

The Geographic Bias of Growth and Structural Change

James Macek

November 11, 2021

Abstract

1 Introduction

Contemporary models of spatial equilibrium imply that productivity growth, when uniform across locations, does not change the overall distribution of economic activity in space. In this paper, I argue that this feature masks an important mechanism by which growth concentrates economic activity even when the spatial distribution of productivity remains unchanged.

The theory rests on two ideas. Firstly, growth changes the consumption basket; shifting expenditures away from food and toward other goods. This manifests in general equilibrium as *structural change*, where agricultural employment falls relative to other sectors as the economy grows. On the other hand, demand for agriculture shapes geography insofar as the returns to concentrating it in space are dwarfed by other sectors. Taken together, these imply that workers leaving agriculture during the structural change process are absorbed by regions with high population density. This reallocation increases the dispersion of population density across space.

This link between structural transformation and urbanization has previously been explored in [Michaels et al. \(2012\)](#). They build a theory and rationalize rising spatial concentration in part by appealing to differences in land intensity across sectors. As a result, population dense regions specialize outside of agriculture under trade and, conditional on specializing, tend to have more dispersion in population densities. I build on this research in two ways. As a small contribution, I provide conditions on the strength of agglomeration and productivity growth and explicitly prove this relationship holds. Secondly, and most importantly, I measure the magnitude of this effect by taking a spatial equilibrium model to Chinese data from 2000-2005. The theory reveals how to make this measurement.

The structural model allows for falling migration, regional trade costs and changing relative productivity as alternative explanations for the reallocation of workers across space. I find that productivity growth, when uniform across locations and set to the inferred average level, accounts for approximately 38 percent of the total rise in population density dispersion predicted by the

model¹. In addition, the model is not calibrated to match the observed rise in population density dispersion, but accounts for over half of it.

Why study China? During this period, the country saw a massive reallocation of workers toward the population-dense coastal provinces, such as Beijing, Shanghai and Guangdong. Such mobility was and is made exceptionally difficult by a unique institution– the Hukou system (Chan, 2019) (Meng, 2012). However, by the 2000’s the Chinese government introduced significant reform, and there is some quantitative evidence that these reforms caused migration inflows (Fan, 2018). At the very least, migration costs (measured by observing spatial differences in real income relative to migration flows) fell significantly during the period (Tombe and Zhu, 2019). These falling migration costs can increase the dispersion of population density as more workers arbitrage away spatial differences in real income. Moreover, Tombe and Zhu (2019) also find that falling trade costs during this period were also significant; and falling trade costs are key in determining the spatial distribution of economic activity in the economic geography literature. How does the impact of the theory stack up against these two variables, which themselves are largely influenced by policy? China provides a unique opportunity to answer this question.

At the same time, China also experienced rapid growth and structural change during this period. These facts alone distinguish it as a country where the geographic bias of growth could play a major role in shaping the population distribution.

The strategy behind calibrating this model starts with the fact that *relative* productivity growth across space is identified using regional trade data and a gravity model. However, this procedure is uninformative about *aggregate* growth– that is, a scale factor that determines productivity in absolute terms. I propose an estimation strategy that disciplines this growth using information contained in falling agricultural spending. First, I estimate income and substitution elasticities over agricultural goods using GDP data and prices inferred from the gravity model. Given these estimates, I choose the level of aggregate

¹The model I use is nonlinear, so there is some ambiguity surrounding the allocation of marginal effects between productivity growth, trade and migration costs. I report the Shapely decomposition here.

productivity growth that matches the fall in average agricultural spending observed in the data.

For other parameters, calibration uses recent methodology in the trade and spatial literature. Trade costs are mapped from provincial trade data and standard estimates of the trade elasticity. Combined with the calibrated productivity, these are used to construct measures of real income per capita implied by the model. Migration frictions are then inferred from the discrepancy between differences in real income and the observed movement of workers from their hukou.

The paper is organized as follows. I review the literature in Section 2, provide a description of the data and some key motivating facts in Section 3, present a theory and prove it can explain these facts in Section 4, and employ the structural model in Sections 5, 6 and 7. Section 8 concludes.

2 Literature

This paper lies at the heart of a budding literature on the interplay between structural change and geographic outcomes. The earliest of these frames structural change as the driver for spatial convergence in US earnings before 2000 (Caselli and Coleman II, 2001), which has subsequently been applied in China after 2000 (Hao et al., 2020). In contrast to these papers, I focus on the population distribution rather than falling inequality. As a result, they do not emphasize the weaker returns to concentrating agriculture in space as I do.

Another strand of literature throws international trade integration into the mix. Gollin et al. (2016) show that specialization in non-industrial sectors under trade lead to higher non-traded goods employment in "consumption cities"; though space is not directly modeled. Fajgelbaum and Redding (2014) show that regions more integrated into world markets exhibit low agricultural employment shares because of both differences in land intensity and rising employment in non-traded sectors. Data limitations do not allow me to consider the Balassa-Samuelson effect central to these papers. In addition, while I don't focus on international trade in the theory, it is empirically a force for suppressing agricultural employment in the coastal provinces because they import a

significant amount of food².

This literature also highlights dynamic productivity diffusion in space as drivers of structural change (Desmet and Rossi-Hansberg, 2014) (Delventhal, 2019). I abstract from the main mechanisms in these papers by using a simple model of local increasing returns.

Of this body of work, the most closely related are Michaels et al. (2012), Murata (2008) and Eckert and Peters (2018), whom explicitly study the effect of structural change on geography. I build on the mechanism driving spatial concentration in the first two by linking theory to a structural model and assessing it quantitatively against competing hypotheses. The third contains a negative result; that the spatial reallocation of workers accounted for essentially zero of US structural transformation since 1880. In contrast, I ask how falling agricultural employment accounts for total spatial reallocation toward dense locales. I also emphasize the link between density and specialization outside of agriculture, which they do not. In the following section, I show that the empirical observation in Eckert and Peters (2018) holds in China, and that it is consistent with structural change accounting for sizable portion of rising concentration in the period that I study.

This paper is nested within the broader literature on structural change, growth and the agricultural productivity gap³. The general concern of these papers are macroeconomic outcomes, and use models where there are multiple sectors but no notion of space within countries⁴. However, the models that give rise to structural change are central to the equilibrium geography that I study here. In particular, I use the non-homothetic CES preferences of Comin et al. (2021) and Lewis et al. (2021) and adapt their estimation strategy for to use with prices inferred from trade data.

²In the context of China, local specialization under trade is in part determined by proximity to the eastern coast, highlighted in Coşar and Fajgelbaum (2016).

³See Herrendorf et al. (2014) for a recent review of the structural change literature. This includes Ngai and Pissarides (2007), Boppart (2014), Comin et al. (2021), Matsuyama (1992, 2009, 2019), Uy et al. (2013), Bustos et al. (2016), Bustos et al. (2020), Swiecki (2017), Lewis et al. (2021), Cravino and Sotelo (2019) and Duarte and Restuccia (2010).

⁴A notable exception is Karádi and Koren (2017), which links a disaggregated equilibrium in a monocentric city model to the aggregate production function to facilitate development accounting. While the research question is fundamentally different, their model also features a version of the spatial bias of growth – see Proposition 3.

While the model and methodology in this paper may come from elsewhere, I study a central question in economic geography – how is the distribution of economic activity in space determined? I build on the models of [Krugman \(1991\)](#), [Puga \(1999\)](#), [Ottaviano et al. \(2002\)](#), [Murata \(2008\)](#), [Davis \(1998\)](#) and subsequent work to answer this question. I also draw from recent spatial models that are rich enough to have an empirical implementation, including [Allen and Arkolakis \(2014\)](#), [Redding \(2016\)](#), [Nagy \(2020\)](#), [Allen and Donaldson \(2020\)](#), [Desmet et al. \(2018\)](#), [Sotelo \(2020\)](#), [Eckert and Peters \(2018\)](#) and importantly [Tombe and Zhu \(2019\)](#), [Hao et al. \(2020\)](#), [Ma and Tang \(2020\)](#) and [Fan \(2019\)](#) in the context of China. Each employ a scheme to map observed data to unobservable parameters, which I use extensively for productivity, trade and migration costs. In particular, I draw heavily from the methodology and data choices in [Tombe and Zhu \(2019\)](#). The point of departure is the use of homothetic preferences⁵, so that any spatially uniform scaling of productivity do not change the equilibrium outcomes that I study. I generalize these empirical implementations to a model falling outside this common class.

Surrounding this central question is an empirically oriented literature in which transportation infrastructure is the main dependent variable. These include [Baum-Snow \(2007\)](#), [Duranton and Turner \(2011\)](#), [Duranton and Turner \(2012\)](#), [Baum-Snow \(2020\)](#), [Herzog \(2021\)](#), [Faber \(2014\)](#), [Bartelme \(2018\)](#) and in particular [Baum-Snow et al. \(2017\)](#) and [Baum-Snow et al. \(2020\)](#) in the Chinese context. I compare the magnitude of the spatial bias of growth against the effects of falling trade costs (among other benefits) considered in these papers. Moreover, data constraints force me to consider a large level of spatial aggregation than what is typical. I hope to consider a model with prefectures instead of provinces in future work.

3 Data and Facts

The data reveal key features of China’s growth that link geographic concentration and structural change. Firstly, I document rising spatial concentration that is *purely* driven by employment growth outside of agriculture. I also show that locations with higher population density tend to have lower agricultural employment shares and a revealed comparative advantage outside of agricul-

⁵With the exception of [Hao et al. \(2020\)](#).

ture in the cross section. These relationships have a clear grounding in the theory I present in the following section.

To construct these facts, I draw on multiple data sources at the provincial level, excluding Tibet. These concern provincial and international trade, migration, capital, agricultural spending and GDP data disaggregated by the primary and non-primary industries. In the Chinese census and China Statistical Yearbooks (CSY), the primary industry is an aggregate of farming, animal husbandry, forestry and fishing. I follow [Hao et al. \(2020\)](#) by defining the primary industry as the agricultural sector, and all other economic activity as an aggregated non-agricultural sector.

I exploit a model of trade to calibrate productivity, which requires provincial and international trade data. These flows come from the 2002 and 2007 regional input-output tables, which are disaggregated by sector. These are the same data used in [Fan \(2019\)](#) and [Tombe and Zhu \(2019\)](#). However, this level of disaggregation comes at a disadvantage. Flows are limited to rail transport, which misses a bulk of trade.

To capture imperfect labour mobility, I use migration data from the 2000 and 2005 Chinese population censuses. These data map individuals residing in a province and sector to their hukou. I measure migration flows as the fraction working in a sector or province different from the hukou, in the same way as [Tombe and Zhu \(2019\)](#) and [Hao et al. \(2020\)](#). Without loss of generality, I use the term “agricultural” in reference to the rural hukou. Interpreting the movement of workers across sectors but not space as migration is not necessary; the model has no notion of space *within* provinces.

Value added and labour employment data come from the CSY. I use the capital data from [Hao et al. \(2020\)](#), whom construct it using observed investment. I do not observe good measures of land use, so I infer it as the allocation that equalizes the return to land within provinces. I also assume that the entire official landmass is used for either production or consumption, which is also reported in the CSY.

