lambda_s < 0$, the opposite is the case and technologically backward locations fall further behind.[8]

Our interpretation of Z_{rAt} and Z_{rMt} is intentionally broad. A region's "benefit of backwardness" could be due to actual spatial technology diffusion, where lagging localities adopt existing techniques and catch up to the technological frontier. But catch-up growth could also be driven by infrastructure investments, capital deepening, or other institutional changes that spatially diffuse with a time lag and reach less productive locations at later stages of economic development. In Section 4.3 below, we provide direct empirical evidence for this pattern of catch-up growth for a variety of development indicators.

Importantly, equation (5) does *not* hardwire any specific relationship between local productivity *growth* and the current level of sectoral specialization. Whether agriculturally specialized locations experience faster growth depends on *why* they specialize in the agricultural sector. If agricultural locations have, on average, lower physical productiv-

[8] If $\lambda_s > 0$, equation (5) implies regional productivity differences disappear in the long run. This assumption is for simplicity only. Suppose equation (5) were given by $d \ln Z_{rst} = g_s + \lambda_s \ln \left(\overline{Z}_{st} / Z_{rst} \right) - \mu_{rs}$, where $\mu_{rs} \geq 0$. Then, $Z_{rst} \to e^{-\mu_{rs}/\lambda_s} \overline{Z}_{st}$.For the case of $\mu_{rs} = 0$, we recover $Z_{rst} \to \overline{Z}_{st}$. In our empirical application, which covers a 40-year period, this long-run result is not consequential.

ity Z_{rAt} or Z_{rMt}, they benefit from catch-up growth. However, a comparative advantage in agriculture is also consistent with an absolute advantage in both sectors or could be entirely due to an abundance of agricultural land T_r. As a result, our model does not mechanically produce a systematic relationship between sectoral specialization and future growth but delivers it as an outcome of our structural estimation.

For simplicity, we assumed the diffusion process in (5) does not reflect any geographic attributes. For example, we could have assumed productivity growth in location r depends on the productivity gap and the geographical distance from the technological frontier. However, local productivity growth will be correlated across labor markets if the initial cross-sectional distribution of productivity, $\{Z_{rst}\}_{rs}$, is spatially correlated. Similarly, we take the process in (5) as exogenous and structurally estimate g_s and λ_s. Although microfounding (5) would be interesting (e.g., in the spirit of Acemoglu et al. (2006)), we focus instead on understanding the quantitative implications of catch-up growth rather than its fundamental source.

2.3 Aggregate Demand and Spatial Welfare

To compute the equilibrium, we need to characterize workers' expected utility $\mathcal{V}_{rt}$ and the aggregate demand system. As we show in Section A.2.3 in the Appendix, the combination of PIGL preferences and the Fréchet distribution of individual income allows us to derive closed-form expressions for these objects, despite the fact that consumer demand is non-homothetic.

First, the *aggregate* expenditure share on agricultural goods in region r, ϑ_{rA}, is given by

$$\vartheta_{rA} \equiv \frac{\int \vartheta_A (y, p_r) y dF_r(y)}{\int y dF_r(y)} \quad = \quad \phi + \nu^{\mathrm{RC}} \left(\overline{w}_r / P_{rM}^{1-\phi}\right)^{-\eta},$$

where $\nu^{RC} = \nu \frac{\Gamma_{\zeta/(1-\eta)}}{\Gamma_\zeta}$ is a composite parameter that depends on the underlying micro preference parameter ν, the second moment of the income distribution ζ, and the Engel elasticity η. Hence, the aggregate demand system is akin to the one generated by a representative agent who earns the average wage, $\overline{w}_{rt}$, and has a preference parameter ν^{RC}.

Importantly, the aggregate demand system is still non-homothetic: an increase in average income reduces the aggregate spending share on agricultural goods. As we show in Section 2.5, such demand shifts put downward pressure on wages in agriculturally specialized locations, making aggregate growth urban-biased.

Second, we can also derive an intuitive expression for $\mathcal{V}_r$:

$$\mathcal{V}_r = \int V(y, p_r) dF_r(y) = \frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \left(\overline{w}_r / P_{rM}^{1-\phi}\right)^{\eta} - \nu \ln (1/P_{rM}). \tag{6}$$

Expected utility in region r resembles the indirect utility of a representative agent who earns average income $\overline{w}_{rt}$ and has a "taste" parameter $\Gamma_{\zeta/\eta}$ determining the relative importance of real income and relative prices.

2.4 Equilibrium Wages and Equilibrium Industrialization

Our theory permits an explicit characterization of equilibrium wages and agricultural employment shares across space. These two objects are our main outcomes of interest, and we relegate a discussion of the full equilibrium system to Appendix A.1. The equilibrium of our model is defined as follows:

Definition. *Let* $\{L_{r0}, Z_{rA0}, Z_{rM0}\}_r$ *be the initial distribution of workers and productivity, and let* $\{\overline{Z}_{At}, \overline{Z}_{Mt}\}_t$ *be a path of the technological frontier. An equilibrium is a sequence of prices* $\{P_{rAt}, P_{rMt}\}_{rt}$*, wages* $\{w_{rAt}, w_{rMt}\}_{rt}$*, rental rates* $\{R_{rt}\}_{rt}$*, non-agricultural varieties* $\{N_{rt}\}_{rt}$*, employment allocations* $\{H_{rAt}, H_{rEt}, H_{rPt}\}_{rt}$*, local employment* $\{L_{rt}\}_{rt}$*, individual consumption* $\{c^i_{rAt}, [c^i_{rMt}(\omega)]_\omega\}^i_{rt}$*, and productivity processes* $\{Z_{rAt}, Z_{rMt}\}_{rt}$*, 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 law of motion (5).*

To derive an explicit formulation of equilibrium wages and employment shares, we exploit a convenient representation of aggregate revenue in the non-agricultural sector. Under free entry, the mass of firms is proportional to non-agricultural production labor who receive a fixed fraction of sectoral revenue. As a result, a location's non-agricultural revenue, $\mathcal{R}_{rM}$, is given by

$$\mathcal{R}_{rM} = \tilde{f}_E \mathcal{D}_r^{\frac{1}{\sigma}} Z_{rM}^{\frac{\sigma-1}{\sigma}} H_{rM}, \text{ where } \quad \mathcal{D}_r \equiv \sum_j \tau_{rjM}^{1-\sigma} P_{jM}^{\sigma-1} \vartheta_{jM} \Gamma_\zeta L_j \overline{w}_j, \tag{7}$$

where $\mathcal{D}_r$ is a measure of the effective *demand* for non-agricultural products in region r and $\tilde{f}_E$ is an inconsequential composite constant. Hence, non-agricultural revenue takes the form of a constant-returns-to-scale production function, with revenue TFP being a combination of physical productivity, Z_{rM}, and the endogenous demand term, $\mathcal{D}_r$. The presence of $\mathcal{D}_r$ highlights the link between structural change and sectoral revenue productivity: as incomes rise and spending shifts toward non-agricultural goods, revenue productivity in the non-agricultural sector increases.

Using the representation in equation (7), local wages and employment shares can be expressed in the following way.

Proposition 1. *Let* $\ell_r \equiv L_r/T_r$ *denote employment density in region* r *and define the following* *"effective" sectoral productivity terms in region r:*

$$\mathcal{Z}_{rM} \equiv Z_{rM}^{\frac{\sigma-1}{\sigma}} \tilde{f}_E^{-1} \mathcal{D}_r^{\frac{1}{\sigma}} \quad \text{and} \quad \mathcal{Z}_{rA} \equiv Z_{rA} \left(\Gamma_\zeta \ell_r\right)^{-\alpha}. \tag{8}$$

Local average wages $\overline{w}_{rt}$ *and agricultural employment shares* s_{rAt} *are then determined by*

$$1 = \left(\frac{\mathcal{Z}_{rM}}{\overline{w}_r}\right)^{\zeta} + \left(\frac{\mathcal{Z}_{rA}}{\overline{w}_r}\right)^{\frac{\zeta}{\alpha(\zeta-1)+1}}; \qquad \frac{s_{rA}^{1+(\zeta-1)\alpha}}{1-s_{rA}} = \left(\frac{\mathcal{Z}_{rA}}{\mathcal{Z}_{rM}}\right)^{\zeta}. \tag{9}$$

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

Proposition 1 shows local wages and sectoral specialization are fully determined from two sufficient statistics, $\mathcal{Z}_{rM}$ and $\mathcal{Z}_{rA}$, that we refer to as "effective" sectoral productivity. Whereas both $\mathcal{D}_r$ and ℓ_r are endogenous and intrinsically linked to the way locations spatially interact on the market for goods ($\mathcal{D}_r$) and in terms of inter-regional migration (ℓ_r), Proposition 1 shows that as far as wages and sectoral specialization are concerned, they are isomorphic to physical sectoral productivity Z_{rs}.

Proposition 1 is important both conceptually and in terms of its measurement implications. On the conceptual side, it highlights that local growth and industrialization can be driven by three distinct channels: growth in physical productivity Z_{rs}, changes in employment density ℓ_r through migration or employment growth, and shifts in non-agricultural demand $\mathcal{D}_r$. As such, Proposition 1 highlights the value of jointly studying sectoral reallocation and technological catch-up growth, because both shape the spatial distribution of wages and sectoral specialization.

At the same time, Proposition 1 also highlights an important measurement challenge. To identify the spatial distribution of physical productivity Z_{rs} from data on wages and employment shares, taking into account the spatial linkages between locations is important. For example, a region can have a comparative advantage in agriculture either because of high relative productivity Z_{rA}/Z_{rM} or because of abundant land supply and little non-agricultural demand. Similarly, high wages can either reflect physical productivity or employment density and market access.

Whether regional wages and employment shares reflect physical productivity or differences in employment density or demand is also important to understand regions' ability to benefit from catch-up growth. If differences in physical productivity drive most of the variation in $\mathcal{Z}_{rM}$, rural locations benefit from catch-up growth because their Z_{rM} is low by virtue of them being specialized in the agricultural sector and being poor. If, by contrast, most of the variation in revenue productivity is due to market access, $\mathcal{D}_r$, the potential for a rural location to benefit from productivity convergence is limited. Similarly, if most of the variation in agricultural effective productivity, $\mathcal{Z}_{rA}$, is driven by differences in employment densities, ℓ_r, the agricultural productivity distribution is compressed, and the potential for catch-up growth is minimal. Within the context of our structural model, we can separately identify Z_{rs} from $\mathcal{D}_r$ and ℓ_r and therefore estimate the extent of catch-up growth in a model-consistent way.

In addition, equation (9) also stresses that wages depend on the substitutability of workers across sectors (ζ) and the agricultural land intensity (α). The reason is that decreasing returns in agriculture imply that the marginal product of labor depends on the quantity of agricultural labor. To see this directly, note the sectoral factor prices w_{rM} and w_{rA} can be expressed as follows:

$$w_{rM} = \mathcal{Z}_{rM}; \qquad 1 = \left(1 + \left(\frac{\mathcal{Z}_{rM}}{w_{rA}}\right)^{\zeta}\right)^{\frac{\zeta-1}{\zeta}} \left(\frac{\mathcal{Z}_{rA}}{w_{rA}}\right)^{\frac{1}{\alpha}}.$$

Non-agricultural wages depend only on Z_{rM} and $\mathcal{D}_r$ and are independent of sectoral labor supply. By contrast, wages in the agricultural sector respond to sectoral labor supply, which is reflected in their dependence on the effective productivity of the local non-agricultural sector, $\mathcal{Z}_{rM}$, with which it competes for workers. The sectoral supply elasticity ζ appears because it shapes how much agricultural wages have to rise to lure workers away from non-agriculture. Decreasing returns are hence the reason why the employment density appears in Proposition 1 and why the elasticity of substitution across *sectors* helps shape the *spatial* distribution of wages and employment shares.

2.5 Rural Convergence: Incidence vs. Exposure

Proposition 1 highlighted the static determinants of the cross-section of wages and employment shares across regions and how they depend on the same two effective productivity terms. We now leverage these results to study how they translate into *changes* in wages and agricultural employment shares across space and what the empirical patterns documented in Section 1 reveal about the underlying mechanisms driving these changes.

Consider a single region r that takes aggregate prices as given. The following Proposition describes the determinants of local wage *growth* and local industrialization, that is, *changes* in the agricultural employment share.

Proposition 2. *Local wage growth and local industrialization are given by*

$$\begin{aligned} d\ln \overline{w}_{rt} &= \phi(s_{rA})\, d\ln \mathcal{Z}_{rMt} + (1-\phi(s_{rA}))\, d\ln \mathcal{Z}_{rAt} \\ ds_{rAt} &= \psi(s_{rA})\,(d\ln \mathcal{Z}_{rMt} - d\ln \mathcal{Z}_{rAt}), \end{aligned}$$

where the two exposure elasticities are given by

$$\phi_r \equiv \phi(s_{rA}) = \frac{(\gamma+1)(1-s_{rA})}{\gamma(1-s_{Ar})+1}; \qquad \psi_r \equiv \psi(s_{rA}) = -\frac{s_{rAt}(1-s_{rAt})\zeta}{\gamma(1-s_{rAt})+1}, \tag{10}$$

with $\gamma \equiv \alpha(\zeta-1)$. *The regional incidence of effective productivity growth can be decomposed as*

$$d\ln \mathcal{Z}_{rMt} = \frac{\sigma-1}{\sigma} d\ln Z_{rMt} + \frac{1}{\sigma} d\ln \mathcal{D}_{rt}; \qquad d\ln \mathcal{Z}_{rAt} = d\ln Z_{rAt} - \alpha d\ln \ell_{rt}.$$

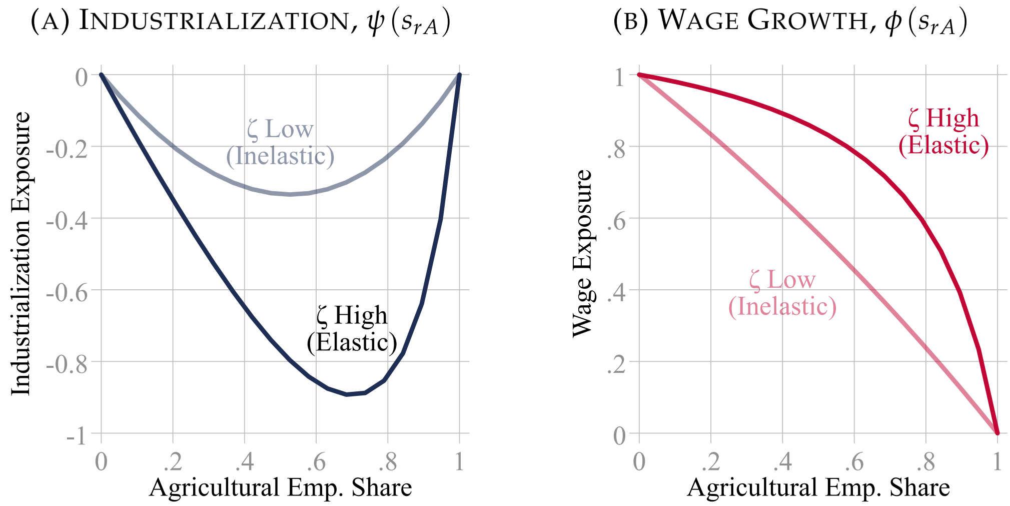
Proof. See Section A.3.2 in the Appendix. □

Proposition 2 highlights that local wage growth and industrialization vary across space for two reasons. First, regions differ in their *exposure* to changes in effective productivity, and the agricultural employment share s_{rA} emerges as the sufficient statistic for the regional heterogeneity in exposure. Second, effective productivity itself might grow faster in some regions than in others, $d\ln \mathcal{Z}_{rst} \neq d\ln \mathcal{Z}_{r'st}$; that is, locations can differ in the *incidence* of growth.

Proposition 2 thus formalizes the two narratives about spatial growth we highlighted in the beginning: unbalanced productivity growth versus regional specialization in declining industries. At the same time, it highlights that the relevant notion of local productivity growth does not merely encompass technological efficiency but that it captures all factors that influence effective productivity. As a result, differences in the incidence of growth could be due to (i) technological catch-up ($d\ln Z_{rMt}$ and $d\ln Z_{rAt}$), as well as (ii) local employment growth $(d\ln \ell_{rt})$ and (iii) differential changes in non-agricultural demand $(d\ln \mathcal{D}_{rt})$.

The regional differences in exposure are summarized by the two exposure elasticities $\phi_r \in (0,1)$ and $\psi_r \in (-1,0)$ shown in Figure 2. Growth in the average wage is a linear combination of effective productivity growth in each sector, with the weight of the non-agricultural sector given by ϕ_r. As shown in the right panel, $\phi'(\cdot) < 0$; that is, industrial areas benefit especially from non-agricultural effective productivity growth and rural locations particularly from effective productivity growth in agriculture. The left panel shows that the industrialization elasticity, ψ_r, is a U-shaped function of the agricultural employment share. Changes in comparative advantage, $\mathcal{Z}_{rMt}/\mathcal{Z}_{rAt}$, therefore induce industrialization everywhere, but especially at intermediate levels of agricultural specialization. Intuitively, the most urban locations cannot reduce their agricultural employment share, because they already are almost fully industrialized, whereas the most rural counties have such a strong comparative advantage in the agricultural sector that labor reallocation is limited.

FIGURE 2: SPATIAL HETEROGENEITY IN EXPOSURE



Notes: The figure shows the exposure elasticities $\phi(s_{rA})$ and $\psi(s_{rA})$ given in Proposition 2 as a function of the agricultural employment share. We depict the case of relative inelastic supply (low ζ) as a darker line and the case of relative elastic supply (high ζ) as a more lightly shaded line.

Overall, Proposition 2 highlights that the *sectoral* origins of growth have direct *spatial* implications, similar to the logic of "Bartik"-style instruments. But it goes beyond that by showing the supply elasticity ζ is a key determinant of the strength of such exposure effects. The reason is that ζ captures the ease of sectoral reallocation and hence the ability of local labor markets to adjust to changes in the economic environment. Expectedly, the left panel shows that the higher the sectoral labor elasticity, the stronger the sectoral reallocation induced by changes in comparative advantage and the more pronounced the U-shape. Furthermore, the right panel shows non-agricultural productivity growth becomes a more important determinant of local wage growth, the higher ζ. Differences in exposure *across* locations are thus a symptom of the imperfect sectoral substitutability of workers *within* locations.[9]

Figure 2 and the empirically observed U-shaped pattern of industrialization in Figure 1 therefore already hint at the importance of exposure forces in shaping the spatial pattern of industrialization. At the same time, the strong pro-rural pattern of wage growth points to growth in the effective productivity in agriculture, $\mathcal{Z}_{rAt}$, to which agricultural regions are most exposed.

To understand why our finding of rural-biased wage growth implies an important role for heterogeneity in incidence, suppose all locations experienced the same rate of effective productivity growth. This case would arise in the absence of catch-up ($\lambda_A = \lambda_M = 0$) and regional migration ($d \ln \ell_{rt} = 0$), and if trade were free ($\mathcal{D}_{rt} = \mathcal{D}_t$). Letting ι_A and ι_M denote the common growth rates of $\mathcal{Z}_{rAt}$ and $\mathcal{Z}_{rMt}$, Proposition 2

[9] Note that in the limit, as labor becomes freely substitutable across sectors, the regional heterogeneity in wage exposure disappears entirely, that is, $\lim_{\zeta \to \infty} \phi_r = 1$.

implies

$$d\ln\overline{w}_{rt} = \iota_A + \phi\left(s_{rA}\right)\left(\iota_M - \iota_A\right); \qquad ds_{rAt} = \psi\left(s_{rA}\right)\left(\iota_M - \iota_A\right). \tag{11}$$

Because $\psi\left(s_{rA}\right) < 0$, the agricultural employment share declines if and only if $\iota_M > \iota_A$. This, however, also implies wage growth in rural locations should have been *lower* because $\phi'\left(s_{rA}\right) < 0$. Hence, if effective productivity growth had been balanced, the sectoral reallocation away from agriculture should have led to *urban-biased* growth and to a rise in wage inequality. This scenario is, of course, sharply at odds with the dramatic extent of rural wage convergence documented in Section 1.

To rationalize the observed pattern of rural-biased wage growth, agricultural labor markets must have experienced faster growth in effective productivity. Proposition 2 highlights such faster wage growth can be achieved in three ways: (i) rural out-migration, so that employment density ℓ_{rt} falls in rural locations; (ii) higher growth in non-agricultural demand $\mathcal{D}_{rt}$ in former agricultural locations; and (iii) faster physical productivity growth Z_{rMt} and Z_{rAt} through catch-up. Separately identifying these channels requires a structural model. As we show in our quantitative exercise, the productivity channel via catch-up played the dominant role in saving rural America. In its absence, rural labor markets would have fallen further behind and living standards would have diverged. Hence, both exposure and incidence differences are essential ingredients of the observed patterns of regional convergence in wages and employment shares.

3. Structural Estimation

We now estimate the structural parameters of our model. Our key finding in this section is that when accounting for differences in regional exposure, local employment growth, and region-specific shifts in demand, strong rural catch-up growth in productivity is required to rationalize the patterns of regional convergence in the data. In Section 4, we quantify the role of this rural growth premium for local wage growth and industrialization.

3.1 Data Description

We assume a period in the model corresponds to 20 years in the data and obtain total employment by sector and county from the U.S. Census Bureau's Decennial Full Count Census files for 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 supplement these data with information on average earnings at the county level from the Census of Manufacturing and average values of farmland and buildings per acre for each decade from the Census of Agriculture (both via NHGIS; see Manson, Schroeder, Van Riper, and Ruggles (2017)).[10] Finally, we use longitudinal data at the individual level from the linked version of the Decennial Census data to measure migration flows across commuting zones (via IPUMS; see Ruggles et al. (2015)).

We interpret locations in the model as commuting zones and spatially harmonize the data using the crosswalk in Eckert et al. (2020b). We drop data from states that were not in the Union by 1880. Our final sample consists of a balanced panel of 495 commuting zones in 1880, 1900, and 1920 (see B.1 in the Appendix for a map).

In addition, we rely on time-series data from the "Historical Statistics of the United States" (see Carter, Gartner, Haines, Olmstead, Sutch, Wright et al. (2006)) on real GDP per capita and the sectoral price indices. In Appendix B.1, we provide more details on data sources, data construction, and sample selection.

3.2 Estimation Strategy

We quantify our model using a combination of structural estimation and model inversion. We estimate eight structural parameters within the model: the two catch-up parameters (λ_M, λ_A), the labor-supply elasticity ζ, three preference parameters (ν, η, ε), and the growth rates of the sectoral productivity frontiers (g_M, g_A). We do so by using 11 empirical moments. In addition, we estimate migration and trade costs from the gravity relationships of trade and migration flows outside of the model. Finally, given these structural parameters, we invert our model to infer the distribution of local fundamentals, that is, initial productivity in 1880, $[Z_{rM1880}, Z_{rA1880}]_r$, the endowment of agricultural land $[T_r]$, and local amenities $[B_r]$, to perfectly rationalize the data on wages, total employment, land rents, and sectoral employment shares in 1880.

Combining model inversion with structural estimation in this way has several virtues. The inversion part ensures the crucial correlation between regions' agricultural specialization and sector-specific productivity is directly inferred from the data. The reason is that the initial productivity distribution is chosen so that the model matches the observed data on employment shares, factor prices, and total employment in 1880. Hence, we do not assume agricultural regions are necessarily technologically backward, but let the data flexibly inform this correlation. Parameterizing the productivity process (instead of inferring new productivity terms for each cross-section) then makes the model amenable to a counterfactual exercise of what the structural transformation in the US had looked like in the absence of productivity convergence. It also allows us to study whether such a parsimonious (and, we think, natural) productivity process can

[10] In the model, average earnings in manufacturing exactly coincide with average regional earnings, $\overline{w}_{rt}$, which we compute as manufacturing payroll divided by manufacturing employment. To the best of our knowledge, no data on agricultural wages exist at the county level.

quantitatively rationalize the extent of rural convergence documented in Section 1.[11]

In Table 3 below, we provide an overview of all the parameters of our model and the empirical moments we use for identification. Despite calibrating most parameters jointly, we discuss our calibration strategy for particular structural parameters in terms of the most informative empirical moments.

Aggregate Productivity Growth (g_A and g_Z) and Consumer Preferences (η, ν, and ϕ) We estimate the growth rates of the agricultural and non-agricultural frontier, g_A and g_{NA}, and consumers' preferences, η and ν, to ensure the model matches three macroeconomic time-series moments: (i) aggregate GDP growth between 1880 and 1920, (ii) the change in the relative price of agricultural goods between 1880 and 1920, and (iii) the evolution of the agricultural employment share. Given our strategy of matching the 1880 cross section exactly, we also match the aggregate agricultural employment share in 1880 by construction. We thus estimate four parameters by targeting six macroeconomic moments (two growth rates, 1880-1900 and 1900-1920, for each of the three outcomes). The remaining preference parameter, ϕ, corresponds to the asymptotic spending share on agricultural value added for very high incomes. We set $\phi = 0.01$, which is close to the agricultural employment share in the US in 2020.

Regional Fundamentals: $[T_r]$, $[B_r]$, and $[Z_{rA1880}, Z_{rM1880}]$ We choose regions' sectoral productivity in 1880, $[Z_{rA1880}, Z_{rM1880}]_{rs}$, and land endowments, $[T_r]$, to *exactly* match the distribution of average earnings $\{\overline{w}_{r1880}\}_r$, agricultural employment shares $\{s_{rA1880}\}_r$, and land rents $\{R_{r1880}\}$, given the observed level of employment $\{L_{r1880}\}$.

Specifically, local wages and employment shares identify the productivity level in manufacturing Z_{rM1880} and the *combined* agricultural productivity index $Z_{rA1880}T_r^{\alpha}$. To separately identify Z_{rA1880} from T_r, we then use local land rents (relative to the prevailing wage). Note our identification strategy only uses *static* equilibrium conditions; it does not assume the economy is in steady-state, nor is it impacted by our particular assumptions of the convergence process.

Table 2 shows the relationship between the inferred productivity terms and agricultural employment shares. Agricultural regions in 1880 had *both* low agricultural productivity, Z_{rA1880}, and low manufacturing productivity, Z_{rM1880}. Agricultural specialization was thus a reflection of comparative advantage in agriculture and not of absolute agricultural advantage. This is seen in column 3, which shows a positive, albeit statistically insignificant, correlation between agricultural employment shares and the relative productivity of the agricultural sector. The fact that rural locations were technologically behind the frontier in both industries implies they benefited from catch-up growth in both sectors. For each sector separately, we set the level of the economy's

[11]This restriction also gives us additional degrees of freedom that we use to estimate the elasticity of substitution ζ.

TABLE 2: DETERMINANTS OF AGRICULTURAL SPECIALIZATION

	$\ln Z_{rA1880}$	$\ln Z_{rM1880}$	$\ln \frac{Z_{rA1880}}{Z_{rM1880}}$	$\ln \ell_{r1880}$	$\ln B_{r1880}$
s_{rA1880}	-0.231***	-0.257***	0.0368	-0.133***	-0.115***
	(0.0110)	(0.0114)	(0.0292)	(0.00630)	(0.0304)
Observations	495	495	495	495	495
Adjusted R^2	0.430	0.515	0.003	0.493	0.045

Notes: Notes: The table reports the results of a set of bivariate regression $s_{rA1880} = \alpha + \beta x_r + u_r$, 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 in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively.

technological frontier, $[\overline{Z}_{s1880}]_s$, to the highest regional productivity level in 1880. Table 2 also shows rural labor markets are land-abundant; that is, their employment density ℓ_{rt} is low (column 4). This pattern is implied by the fact that, empirically, agricultural land rents in rural regions are relatively low compared with more urban locations.

We estimate local amenities $[B_r]_r$ by requiring that the observed level of employment $\{L_{r1880}\}$ is consistent with individuals' spatial labor-supply decisions.[12] Because our economy features distance-specific moving costs, the employment distribution is a dynamic state variable. Hence, given the observed factor prices in 1880, local employment $\{L_{r1880}\}$ depends on both the amenity vector B_r, initial employment in 1860, and differential employment growth (through migration and fertility and migration) between 1860 and 1880. Because of the territorial expansion of the US, we choose to not rely on the data in 1860. Instead, we estimate B_r so that the employment distribution would be stationary *if* productivity and aggregate employment stocks were to remain constant. Intuitively, we ensure spatial reallocation in our model is driven by changing factor prices and future employment growth, rather than transitional employment dynamics that originate prior to 1880. Table 2 shows amenities are lower in rural regions, indicating such regions are sparsely populated even given their low wages (column 5).