Food spending shares are split by urban and rural areas within each province using data from *China Data Online*. I ascribe rural spending shares to agricul-

tural workers and urban to workers outside agriculture, similarly to [Hao et al. \(2020\)](#). This is not a perfect match, because workers in rural areas within a province may be employed outside of agriculture. However, it is a sensible one – an overwhelming majority of those holding rural hukou are employed in the primary sector. These food spending shares appear quite large (on average 30 percent in 2000) because I use after tax income. This potentially ignores income effects stemming from the provision of goods and services by the government sector.

The primary sector and rising spatial concentration

How much has spatial concentration risen from 2000 to 2005? When measuring it as the coefficient of variation of employment density across provinces, it increased from 1.1 to 1.218, representing an approximate 10 percent rise in just 5 years⁶. To put in perspective, I compare this to the US. At the Minor Civil Division level of aggregation, the same statistic increased from 0.3716 to 0.506 from 1880-2000 – representing a growth of 36 percent in a period of 120 years ([Michaels et al., 2012](#)). Extrapolating the observation in China to this longer time period implies a growth of almost 1000 percent⁷.

The rising spatial concentration can be illustrated with a regression of employment growth on initial employment density. If employment growth is on average larger for more population dense locations, then population density dispersion will tend to rise. A plot of the regression and the data is reported in Figure 1. The positive coefficient is clearly driven by the inclusion of exceptionally fast growing provinces, like Shanghai, Beijing and Guangdong.

Which sectors are driving the rising concentration? To explore this, I decompose the total employment growth rate g_i into a *agriculture* (a) and *non-agriculture* (n) components. These are related in the following way,

$$g_i = g_{i,a}s_{ia,2000} + g_{i,n}s_{in,2000}$$

⁶When measuring this as employment weighted population density dispersion, it increased from 0.714 to 0.806.

⁷This extrapolation is simply for comparison. The rise in population density dispersion was transitory – it vanished completely from 2005-2010. Explaining this observation is outside the scope of the paper.

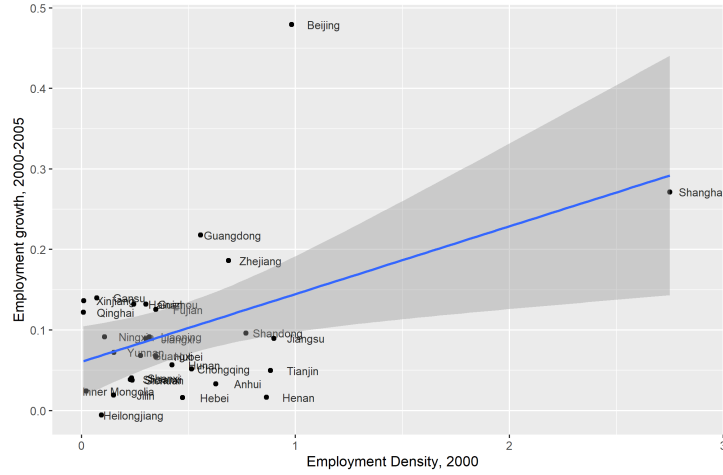


Figure 1: Relationship between average employment growth and employment density by Chinese province. Density is measured in millions per 1000 square miles. Slope estimate for unweighted and employment-weighted is 0.0841 and 0.0633, respectively. Unweighted and employment weighted regressions have a p value of < 0.01 and 0.08.

where $g_{i,s}$ and $s_{i,s,2000}$ are the growth rates of employment and the share of employment in 2000 in sector s and province i , respectively. I regress both of these components against initial employment density separately, noting that the coefficients on each regression must add up to the slope reported in Figure 1. I find that the slope for the agricultural component is negative, implying that the relationship is entirely driven by activity outside of the primary sector. A plot of both regressions is reported in Figure 2.

The facts shown in Figures 1 and 2 are readily explained by the interplay between structural change and the differential returns in space across sectors. Structural change causes outflows of labour from the agriculture, and these outflows tend to have a relationship with population density that is relatively weaker in magnitude⁸. These workers tend to get absorbed in provinces with a comparative advantage outside of agriculture. But, comparative advantage is in part determined by density insofar as the local supply of land in production is inelastic. The question remains – is the relationship between density and comparative advantage a feature of the data? I document below.

⁸The agriculture slope estimate in Figure 2 has less than half the magnitude than that of the non-agriculture component.

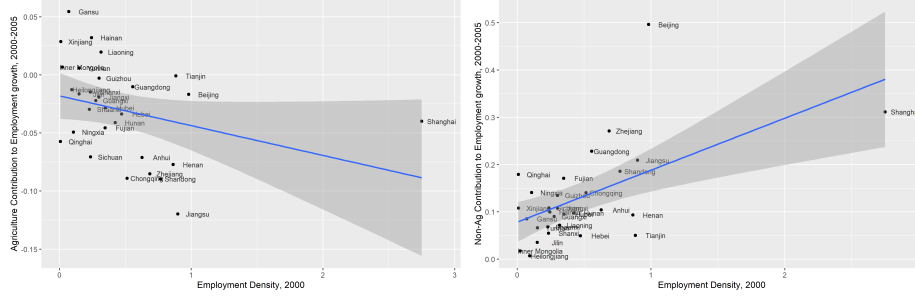


Figure 2: Relationship between employment growth by sector and employment density by Chinese province. Density is measured in millions per 1000 square miles. Employment-weighted slope estimates are -0.0593 and 0.1226 for the agricultural and non-agricultural components, with similar magnitudes for the unweighted regressions. Non-ag weighted and unweighted regressions have a p-value of < 0.01 and 0.02, respectively.

Before continuing, I note an additional caveat. These results may appear at odds with the motivating evidence in [Eckert and Peters \(2018\)](#), but I show that they are not. I repeat their additive decomposition of aggregate structural change into “spatial reallocation” and “regional transformation” components, defined by

$$s_{a,2005} - s_{a,2000} = \sum_i \left[s_{ia,2000}(l_{i,2005} - l_{i,2000}) + (s_{ia,2005} - s_{ia,2000})l_{i,2005} \right] \quad (1)$$

where s_{at} is the aggregate employment share in agriculture at time t , and l_{it} the fraction of employment in province i at time t , respectively. Intuitively, if the spatial reallocation term $\sum_i s_{ia,2000}(l_{i,2005} - l_{i,2000})$ is relatively large, then most of the structural transformation comprises of workers moving across provinces. I find that the term accounts for only 5 percent of the total change in the share of agricultural employment. This is approximately the same number they find in the US between 1880 and 2000. I reiterate that, while spatial reallocation may not account for much of China’s structural transformation, structural transformation may significantly affect spatial reallocation toward dense provinces. This is precisely what Figure 2 suggests.

Structural change, specialization and comparative advantage

How fast was structural change during this period? The aggregate employment share in agriculture fell from 52.9 percent to 44.8 percent. This decrease

is in part reflected in the falling aggregate food expenditure shares implied by the provincial data. Food represented 31.8 percent of the aggregate budget in 2000, and decreased to 28.2 percent in 2005.

How has structural change varied by density, and does this reflect comparative advantage? To examine the relationship in the cross section, I plot the 2000 agricultural employment share against employment density in Figure 7 of Appendix B. The relationship is strong. Each additional million people per thousand square miles is associated with a 0.19 decrease in the share of agriculture employment in the range of densities observed in the data. This relationship remains stable in transition to 2005.

The relation in Figure 7 can be rationalized in the absence of trade. Spatial differences in equilibrium real income could induce this pattern whenever migration is costly. Instead, I consider a more direct measure – the Revealed Comparative Advantage (RCA) as in Balassa (1963). As is well known, it is defined as follows for province i and sector s ,

$$RCA_{i,s} = \frac{E_{i,s} / \sum_{s' \in \{a,n\}} E_{i,s'}}{\sum_{i' \in \mathbf{P}} E_{i',s} / \sum_{s' \in \{a,n\}, i' \in \mathbf{P}} E_{i',s'}}$$

where $E_{i,s}$ are exports and \mathbf{P} is the set of province indices and a rest-of-the-world aggregate. Exports in this case are defined to include regional trade. In words, RCA measures how intensively a province exports goods in s relative to how intensively that good is exported worldwide. I plot a regression of RCA in agriculture on employment density in 2000, see Figure 3. It reveals that density accounts for about 30 percent of the variation in revealed comparative advantage in the data. In addition, this correlation is supplemented by evidence in Baum-Snow et al. (2020). They show that improved road infrastructure (which presumably increases regional trade integration) caused smaller prefectures to specialize in agriculture.

In short, density matters for trade patterns. Like the previous fact, this is grounded in theory. Chinese agriculture uses land more intensively with a cost share of 0.26 relative to 0.01 in non-agriculture (Hao et al., 2020). If the supply of land for use in agriculture is inelastic and agglomeration forces do not vary too widely by sector, this means that high land prices in dense locations will crowd out agricultural activity.

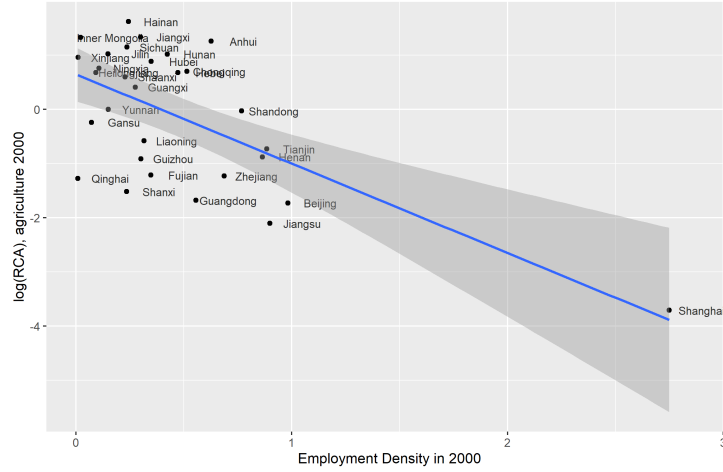


Figure 3: Relationship between revealed comparative advantage and employment density by Chinese province. Density is measured in millions per 1000 square miles. The slope estimate is -1.899 with a p-value of < 0.0001 , and the regression has an R^2 of 0.3. Similar results hold in the 2005 cross-section.

4 Theory

Explaining these facts and guiding the methodology behind the empirical work, I introduce a theory that captures when and how the geographic bias of growth will occur, and how to measure it. At the heart of this theory is the idea that concentrating agriculture in space comes with a penalty that is larger than other sectors. As a result, the intensity of agricultural consumption (or equivalently, employment in general equilibrium) will inversely affect how concentrated the geography will be. Structural change thus increases concentration.

The theory assumes exogenous spatial variation in productivity. I make this assumption in part because productivity differences are evident, and persist in the time period I study. Moreover, the inferred differences across provinces cannot be explained by population density using typical estimates of the agglomeration elasticity⁹. I use productivity differences solely to generate spatial

⁹See [Combes and Gobillon \(2015\)](#) for a review of the literature on the estimation of agglomeration elasticities.

variation in economic activity.

To this end, define the set of locations $\{1, 2\}$ indexed by i , a set of sectors $\{a, n\}$ indexed by s , and assume there are a measure L workers to be allocated to each of these sector-location pairs. Each location is endowed with H units of land that can be allocated to each sector, but are immobile across space¹⁰.

Production

Each sector-location uses a Cobb-Douglas technology

$$Y_{i,s} := A_{i,s} L_{i,s}^{1-\beta_s} H_{i,s}^{\beta_s}$$

where $A_{i,s} := a_{i,s} \left(\frac{L_{i,s}}{H_{i,s}} \right)^{\alpha_s}$ captures economies of density that are not internalized by the representative firm. I assume land is used more intensively in agriculture, $1 > \beta_a > \beta_n > 0$, and positive economies of density $\alpha_s > 0$. To abstract from Ricardian motives for specialization, I assume that location 1 and 2 have productivity ca_s and a_s in each sector s and $c > 1$, respectively. In other words, there are no differences in relative productivity across sectors. However, differences in relative marginal costs will arise. This is because location 1 must have more workers in an equilibrium where mobility is free – and thus produce at a higher density¹¹.

For simplicity, I assume workers in sector s are paid the rents of all employed factors equally. In other words, leaving the sector or location changes the stock of land that workers are entitled to earn rents from. I maintain this assumption in the empirical model and defer a discussion about its implications. I also assume land are perfectly mobile across sectors within a location. Relaxing this assumption to include completely immobile land does not change the results.

Abstracting away from labour mobility for now is useful. Let $\phi_{i,s}$ be nominal output per worker in i, s . By properties of the technology, we have that $(1 - \beta_s)\phi_{i,s} = w_{i,s}$, where w are the rents paid to labour. Coupled with perfect land

¹⁰Equivalently, the model could be solved with arbitrary land endowments. Then, I can normalize land in each location to 1 and express the labour allocation relative to its total land endowment.

¹¹This assumption mirrors that of larger spatial variance of (log) productivity in non-agriculture used in [Michaels et al. \(2012\)](#). Here, I make the assumption that the spatial variance in productivity are equal across sectors, which is isomorphic to eliminating Ricardian comparative advantage.

mobility, this means that the allocations of land satisfy

$$\frac{H_{i,a}}{H_{i,n}} = \frac{\beta_a \phi_{i,a} L_{i,a}}{\beta_n \phi_{i,n} L_{i,n}}$$

and in particular land market clearing implies,

$$H_{i,s} = \frac{\beta_s \phi_{i,s} L_{i,s}}{\beta_a \phi_{i,a} L_{i,a} + \beta_n \phi_{i,n} L_{i,n}} H$$

So that the *employment density* as a function of labour and output per worker satisfies

$$\frac{L_{i,s}}{H_{i,s}} = \frac{\beta_a \phi_{i,a} L_{i,a} + \beta_n \phi_{i,n} L_{i,n}}{\beta_s \phi_{i,s} H} \quad (2)$$

I assume that output markets are perfectly competitive. Prices are then equal to marginal cost, which can easily be shown to satisfy

$$p_{i,s} = (1 - \beta_s) \phi_{i,s}^{1-\Omega_s} \left[\frac{\beta_a \phi_{i,a} L_{i,a} + \beta_n \phi_{i,n} L_{i,n}}{\beta_s H} \right]^{\Omega_s} a_{i,s}^{-1} \quad (3)$$

where $\Omega_s := \beta_s - \alpha_s$ is the negative *effective returns to density* in sector s . This describes the penalty (in terms of prices) that arises whenever good s is produced at high density. I make the assumption that $\Omega_a > 0$ and $\Omega_n \leq 0$ so that there are net decreasing returns in the a sector, and weakly increasing returns in n . In the empirical model, I use $\beta_a = 0.26$ and $\beta_n = 0.01$ as in [Hao et al. \(2020\)](#). If the returns to scale in both sectors is $\alpha_s = 0.05$, a value that implies an effective agglomeration elasticity of $\Omega_n = 0.04$ which is reasonable given estimates in the literature, then this is easily satisfied. This would only be violated if the spatial returns to agriculture were extremely large.¹²

Assumption 1. $\Omega_a > 0$ and $\Omega_n \leq 0$.

Regional trade

To make the theory simple, I consider equilibria where there is free trade. Let $\theta > 0$ be the trade elasticity associated with each sector. Prices $P_{i,s}$ are equal-

¹²The assumption that there is increasing returns to scale can be relaxed and replaced with $\Omega_a < \Omega_n$ so that n has increasing returns relative to a . This requires an additional assumption to rule out equilibria at extreme parameter values, so I omit it for brevity. I don't view this as an issue—there is overwhelming evidence of agglomeration economies.

ized across locations and are assumed to satisfy¹³

$$P_{i,s} = \left[\sum_{i \in \{1,2\}} p_{i,s}^{-\theta} \right]^{-\frac{1}{\theta}}$$

where $\left[\frac{p_{i,s}}{P_{i,s}} \right]^{-\theta}$ is the fraction of spending in on s goods from location i .

Preferences

Individual preferences are generalized CES as in [Comin et al. \(2021\)](#) and [Matsuyama \(2019\)](#). $U(c)$ is implicitly defined by the relation

$$\sum_{s \in \{a,n\}} \gamma_s^{\frac{1}{\eta}} \left(\frac{c_s}{U \epsilon_s} \right)^{\frac{\eta-1}{\eta}} = 1 \quad (4)$$

for $\epsilon_k > 0$ and $\eta \in [0, 1) \cup (1, \infty)$. Following these papers, I normalize ϵ_n and γ_n to one as doing so involves a monotone transformation of U . Let $\omega_{i,s}$ denote the fraction of income spent on a by workers in i as a function of prices and income. It is well known that ω satisfies the implicit relation

$$\log(\omega_{i,s}) - \log(1 - \omega_{i,s}) = (1 - \eta) \log \left[\frac{P_{i,a}}{P_{i,n}} \right] + (1 - \eta)(\epsilon_a - 1) \log [U_{i,s}] + \log(\gamma_a) \quad (5)$$

Equation (5) is instructive. If $\epsilon_a < 1$ when $\eta < 1$ or $\epsilon_a > 1$ when $\eta > 1$, an increase in the standard of living (measured by the utility function) decreases the agricultural spending share whenever relative prices are fixed. I maintain this standard assumption throughout the paper. If $\epsilon_a = 1$, the preferences become the standard homothetic CES.

The isoelastic nature of this utility function will also be useful in characterizing how agricultural spending will change as an economy grows. In fact, it only depends on the growth rates of the prices of each sector. (5) can be instead written as

$$\log(\omega_{i,s}) - \epsilon_a \log(1 - \omega_{i,s}) = (1 - \eta) \log \left[\frac{P_{i,a}}{P_{i,n}} \right] + (1 - \eta)(\epsilon_a - 1) \log \left[\frac{\phi_{i,s}}{P_{i,n}} \right] + \log(\gamma_a) \quad (6)$$

¹³This relationship can be derived using an Eaton-Kortum or Armington model. I say nothing about the microfoundations here as it is inconsequential for the spatial bias of growth.

Suppose for the sake of this argument that ϕ_i is normalized to one. What are the permissible (negative) growth rates of sector prices g_s such that ω_i must necessarily fall, given that these growth rates must equal across space? It follows directly from (6) that $g_a < \epsilon_a g_n$ when $\eta > 1$ or $g_a > \epsilon_a g_n$ when $\eta < 1$. Intuitively, strong income effects (ϵ_a is small or large when $\eta < 1$ or $\eta > 1$, respectively) increases the range of growth rates amenable to falling agricultural spending.

In what directly follows, I assume growth rates of the TFP term $a_{i,s}$ lie within this range. It is perhaps not surprising that this holds in the data, because I estimate ϵ_a and η to partially fit the observed fall in food spending.

Assumption 2. *Let g_s be the growth rate of $a_{i,s}$. Then $g_a < \epsilon_a g_n$ when $\eta > 1$ or $g_a > \epsilon_a g_n$ when $\eta < 1$.*

Measuring population density dispersion

How do we capture the idea that economic activity may be more or less even in space? There is some ambiguity surrounding the quantitative description of population density dispersion that I deal with here. On one hand, I could choose a plot of land uniformly at random and record its population density. The population density dispersion would then be defined as the coefficient of variation of this random variable. In the theory, these plots correspond to the locations defined above. It is clear that the density dispersion, when defined in this way, is strictly increasing in the population of the largest location holding the total population fixed at L .

Alternatively, I could choose a *worker* uniformly at random and record the population density in the location for which they live. This would map to a population-weighted measure of density dispersion. However, this measure is generally *not* increasing on the size of the largest location. This can be easily illustrated – if all workers are concentrated in location 1, then there is no variation in the weighted density because all workers face the same density.

In the following section, I show that growth tends to increase density dispersion when it is not weighted by population. In the structural model, I compute both measures of density dispersion and show that they increase due to the spatial bias of growth.

Free trade equilibrium

I now assume perfect labour mobility. I search for a spatial equilibria on the *interior*, that is, where each location is populated and output per worker equalizes. Assuming $1 + \theta\Omega_n > 0$, a lemma in Appendix A shows that all equilibria must be interior whenever there is nonzero total employment in either sector. In other words, the increasing returns in sector n cannot be strong enough to support complete specialization¹⁴. I abstract away from pathological equilibria where there is no agricultural or non-agricultural production in any location.

Free mobility and trade imply that ω_i equalizes across locations, and equals the fraction of total employment in agriculture

$$\omega = \frac{L_{1a} + L_{2a}}{L} \quad (7)$$

and the following must be satisfied in order for the goods market to clear,

$$\frac{L_{1,s}}{L_{2,s}} = c^\theta \left[\frac{\beta_a L_{1,a} + \beta_n L_{1,n}}{\beta_a L_{2,a} + \beta_n L_{2,n}} \right]^{-\theta\Omega_s}. \quad (8)$$

The relation relates relative employment to the relative density at which production takes place. This implies the following relationship between relative employment in each sector,

$$\frac{L_{1,n}}{L_{2,n}} = c^{\theta \frac{\Omega_a - \Omega_n}{\Omega_a}} \left[\frac{L_{1,a}}{L_{2,a}} \right]^{\frac{\Omega_n}{\Omega_a}} \quad (9)$$

Equations (7), (8), (9) and labour market clearing define a map between relative employment in 1 and the level of spending ω . This means that these variables do not change unless ω does. Moreover, (9) immediately implies that location 1 is specialized in sector n (has a higher relative employment in n) whenever relative employment in agriculture is less than c^θ . I defer showing that this condition must hold in equilibrium to the proof in Appendix A. (9) is also instructive about the spatial bias of growth. If there is aggregate disemployment in agriculture and location 1 is specialized outside of the sector, then it mechanically absorbs a majority of those outflowing workers. This is

¹⁴Simonovska and Waugh (2014) pin the estimates of the trade elasticity at $\theta = 4$, and there isn't much evidence that it differs in the agricultural sector (Tombe, 2015). Combined with a reasonable value of $\Omega_n = -0.03$, this assumption holds in the structural model.

precisely the story suggested by Figure 2.

There is one catch – small changes in productivity, when satisfying Assumption 2, may not decrease the share to total employment in agriculture ω . Instead, it may shuffle around the distribution of workers and employment densities in equilibrium, which additionally affect relative prices. With this in mind, the following proposition *does not* describe a monotone relationship between the population density of location 1 and growth. Instead, I show that the relationship holds *when growth is large enough*. The definition of large is left to empirical work.