Technological Catch-Up (λ_A and λ_M) and Skill Substitutability (ζ) The key empirical pattern motivating our analysis is the presence of rural convergence documented in Section 1: the positive relationship between agricultural specialization and wage growth, and the *U*-shaped relationship of agricultural specialization and subsequent industrialization. In Section 2.5, we showed theoretically that the strength of technological catch-up (λ_s) and the sectoral substitutability of skills (ζ) are important determinants of these patterns. In addition, all else equal, changes in population density affect wage growth across regions, due to decreasing returns in agriculture.

We exploit these facts to estimate $(\zeta, \lambda_A, \lambda_M)$ by indirect inference. The hallmark of indirect inference is the use of an auxiliary model to capture aspects of the data upon

[12] Note also that such calibrated amenities implicitly control for differences in the size of commuting zones. For given wages, commuting zones with a larger area and correspondingly larger total employment are associated with a higher amenity term.

which to base the estimation. Indirect inference chooses the parameters of the economic model so that the parameter estimates of the auxiliary model are as close as possible when using the actual versus model-generated data. Importantly, indirect inference does not require that the auxiliary model be correctly specified (see Smith (2008)). For our auxiliary model, we use the relationships between agricultural specialization and subsequent wage growth, industrialization, and population growth as summarized by the following regressions:

$$d\ln\bar{w}_{rt} = \delta_t + \delta_{j(r)} + \beta_w s_{rAt} + \nu_{rt}; \tag{12}$$

$$d\ln\ell_{rt} = \delta_t + \delta_{j(r)} + \beta_\ell s_{rAt} + \nu_{rt}; \tag{13}$$

$$ds_{rAt} = \delta_t + \delta_{j(r)} + \beta_{s_A} s_{rAt} + \gamma_{s_A} s_{rAt}^2 + \nu_{rt}. \tag{14}$$

Here, δ_t and $\delta_{j(r)}$ are period fixed effects and state fixed effects, and time differences are taken over 20-year intervals. For the case of local industrialization, ds_{rAt}, we estimate a quadratic relationship to capture the U-shape documented in Figure 1. We match the four coefficients β_w, β_ℓ, β_{s_A}, and γ_{s_A} in our estimation. The estimates of β_w, β_{s_A}, and γ_{s_A} are reported in column 1 of Table 1. The estimate for β_ℓ, which is given by -0.36, is reported in Table A.1 in the Appendix.

In addition to the parameters β_{s_A} and γ_{s_A}, we also target the change in agricultural employment shares between 1880 and 1920 among the most rural locations. In doing so, we force our model to match the "trough" of the U-shaped relationship of local industrialization and initial agricultural specialization. Specifically, we target the change in the agricultural employment share between 1880 and 1920 among locations with at least 80% of their 1880 workforce in agriculture.

The three regressions are informative about λ_A, λ_M, and ζ because rural locations have an absolute disadvantage in both sectors and thus benefit from catch-up growth in both industries. Hence, wage growth increases in both λ_A and λ_M. At the same time, λ_A and λ_M have opposite effects on rural industrialization: if most catch-up growth occurs in agriculture (non-agriculture), agricultural specialization would increase (decrease) in rural regions. In addition, the higher β_ℓ, the higher λ_A because increasing population density depresses agricultural wage growth, requiring stronger technological convergence to match the observed patterns of wage convergence. Finally, Figure 2 shows a larger supply elasticity, ζ, leads to a more pronounced U-shape in industrialization.

Spatial Labor Supply In the model, three parameters shape spatial labor supply. The first is migration costs. We parameterize migration costs as a function of distance. Denoting the geographic distance between regions r and r' by $d_{rr'}$, migration costs are given by $\mu_{rr'} = d_{rr'}^{-\kappa}$. The distance elasticity κ is hence an important determinant of local labor supply. To estimate κ, we use the following log linear relationship for inter-regional migration flows implied by our model:

$$\log m_{rr't} = \delta^{o}_{rt} + \delta^{d}_{r't} - \kappa\epsilon \log d_{rr'}. \tag{15}$$

In equation (15), δ^{o}_{rt} and $\delta^{d}_{r't}$ are origin and destination fixed effects, respectively, that are functions of endogenous location-specific objects and parameters. We estimate equation (15) using commuting-zone-to-commuting-zone migration flows that we constructed with the linked Census data and find $\kappa\varepsilon \approx 2.8$ (see Appendix B.2.2), consistent with Allen and Donaldson (2020), who find a distance elasticity of 2.16 across counties during the same time period in the US.

Second, the sensitivity of migration flows with respect to local factor prices is governed by the dispersion of location preference shocks, ε. Equation (4) implies the partial elasticity of migration flows from r to r' with respect to wages in r' is given by

$$\frac{\partial \ln m_{rr'}}{\partial \ln \overline{w}_{r'}} = \varepsilon\eta \left(1 + \nu \frac{\ln \left(P_{r'A}/P_{r'M}\right)}{\frac{1}{\eta}\Gamma_{\zeta/\eta}\left(\overline{w}_{r'}/\left(P^{\phi}_{r'A}P^{1-\phi}_{r'M}\right)\right)^{\eta} - \nu \ln\left(P_{r'A}/P_{r'M}\right)} \right). \tag{16}$$

Hence, in addition to ε, this elasticity also depends on the Engel elasticity η, the taste parameter ν, and a set of endogenous variables. We target an average labor-supply elasticity of two, a consensus estimate in the literature (see, e.g., Allen and Donaldson (2020), Monte, Redding, and Rossi-Hansberg (2018) or Peters (2022)). In addition, ε obviously also affects the above-mentioned relationship between specialization and population growth, β_ℓ.

Third, spatial labor supply depends directly on the vector of exogenous employment growth in each region, $\{n_{rt}\}_{r,t}$, which captures all employment growth not due to worker in- or out-migration from or to other regions. Such exogenous sources of local employment growth capture differences in local demographics, namely, fertility and mortality rates, and international migration flows, which are unbalanced across space. Both of these aspects are quantitatively important in the context of our application. We introduce a new methodology to infer $\{n_{rt}\}_{r,t}$ from observed data on county-level immigration, births, and age distributions, which we describe in detail in Appendix B.2.1. In essence, we choose n_{rt} to match the net effect of the cross-sectional variation in immigration and fertility rates for each commuting zone, as well as the overall aggregate rate of employment growth between 1880 and 1920. Importantly, because workers at the beginning of each period have the option to migrate before becoming economically active, employment growth in each location remains endogenous in our theory.

TABLE 3: STRUCTURAL PARAMETERS AND MODEL FIT

Structural Parameters			Estimation Method		
	Description	Value	Panel A: In-Model (Moment, Data, Model)		
ζ	Labor Supply Elasticity	6.9	γ_{sA} in regression (14)	0.45	0.51
			$E[s_{rA1920} - s_{rA1880} \vert s_{rA1880} > 0.8]$	-0.20	-0.20
λ_A	Catch-Up in Agricult.	0.21	β_w in regression (12)	0.25	0.16
			β_l in regression (13)	-0.36	-0.04
g_A	Growth of Agricult. Frontier	0.07	Ag. Empl. Share 1900	0.39	0.35
			Ag. Empl. Share 1920	0.26	0.25
λ_M	Catch-Up in Non-agricult.	0.05	β_{sA} in regression (14)	-0.48	-0.57
g_M	Growth of Non-agricult. Frontier	0.09	GDP growth 1880-1900	1.43	1.50
			GDP growth 1900-1920	2.04	2.05
ϵ	Location Taste Heterogeneity	3.80	Avg. Migration Elasticity	2	2.03
η	Engel Elasticity	0.93	Rel. price P_M/P_A 1900	0.94	1.01
ν	PIGL preference parameter	0.12	Rel. price P_M/P_A 1920	0.89	0.87
			Panel B: Out-of-Model (Strategy)		
κ	Migration Cost Distance Elasticity	2.8	Gravity relationship of migration flows		
θ	Trade Costs Distance Elasticity	1.35	Gravity relationship of trade flows		
			Panel C: Exogenously-Set (Source)		
σ	Elastictiy of Substitution Mfg Good	6	NA		
ρ	Amenity Congestion Elasticity	0.15	Allen and Donaldson (2020)		
α	Land Share in Production Function	0.4	Valentinyi and Herrendorf (2008)		
ϕ	Asy. Exp. Share on Agricult. Goods	0.01	NA		

Notes: The table contains the values for all structural parameters and targeted moments of our model. The eight parameters in the upper panel are estimated within the model, targeting the 11 moments on the right. The two distance elasticities are estimated from gravity equations outside of the model. The remaining four parameters are set exogenously.

Other Parameters As is common in the literature, we parameterize trade costs as power functions of distance so that trade costs in both sectors are $\tau_{rr'} = d_{rr'}^{-\theta}$. For the elasticity of trade flows to distance, $(1-\sigma)\theta$, Allen and Donaldson (2020) report an estimate of -1.35.[13] We take the remaining parameters from various sources in the literature. Most related papers assume an elasticity of substitution σ between 3 and 8; we set $\sigma = 6$. For the agricultural land share, we follow Valentinyi and Herrendorf (2008) and set α to 0.4. We also borrow 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.

3.3 Estimates and Model Fit

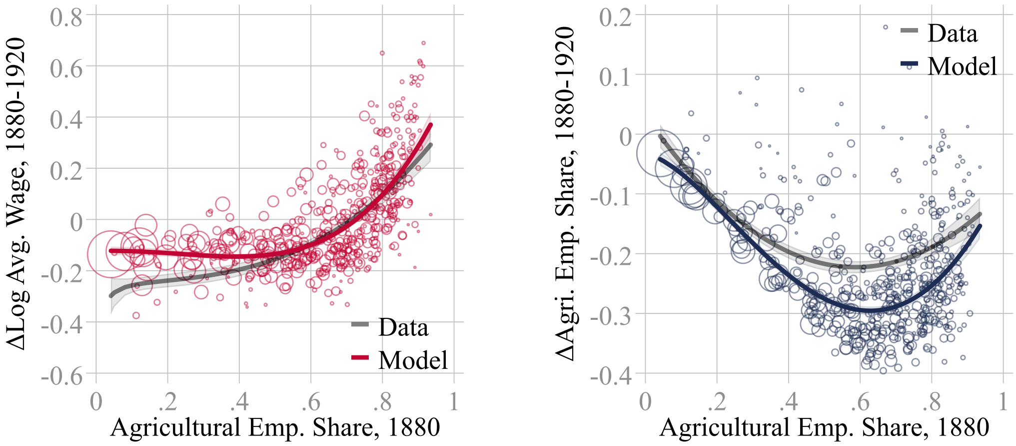
Table 3 presents our parameter estimates and their associated moments in the calibrated model and the data. We differentiate parameters estimated within the model (Panel A), parameters estimated outside the model (Panel B), and parameters that are set exogenously (Panel C).

Overall, our model is able to successfully capture the most important empirical features of spatial structural change in the US between 1880 and 1920. First, the calibrated model produces the time-series patterns of the three aggregate "macro" moments: it successfully captures the large decline in agricultural employment, the increase in GDP

[13] Monte et al. (2018) find a similar elasticity of -1.29. Disdier and Head (2008) show this elasticity is roughly constant in international trade data in the 20th century.

FIGURE 3: RURAL CONVERGENCE – MODEL AND DATA

(A) WAGE GROWTH (B) INDUSTRIALIZATION



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

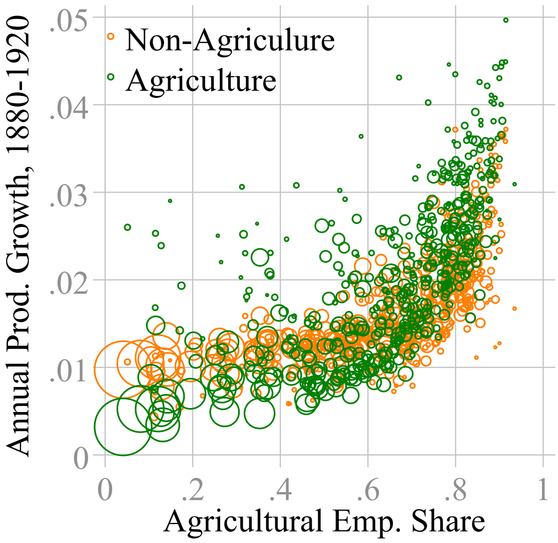
per capita, and the small increase in the relative price of agricultural goods between 1880 and 1920. These time-series moments are mostly informed by the rates of aggregate productivity growth and preference parameters. We estimate that the productivity frontier in non-agriculture ($\overline{Z}_{Mt}$) grew at a rate of 0.09, and the frontier in agriculture ($\overline{Z}_{At}$) grew at a rate of 0.07 over a 20-year period. The estimates of the preference parameters imply an important role of the demand-side non-homotheticities: we find an Engel elasticity η of 0.93 and $\nu = 0.12$, which implies agricultural value added is a necessity.[14]

Second, and most importantly, the calibrated model matches the patterns of rural convergence documented in Section 1. The cross-sectional estimates of the parameters β_w, β_{sA}, and γ_{sA} from the two regressions in equations (12) and (14) are similar in the model and the data. In Figure 3, we replicate the non-linear relationships in both the data (grey) and our model (red and blue, respectively). Our model reproduces the rural bias of wage growth (left panel) and the *U*-shape of industrialization (right panel) very well.

To fit these patterns of rural convergence, our estimates imply an important role for catch-up growth. Recall local productivity growth depends both on a region's distance to the frontier (i.e., $\overline{Z}_{s1880}/Z_{rs1880}$) and the catch-up parameters λ_A and λ_M. Our estimates of $\lambda_A = 0.21$ and $\lambda_M = 0.05$ indicate significant catch-up growth and spatial convergence between 1880 and 1920. Moreover, because we estimate sectoral productivity in 1880 to be negatively correlated with the agricultural employment share, rural labor markets were the main beneficiaries of such catch-up growth.

[14]In Section B.2.4 in the Appendix, we compare this estimate from time-series data with cross-sectional estimates.

FIGURE 4: RURAL CATCH-UP GROWTH



Notes: The figure 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.