Proposition 1. *Suppose Assumptions 1 and 2 hold, along with $1 + \theta\Omega_n > 0$. Then the equilibrium path of ω converges to 0 and $L_{1a} + L_{1n}$ converges to its maximum possible value in any equilibrium where $\frac{a_{1,s}}{a_{2,s}} = c$ for every s . Moreover, location 1 is specialized in sector n $\frac{L_{1,n}}{L_{2,n}} > \frac{L_{1,a}}{L_{2,a}}$ and has a higher population density $\frac{L_{1,a} + L_{1,n}}{H}$ in any of these equilibria.*

Proof. See Appendix A. □

How does Proposition 1 guide the empirical work? It tells us that population density dispersion can increase when productivity growth is uniform across space. This effect *cannot be described in a model where agricultural spending shares are exogenous*. On this basis, I measure this effect in the data by scaling productivity uniformly and observing how dispersion changes, holding all alternative variables affecting the population distribution fixed. This scaling factor should be some central measure of observed productivity growth across provinces and the rest of the world. In the following sections, I detail a procedure to arrive at it.

5 Extending the model

I extend the theory to allow for multiple provinces, trade and migration frictions, international trade and idiosyncratic productivity differences to accompany a range of alternative explanations for the observed rise in spatial concentration. The end goal is to use a general equilibrium model as an experimental environment. It provides a way infer counterfactual equilibria in 2005 where any combination of these explanatory factors are present. I use these equilibrium outcomes to construct marginal effects associated with each explanatory

factor. These equilibria will be computed by solving the model in changes using a variation of “hat algebra” as in [Dekle, Eaton, and Kortum \(2007\)](#). Before arriving at this, I complete the set up of the model. Unless otherwise stated, this model will share features from the theoretical one above.

Let \mathbf{P} be the set of provinces and a rest of the world aggregate, indexed by i and RoW , respectively. Let $s \in \{a, n\}$ index sectors and $t \in \{2000, 2005\}$ index time. Each province-sector is exogenously endowed with land H_i , capital K_{ist} , stock of individuals with hukou status M_{ist} ¹⁵, and total factor productivity a_{ist} . To produce a good, each province-sector combines labour, capital and land with a Cobb-Douglas technology. The technology has cost shares β_s^L, β_s^K and β_s^H for labour, capital and land, respectively. There are external increasing returns to employment density, measured in workers per unit of land, with elasticity α_s ¹⁶. Each of these parameters is assumed time-invariant.

As in [Caliendo and Parro \(2014\)](#), each sector can use goods from other sector-locations in production. A local good in s is a CES aggregate over goods from each i with elasticity of substitution $\theta_s + 1$. Firms in sector k use the local good in each s with cost shares β_{sk} .

Preferences

I make an augmentation to the non-homothetic CES preferences above. Recall that γ_a in Equation (6) is a preference parameter that increases spending on agriculture independent of real income or relative prices. In order to fit idiosyncratic variation in agricultural spending across provinces and sectors, I allow the preference parameter γ_a to *vary with each sector-location pair*. Crucially, this parameter is time invariant. When migrants leave to work in a new location or sector, they inherit this γ_a . I do not interpret this assumption as a literal changing of preferences, rather than a time-persistent and unobserved characteristic of the goods available to workers in that location and sector. Lastly,

¹⁵I use the sector subscript a and n for registrant stock M_{ist} simply as notation to refer to those holding a rural or urban hukou, respectively. The stock of hukou registrants M_{ist} are inferred by inverting the migration flow matrix and using the employment data.

¹⁶Keeping the capital stock exogenous and not allowing for factor biased technical change rules out the mechanisms considered in [Bustos et al. \(2016\)](#) and [Bustos et al. \(2020\)](#). The latter highlights the role of structural change and urban investment. I am forced to ignore this, in particular, because I don’t observe spatial trade imbalances. More on this to follow.

preference parameters η and ϵ_a will remain fixed across time, space and sectors. These will be estimated.

To capture an additional congestion force, I also allow workers to consume land commanding a constant share of income $1 - \nu$. Preferences can then be represented by utility function proportional to

$$U_{i,s}^\nu [H_{i,s}^c]^{1-\nu} \quad (10)$$

where $H_{i,s}^c$ is the consumption of land by workers in (i, s) , and $U_{i,s}$ is defined implicitly by equation (4) under parameters $\{\gamma_{is}, \epsilon_a, \eta\}$. In what follows, let V_{ist} be the indirect utility function associated with (10) as a function of income, goods prices and land prices faced by workers in (i, s) at time t . Each worker is paid the entire value added in the sector and location for which they work and is also rebated their spending on this local housing.

Trade

Trade costs take the canonical iceberg form. Shipping one unit of a good from location i to location j in sector s and time t requires producing $\tau_{ij,st}$ units of the good, varying by sector. Within a location, trade is free. Data constraints require me to forgo the cost of shipping food within a location, so that agriculture and non-agricultural workers face the same prices. Trade costs change the price index in Equation (3) to instead satisfy

$$P_{i,st} = \left[\sum_{j \in \mathbf{P}} \tau_{ji,st}^{-\theta_s} p_{j,st}^{-\theta_s} \right]^{-\frac{1}{\theta_s}} \quad (11)$$

with the share of spending in sector s in i allocated to j satisfying $\pi_{ji,st} = \tau_{ji,st}^{-\theta_s} p_{ji,st}^{-\theta_s} P_{i,st}^{\theta_s}$, which are observed in the data. Since these cost shares do not distinguish firms from consumers, I assume both groups allocate spending in sector s equally to all locations. I normalize $\tau_{ii,st} = 1$ noting that only relative trade costs are identified.

Land allocation

Lack of data requires me to assign land use. In particular, I assume land H_i is allocated to production in each sector H_{is} and for consumption H_{is}^c to equalize

the rate of return on land in each province r_{it} . I can drop this assumption in place of zero land mobility, and I find even stronger results.

Migration

I model costly migration similarly to [Tombe and Zhu \(2019\)](#) and [Hao et al. \(2020\)](#), with some key differences. Each worker is endowed with a hukou status (j, k) , and must choose Chinese province-sector pair to work. There is no international migration. The fraction of workers $m_{jkis,t}$ (observed in the data) with hukou status (j, k) who choose to work in (i, s) is given by the constant-elasticity regional supply curve¹⁷

$$m_{jkis,t} = \frac{(V_{ist}/f_{jkis,t})^\kappa}{\sum_{i' \in \mathbf{P} \setminus \{RoW\}, s' \in \{a,n\}} (V_{i's't}/f_{jki's',t})^\kappa} \quad (12)$$

where $\kappa > 0$ is the elasticity of migration with respect to (discounted) real income, and $f_{jkis,t}$ are exogenous migration frictions that discount real income earned when migrating. Similar to trade costs, I make the normalization $f_{isis,t} = 1$. This supply curve is useful because the migration flows in 2005 can be written as a function of only the proportional change in real income $\frac{V_{is2005}}{V_{is2000}}$, frictions $\frac{f_{jkis,2005}}{f_{jkis,2000}}$ and initial migration flows $m_{jkis,2000}$, which I use to solve the model.

There is one key issue with using non-homothetic preferences in conjunction with this regional labour supply curve. From Equation (12), the size of differences in real income across space matter for determining migration flows. This implicitly demands a cardinal interpretation of the utility function. In the following section and in Appendix C, I derive a cardinal measure from the compensating variation, rather than using the proportional change in real income measured by the utility function $\frac{V_{is2005}}{V_{is2000}}$. Here, I define compensating variation as the fraction of income deducted in 2005 required to be as well off as in 2000. When preferences are homothetic, this measure corresponds to a transformation of both the equivalent variation *and* the measured growth in the utility function ([Samuelson and Swamy, 1974](#)). Unfortunately, this useful property does not hold here.

¹⁷This regional supply curve can be derived in a model where there are non-pecuniary and idiosyncratic Fréchet shocks to real income discounted by the migration friction f . See, for example, [Redding \(2016\)](#).

There is another point of departure from [Tombe and Zhu \(2019\)](#). Informed by Chinese institutions, they model the decision to migrate as additionally forgoing rights to the income of communally owned land and capital. Like in the theory, I do not consider this feature – migrants working in (i, s) are paid the entire value added of production and the rents paid on their own housing consumption. While this pecuniary cost of migration will be reflected in the calibrated frictions in my model, they do not adjust with local prices of land and capital. I make this assumption because the utility function I use does not allow for a clean aggregation of agents with differing incomes.

Solving the model in changes

I use the conventional “hat” notation. For any quantity g , define $\hat{g} := \frac{g_{2005}}{g_{2000}}$ to be its value in 2005 relative to 2000. The objective is to solve for the equilibrium in 2005 given changes in trade costs $\tau_{ij,s}$, migration frictions f_{isk} , productivity \hat{a}_{is} , and capital \hat{K}_{is} , as well as observed initial migration flows $m_{kij,2000}$, trade flows $\pi_{ji,s,2000}$, food spending shares $\omega_{is,2000}$, value added per worker $\phi_{is,2000}$, initial employment distribution $L_{is,2000}$, hukou registrants M_{ist} and the set of parameters $\{\theta_s, \epsilon_a, \eta, \beta_s^L, \beta_s^K, \beta_s^H, \beta_{sk}, \nu, \alpha_s, \kappa\}$. As a normalization, I set labour, capital and land used internationally to 1 and assume they are completely immobile across sectors. I assume the international aggregate spends a share of income in agriculture equal to world value added – approximately 4 percent. For a formal definition of equilibrium and a proof that these variables identify the set of equilibria, see [Appendix C](#).

Before continuing, I note two key caveats behind solving the model. Firstly, one of the assumptions that identifies equilibrium is balanced trade – both internationally and regionally. This is clearly strong. By solving the model in changes, it requires comparing data on value added per worker in the 2000 where trade is *not* balanced to a 2005 model-implied equilibrium where it is. To add sense to the comparison, I instead choose the initial distribution of value added per worker $\phi_{is,2000}$ that uniquely balances trade given observed trade flow data. I do so only to solve the model, but not to perform calibration in the following section. It greatly improves the model’s ability to match the rise in population density dispersion, with little change to other results. For details

on this procedure, see Appendix C¹⁸.

Secondly, multiple equilibria in this model cannot be ruled out ex-ante. This is because the model features non-homothetic preferences, increasing returns to scale and potential complementarity between goods from different sectors¹⁹. In Appendix C, I also detail a process that selects from the set of equilibria given some initial distribution of value added per worker and employment.

The question remains – if changes in costs τ and f , as well as productivity a are not directly observed in the data, how can one compute the counterfactual equilibrium in 2005? I use the model to infer these quantities. This is the purpose of the following section.

6 Calibration

Trade costs

Making the implicit assumption that trade costs are symmetric across locations, I recover them using a simple Head-Riess index. However, the matrix of trade flows in agriculture at the province level is not strictly positive, making the index ill-defined. Following Tombe and Zhu (2019), I aggregate provinces to fall in a set of 9 regions such that the implied matrix is strictly positive. Then, I choose trade costs between provinces (i, j) in sector s to satisfy

$$\tau_{ij,st} = \left[\frac{\pi_{i'i',st} \pi_{j'j',st}}{\pi_{j'i',st} \pi_{i'j',st}} \right]^{\frac{1}{2\theta_s}} \quad (13)$$

where i' is the region associated with province i .

Productivity

To calibrate total factor productivity a_{is} , I start with identifying the change relative marginal costs $c_{i,s}$ assuming trade must be balanced in 2005, given initial trade shares $\pi_{ji,s,2000}$, observed agricultural spending $\omega_{is,2005}$ and gross

¹⁸This procedure is similar to the one tested in the empirical work of Allen, Arkolakis, and Takahashi (2020). See page 423, footnote 31.

¹⁹The demand system defining general equilibrium in this model does not satisfy any notion of gross substitution, including that of Berry et al. (2013).

output $Y_{is,2005}$, as well as changes in trade costs $\tau_{ij,s}$. Using some hat algebra, this implies that observed gross output $Y_{is,2005}$ in sector (i, s) equals total payments made on the goods

$$\sum_{j \in \mathbf{P}, k \in \{a, n\}} \left[\omega_{jk,2005}^s v_k Y_{jk,2005} + \beta_{sk} Y_{jk,2005} \right] \frac{\pi_{ij,s,2000} \tau_{ij,s}^{\hat{c}_{i,s} - \theta_s} \hat{c}_{i,s}^{-\theta_s}}{\sum_{l \in \mathbf{P}} \pi_{lj,s,2000} (\tau_{lj,s}^{\hat{c}_{l,s} - \theta_s} \hat{c}_{l,s}^{-\theta_s})} \quad (14)$$

where v_k is the value added share in sector k and $\omega_{jk,2005}^s = \omega_{jk,2005}$ whenever $s = a$, or $1 - \omega_{is,2005}$ otherwise.

Redding (2016) shows that there is a unique solution (up to a scale factor in each sector) to (14) in $\hat{c}_{i,s}$ provided the trade shares are strictly positive. As mentioned before, this is not the case for the agricultural sector. In future drafts, I plan on providing a proof that a solution must be unique provided the matrix of trade flows is connected, which holds in the data. This proof will be based on work in Berry et al. (2013).

I choose the solution $\hat{c}_{i,s}$ to equation (14) under some arbitrary normalization. I use it to construct prices $\hat{P}_{i,s}$ using the hat-algebra analogue of equation (11). Then, \hat{A}_{is} can be solved using the marginal cost equation

$$\hat{c}_{is} = \frac{\hat{\phi}_{is}^{v_s} \left[\frac{\hat{L}_{is}}{\hat{H}_{is}} \right]^{\beta_s^H} \left[\frac{\hat{L}_{is}}{\hat{K}_{is}} \right]^{\beta_s^K} \hat{P}_{ia}^{\beta_{as}} \hat{P}_{in}^{\beta_{ns}}}{\hat{A}_{is}} \quad (15)$$

which determines \hat{A}_{is} up to a sector specific scale factor, which I choose so that it is on average 1 across provinces. \hat{L}_{is} , \hat{K}_{is} and $\hat{\phi}_{is}$ are observed in the data, and \hat{H}_{is} is chosen to equate the return on land within each province given this output and factor data. Note that \hat{A}_{is} includes productivity gains from agglomeration.

Let's take stock. Relative productivity growth will be useful for determining the endogenous spatial concentration in this model. However, in order to capture the spatial bias of growth, I need to assign a value to that scale factor to measure how much TFP grew in absolute terms. To do so, I exploit information about this growth contained in the observed fall in average agricultural spending. This relationship is mediated by the real income elasticity ϵ_a , and

relative price elasticity η . It makes sense to estimate them on the same data I use to calibrate the model.

The estimation strategy will leverage a key fact in the macro-development literature. Variation in real incomes of workers across provinces and sectors is going to be very informative about ϵ_a . This echoes the fact that the agricultural sector accounts for a large amount of cross-country variation in real incomes, and that the relative employment in the sector is strongly correlated with real income (Restuccia, Yang, and Zhu, 2008). With this in mind, I repeat the calibration exercise in equations (14) and (15) *in levels* using data in 2000, and again *in changes* using data from 2005 – 2010. I use this additional information to construct regional prices in the years 2000, 2005 and 2010, yielding a spatial panel of prices across provinces $P_{is,t}$, identified up to a scaling factor λ_{st} . The true value for prices is thus $\lambda_{st}P_{is,t}$.

With the data in hand, I write out a model to estimate ϵ_a and η . Manipulating Equation (6), we have a relationship between agricultural spending shares, prices and output per worker

$$\log \left[\frac{\omega_{is,t}}{(1 - \omega_{is,t})^{\epsilon_a}} \right] = \tilde{\eta} \log \left[\frac{\lambda_{at}P_{ia,t}}{\lambda_{nt}P_{in,t}} \right] + \tilde{\epsilon}_a \log \left[\frac{\phi_{is,t}}{\lambda_{nt}P_{in,t}} \right] + \log(\gamma_{is}) \quad (16)$$

where $\tilde{\eta} := 1 - \eta$ and $\tilde{\epsilon}_a := (1 - \eta)(\epsilon_a - 1)$. Notice how the λ_{st} reduce to time and sector varying parameters. I define the time fixed effect $\tilde{\lambda}_t := \tilde{\eta} \log[\frac{\lambda_{at}}{\lambda_{nt}}] - \tilde{\epsilon}_a \log(\lambda_{nt})$. I also introduce an additional fixed effect $\tilde{\gamma}_a$ capturing unobserved differences in agricultural spending made by farmers. Lastly, to fit idiosyncratic variation in spending shares across time, I add an additional error term $\log(u_{is,t})$ to equation (16). I assume this additional error term is pure measurement error. These assumptions define a parametric regression model

$$\omega_{is,t} = F(\{P, \phi\}_{is,t}; \epsilon_a, \eta, \lambda_{2000}, \lambda_{2005}, \lambda_{2010}, \tilde{\gamma}_a) + v_{ist} \quad (17)$$

where F is implicitly defined by equation (16) net of the error term²⁰. I estimate the model with nonlinear least squares and report along with bootstrapped confidence intervals in Figure 4²¹.

²⁰The regression error is defined as $v_{ist} = F(\{P, \phi, \gamma, u\}_{is,t}) - F(\{P, \phi\}_{is,t})$. This is the difference between the parametric model F when the errors are present relative to where they are not. I implicitly make the assumption on the distribution of u_{ist} and γ_{is} such that $\mathbf{E}(v_{ist}) = 0$.

²¹The minimization problem associated with this least squares is not globally convergent. For the point estimates, I simulate the result using 50 randomly generated initial conditions, and find

Mostly all of these estimates have a grounding in the literature. Firstly, the model fits a declining relationship between real income and agricultural spending (see Equation 5), and does so for all values in their respective 99 percent confidence intervals. Secondly, point estimates of the time fixed effects are declining, mirroring the observed fall in agricultural spending (though these are imprecisely estimated). However, I estimate $\hat{\eta} > 1$, and this hypothesis rejects at the 1 percent level. This is at odds with substitution-driven structural change highlighted in [Ngai and Pissarides \(2007\)](#). On the other hand, the relatively small value of $\hat{\eta}$ suggests that most of the structural change I observe here is through income effects, which is corroborated by [Comin et al. \(2021\)](#)²². I also find that the model-implied measures of real GDP per capita under these estimates correlate strongly with those constructed in [Tombe and Zhu \(2019\)](#); see Figure 8 of Appendix B.

ϵ_a	η	λ_{2000}	λ_{2005}	λ_{2010}	$\tilde{\gamma}_a$
2.3371	1.1976	-0.0387	-0.0710	-0.1184	-0.0680
[1.351, 5.531]	[1.102, 1.289]	[-0.6355, 1.3855]	[-0.2456, 0.0753]	[-0.3504, 0.1090]	[-0.4327, 0.1719]

Figure 4: 99 percent BCa confidence intervals with 1000 repetitions are reported

As an additional robustness check, I estimate ϵ_a and η using a similar procedure to [Comin et al. \(2021\)](#) and observe if $\hat{\eta}$ changes. I report the results in Figure 9 of appendix B. A few additional assumptions are made in their specification, which I also detail in the appendix. I find an even larger value of $\hat{\eta} = 1.698$, which is orthogonal to their robust finding that $\eta < 1$ on international data. I defer a discussion about potential endogeneity issues and explanations of this result for future drafts.

With these estimates in hand, I detail a procedure to calibrate the scale factor associated with \hat{A}_{is} . Taking the time differences between 2000 and 2005 in Equation (16) and rearranging yields, at the point estimates of $\tilde{\eta}$ and $\tilde{\epsilon}_a$,

$$\log \left[\frac{\hat{\omega}_{is}}{(1 - \hat{\omega}_{is})^{\epsilon_a}} \right] - \tilde{\eta} \log \left[\frac{\hat{P}_{ia}}{\hat{P}_{in}} \right] - \tilde{\epsilon}_a \log \left[\frac{\hat{\phi}_{is}}{\hat{P}_{in}} \right] = \tilde{\eta} \log \left[\frac{\hat{\lambda}_a}{\hat{\lambda}_n} \right] - \tilde{\epsilon}_a \log \left[\hat{\lambda}_n \right] \quad (18)$$

that it represents the minimum value of these. However, this is infeasible when bootstrapping. The confidence intervals reported here are constructed using the point estimate as an initial condition.

²²In particular, see the variance decomposition in Table 4 of [Comin et al. \(2021\)](#).

where $\hat{\lambda}_n$ and $\hat{\lambda}_a$ are precisely the scale factors needed for the calibration. I choose values of $\hat{\lambda}_a$ and $\hat{\lambda}_n$ such that (18) holds on average over provinces and sectors. One problem is that there are multiple choices that satisfy this – there are two parameters and one moment. I look to other data sources for information on average relative prices. Using the Groningen 10-Sector Database (Timmer et al., 2015), I infer that the aggregate relative price of agriculture in China grew approximately 9 percent in the period, so I choose $\frac{\hat{\lambda}_a}{\hat{\lambda}_n}$ so that the average increase in relative prices match. Then, I choose $\hat{\lambda}_n$ to match the average fall in the left hand side of the equation. This pins down the average growth of \hat{A}_{in} and \hat{A}_{ia} across provinces, and thus \hat{a}_{ia} and \hat{a}_{ia} after accounting for agglomeration forces at observed changes in employment density.

The implied results make sense. I find that average TFP growth in China was approximately 24 and 20 percent in agriculture and non-agriculture, respectively. This may seem low, considering that China officially grew almost 80 percent during this period. However, these numbers ignore input-output linkages, changes in capital per worker and falling trade costs. Taking these into account, I measure productivity growth as the provincial-average price index in s relative to the average value added per worker, and find that agriculture and non-agriculture grew approximately 50 and 68 percent, respectively. These are moderately shy of the official GDP statistics.

Migration costs

The calibration of \hat{a}_{is} allows for the construction of changes in welfare across locations and sectors. Recall that I measure the change in welfare deriving from the compensating variation. Let \hat{W}_{is} be this change in welfare. Coupled with observed migration flows, these can be used to infer the change in migration frictions \hat{f}_{isjk} . Using Equation (12) and some hat-algebra the following holds,

$$\frac{m_{isjk,2005}}{m_{isis,2005}} = \left[\frac{\hat{W}_{jk}}{\hat{W}_{is}} \right]^\kappa [\hat{f}_{isjk}]^{-\kappa} \frac{m_{isjk,2000}}{m_{isis,2000}} \quad (19)$$

Which determines the value of \hat{f} given constructed welfare changes and observed migration flows in each year. I define $\hat{W}_{is} := \frac{1}{1-CV_{is}}$, where CV_{is} is the fraction of income deducted in 2005 required to make a worker in (i, s) as well off as they were in 2000. Using the expenditure function associated with the

utility function in Equation (10), this can be shown to solve

$$\hat{W}_{is} = \frac{1}{1 - CV_{is}} = \frac{\hat{\phi}_{is}}{\left[\sum_{k \in \{a,n\}} \omega_{is,2000}^k \hat{P}_{ik}^{1-\eta} \right]^{\frac{\nu}{1-\eta}} \hat{r}_i^{1-\nu}} \quad (20)$$

where \hat{r}_i is the change in the price of land in i . Making this monotone transformation is useful because the welfare index \hat{W}_{is} looks identical to the homothetic CES index with elasticity of substitution η . This justifies the use of the income elasticity of migration κ estimated in [Tombe and Zhu \(2019\)](#), who also use a homothetic welfare index. See Appendix C for a derivation.