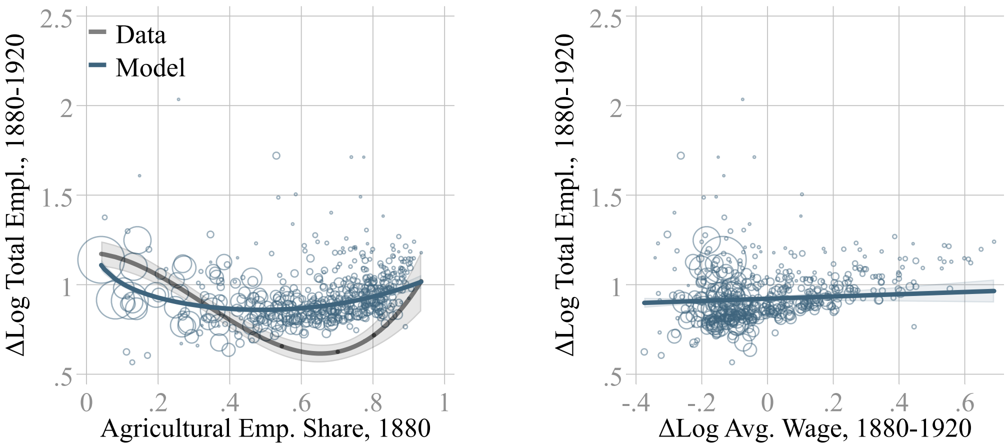
In Figure 4, we show the implied heterogeneity in productivity growth across regions. In the four decades following 1880, rural labor markets experienced a growth premium of around two percentage points. The similarity in productivity growth in both sectors reflects the combination of two aspects of our theory. First, there is less regional dispersion in agricultural productivity, reducing the opportunities for productivity catch-up. Second, our structural estimation showed $\lambda_A > \lambda_M$; that is, the process of catch-up is faster in agriculture (which, in turn, might be why agricultural productivity in 1880 is less dispersed). In terms of their regional growth implications, these two forces roughly balance out.

In Figure 5, we turn to the implications for spatial mobility. In the left panel, we show the cross-sectional relationship between local employment growth and initial agricultural specialization. The empirical relationship is non-monotone: employment growth is faster in locations with very low and locations with very high agricultural employment shares. Our calibrated model captures this qualitative relationship and reproduces the U-shape of employment growth in the data. However, we overestimate employment growth for commuting zones in the intermediate range of agricultural employment shares.

The positive relationship between agricultural employment shares and employment growth among regions with high agricultural employment shares comes as no surprise, because our model replicates the faster wage growth in agricultural locations and generates an empirically reasonable migration elasticity of two. The negative correlation between agricultural employment shares and employment growth among regions with high agricultural employment shares may be more surprising. One reason our

FIGURE 5: LOCAL EMPLOYMENT GROWTH

(A) EMPLOYMENT GROWTH (B) SPATIAL LABOR SUPPLY



Notes: In the left panel, we show the relationship between employment growth and the agricultural employment share. We show the data in grey and our model in orange. The size of the markers reflects the relative size of different commuting zones. The solid lines show the best non-linear fit. In the right panel, we display the correlation between wage growth and employment growth in the model.

calibrated model replicates the U-shape is the exogenous component of population growth n_{rt}.[15] However, even in the absence of this exogenous population-growth component, the model can generate employment growth in locations with low relative wage growth: if regions with low agricultural employment shares purchase agricultural goods from regions that benefit from catch-up growth, real wages can rise and stimulate local employment growth.

To see this effect, consider the right panel of Figure 5, which, for the data generated by our model, shows a positive correlation between wage growth and employment growth. However, the relationship is noisy because goods prices change at different rates, due to trade costs. In addition, the current employment distribution matters directly for future employment growth, because of moving costs.

4. THE DRIVERS OF RURAL CONVERGENCE

With the calibrated model in hand, we now quantify the role of the exposure versus the incidence channel in generating the observed patterns of rural convergence. We also estimate the relative importance of physical productivity growth versus changes in demand and employment density. Finally, we provide a set of concrete empirical examples of how rural locations caught up with the technological frontier.

4.1 The Importance of Catch-Up Growth

Our structural estimation showed rural labor markets experienced a productivity growth premium of about two percentage points in each sector (cf. Figure 4). To

[15]We show in the Appendix that the exogenous component of population growth induces a U-shape; see Figure A.3.

TABLE 4: THE MACRO CALIBRATION

	TECHNOLOGY PARAMETERS				MACRO MOMENTS					
	Agri.		Non-Agri.		Agri. Emp. Share		GDP pc		P_M/P_A	
Calibration	g_A	λ_A	g_M	λ_M	1900	1920	1900	1920	1900	1920
No-Catch-up	0.41	0	0.34	0	0.34	0.24	1.41	2.03	0.98	0.88
Baseline	0.07	0.21	0.08	0.05	0.34	0.25	1.5	2.05	1.01	0.87

Notes: The table reports the technology parameters and the macro moments for the baseline model and the "macro-calibration." All other parameters are the same in both calibrations and reported in Table 3.

understand their quantitative importance, we now shut down these differences in the incidence of technological change. In this counterfactual, the spatial heterogeneity of wage growth and industrialization therefore result from differences in exposure and differential changes in non-agricultural demand and employment densities.

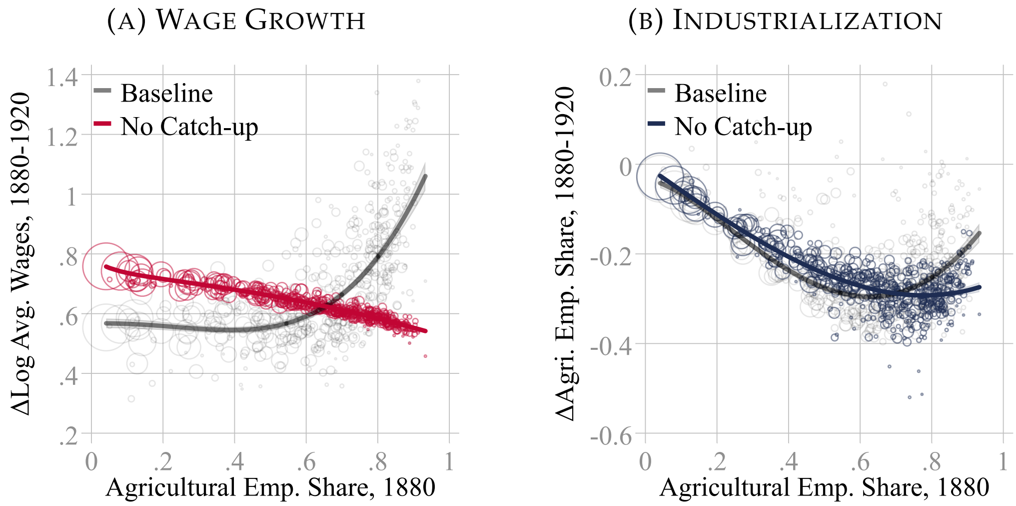
Specifically, we assume the productivity distribution is stationary, that is, $\lambda_A = \lambda_M = 0$, and re-estimate the growth rates of the technological frontiers, g_A and g_M, to match the growth of aggregate income per capita and the change in relative prices since 1880, keeping all other parameters the same. In the resulting counterfactual economy, local labor markets are spatially segmented and differentially exposed to sectoral reallocation, but the rate of technological progress is the same in all regions. We refer to this parameterization as our model's "no catch-up" calibration.

We report the resulting parameters of the productivity process (columns 1 - 4) and the implied macro moments (columns 5 - 10) in Table 4. All parameters except g_s and λ_s are held fixed. Note that the location fundamentals, that is, initial productivity, the land endowment, and local amenities in 1880, are *exactly* the same in both calibrations, because they are estimated from static equilibrium conditions and are therefore independent of g_s and λ_s.

Table 4 shows the overall rate of frontier productivity growth is substantially faster in the macro calibration to compensate for the absence of catch-up growth. In terms of the macro moments, however, both calibrations are almost indistinguishable and replicate the time-series patterns of the structural transformation equally well. Note in particular that the "no catch-up" economy still experiences the structural transformation: the decline in the agricultural employment share is very similar to our baseline calibration and hence the data.

In contrast to these aggregate patterns, however, Figure 6 shows the "no catch-up" calibration has strikingly counterfactual predictions for the process of rural convergence. The left panel shows catch-up growth is essential to rationalize the empirically observed features of rural wage growth, both quantitatively and qualitatively. In the absence of productivity catch-up, growth would have been urban biased, and rural labor markets would have fallen even further behind their urban counterparts.

FIGURE 6: THE ROLE OF RURAL PRODUCTIVITY CATCH-UP



Notes: In the left (right) panel, we show local wage growth (local industrialization) as a function of the initial agricultural employment share. We depict the baseline calibration in grey and the macro calibration with no catch-up in red and blue, respectively. The size of the markers reflects the relative size of different commuting zones.

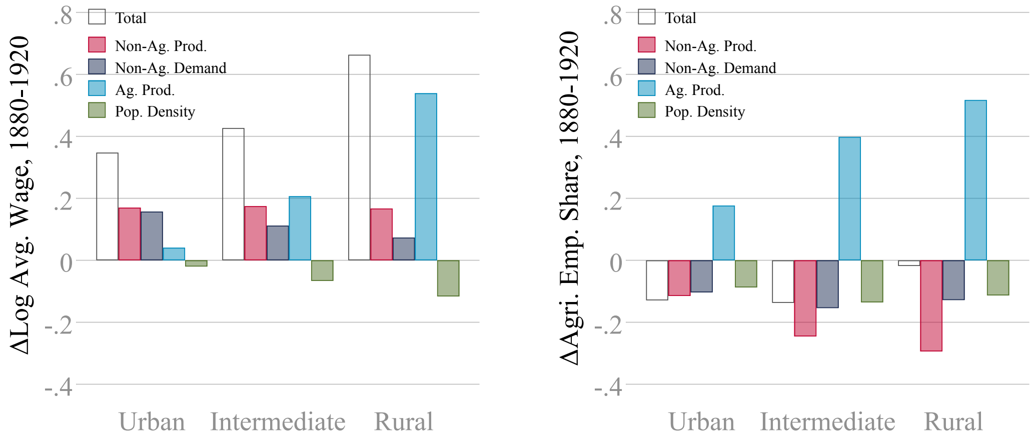
This pattern resembles the theoretical results in equation (11), which showed analytically that falling agricultural employment would lead to urban-biased growth, albeit for the special case of no trade costs and no migration. Quantitatively, Figure 6 shows the exposure channel leads to an economically meaningful urban bias of the structural transformation: in the absence of catch-up growth, urban locations would have experienced roughly 15% faster wage growth than rural locations. Hence, catch-up growth not only allowed rural America to close the gap in living standards but *saved* it from falling further behind.

The right panel summarizes the implications for local industrialization. In contrast to the spatial bias of wage growth, regional differences in exposure go a long way toward explaining the spatial heterogeneity in industrialization: for a large number of regions, the relationship between initial agricultural specialization and subsequent industrialization is quite similar to our baseline calibration. The main difference arises for the very rural labor markets whose agricultural employment share exceeds 75%. Both in the data and in our model, these markets experience *less* industrialization - the *U*-shape. Without catch-up growth, the model fails to generate the full *U*-shape, highlighting the importance of the interaction of differences in exposure with catch-up growth in generating the patterns of spatial structural change we documented in Section 1.

The quantitative importance of the exposure channel has important consequences for our understanding of spatial growth. First, it suggests technological catch-up forces must have been substantial, because they not only generated wage convergence but also had to overcome the secular urban bias of the exposure channel. Estimates of the

FIGURE 7: THE MECHANISMS OF SPATIAL STRUCTURAL CHANGE

(A) WAGE GROWTH (B) INDUSTRIALIZATION



Notes: The figure reports the decomposition of local wage growth, $d\ln\overline{w}_{rt}$, and local industrialization, ds_{rA}, (see Proposition 2) into non-agricultural demand $(\phi_r(1/\sigma)d\ln\mathcal{D}_{rt})$, local productivity growth $(\phi_r((\sigma-1)\sigma)d\ln Z_{rMt}$ and $(1-\phi_r)d\ln Z_{rAt})$, and changes in local employment density $(-\phi_r\alpha d\ln\ell_{rt})$. We define urban (rural) locations as regions in the lower (upper) quartile of the distribution of agricultural employment share in 1880 and intermediate locations in the interquartile range. We refer to all commuting zones in the interquartile range as "intermediate."

strength of technological catch-up that do not explicitly account for exposure difference are therefore bound to underestimate the strength of technological convergence. Second, it shows wage convergence and industrialization do not necessarily go hand in hand. In the absence of catch-up growth, wage growth in rural America would have been slower but industrialization faster.

4.2 The Sources of Rural Convergence

The previous section focused on the separate roles of exposure versus incidence. Now, we decompose effective productivity growth in each sector into the distinct roles of technological catch-up and the forces of demand and employment density and study their spatial impact.

Propositions 1 and 2 highlight that local wage growth and industrialization are fully determined from four factors: sectoral productivity growth ($d\ln Z_{rAt}$ and $d\ln Z_{rMt}$), demand growth ($d\ln\mathcal{D}_{rt}$), and changes in employment density $(d\ln\ell_{rt})$. In Figure 7, we implement the formal decomposition in Proposition 2 in our calibrated model.[16] Specifically, for each commuting zone, we compute the impact of each of the four components multiplied by their respective exposure term, separately for local wage growth and local industrialization. We then aggregate these results 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.

[16]Proposition 2 relies on a first-order approximation. In Section B.2.5 in the Appendix, we compare these predictions with the full non-linear solution in our model and show they are very close.

The left panel of Figure 7 presents the decomposition of local wage growth. The white bars represent total wage growth in each group of commuting zones, and thus exhibit the previously documented pattern of rural-biased wage growth. The remaining bars show regions differed substantially in *why* their wages grew.

In rural labor markets, agricultural productivity growth, shown in blue, was the dominant factor. By contrast, industrial productivity growth, shown in red, played a much smaller role. Recall this within-region discrepancy does *not* reflect that agricultural productivity grew faster. Figure 4 showed that productivity growth was actually quite similar in both sectors. The difference rather reflects the regional exposure: because rural locations only have a small share of their workforce outside of agriculture, the non-agricultural exposure elasticity, $\phi(s_{rA})$, is small and wage growth is mostly driven by agricultural productivity growth. This also explains why rising non-agriculture demand, $(\mathcal{D}_{rt})$, shown in grey, only had a small effect on wages in rural America. Finally, local employment growth, shown in green, reduced wage growth in rural locations, whose sectoral structure exposes them to decreasing returns in the agricultural sector.