7 The Main Counterfactual

I allow for four different explanations behind the rise in population density dispersion – falling migration costs, falling trade costs, spatial variation in productivity growth and finally spatially uniform productivity growth. The theory directly speaks to the last of these forces. To measure the contribution of each effect to the rise in density dispersion, I calculate an equilibrium population density dispersion where any combination of each of these variables are present and set to their calibrated values. In the case where the variable is not present, its value is set to 1. For example, $\tau_{ij,s} = 1$ represents trade costs remaining at their 2000 values. I ascribe to each variable its Shapely value, which is the unweighted average of these marginal effects relative to the total contribution.

I define “relative” and “uniform” productivity growth as follows. When both are present, productivity and capital growth are equal to its calibrated value. When only relative growth is present, I scale TFP and capital growth down by the world average in each sector, preserving relative differences. When only uniform growth is present, TFP and capital growth are set to the world average in each sector. I report the results in Figure 6 and a description of the additional calibrated parameters in Figure 5.

The results suggest that the spatial bias of growth matters. It accounts for 38 percent of the rise in population density dispersion predicted by the model, which is slightly larger than the contribution ascribed to changing trade costs.

Moreover, this represents a lower bound, since falling trade and migration frictions also increase real income and put downward pressure on agricultural spending. This means that some of the spatial bias of growth is contained in these falling frictions, but there is no clear way to parse them.

The other explanatory factors also behave as expected. Falling migrations costs (in part a manifestation of the hukou reforms) increased the relative size of population dense regions by making spatial arbitrage easier. The magnitude of this force is relatively massive and directly reflects the magnitude of migration in this period. Falling trade costs also increased density dispersion, which reflects one view in a contested literature. This result appears to be in line with empirical evidence in [Baum-Snow et al. \(2020\)](#). The contribution of relative growth is large and negative, suggesting considerable spatial mean reversion in productivity disproportionately benefiting workers in hinterland provinces. This mean reversion is not enough to overpower the other three forces.

There are some drawbacks to these results. The model predicts density dispersion increasing from 1.0646 to 1.1259, while in actuality it increased from 1.1 to 1.218, representing just over half of the variation. This is due to the model tending to ascribe a larger agricultural sector than what is actually observed in the data, which is itself probably driven by the assumption of spatially balanced trade and the way agricultural spending shares are constructed. Moreover, there is substantial heterogeneity in marginal effects, making it difficult to assign unambiguous weights in the Shapely decomposition. I report a table of equilibrium outcomes under the presence of each explanatory factor in [Figure 10 of Appendix B](#).

8 Conclusion

TBD

Parameter	Value	Source
(β_a^L, β_n^L)	(.27, .19)	Hao et al. (2020) Labour share of income
(β_a^K, β_n^K)	(.06, .15)	Hao et al. (2020) Capital share of income
(β_a^H, β_n^H)	(.26, .01)	Hao et al. (2020) Land share of income
$(\beta_{aa}, \beta_{na}^K)$	(.06, .15)	Hao et al. (2020) Agriculture input shares
(β_{an}, β_{nn})	(.25, .61)	Hao et al. (2020) Non-agriculture input shares
ν	0.87	Tombe and Zhu (2019) Share of expenditure on goods
ϵ_a	2.33	Section 6
η	1.19	Section 6
κ	1.5	Tombe and Zhu (2019) income elasticity of migration
\hat{a}_{is}	Multiple	Section 6
$\tau_{ij,s}$	Multiple	Section 6
$f_{is,jk}$	Multiple	Section 6
\hat{K}_{is}	Multiple	Hao et al. (2020) Capital data
(θ_a, θ_n)	(4, 4)	Tombe (2015) Trade elasticity
(α_a, α_n)	(0.05, 0.05)	Combes and Gobillon (2015) To match $\beta_n^H - \alpha_n = -0.04$

Figure 5: Calibrated parameters.

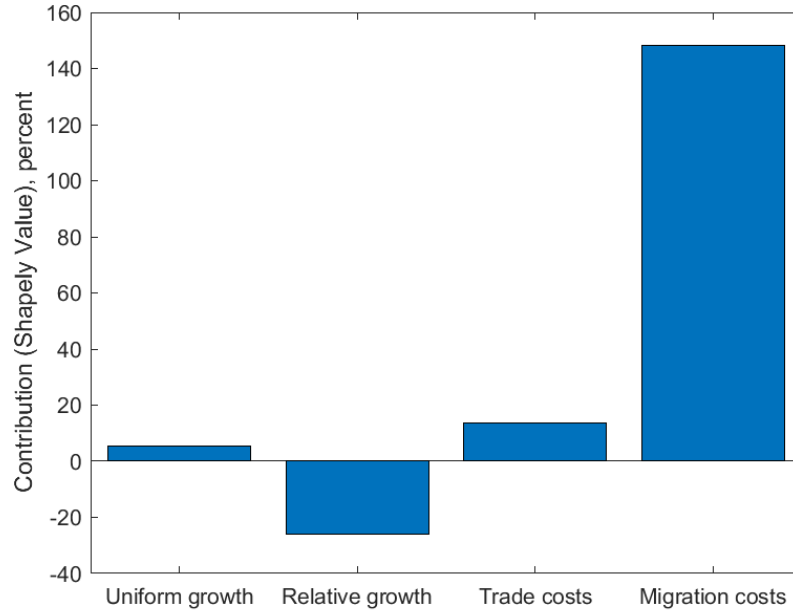


Figure 6: Shapely decomposition of the rise in density dispersion. All values add up to 100 percent. Growth in TFP and capital per worker, when uniform across provinces and set to the observed sector average, accounts for approximately 38 percent.

References

- Allen, T. and C. Arkolakis (2014, 05). Trade and the Topography of the Spatial Economy *. *The Quarterly Journal of Economics* 129(3), 1085–1140.
- Allen, T., C. Arkolakis, and Y. Takahashi (2020). Universal gravity. *Journal of Political Economy* 128(2), 393–433.
- Allen, T. and D. Donaldson (2020, November). Persistence and path dependence in the spatial economy. Working Paper 28059, National Bureau of Economic Research.
- Balassa, B. (1963). An empirical demonstration of classical comparative cost theory. *The Review of Economics and Statistics* 45(3), 231–238.
- Bartelme, D. (2018). Trade costs and economic geography: Evidence from the u.s.
- Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics* 122(2), 775–805.
- Baum-Snow, N. (2020, 12). Urban Transport Expansions and Changes in the Spatial Structure of U.S. Cities: Implications for Productivity and Welfare. *The Review of Economics and Statistics* 102(5), 929–945.
- Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang (2017, 07). Roads, Railroads, and Decentralization of Chinese Cities. *The Review of Economics and Statistics* 99(3), 435–448.
- Baum-Snow, N., J. V. Henderson, M. A. Turner, Q. Zhang, and L. Brandt (2020). Does investment in national highways help or hurt hinterland city growth? *Journal of Urban Economics* 115, 103124. Cities in China.
- Berry, S., A. Gandhi, and P. Haile (2013). Connected substitutes and invertibility of demand. *Econometrica* 81(5), 2087–2111.
- Boppart, T. (2014, 11). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica* 82, 2167–2196.
- Bustos, P., B. Caprettini, and J. Ponticelli (2016, June). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review* 106(6), 1320–65.
- Bustos, P., G. Garber, and J. Ponticelli (2020, 01). Capital Accumulation and Structural Transformation*. *The Quarterly Journal of Economics* 135(2), 1037–1094.
- Caliendo, L. and F. Parro (2014, 11). Estimates of the Trade and Welfare Effects of NAFTA. *The Review of Economic Studies* 82(1), 1–44.
- Caselli, F. and W. J. Coleman II (2001). The u.s. structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy* 109(3), 584–616.
- Chan, K. W. (2019). China’s hukou system at 60: continuity and reform. In *Handbook on Urban Development in China*. Cheltenham, UK: Edward Elgar Publishing.
- Chen, Y., J. V. Henderson, and W. Cai (2017). Political favoritism in china’s capital markets and its effect on city sizes. *Journal of Urban Economics* 98, 69–87. Urbanization in Developing Countries: Past and Present.
- Coşar, A. K. and P. D. Fajgelbaum (2016, February). Internal geography, international trade, and regional specialization. *American Economic Journal: Microeconomics* 8(1), 24–56.
- Combes, P.-P. and L. Gobillon (2015). Chapter 5 - the empirics of agglomeration economies. In G. Duranton, J. V. Henderson, and W. C. Strange (Eds.), *Handbook of Regional and Urban Economics*, Volume 5 of *Handbook of Regional and Urban Economics*, pp. 247–348. Elsevier.

- Comin, D., D. Lashkari, and M. Mestieri (2021, 01). Structural Change With Long-Run Income and Price Effects. *Econometrica* 89(1), 311–374.
- Cravino, J. and S. Sotelo (2019, 07). Trade-induced structural change and the skill premium. *American Economic Journal: Macroeconomics* 11(3), 289–326.
- Davis, D. R. (1998). The home market, trade, and industrial structure. *The American Economic Review* 88(5), 1264–1276.
- Dekle, R., J. Eaton, and S. Kortum (2007, May). Unbalanced trade. *American Economic Review* 97(2), 351–355.
- Delventhal, M. (2019). The Globe as a Network: Geography and the Origins of the World Income Distribution. 2019 Meeting Papers 840, Society for Economic Dynamics.
- Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2018). The geography of development. *Journal of Political Economy* 126(3), 903–983.
- Desmet, K. and E. Rossi-Hansberg (2014, 04). Spatial development. *American Economic Review* 104(4), 1211–43.
- Duarte, M. and D. Restuccia (2010, 02). The Role of the Structural Transformation in Aggregate Productivity*. *The Quarterly Journal of Economics* 125(1), 129–173.
- Duranton, G. and M. A. Turner (2011, October). The fundamental law of road congestion: Evidence from us cities. *American Economic Review* 101(6), 2616–52.
- Duranton, G. and M. A. Turner (2012, 03). Urban Growth and Transportation. *The Review of Economic Studies* 79(4), 1407–1440.
- Eckert, F. and M. Peters (2018). Spatial Structural Change. 2018 Meeting Papers 98, Society for Economic Dynamics.
- Faber, B. (2014, 03). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies* 81(3), 1046–1070.
- Fajgelbaum, P. and S. J. Redding (2014, 06). Trade, structural transformation and development: Evidence from argentina 1869-1914. Working Paper 20217, National Bureau of Economic Research.
- Fan, J. (2018). Hukou reforms in chinese cities , 1997-2010 : Data and facts.
- Fan, J. (2019, July). Internal geography, labor mobility, and the distributional impacts of trade. *American Economic Journal: Macroeconomics* 11(3), 252–88.
- Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization with and without industrialization. *Journal of Economic Growth* 21(1), 35–70.
- Hao, T., R. Sun, T. Tombe, and X. Zhu (2020). The effect of migration policy on growth, structural change, and regional inequality in china. *Journal of Monetary Economics* 113, 112–134.
- Herrendorf, B., R. Rogerson, and Ákos Valentinyi (2014). Chapter 6 - growth and structural transformation. In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 2 of *Handbook of Economic Growth*, pp. 855–941. Elsevier.
- Herzog, I. (2021). National transportation networks, market access, and regional economic growth. *Journal of Urban Economics* 122, 103316.
- Karádi, P. and M. Koren (2017). Cattle, steaks and restaurants: Development accounting when space matters. Technical report, Central European University.
- Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy* 99(3), 483–499.

- Lewis, L. T., R. Monarch, M. Sposi, and J. Zhang (2021, 07). Structural Change and Global Trade. *Journal of the European Economic Association*. jvab024.
- Ma, L. and Y. Tang (2020). Geography, trade, and internal migration in china. *Journal of Urban Economics* 115, 103181. Cities in China.
- Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory* 58(2), 317–334.
- Matsuyama, K. (2009, 05). Structural Change in an Interdependent World: A Global View of Manufacturing Decline. *Journal of the European Economic Association* 7(2-3), 478–486.
- Matsuyama, K. (2019, 03). Engel’s law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica* 87, 497–528.
- Meng, X. (2012, 11). Labor market outcomes and reforms in china. *Journal of Economic Perspectives* 26(4), 75–102.
- Michaels, G., F. Rauch, and S. J. Redding (2012, 03). Urbanization and Structural Transformation *. *The Quarterly Journal of Economics* 127(2), 535–586.
- Murata, Y. (2008). Engel’s law, petty’s law, and agglomeration. *Journal of Development Economics* 87(1), 161–177.
- Nagy, D. K. (2020, February). Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion. Working Papers 1172, Barcelona Graduate School of Economics.
- Ngai, L. R. and C. A. Pissarides (2007, 03). Structural change in a multisector model of growth. *American Economic Review* 97(1), 429–443.
- Ottaviano, G., T. Tabuchi, and J. Thisse (2002). Agglomeration and trade revisited. *International Economic Review* 43(2), 409–436.
- Puga, D. (1999). The rise and fall of regional inequalities. *European Economic Review* 43(2), 303–334.
- Redding, S. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics* 101(C), 148–167.
- Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics* 55(2), 234–250.
- Samuelson, P. A. and S. Swamy (1974). Invariant economic index numbers and canonical duality: Survey and synthesis. *The American Economic Review* 64(4), 566–593.
- Simonovska, I. and M. E. Waugh (2014). The elasticity of trade: Estimates and evidence. *Journal of International Economics* 92(1), 34–50.
- Sotelo, S. (2020). Domestic trade frictions and agriculture. *Journal of Political Economy* 128(7), 2690–2738.
- Swiecki, T. (2017, March). Determinants of Structural Change. *Review of Economic Dynamics* 24, 95–131.
- Timmer, M., G. de Vries, and K. de Vries (2015). *Patterns of structural Change in Developing Countries*, pp. 65–83. Routledge Handbooks. Routledge.
- Tombe, T. (2015, July). The missing food problem: Trade, agriculture, and international productivity differences. *American Economic Journal: Macroeconomics* 7(3), 226–58.
- Tombe, T. and X. Zhu (2019, 05). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review* 109(5), 1843–72.
- Uy, T., K.-M. Yi, and J. Zhang (2013). Structural change in an open economy. *Journal of Monetary Economics* 60(6), 667–682.

A Proofs of Propositions

Proof of Proposition 1

I start by showing that employment in both locations must be positive if there is positive employment in both sectors under Assumption 1 and $1 + \theta\Omega_n > 0$. As mentioned before, I abstract away from equilibria where there is zero total employment in either sector.

Lemma A.1. *Assume 1 and $1 + \theta\Omega_n > 0$. Consider the set of equilibria where there is positive total employment in both sectors. Then employment is positive in all sector-locations and nominal output per worker equalizes across space.*

Proof. Note that $1 + \theta\Omega_a > 0$ holds by these assumptions. Market clearing when nominal output per worker is $\phi_{i,s}$ implies, as an analogue to equation (8),

$$\frac{\phi_{1,s}L_{1,s}}{\phi_{2,s}L_{2,s}} = c^\theta \left[\frac{\phi_{1,s}}{\phi_{2,s}} \right]^{-\theta(1-\Omega_s)} \left[\frac{\beta_a\phi_{1,a}L_{1,a} + \beta_n\phi_{1,n}L_{1,n}}{\beta_a\phi_{2,a}L_{2,a} + \beta_n\phi_{2,n}L_{2,n}} \right]^{-\theta\Omega_s} \quad (21)$$

for every s . Define $Y_{i,s} = \phi_{i,s}L_{i,s}$. Fix any sequence of $\phi_{i,s} > 0$ and $Y_{i,s} > 0$ such that (21) holds and $\frac{Y_{1,s}}{Y_{2,s}} \rightarrow 0$ for at least some s . Also assume without loss that $\sum_{i,s} Y_{i,s} = 1$. I will show that this implies that relative nominal output per worker $\frac{\phi_{1,s}}{\phi_{2,s}}$ grows without bound, making a spatial equilibrium impossible. (21) can be rewritten as

$$\left[\frac{\phi_{1,s}}{\phi_{2,s}} \right]^{1+\theta(1-\Omega_s)} = c^\theta \left[\frac{\beta_a Y_{1,a} + \beta_n Y_{1,n}}{\frac{Y_{1,s}}{Y_{2,s}}(\beta_a Y_{2,a} + \beta_n Y_{2,n})} \right]^{-\theta\Omega_s} \left[\frac{Y_{1,s}}{Y_{2,s}} \right]^{-(1+\theta\Omega_s)} \quad (22)$$

The second term $\left[\frac{\beta_a Y_{1,a} + \beta_n Y_{1,n}}{\frac{Y_{1,s}}{Y_{2,s}}(\beta_a Y_{2,a} + \beta_n Y_{2,n})} \right]$ on the right hand side of this equation defines a sequence bounded within some interval $(0, b)$ for $a > 0$ and b finite. Since $\Omega_a > 0$ this implies that relative wages in the agricultural sector diverge as relative employment $\frac{Y_{1,a}}{Y_{2,a}} \rightarrow 0$. What remains is to show that relative wages in the n sector also behave in this way if $\frac{Y_{1,n}}{Y_{2,n}} \rightarrow 0$. The second term in (22) for the n sector can be written as

$$\frac{\beta_a}{\beta_a + \beta_n \frac{Y_{2,n}}{Y_{2,a}}} + \frac{\beta_n}{\beta_a + \beta_n \frac{Y_{2,a}}{Y_{1,a}} \frac{Y_{1,n}}{Y_{2,n}}}$$

Suppose for contradiction that $\frac{\phi_{1,n}}{\phi_{2,n}}$ remains bounded. A necessary condition is then that this term tends to 0. But the only way that this is possible is if $\frac{Y_{1,a}}{Y_{2,a}} \rightarrow 0$ which we have already ruled out as supporting a spatial equilibrium. Hence, there is positive output in each sector-location. Under free mobility, this means that nominal output per worker equalizes across all sector-locations. \square

In what follows, I search for equilibria that are interior, so that output per worker is equal across space and sectors and is normalized to 1.

I solve equations (7), (8), and (9) to arrive at a map between relative employment in each sector as a function of total agricultural employment ω . Let $\frac{L_{1,a}}{L_{2,a}} = x$. Then the equilibrium level of x is implicitly defined by the relation

$$x = c^\theta \left[\nu(x, \omega)x + (1 - \nu(x, \omega))c^{\theta \frac{\Omega_a - \Omega_n}{\Omega_a}} x^{\frac{\Omega_n}{\Omega_a}} \right]^{-\theta \Omega_a} \quad (23)$$

where $\nu(x, \omega)$ is the share of land in agriculture in location 2. Define the function $E(x, \omega)$ as the right hand side of the equation. Then the equilibrium value of x given ω is a fixed point of E given ω . Moreover, it can be shown that this relation defines x as a continuous function of ω . It can also be shown that, when $x \in [0, c^\theta]$, $E(x, \omega)$ is bounded by the functions

$$c^{\theta(1-\theta(\Omega_a - \Omega_n))} x^{-\theta \Omega_n} \leq E(x, \omega) \leq c^\theta x^{-\theta \Omega_a} \quad (24)$$

Since $\Omega_a > \Omega_n$. The set of fixed points of E must then be bounded by the minimum fixed point of $c^{\theta(1-\theta(\Omega_a - \Omega_n))} x^{-\theta \Omega_n}$ and the maximum fixed point of $c^\theta x^{-\theta \Omega_a}$ provided these fixed points occur in the interval $[0, c^\theta]$. These correspond to the unique fixed points $c^{\frac{\theta(1-\theta(\Omega_a - \Omega_n))}{1+\Omega_n\theta}}$ and $c^{\frac{\theta}{1+\Omega_a\theta}}$, respectively, which both occur in the interval under the maintained assumption that $1 + \Omega_n\theta > 0$. This leads to the following lemma.

Lemma A.2. *Suppose Assumption 1 and $1 + \Omega_n\theta > 0$. Then $\frac{L_{1,n}}{L_{2,n}} > \frac{L_{1,a}}{L_{2,a}}$ in every equilibrium, and in all of such equilibria, $\frac{L_{1,a}}{L_{2,a}}$ lies in the interval $\left[c^{\frac{\theta(1-\theta(\Omega_a - \Omega_n))}{1+\Omega_n\theta}}, c^{\frac{\theta}{1+\Omega_a\theta}} \right]$. Lastly, $\frac{L_{1,a}}{L_{2,a}}$ attains this lower bound when $\omega = 0$.*

Proof. The second statement follows from the arguments above. The first follows from the fact that $c^{\theta \frac{(\Omega_a - \Omega_n)}{\Omega_a}} x^{\frac{\Omega_n}{\Omega_a}} > x$ if and only if $x < c^\theta$. See Equation (9).

Lastly, $\omega = 0$ implies that $\nu(x, \omega) = 0$ and thus $E(x, \omega) = c^{\theta(1-\theta(\Omega_a-\Omega_n))}x^{-\theta\Omega_n}$, whose fixed point in x attains the lower bound. \square

We can use the lemma to bound the possible densities that can occur in equilibrium, which bound the permissible prices, to show that ω must lie on the interior of $[0, 1]$. To this end, density in sector-location 1, s can be written as, for $x = \frac{L_{1,a}}{L_{2,a}}$

$$d_{1,s}(\omega) = \frac{\beta_a \frac{x}{x+1} \omega + \beta_n \frac{l(x)}{l(x)+1} (1-\omega)}{\beta_s} \frac{L}{H} \quad (25)$$

where $l(x) = c^{\theta \frac{(\Omega_a - \Omega_n)}{\Omega_a}} x^{\frac{\Omega_n}{\Omega_a}}$ as in Equation (9). This expression must lie in some strictly positive compact interval for every $\omega \in [0, 1]$ by the previous lemma. The same argument holds for location 2. Thus the set of permissible equilibrium log-price indices $\log(P_s)$ are bounded, as they are continuous functions of density in both locations on some strictly positive compact interval. This means that some value of ω must exist that solves Equation (6) by the Intermediate Value Theorem. This must be the equilibrium ω .