These patterns differ dramatically in urban areas. Revenue productivity growth in the non-agricultural sector played a dominant role for wage growth, and almost half of all wage growth stemmed from increased demand. Even though manufacturing productivity growth was slower in urban areas (see Figure 4), their outsized exposure to this sector implies the total impact is comparable to rural locations. By contrast, rising agricultural productivity did not meaningfully affect wages in urban labor markets given their small agricultural employment share. Finally, increased employment density had a negligible effect, given that - in our model - non-agricultural production is subject to constant returns to scale and accounts for the bulk of employment in urban labor markets.

Overall, this decomposition highlights the pivotal nature of agricultural productivity growth in spurring wage growth in rural regions. It also highlights the importance of exposure versus incidence. The wage impact of non-agricultural productivity growth is roughly balanced across regions, because exposure and incidence are inversely correlated: productivity growth is faster in rural regions where exposure is lower. By contrast, rural regions are both more exposed to and benefit from faster agricultural productivity growth, making it a powerful source of rural wage convergence.

The right panel of Figure 7 displays the same decomposition for local industrialization. Rural locations industrialized because of rising manufacturing productivity, rising non-agricultural demand, and increasing employment density. Quantitatively, non-agricultural productivity growth played the most important role. By contrast, productivity growth in the agricultural sector was a strong counteracting force that kept workers in agriculture, especially in the most rural labor markets. Intermediate locations saw a slightly faster decline in agricultural employment shares (the "U-shape"), primarily due to a less pronounced increase in agricultural productivity. Finally, in urban centers, rising demand, nonagricultural productivity growth, and rising employment density were equally important contributors to the modest decline in agricultural employment they experienced between 1880 and 1920.

These findings emphasize the fundamentally different roles of agricultural and non-agricultural productivity growth. Agricultural productivity growth was essential for the convergence of wages across space but, all else equal, would have led to divergence in agricultural employment shares. Non-agricultural productivity growth, on the other hand, was the main engine behind the industrialization of the rural hinterland and the convergence in agricultural employment shares we observed in the data. Furthermore, Figure 7 highlights the interaction between agricultural and non-agricultural productivity growth that gives rise to the U-shape in industrialization.

4.3 Direct Evidence on Rural Catch-Up

The analysis above established the key role of faster productivity growth in rural America. Our theory and quantitative analysis summarizes all factors leading to catch-up growth in the reduced-form process of productivity convergence. We view this parameterization as a modeling device for various technological and institutional developments in the US between 1880 and 1920 that benefited rural locations.

In this section, we complement these model-based estimates with direct empirical evidence for the presence of faster rural productivity growth. In Table 5, we provide evidence for such developments from multiple data sources. Specifically, we run a set of bivariate regressions where we regress the growth of different outcomes y_{rt} between 1880 and 1920 against the agricultural employment share in 1880:

$$y_{r1920} - y_{r1880} = \delta + \beta s_{rAt} + \nu_{rt}.$$

We differentiate between outcomes we expect to be correlated with general productivity growth and those we expect are correlated with sector-specific productivity growth.

In columns 1-3, we report three examples of sector-neutral developments that benefited rural locations. In column 1, we show rural locations experienced faster financial development, as measured by the growth of the number of banks per capita. In the second column, we provide evidence for the pronounced catch-up in educational attainment, proxied by the share of children attending school. We find the school attendance rate increased much faster in agriculturally specialized labor markets between 1880 and 1920. Finally, in the third column, we document that rural locations experienced faster market integration. More specifically, there is a positive relationship between agricultural employment shares in 1880 and changes in the inverse of the transportation-cost-weighted distance to all other commuting zones, using transportation cost estimates from Donaldson and Hornbeck (2016). Hornbeck and Rotemberg (2021) show directly that improved railroad access led to local productivity growth in rural locations in the US.

In the remaining four columns, we present additional evidence for sector-specific factors. In particular, rural locations saw faster growth in the capital stock in both sectors (columns 4 and 6) and experienced a faster increase in scale: the growth in both average farm and firm size is positively correlated with the initial agricultural employment share (columns 5 and 7).[17]

We view these results as an empirical description of the general transformation of rural labor markets between 1880 and 1920. Rising educational attainment, changes in the scale and capital intensity of production, improved transport systems, and financial deepening are often seen as markers of economic development across countries. Table 5 shows the same patterns were also present across local labor markets in the US during the first phase of the structural transformation.

Table 5 shows a diverse set of institutional developments (e.g., the spread of schooling), infrastructure measures (e.g., the railway system), and capital deepening (e.g., tractors in agriculture) all appear to have systematically benefited more agricultural regions between 1880 and 1920. The diversity of these factors makes modelling them all individually difficult. Hence, we chose to summarize them in the two productivity scalars Z_{rA} and Z_{rM}. Although performing a more structural decomposition of local productivity growth into its different components would be interesting, our approach allows us to directly infer the productivity process required to explain rural convergence within a model of structural change.

The main takeaway from this section is that between 1880 and 1920, many factors came together to benefit rural regions and help them counteract the adverse effects of their exposure to the US economy's shift away from agriculture.

CONCLUSION

Economic growth systematically reallocates resources out of the agricultural sector. Such structural change has adverse effects on regions specialized in agriculture and depresses their wage growth. At the same time, agricultural regions are usually poor and hence are the natural beneficiaries of the usual forces of technological catch-up. In this paper, we provided a novel framework of spatial structural change that allows us to distinguish between the regional impacts of sectoral reallocation and the consequences of technological convergence. We found catch-up growth played a crucial role; without

[17] Desmet and Rossi-Hansberg (2009) provide direct estimates of manufacturing TFP convergence across regions in the US between 1900 and 1920; their estimates are of a similar magnitude (cf. Figure 6 in their paper).

TABLE 5: DRIVERS OF RURAL CATCH-UP

	Growth in...						
	General			Sector-specific Factors			
	Banks pc	School Attendance	Market Access	Agri. Machinery	Farm Size	Non-agri. Machinery	Plant Size
s_{rA1880}	0.118***	0.007***	0.005***	0.032***	0.012***	0.039***	0.032***
	(0.006)	(0.001)	(0.000)	(0.002)	(0.003)	(0.009)	(0.007)
Obs.	495	495	495	495	495	495	495
R^2	0.725	0.331	0.424	0.359	0.107	0.090	0.242

Notes: The dependent variables are the growth rate in the number of banks per capita from Jaremski and Fishback (2018) (column 1), the change in the share of children attending school from the Decennial Census (column 2), the inverse transportation cost weighted distance to all other commuting zones using transportation cost estimates from Donaldson and Hornbeck (2016) (column 3), and the growth rates of the sectoral capital stocks and average employment per farm/firm from the Census of Manufacturing (columns 6 and 7) and the Census of Agriculture (columns 4 and 5). All regressions are employment weighted.

it, the structural transformation would have led to decisively urban-biased wage growth and regional divergence in the US economy between 1880 and 1920.

Our paper suggests several important directions for future research. First, a more systematic comparison of the first and second phases of the structural transformation could produce a better understanding of the striking differences between these episodes. In contrast to the regional convergence during the transition away from agriculture, spatial inequality has increased during the ongoing transition away from manufacturing (see, e.g., Austin, Glaeser, and Summers (2018) and Chatterjee and Giannone (2021)). Our theory suggests these differences might reflect a strong exposure channel against the backdrop of weakened catch-up forces. Today, banks and high schools have long reached most areas of the US, the railroad network and interstate highway system are complete, and the human capital that has replaced physical capital may be harder to move into remote regions; in short, "old-fashioned" catch-up growth may have run out of steam.

Second, comparing the historical US experience with that of today's developing countries that are currently going through their first structural transformation would be interesting. Are similar convergence forces at play, or was the US case exceptional? Such a comparative analysis could provide new insight into the nature of economic development and help distinguish between different theories of economic growth.

Finally, although we did not explicitly model capital and its adoption across space, quantitatively assessing the role of spatial capital deepening versus productivity growth would be interesting. The tabulated data from Manson et al. (2017) provide an avenue to observe certain types of capital empirically, and recent theoretical advances show how to incorporate capital into spatial models; see Kleinman, Liu, and Redding (2021).

We hope our framework of *spatial structural change* will serve as a starting point to make these and other types of research possible.

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," .

Acemoglu, D., P. Aghion, and F. Zilibotti (2006): "Distance to frontier, selection, and economic growth," *Journal of the European Economic Association*, 4, 37–74.

Acemoglu, D. and V. Guerrieri (2008): "Capital deepening and nonbalanced economic growth," *Journal of Political Economy*, 116, 467–498.

Akcigit, U., H. Alp, and M. Peters (2021): "Lack of selection and limits to delegation: firm dynamics in developing countries," *American Economic Review*, 111, 231–75.

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. A., E. L. Glaeser, and L. H. Summers (2018): "Jobs for the heartland: Place-based policies in 21st century America," Tech. rep., National Bureau of Economic Research.

Autor, D. H. and D. Dorn (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 103, 1553–97.

Barro, R. J. and X. Sala-i Martin (1991): "Convergence across states and regions," *Brookings Papers on Economic Activity*, 107–182.

——— (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

Blanchard, O. J., L. F. Katz, R. E. Hall, and B. Eichengreen (1992): "Regional evolutions," *Brookings papers on economic activity*, 1992, 1–75.

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.

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 (2021): "The economic geography of global warming," Tech. rep., National Bureau of Economic Research.

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 (2009): "Spatial growth and industry age," *Journal of Economic Theory*, 144, 2477–2502.

——— (2014): "Spatial Development," *American Economic Review*, 104, 1211–43.

Disdier, A.-C. and K. Head (2008): "The puzzling persistence of the distance effect on bilateral trade," *The Review of Economics and Statistics*, 90, 37–48.

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 (2020a): "Skilled scalable services: The new urban bias in economic growth," .

Eckert, F., A. Gvirtz, J. Liang, and M. Peters (2020b): "A method to construct geographical crosswalks with an application to US counties since 1790," Tech. rep., National Bureau of Economic Research.

Fajgelbaum, P. D., E. Morales, J. C. Suárez Serrato, and O. Zidar (2019): "State taxes and spatial misallocation," *The Review of Economic Studies*, 86, 333–376.

Fajgelbaum, P. D. and S. J. Redding (2022): "Trade, structural transformation, and development: Evidence from Argentina 1869–1914," *Journal of Political Economy*, 130, 1249–1318.

Fan, T., M. Peters, and F. Zilibotti (2022): "Growing like India: The unequal effects of service-led growth," Tech. rep., National Bureau of Economic Research.

Farrokhi, F. and H. S. Pellegrina (2020): "Global trade and margins of productivity in agriculture," Tech. rep., National Bureau of Economic Research.

Herrendorf, B., R. Rogerson, and Á. 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 (2021): "Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies," .

JAREMSKI, M. AND P. V. FISHBACK (2018): "Did inequality in farm sizes lead to suppression of banking and credit in the late nineteenth century?" *The Journal of Economic History*, 78, 155–195.

KLEINMAN, B., E. LIU, AND S. J. REDDING (2021): "Dynamic spatial general equilibrium," Tech. rep., National Bureau of Economic Research.

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

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. http://doi.org/10.18128/D050.V12.0*.

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

MONTE, F., S. J. REDDING, AND E. ROSSI-HANSBERG (2018): "Commuting, migration, and local employment elasticities," *American Economic Review*, 108, 3855–90.

NAGY, D. K. (2023): "Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion," *Princeton University, mimeograph*.

NGAI, L. R. AND C. A. PISSARIDES (2007): "Structural change in a multisector model of growth," *American Economic Review*, 97, 429–443.

PELLEGRINA, H. S. AND S. SOTELO (2021): "Migration, Specialization, and Trade: Evidence from Brazil's March to the West," Tech. rep., National Bureau of Economic Research.

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

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

REDDING, S. J. AND D. M. STURM (2008): "The costs of remoteness: Evidence from German division and reunification," *American Economic Review*, 98, 1766–97.

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. https://doi.org/10.18128/D010.V7.0*.

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

SMITH, A. (2008): "Indirect Inference," *The New Palgrave Dictionary of Economics, 2nd Edition (forthcoming)*.

SOTELO, S. (2020): "Domestic trade frictions and agriculture," *Journal of Political Economy*, 128, 2690–2738.

TOLBERT, C. M. AND M. SIZER (1996): "US commuting zones and labor market areas: A 1990 update," .

VALENTINYI, Á. AND B. HERRENDORF (2008): "Measuring factor income shares at the sectoral level," *Review of Economic Dynamics*, 11, 820–835.

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 The Equilibrium System

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.1) given by

$$L_{jt} = \sum_r m_{rjt} L_{rt}^{Y} = \sum_r m_{rjt} n_{rt-1} L_{rt-1}, \tag{A.1}$$

where m_{rjt} is given in equation (4). Together with the expression for expected utility $\mathcal{V}_{rt}$ given in (6), equation A.1 determines the spatial supply function as a function of local wages $\overline{w}_{rt}$ and local prices $\{P_{rAt}, P_{rMt}\}_r$.

2. **Labor market clearing in agriculture:** the agricultural labor market clears when labor demand (LHS) equals labor supply (RHS)

$$w_{rAt}^{-\frac{1}{\alpha}} Z_{rAt}^{\frac{1}{\alpha}} T_r = \Gamma_\zeta L_{rt} \left(\frac{w_{rAt}}{\overline{w}_{rt}} \right)^{\zeta - 1}. \tag{A.2}$$

Equation A.2 determines the scaled skill prices in the agricultural sector, $w_{rA} = \frac{1}{1-\alpha} \tilde{w}_{rA}$, as

$$w_{rAt}^{\zeta - 1 + \frac{1}{\alpha}} = \overline{w}_{rt}^{\zeta - 1} Z_{rAt}^{\frac{1}{\alpha}} \frac{T_r}{\Gamma_\zeta L_{rt}} \tag{A.3}$$

Market clearing for non-agricultural products: For non-agricultural products, sales of firm ω located in region r are given by

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

The mass of non-agricultural 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:

$$\frac{\sigma}{\sigma-1} w_{rMt} H_{rPt} = N_{rt} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} w_{rMt}^{1-\sigma} Z_{rt}^{\sigma-1} \sum_j \left(\frac{\tau_{rjM}}{P_{jMt}}\right)^{1-\sigma} \vartheta_{jMt} \Gamma_\zeta L_{jt} \overline{w}_{jt},$$

where we used that total payments to production workers are a constant fraction, $\frac{\sigma-1}{\sigma}$, of total sales. We denote by H_{rPt} and H_{rEt} the total mass of non-agricultural workers engaged in production and entry, respectively, so that $H_{rPt} + H_{rEt} = H_{rMt}$. The mass of local varieties, N_{rt}, itself is determined from free entry as $N_{rt} = \frac{1}{f_E} H_{rEt} = \frac{1}{\sigma f_F} H_{rMt}$ and $H_{rPt} = \frac{\sigma-1}{\sigma} H_{rMt}$. Hence,

$$w_{rMt} = \frac{1}{\sigma f_F} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} w_{rMt}^{1-\sigma} Z_{rMt}^{\sigma-1} \sum_j \left(\frac{\tau_{rjM}}{P_{jMt}}\right)^{1-\sigma} \vartheta_{jMt} \Gamma_\zeta L_{jt} \overline{w}_{jt},$$

which implies that

$$w_{rMt}^{\sigma} = \frac{1}{\sigma f_F} \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} Z_{rMt}^{\sigma-1} \mathcal{D}_{rt}, \tag{A.4}$$

where $\mathcal{D}_{rt} = \sum_j \tau_{rjM}^{1-\sigma} P_{jMt}^{\sigma-1} \left(1 - \vartheta_{jAt}\right) \Gamma_\zeta L_{jt} \overline{w}_{jt}$ was defined in equation (7) and the non-agricultural spending share ϑ_{rMt} is given in equation A.10.

These equations fully determine the equilibrium. In particular, upon substituting for ϑ_{rMt} and $\mathcal{V}_{rt}$, equations (A.1), (A.2), and (A.4) are $3 \times R$ equations in the $3 \times R$ unknowns $\{w_{rAt}, w_{rMt}, L_{rt}\}$.

A.2 Additional Derivations

In this section, we present additional derivations omitted from the body of the paper.

A.2.1 Labor Supply

We denote total payments per efficiency unit of labor in region r and sector s by w_{rs}. In non-agriculture, these payments reflect only the wage per efficiency unit. In agriculture, they also include the payments to land which are redistributed to workers, so:

$$w_{rA} \equiv \tilde{w}_{rA} + \frac{\alpha}{1-\alpha} \tilde{w}_{rA} = \frac{1}{1-\alpha} \tilde{w}_{rA},$$

where $\tilde{w}_{rA}$ denotes the wage per efficiency unit in agriculture in region r.

Individual workers learn the amount of efficiency units of labor they can supply to either sector once they arrive in a location. 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, $P\left(z_s^i \le z\right) = F_s\left(z\right) = e^{-z^{-\zeta}}$, where ζ the dispersion of efficiency units across workers in sector s.

Worker i then chooses their sector so as to maximize their labor income, $y_r^i = \max_s \left\{w_{rs} z_s^i\right\}$. A standard set of arguments implies the following analytical expressions for the key objects of our theory.

1. Sectoral employment shares are

$$s_{rs} = \left(\frac{w_{rs}}{\overline{w}_r}\right)^{\zeta} \quad \text{where} \quad \overline{w}_r = \left(\sum_s w_{rs}^{\zeta}\right)^{1/\zeta}. \tag{A.5}$$

2. The aggregate amounts of sectoral human capital are

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

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

3. Total sectoral earnings are

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

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

$$F_r(y) \equiv P\left(y_r^i \le y\right) = e^{-\left(\sum_s w_{rs}^{\zeta}\right) y^{-\zeta}} = e^{-\left(y/\overline{w}_r\right)^{-\zeta}}. \tag{A.6}$$

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

A.2.2 The PIGL Demand Function

Consider the indirect utility function given in equation (1). Roy's Identity implies that sectoral expenditure shares are given by the following formula:

$$\vartheta_s \equiv \vartheta_s\left(y, P_{rA}, P_{rM}\right) = -\frac{\frac{\partial V\left(y, P_{rA}, P_{rM}\right)}{\partial p_s} P_{rs}}{\frac{\partial V\left(y, P_{rA}, P_{rM}\right)}{\partial y} y}. \tag{A.7}$$

We compute the numerator and denominator separately and then combine them. The numerator can be written:

$$\begin{aligned}\frac{\partial V\left(y,P_{rA},P_{rM}\right)}{\partial P_{rA}}P_{rA} &= \frac{\partial}{\partial P_{rA}}\left[\frac{1}{\eta}\left(\frac{y}{P_{rA}^{\phi}P_{rM}^{1-\phi}}\right)^{\eta}-\nu\ln\left(\frac{P_{rA}}{P_{rM}}\right)\right]P_{rA}\\ &= -\phi\left(\frac{y}{P_{rA}^{\phi}P_{rM}^{1-\phi}}\right)^{\eta}-\nu,\end{aligned}$$

while the denominator has the following expression:

$$\frac{\partial V\left(y,P_{rA},P_{rM}\right)}{\partial y}y=\left(\frac{y}{P_{rA}^{\phi}P_{rM}^{1-\phi}}\right)^{\eta}.$$

Combining the two previous derivatives using the expression in equation A.7 yields the following expressions of the sectoral expenditure shares:

$$\vartheta_A=\phi+\nu\left(\frac{y}{P_{rA}^{\phi}P_{rM}^{1-\phi}}\right)^{-\eta}\qquad \vartheta_M=(1-\phi)-\nu\left(\frac{e}{P_{rA}^{\phi}P_{rM}^{1-\phi}}\right)^{-\eta}. \tag{A.8}$$

The Allen-Uzawa elasticity of substitution is given by

$$\varrho=\frac{\frac{\partial^2 e\left(P_{rA},P_{rM},V\right)}{\partial P_{rA}\partial P_{rM}}e\left(P_{rA},P_{rM},V\right)}{\frac{\partial e\left(P_{rA},P_{rM},V\right)}{\partial P_{rA}}\frac{\partial e\left(P_{rA},P_{rM},V\right)}{\partial P_{rM}}},$$

where $e\left(P_{rA},P_{rM},V\right)$ is the expenditure function given by

$$e\left(P_{rA},P_{rM},V\right)=\left(V+\nu\ln\left(P_{rA}/P_{rM}\right)\right)^{1/\eta}\eta^{1/\eta}P_{rA}^{\phi}P_{rM}^{1-\phi}. \tag{A.9}$$

Using equation (A.9), one can show that

$$\varrho=1-\eta\frac{\left(\vartheta_A-\phi\right)\left(\vartheta_M-(1-\phi)\right)}{\vartheta_A\vartheta_M}=1+\eta\frac{\left(\vartheta_A-\phi\right)^2}{\vartheta_A\left(1-\vartheta_A\right)}.$$

A.2.3 PIGL Aggregation

In this section we derive the aggregate demand system and the expression for special welfare introduced in Section 2.3.

Aggregate Demand

Let $F_r(y)$ be the distribution of income derived in equation (A.6). Integrating over the sectoral expenditure shares of individual workers in region r in equation A.8 yields an expression for a region's aggregate expenditure share:

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

Given that $F_r(y) = e^{-(y/\overline{w}_r)^{-\zeta}}$, we have that

$$P\left(y^{1-\eta} < m\right) = P\left(y < m^{\frac{1}{1-\eta}}\right) = e^{-\left(\frac{m^{\frac{1}{1-\eta}}}{\overline{w}_r}\right)^{-\zeta}} = e^{-\left(\frac{m}{\overline{w}_r^{1-\eta}}\right)^{-\frac{\zeta}{1-\eta}}}.$$

Hence,

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

so that

$$\vartheta_{rA} = \phi + \nu \frac{\Gamma_{\frac{\zeta}{1-\eta}}}{\Gamma_\zeta} \left(\frac{\overline{w}_r}{p_{rA}^{\phi} p_{rM}^{1-\phi}}\right)^{-\eta} = \phi + \nu^{RC} \left(\frac{\overline{w}_r}{p_{rA}^{\phi} p_{rM}^{1-\phi}}\right)^{-\eta}, \tag{A.10}$$

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

Indirect Utility

Using the indirect utility function in equation (1), we derive the following expression for the, expected utility in region r:

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

Workers effectively draw their income from the Fréchet distribution in equation (A.6) upon arriving in their region of choice. We use the properties of the Fréchet distribution to show that y^{η} itself is drawn from a Fréchet distribution with a shape parameter ζ/η

and scale $\overline{w}_r^{\eta}$:

$$P\left(y^{\eta}<m\right)=P\left(y<m^{\frac{1}{\eta}}\right)=e^{-\left(\frac{m^{\frac{1}{\eta}}}{\overline{w}_r}\right)^{-\zeta}}=e^{-\left(\frac{m}{\overline{w}_r^{\eta}}\right)^{-\frac{\zeta}{\eta}}}.$$

By implication, $\int y^{\eta}dF_r\left(y\right)=\Gamma\left(1-\frac{1}{\zeta/\eta}\right)\overline{w}_r^{\eta}=\Gamma_{\frac{\zeta}{\eta}}\overline{w}_r^{\eta}$,so that

$$E\left[V\left(y,P_{rA},P_{rM}\right)\right]=\frac{1}{\eta}\Gamma_{\frac{\zeta}{\eta}}\left(\frac{\overline{w}_{rt}}{P_{rA}^{\phi}P_{rM}^{1-\phi}}\right)^{\eta}-\nu\ln\left(\frac{P_{rA}}{P_{rM}}\right). \tag{A.11}$$

This is the expression in equation (6).

A.3 Proofs of Propositions

In this section, we present the proofs of Propositions 1 and 2 which were omitted from the body of the paper.

A.3.1 Proposition 1

First, rewrite equation (A.4) as follows

$$w_{rMt}=\left(\frac{1}{f_E}\right)^{\frac{1}{\sigma}}\left(\frac{1}{\sigma}\right)\left(\sigma-1\right)^{\frac{\sigma-1}{\sigma}}Z_{rMt}^{\frac{\sigma-1}{\sigma}}\mathcal{D}_{rt}^{\frac{1}{\sigma}}\equiv\mathcal{Z}_{rMt}, \tag{A.12}$$

Upon defining $\mathcal{Z}_{rAt}\equiv Z_{rAt}\left(\Gamma_{\zeta}\ell_{rt}\right)^{-\alpha}$ where $\ell_{rt}\equiv\frac{L_{rt}}{T_r}$ is population density, equation (A.3) reads

$$w_{rAt}^{\zeta-1+\frac{1}{\alpha}}=\overline{w}_{rt}^{\zeta-1}\mathcal{Z}_{rAt}^{\frac{1}{\alpha}}=\left(w_{rAt}^{\zeta}+w_{rMt}^{\zeta}\right)^{\frac{\zeta-1}{\zeta}}\mathcal{Z}_{rAt}^{\frac{1}{\alpha}}=\left(w_{rAt}^{\zeta}+\mathcal{Z}_{rMt}^{\zeta}\right)^{\frac{\zeta-1}{\zeta}}\mathcal{Z}_{rAt}^{\frac{1}{\alpha}}.$$

Rearranging terms yields

$$1=\left(1+\left(\frac{\mathcal{Z}_{rMt}}{w_{rAt}}\right)^{\zeta}\right)^{\frac{\zeta-1}{\zeta}}\left(\frac{\mathcal{Z}_{rAt}}{w_{rAt}}\right)^{\frac{1}{\alpha}}.$$

Using the definition of $\overline{w}_r=\left(w_{rM}^{\zeta}+w_{rA}^{\zeta}\right)^{1/\zeta}$, it follows that $w_{rA}=\left(\overline{w}_r^{\zeta}-w_{rM}^{\zeta}\right)^{1/\zeta}=$

$\left(\overline{w}_r^\zeta - \mathcal{Z}_{rM}^\zeta\right)^{1/\zeta}$. Hence, $\overline{w}_r$ is determined from

$$\left(\left(\left(\frac{\mathcal{Z}_{rM}}{\left(\overline{w}_r^\zeta - \mathcal{Z}_{rM}^\zeta\right)^{1/\zeta}}\right)^\zeta + 1\right)^{\frac{\zeta-1}{\zeta}}\left(\frac{\mathcal{Z}_{rA}}{\left(\overline{w}_r^\zeta - \mathcal{Z}_{rM}^\zeta\right)^{1/\zeta}}\right)\right)^{\frac{1}{\alpha}} = 1$$

Rearranging terms yields

$$1 = \left(\frac{\mathcal{Z}_{rM}}{\overline{w}_r}\right)^\zeta + \left(\frac{\mathcal{Z}_{rA}}{\overline{w}_r}\right)^{\frac{\zeta}{\alpha(\zeta-1)+1}},$$

which is the first result in Proposition 1.

To derive the second result in Proposition 1, note that sectoral employment shares satisfy $s_{rAt}/\left(1 - s_{rAt}\right) = \left(w_{rAt}/w_{rMt}\right)^\zeta$. As a result, equation (A.3) can be written as $w_{rAt} = s_{rAt}^{-\frac{\zeta-1}{\zeta}\alpha}\mathcal{Z}_{rAt}$, so that

$$\frac{s_{rAt}}{1 - s_{rAt}} = \left(\frac{s_{rAt}^{-\frac{\zeta-1}{\zeta}\alpha}\mathcal{Z}_{rAt}}{\mathcal{Z}_{rMt}}\right)^\zeta = \left(\frac{\mathcal{Z}_{rAt}}{\mathcal{Z}_{rMt}}\right)^\zeta s_{rAt}^{-(\zeta-1)\alpha}.$$

Rearranging terms yields $\frac{s_{rAt}^{1+(\zeta-1)\alpha}}{1-s_{rAt}} = \left(\frac{\mathcal{Z}_{rAt}}{\mathcal{Z}_{rMt}}\right)^\zeta$.

A.3.2 Proposition 2

The wage exposure elasticity $\phi\left(s_{rA}\right)$ The first result in Proposition Proposition 1, directly implies the following:

$$d\ln w_{rMt} = d\ln \mathcal{Z}_{rMt}. \tag{A.13}$$

Taking the total derivative of the expression for the agricultural wage in Proposition 1yields:

$$d\ln w_{rAt} = \frac{\left(1 - s_{rAt}\right)\gamma}{1 + \left(1 - s_{rAt}\right)\gamma}d\ln \mathcal{Z}_{rMt} + \frac{1}{1 + \left(1 - s_{rAt}\right)\gamma}d\ln \mathcal{Z}_{rAt}. \tag{A.14}$$

Finally, we can take the total derivative for the expression for the average income in a location in equation (A.5), and combine it with equations (A.13) and (A.14) to obtain:

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

where $\phi(s_{rAt}) = \frac{(1+\gamma)(1-s_{rAt})}{1+(1-s_{rAt})\gamma}$.

The industrialization elasticity $\psi(s_{rA})$ To derive the change in s_{rAt}, we take the total derivative of equation (A.5) and combined it with the expression for $d\ln\overline{w}_{rt}$ above to obtain:

$$
d\ln s_{rAt} = \zeta\,(d\ln w_{rAt} - d\ln\overline{w}_{rt}) = \zeta\,(1-s_{rAt})\,(d\ln w_{rAt} - d\ln w_{rMt})\,.
$$

Using the expressions in equations (A.14) and (A.13), this implies that

$$
d\ln s_{rAt} = \frac{(1-s_{rAt})\,\zeta}{1+(1-s_{rAt})\,\gamma}\,(d\ln \mathcal{Z}_{rAt} - d\ln \mathcal{Z}_{rMt})\,. \tag{A.15}
$$

Finally, using that $ds_{rAt} = s_{rAt}d\ln s_{rAt}$,yields the expression in Proposition 2.

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 used throughout the paper is the “commuting zone” defined by Tolbert and Sizer (1996). These were introduced into the economics literature by Autor and Dorn (2013). We choose these units since they capture integrated labor market areas within which migration frictions are unlikely to play a role. During the period of our study, county boundaries were subject to substantial changes. To ensure consistent treatment, we use the crosswalk described in Eckert et al. (2020b) to map historical county boundaries to the time-invariant commuting zone delineations of Tolbert and Sizer (1996).

Since data collection by the US Statistical Office only occurred systematically in states that formed part of the Union, we drop data from states that joined the Union after 1870. We chose a cutoff a decade before the start of our period of analysis since the collection of the decennial census took considerable time, so data in the 1880 Census may be incomplete for States that joined the Union less than ten years earlier. As a result, we exclude the following states from our sample and drop them from all our data sets (year of eventual accession to the Union in parentheses): Colorado (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), Hawaii (1959). Figure A.1 shows agricultural employment shares in 1880 for each commuting zone in our final sample.

B.1.1 Full Count Decennial Census, 1880-1920

Source and Description We obtained the full count decennial census micro-data files for the years 1880, 1900, and 1920 from the IPUMS database (see Ruggles, Genadek, Goeken, Grover, and Sobek (2017)). We selected the following variables: state, county, age , school attendance ("school"), 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.

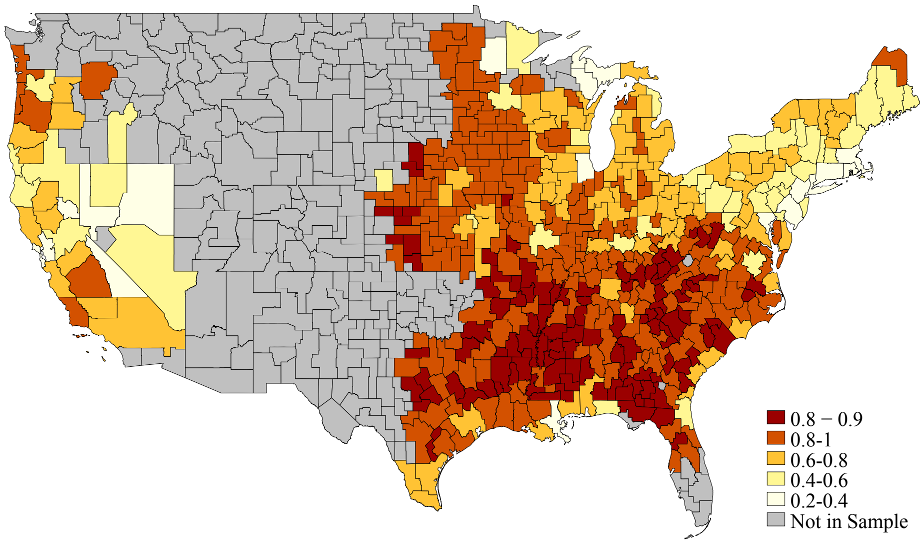
Sample Selection, Processing, and Use In the data, we define different groups of observations used in various parts of the paper. We define "workers" as observations with an industry identifier and age between 20 and 60 years. We define "agricultural workers" as workers who work in Agriculture, Forestry, 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. In Figure A.1, we depict 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}.

"Adolescents" as observations with age between 6 and 18 years who do not work, are white, and are male. We define "adolescents in school" as adolescents who are currently attending school. We use these two groups of workers to compute a measure of the rate of school attendance for each commuting zone.

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

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



Notes: The map shows all commuting zones in the United States. The colors reflects the agricultural employment share bin into which individual commuting zones fall. Darker shades correspond to higher agticultural employment shares. Grey commuting zones are not in our sample since their corresponding state was not part of the Union of US states by 1870.

flows to distance.

B.1.2 Census of Manufacturing

Source and Description We obtained county-level tabulations of the Census of Manufacturing data for the years 1880, 1900, and 1920 from the NHGIS database (see Manson et al. (2017)). We selected the following variables: total manufacturing payroll, total manufacturing employment, number of manufacturing establishments, and capital (real and personal) invested in iron and steel manufacturing establishments.

Sample Selection, Processing, and Use We drop all counties for which manufacturing payroll or employment is zero or missing. We then compute average manufacturing wages in each county by dividing total manufacturing payroll by the number of manufacturing workers. Throughout the paper, we refer to this ratio simply as "average wage" or "earnings ".We compute commuting zone-level average wages by taking the payroll-weighted average across county-level average wages within each commuting zone. In our model average wages are the same in both sectors, so that the average manufacturing wage in the data correspond to the average commuting zone wage in the model, $\overline{w}_{rt}$. We compute average establishment size for each county by dividing total manufacturing employment by the number of manufacturing establishments.

B.1.3 Census of Agriculture

Source and Description We obtained county-level tabulations of the Census of Agriculture for the years 1880, 1900, and 1920 from the NHGIS database (see Manson et al. (2017)). We selected the following variables: average land value per acre, acres of improved farm land, total number of farms, and total value of farm implements and machinery.

Sample Selection, Processing, and Use We drop all counties for which average land value per acre is zero or missing. We compute commuting zone-level average land values per acre by taking the area-weighted average across the land values in all the counties contained in a given commuting zone. We interpret these data in 1880 as land rents in the model and use them to identify the supply of agricultural land in each commuting zone in 1880, T_r. We compute average farm size for each county by dividing the total number of improved acres in farms by the total numbers of farms.

B.1.4 Linked Census Files

Source and Description Economists have written algorithms to match workers across sequential Decennial Census waves based on their names and a variety of other characteristics. IPUMS itself provides a matched file that lists individuals that appear both in 1880 and 1900 (see Ruggles et al. (2017)). In addition, Abramitzky, Boustan, Eriksson, Rashid, and Pérez (2022) provide linked files for various pairs of years. We use the 1880-1900 linked file from IPUMS and from Abramitzky et al. (2022).

Sample Selection, Processing, and Use Both samples only include men, since women's surnames changed frequently making it difficult to match them over time. We only keep observations who are workers according to our definition of workers in the full count Census files.

We use the linked data to compute the share of workers moving from commuting zone r to commuting zone r' between 1880 and 1900. We use the resulting "commuting-zone-to-commuting-zone-migration matrix" to estimate the elasticity of migration flows 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 use the series on real GDP and the series for the price of farm goods and the prices of all commodities other than farm goods.

Sample Selection, Processing, and Use Moments from both the GDP and the price series serve as targets in our estimation. We interpret the price series for farm goods as the price series for agricultural prices in our model, and the series on non-farm commodities as that of manufacturing goods. We target the growth rate of real GDP and relative prices between 1880-1900 and 1900-1920 in our estimation.

B.1.6 Historical Bank Branches

Source and Description We obtained data on the number of private banks for each US county for 1880 and 1910 from Jaremski and Fishback (2018).[18]

Sample Selection, Processing, and Use We merge these data with our Census data on the number of workers in each commuting zone to compute the change in the log of the number of bank branches per worker in each commuting zone. Since there are no branches in many commuting zones in 1880, we add a 1 to each observation in the bank branch data. Our results are robust to simply dropping observations with zero branches in 1880 instead.

B.2 Details on Estimation Moments and Methods

B.2.1 Local Employment Growth

Our model accounts for growth in the local labor force 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 aspects 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.

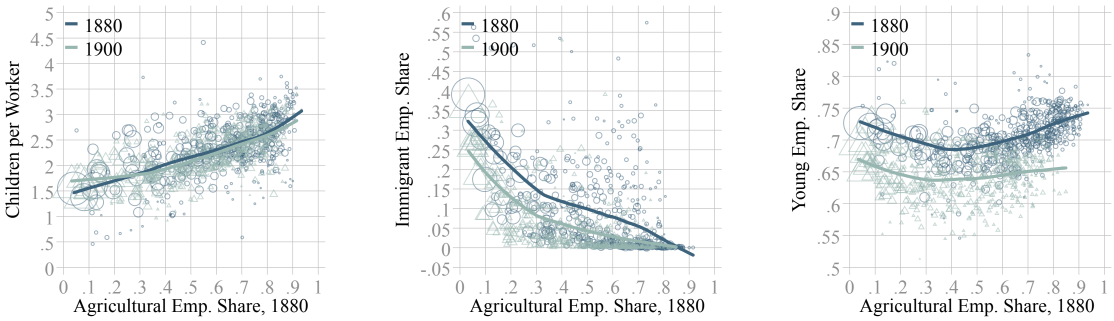
Figure A.2 shows proxies for the three most important sources of local employment growth: births, immigration, and deaths. The rightmost panel shows the number of children per adult ("birth rates"). We measure local "birth rates" as the fraction of children between 0 and 20 relative to the number of working adults aged 20-60. Rural locations have substantially higher birth rates and hence experience faster innate employment growth.

The middle panel shows the the correlation of the share of immigrants in the local workforce and initial agricultural employment shares. We measure immigrants as the share of workers that immigrated in the last 20 years from a foreign country. Immigrants predominantly settled in urban locations, and thus raised the employment of such non-agricultural locations.

[18] We thank Matt Jaremski for his work in compiling these data and his generosity in sharing them with us.

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

(A) FERTILITY (B) IMMIGRATION (C) AGE



Notes: The left panel shows a proxy for the local birth rate, i.e., the share of children between 0 and 20 relative to the number of working adults aged 20-60, as a function of the initial agricultural employment share. The middle panel shows the share of immigrants in the local work force as function of the initial agricultural employment share. The right panel shows the fraction of young workers among the total workforce in each commuting zone. Total workers are defined as all individuals aged 20-60 that have an industry identifier. Young workers are defined as all individuals aged 20-40 that have an industry identifier. The underlying data source for all three panels are the US Decennial Census files for 1880 and 1900. The size of the symbols is proportional to a regions total employment. The graph shows the fit line of a local polynomial regression.

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.

We now show how we use these data to estimate the exogenous component of local employment growth n_{rt} in each region. To do so, recall that we denote by L^Y_{rt} the number of workers in a location at the beginning of period t, i.e., 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 (or remain in) location r during period t. The local rate of exogenous labor force growth, n_{rt}, is thus defined by $L^Y_{rt+1} = n_{rt} L_{rt}$. 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:

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

where $Exit_{rt-1,t}$ is the number of workers exiting 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, δ, 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

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

Let C_{rt} denote the number of children in r at time $t-1$ and I_{rt} denote the number of working immigrants in region r that arrived between $t-1$ and t. Let ι be the fraction of children that join the labor force. 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

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

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

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

where C_{rt}, I_{rt} and L_{rt-1} are observed in the data.

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:

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

Rearranging terms implies that

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

Hence, for a given exit rate δ and labor force participation rate ι 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

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

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

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

$$\text{(A.16)} \qquad \text{Share of old workers}_t = \frac{(1-\delta)\sum_r L_{rt-1}}{\sum_r L^Y_{rt}} = (1-\delta)\frac{\sum_r L_{rt-1}}{\sum_r L_{rt}}.$$

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 δ for any target of the share of old workers. A generation in our model corresponds to 20 years in the data. In calibrating δ, 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.16) implies that δ is given by 0.4.

To calibrate ι, we combine the mortality rate of children with the rate of labor force participation among 20-40 year olds in 1900. which is about 0.5 in our data, which comprises both men and women. As a result, we set $\iota = 0.5$.

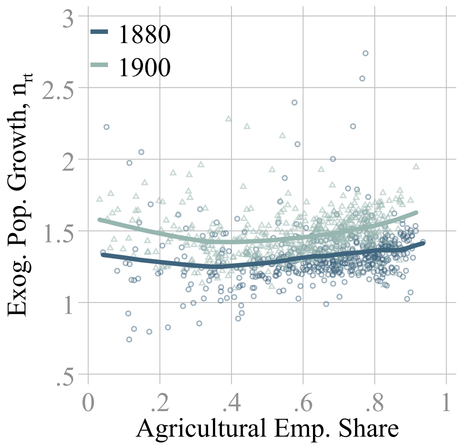
In Figure A.3, we show the calibrated exogenous rate of employment growth, n_{rt}, for the two time periods 1880-1900 and 1900-1920. The figure shows that, on net, exogenous employment growth was slightly higher in agricultural regions. The relationship between agricultural specialization and subsequent 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.

TABLE A.1: EMPLOYMENT GROWTH ACROSS REGIONS: 1880-1900 AND 1900-1920

	EMPLOYMENT GROWTH ($\Delta \log L_{rt}$)				
s_{rAt}	-0.360***	-0.753***	-0.358***	-0.773***	-0.782***
	(0.0309)	(0.0509)	(0.0825)	(0.0228)	(0.0272)
R^2	0.167	0.315	0.138	0.326	0.475
Observations	990	990	990	3910	3910
Geography	CZ	CZ	CZ	County	County
FEs		State	State	State	CZ
Weighted	Yes	Yes		Yes	Yes

Notes: All regression are weighted by initial total employment of the commuting zone (columns 1-3) or county (4-5) and include decade fixed effects. Columns 2-4 also contain state fixed effects and column 5 commuting zone fixed effects.Robust standard errors in parentheses. * , **, and *** denote statistical significance at the 10%, 5% and 1% level respectively.

FIGURE A.3: EXOGENOUS EMPLOYMENT GROWTH



Notes: This figures shows the calibrated rate of exogenous employment growth across commuting zones between 1880-1900 and 1900-1920. In each year there is one region with a rate above 3 which we drop to shows the variation among the remaining observations in more detail.

Finally, in Table A.1 we report the relationship between agricultural employment share and future population growth. The structure of Table A.1 is identical to Table 1 in Section 1. In our quantitative analysis, we calibrate our model to the coefficient reported in the first column.

B.2.2 Migration Gravity Equations

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

$$M_{rr't} = m_{rr't} L_{rt}^{Y} = \frac{\left(\mu_{rr'} \mathcal{V}_{r't} \mathcal{B}_{r't}\right)^{\varepsilon}}{\sum_j \left(\mu_{rj} \mathcal{V}_{jt} \mathcal{B}_{jt}\right)^{\varepsilon}} L_{rt-1} n_{rt-1}.$$

We project the moving cost between two regions on the physical distance between them, i.e., we set $\mu_{rr'} = d_{rr'}^{-\kappa}$., where the parameter κ parameterizes the distance cost of migration. The larger κ, the more the destination utility of areas furhter away is discounted. In our empirical estimation, we set $d_{rr'}\ \forall r = r'$ to the average distance between county centroids within a commuting zone, and $d_{rr'}\ \forall r \neq r'$ to the distance between commuting zone centroids.

Taking logs on both sides and grouping terms then yields:

$$\log M_{rr't} = \alpha_{r't} + \beta_{rt} - \kappa\epsilon \log d_{rr'} \tag{A.17}$$

where

$$\alpha_{r't} = \varepsilon \log(\mathcal{V}_{r't}\mathcal{B}_{r't}) \quad \text{and} \quad \beta_{rt} = \log(L_{rt-1}n_{rt-1}) - \log\left(\sum_{r''}\left(d_{rr''}^{-\kappa}\mathcal{V}_{r''t}\mathcal{B}_{r''t}\right)^{\varepsilon}\right).$$

Since we calibrate the model at the commuting zone level, the indices r and r' refer to commuting zones. Equation A.17 suggests a fixed effect regression of commuting zone migration flows to recover to recover the elasticity of migration flows to distance, $\kappa\epsilon$, 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. Since linking rates are relatively low, the majority of bilateral commuting zone pairs exhibit *no* migration flows between 1880 and 1900. We hence estimate equation (A.17) using Poisson Pseudo Maximum Likelihood (PPML), as proposed by Silva and Tenreyro (2006). More specifically, we estimate the following equation using PPML:

$$M_{rr't} = \exp\left(\alpha_{r't} + \beta_{rt} - \kappa\epsilon \log d_{rr'}\right) + \epsilon_{rr't}. \tag{A.18}$$

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 ($\kappa\varepsilon$) of around 2.75.

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

TABLE A.2: MIGRATION GRAVITY EQUATIONS

	LOG BILATERAL MIGRATION PROBABILITY				
Log Distance	-2.922***	-2.632***	-3.925***	-2.262***	-2.291***
	(0.0327)	(0.0111)	(0.0572)	(0.0306)	(0.0311)
R^2	0.8135	0.8989	0.9296	0.806	0.799
Observations	254762	349230	3983	3983	3935
Year FE			Yes	Yes	Yes
Geography	CZ	CZ	State	State	State
Estimator	PPML	PPML	PPML	OLS	OLS

Notes: 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.

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 estimate is larger than in 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-by-state migration matrix has very few pairs of states with zero flows. Across the two cross-sections of data for 1880-1900 and 1900-1920, about 50 pairs exhibit zero flows.As a result, we can also estimate the gravity regression using simple OLS. instead of PPML. Columns 4 and 5 in Table A.2 report estimates from a regression where we simply replace zeros with 1s and another where we omit all zero-valued pairs of states. Since states further apart are more likely to report zero flows *because* they are further apart, dropping them leads to a smaller estimate of the elasticity at 2.62, the lowest of all our estimates.

Note that our theory only produces an approximate gravity equation for the flows between groups of commuting zones (such as states) because of Jensen's inequality. As a result, the distance elasticities stemming from state-level data do not map directly to the structural parameters $\kappa\varepsilon$. Nevertheless, we find it re-assuring that the state-level regressions estimates are not too dissimilar from the regression estimates using commuting zone. Furthermore, since a fraction of moves in the model happen across commuting zones within the same state, we would expect the state-to-state distance coefficient to be larger than the commuting zone to commuting zone one in model-generated data, too.

B.2.3 Computing Macroeconomic Aggregates

Aggregate GDP Growth

We measure the growth rate of aggregate GDP using the Fisher chained index. The Fisher Index is defined by

$$g_t^F = \sqrt{\frac{P_{ct-1}C_t}{p_{ct-1}C_{t-1}} \times \frac{P_{ct}C_t}{P_{ct}C_{t-1}}},$$

where P_{ct} is the price of the consumption good at time t and C_t is the quantity.

In the context of our model with R regions and two sectors s, we construct the following auxiliary indices:

$$S_{t-1}(P_{t-1}) = \sum_{r=1}^{R}\sum_{s=1}^{S} P_{rst-1}C_{rst-1} \quad \text{and} \quad S_{t-1}(P_t) = \sum_{r=1}^{R}\sum_{s=1}^{S} P_{rst}C_{rst-1}.$$

and

$$S_t(P_{t-1}) = \sum_{r=1}^{R}\sum_{s=1}^{S} P_{rst-1}C_{rst} \quad \text{and} \quad S_t(P_t) = \sum_{r=1}^{R}\sum_{s=1}^{S} P_{rst}C_{rst}.$$

where $P_{rst}(C_{rst})$ denotes the prices (consumption quantities) of sector s goods in region r at time t.

We then combine these expressions into the corresponding Fisher Index as follows:

$$g_t^F = \sqrt{\frac{S_t(P_{t-1})}{S_{t-1}(P_{t-1})} \times \frac{S_t(P_t)}{S_{t-1}(P_t)}}.$$

In our model, C_{rst} can be computed as

$$\begin{aligned} C_{rAt} &= \frac{\vartheta_{rAt}\int y dF_r(y)}{P_{rAt}} = \frac{\vartheta_{rAt}\Gamma_\zeta \overline{w}_{rt}L_{rt}}{P_{rAt}} \\ C_{rMt} &= \frac{(1-\vartheta_{rAt})\int y dF_r(y)}{P_{rMt}} = \frac{(1-\vartheta_{rAt})\Gamma_\zeta \overline{w}_{rt}L_{rt}}{P_{rMt}}. \end{aligned}$$

Relative Prices

To compute the time-series of the relative price of non-agricultural to agricultural goods, we compute chained sectoral price indices and then take their ratio. More specifically, consider sector s and time-period between $t-1$ and t. Let P_{st}^L and P_{st}^P denote the

Laspeyres and Paasche indices, respectively. These are given by

$$P_{st}^{L} = \frac{\sum_r P_{rst} C_{rst-1}}{\sum_r P_{rst-1} C_{rst-1}} \quad P_{st}^{P} = \frac{\sum_r P_{rst} C_{rst}}{\sum_r P_{rst-1} C_{rst}}.$$

The Fisher Index for sector s is then given by $P_{st}^{F} = \sqrt{P_{st}^{L} \times P_{st}^{P}}$.. The time-series of the relative price is then given by $\mathcal{P}_t^{M-A} = P_{Mt}^{F} / P_{At}^{F}$.

B.2.4 Cross-sectional Estimates of the Engel Elasticity η

Targeting the time series of the agricultural employment share implied an estimate of $\eta = 0.93$. However, our model also implies a log-linear relationship between individuals' expenditure share on agricultural products and their total expenditure that can be used to estimate η from cross-sectional microdata:

$$\ln \vartheta_A \left(y, P_{r,M}\right) = \ln\left(-\nu P_{r,M}^{-1}\right) - \eta \ln y, \tag{A.19}$$

where we used our estimate $\phi \approx 0$. We use the 1936 Consumer Expenditure Survey (CEX) by the U.S. Bureau of Labor Statistics obtained from the Inter-university Consortium for Political and Social Research (ICPSR) to provide direct evidence on the log-linear relationship between expenditure shares and total expenditure.[19]

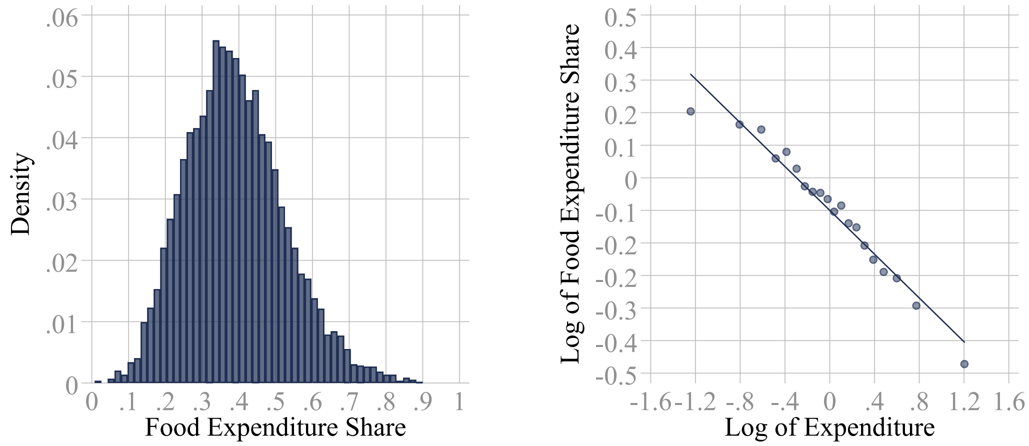
The CEX contains micro data on expenditure of individuals on a large variety of categories and a swath of individual characteristics. We use the household files. We obtain information on households' total expenditure, expenditure on food, urban/rural status, size, interview data, occupation and industry of household head, race of household head, and county of residence.

In Figure A.4, we show that this log-linear relationship is a good description of the data. In the left panel, we display the cross-sectional distribution of food shares. Empirically this variation is substantial, ranging from 5% to 80%. In the right panel, we show the empirical relationship between log expenditure and log food shares as a binned scatter plot. As implied by our theory, the elasticity between food shares and expenditure is indeed essentially constant across the entire range of the distribution of expenditure.

The slope coefficient falls in between $\eta \in (0.315, 0.362)$ depending on which additional controls are chosen, implying that our "macroestimate" of $\eta = 0.93$ is higher than the microestimate that exploits cross-sectional variation.

[19]We note that our theory is written in terms of value added. The expenditure data is in terms of final expenditure data. Herrendorf, Rogerson, and Valentinyi (2013) show that in general there is no direct mapping between the preference parameters of the value added and the final good demand system. However, Fan et al. (2022) show that for the class of PIGL preferences used here, the Engel elasticity η is portable between the final good and value added demand system.

FIGURE A.4: HETEROGENEITY IN FOOD EXPENDITURE SHARES



Notes: The figure shows the cross-sectional distribution of the individual expenditures shares on food (left panel) and the bin scattered relationship between the (log) expenditure share on food and (log) total expenditure (right panel). The relationship in the right panel is conditional on a set of location and family size fixed effects.

B.2.5 Validating the First-Order Approximation

In Figure 7, we reported the decomposition of local wage growth into the four components highlighted in Proposition 2. In the theory outlined in the paper, the underlying first-order approximation that decomposes wage growth into the four margins takes the following form:

(A.20)
$$d\ln \overline{w}_{rt} = \phi_M\left(s_{rA}\right)\left(\frac{1}{\sigma}d\ln \mathcal{D}_t + \frac{\sigma-1}{\sigma}d\ln Z_{rMt}\right) + \phi_A\left(s_{rA}\right)\left(d\ln Z_{rAt} - \alpha d\ln \ell_{rt}\right).$$

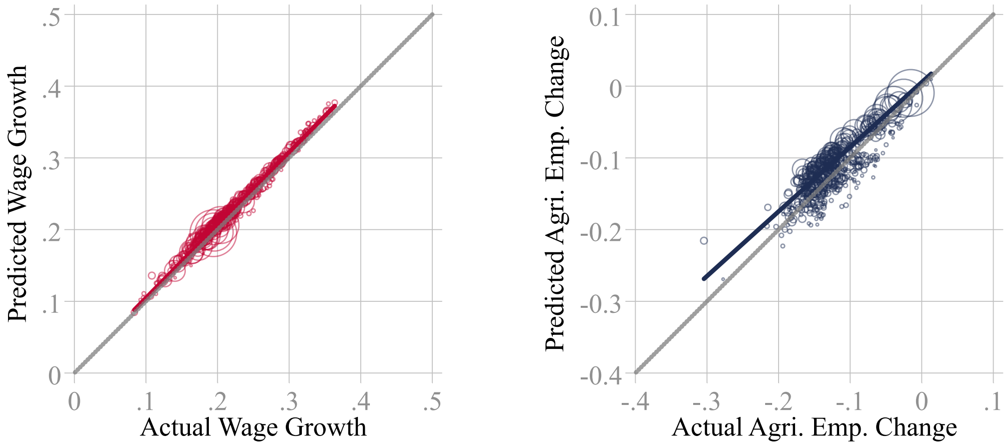
The theory also permits a similar first-order approximation can be derived for the change in local agricultural employment shares

$$ds_{rAt} = \psi\left(s_{rA}\right)\left(d\ln Z_{rAt} - \alpha d\ln \ell_{rt} - \frac{\sigma-1}{\sigma}d\ln Z_{rMt} - \frac{1}{\sigma}d\ln \mathcal{D}_{rt}\right), \tag{A.21}$$

In the quantitative model there is an additional "agricultural demand" term in these approximations that emerges once there are trade costs for agricultural goods, which we abstracted from in the main theory but re-introduce in the quantitative version of the model.

In Figure A.5, we show that equations A.20 and A.21 provide an excellent fit of the data despite being an approximation. This provides justification for using these equations to decompose local wage growth and industrialization in Section 4. Specifically, the left panel shows the correlation between local wage growth based on equation (A.20) and local wage growth stemming from the non-linear solution of the model. If the model were to follow equation (A.20) exactly, the results should lie on a 45 degree line. Each red dot represents a commuting zone, the grey dashed line is a 45 degree line, and the solid red line represents the best fit through the data. The right panel compares the change in the local agricultural employment share based on equation A.21 to the change in the local agricultural employment in the simulated model. The linear fit line is again very close to the grey dashed 45 degree line providing support for using the first-order approximation in our analysis.

FIGURE A.5: ASSESSING THE ACCURACY OF THE FIRST-ORDER-APPROXIMATIONS



Notes: The left panel shows the correlation between actual wage growth in the model and predicted wage growth based on equation A.20. The right panel shows the correlation between actual agricultural employment share changes in the model and predicted agricultural employment share changes based on a first order approximation in the model. Each dot a is a commuting zone. and the size of the dots is proportional to a commuting zone's total employment in 1880. The dashed grey line is a 45 degree line. The solid line in each panel is a weighted fit line using 1880 total employment as weights.