Lemma A.3. *There exists an interior equilibrium.*

With these caveats out of the way, I can finally show that growth in the TFP terms $a_{i,s}$ tend the equilibrium level of $\omega \rightarrow 0$. To this end, the equilibrium level of ω can be written as the solution to Equation (6)

$$\begin{aligned} \log(\omega) - \epsilon_a \log(1 - \omega) &= -(1 - \eta) \log(a_a) + (1 - \eta) \epsilon_a \log(a_n) + \\ &- (1 - \eta) \frac{1}{\theta} \log \left[c^\theta d_{1,a}(\omega)^{-\theta\Omega_a} + d_{2,a}(\omega)^{-\theta\Omega_a} \right] + (1 - \eta) \epsilon_a \frac{1}{\theta} \log \left[c^\theta d_{1,n}(\omega)^{-\theta\Omega_n} + d_{2,n}(\omega)^{-\theta\Omega_n} \right] \end{aligned} \quad (26)$$

where the density terms on the second line are bounded over all ω . By Assumption 2, the terms $-(1 - \eta) \log(a_a) + (1 - \eta) \epsilon_a \log(a_n)$ decrease without bound with respect to time. This immediately implies that all possible ω solving (26) must lie in some interval (a, b) where $a \rightarrow 0$ and $b \rightarrow 0$.

Lemma A.4. *Suppose Assumption 2 holds. Then the set of possible equilibrium ω lie in the interval (a, b) where $a \rightarrow 0$ and $b \rightarrow 0$.*

I can now prove the entire proposition. Total employment in location 1 can

be written in terms of x and ω ,

$$L_{1,a} + L_{1,n} = \left[\frac{x}{x+1} \omega + \frac{l(x)}{l(x)+1} (1-\omega) \right] L \quad (27)$$

I will show that this term approaches a maximized value when $\omega \rightarrow 0$. Lemma (A.2) implies that x attains its minimum value when $\omega = 0$ and also implies $l(x) > x$. Assumption 1 implies that $l(x)$ is strictly decreasing in x , and thus reaches its maximum value when $\omega = 0$. Let x_{min} be the minimum possible value of x given by the interval in Lemma (A.2). Then the following set of inequalities holds

$$\frac{x}{x+1} < \frac{l(x)}{l(x)+1} < \frac{l(x_{min})}{l(x_{min})+1} \quad (28)$$

for every $\omega \in (0, 1)$ and so

$$\left[\frac{x}{x+1} \omega + \frac{l(x)}{l(x)+1} (1-\omega) \right] L < \frac{l(x_{min})}{l(x_{min})+1} L \quad (29)$$

for every $\omega \in (0, 1)$. Since the implicit function $x(\omega)$ is continuous in ω , $L_{1,a} + L_{1,n}$ approaches this maximum value $\frac{l(x_{min})}{l(x_{min})+1} L$ as $\omega \rightarrow 0$. This completes the proof.

B Additional Figures and Tables

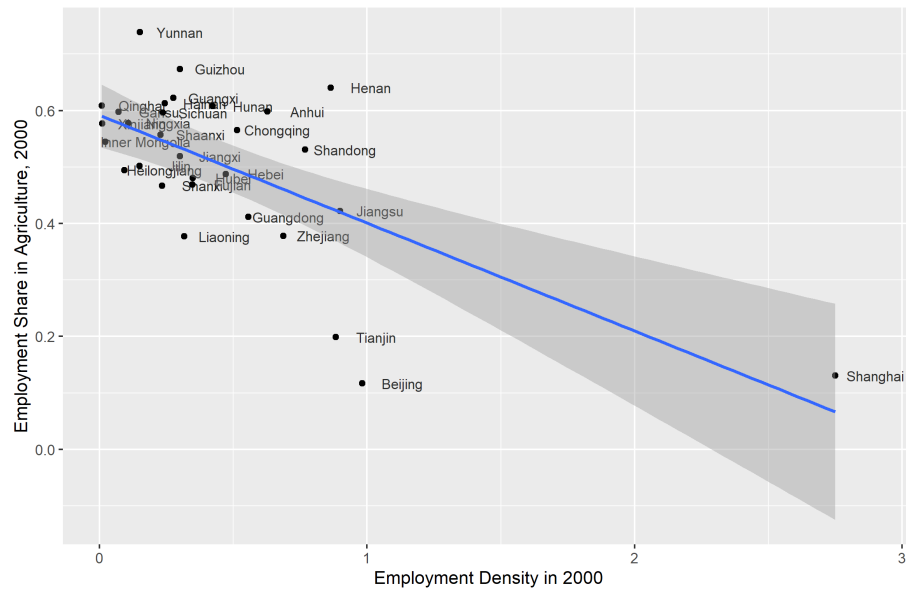


Figure 7: Relationship between primary sector employment and initial employment density by Chinese province. Density is measured in millions per 1000 square miles. The model explains greater than 45 percent of the variation in agriculture employment shares. Similar results hold for the 2005 cross section.

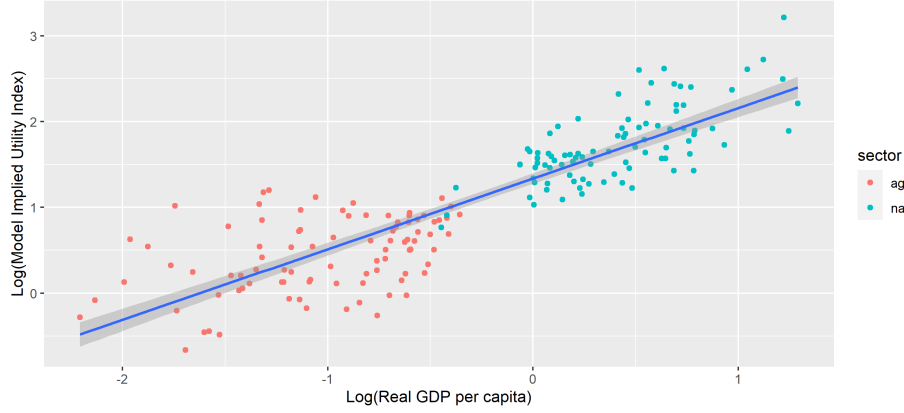


Figure 8: Relationship between model implied real GDP of goods consumption against real income per capita constructed in Tombe and Zhu (2019). Sector "ag" corresponds to agricultural workers. The model implied index is constructed using the relation $\log(U_{ist}) = \log(\phi_{is,t}) - \log(P_{in,t}) + \frac{1}{1-\eta} \log[1 - \omega_{is,t}]$ implied by Equation (6). R^2 of the regression is over 55 percent.

ϵ_a	η	λ_{2000}	λ_{2005}	λ_{2010}
1.875	1.698	0.875	-0.157	0.017
[0.832, 3.459]	[1.437, 2.073]	[-0.2665, 2.2918]	[-0.4774, 0.227]	[-0.5304, 0.6492]

Figure 9: Estimates under the specification used in Comin et al. (2021). Assuming perfect mobility within locations but not across them, as well as no trade, their estimating equation is instead $\log\left[\frac{L_{iat}}{L_{int}}\right] = (1 - \eta)\log\left[\frac{P_{iat}}{P_{int}}\right] + (1 - \eta)(\epsilon_a - 1)\log\left[\frac{\phi_{ist}}{P_{int}}\right] + (\epsilon - 1)\log(1 - \omega_{ist}) + \sum_t \lambda_t + \log(v_{ist})$. I do not introduce sector trade controls that they do, and these do not affect their estimates. 99 percent BCa confidence intervals are reported. This equation is estimated using only the non-agricultural subsample because there is no within-province variation in the dependent variable. These results imply that the entire fall in spending shares can be explained by substitution effects.

	$\Delta M = 0, \Delta R = 0$	$\Delta M = 0, \Delta R = 1$	$\Delta M = 1, \Delta R = 0$	$\Delta M = 1, \Delta R = 1$
$\Delta T = 0, \Delta A = 0$	1.0646	0.9998	1.2039	1.0990
$\Delta T = 0, \Delta A = 1$	1.0953	1.0033	1.2597	1.1046
$\Delta T = 1, \Delta A = 0$	1.0826	1.0121	1.2327	1.1196
$\Delta T = 1, \Delta A = 1$	1.1126	1.0160	1.2860	1.1259

Figure 10: Table of equilibrium population density dispersion, measured in the coefficient of variation of population density across provinces. ΔA corresponds to uniform growth, ΔR corresponds to relative growth, ΔM to changes in migration costs and ΔT to changes in trade costs.

C Equilibrium Description and Computation

Definition

In this appendix, I detail the system of equations in 2005 that is used to construct the counterfactual equilibria. All variables are defined in Section 5, which I repeat here for convenience. As is standard notation, variables with hats denote values in 2005 relative to values in 2000. The endogenous variables determined in equilibrium are output per worker $\hat{\phi}_{is}$, change in employment \hat{L}_{is} , marginal costs \hat{c}_{is} , price indices \hat{P}_{is} , housing production and consumption allocations \hat{H}_{is} and \hat{H}_{is}^c , and rental rates \hat{r}_i , and 2005 spending shares on sector k goods ω_{is}^k and migration flows m_{isjk} . I start with the marginal cost pricing condition, for location i and sector s ,

$$\hat{p}_{is} = \hat{c}_{is} = \frac{\hat{\phi}_{is} v_s \left[\frac{\hat{L}_{is}}{\hat{H}_{is}} \right]^{\beta_s^H} \left[\frac{\hat{L}_{is}}{\hat{K}_{is}} \right]^{\beta_s^K} \hat{P}_{ia}^{\beta_{as}} \hat{P}_{in}^{\beta_{ns}}}{\hat{a}_{is} \left[\frac{\hat{L}_{is}}{\hat{H}_{is}} \right]^{\alpha_s}} \quad (30)$$

where v_s is the value added share in sector s . We have the price index equation, which can be written in hat-algebra form

$$\hat{P}_{is} = \left[\sum_{j \in \mathbf{P}} \pi_{ji,s,2000} \tau_{ji,s}^{\hat{c}_{js}} \hat{c}_{js}^{-\theta_s} \right]^{-\frac{1}{\theta_s}} \quad (31)$$

and a system of equations that implies the goods market clearing and balanced trade, represented in hat-algebra form for all i and s ,

$$\sum_{j \in \mathbf{P}, k \in \{a,n\}} \left[\omega_{jk}^s + \beta_{sk} \frac{1}{v_k} \right] \left[\hat{\phi}_{jk} \hat{L}_{jk} \phi_{jk,2000} L_{jk,2000} \right] \frac{\hat{\phi}_{is} \hat{L}_{is} \phi_{is,2000} L_{is,2000}}{v_s} = \sum_{l \in \mathbf{P}} \pi_{lj,s,2000} (\tau_{lj,s}^{\hat{c}_{ls}} \hat{c}_{ls}^{-\theta_s}) \quad (32)$$

where $\phi_{is,2000}$ is the output per worker uniquely constructed to balance trade in 2000 given trade flow data, with $L_{is,2000}$ and $\pi_{is,2000}$ being employment and trade flow data in 2000, and $\omega_{is,2000}^k$ the observed spending share in sector k in

2000.

Agriculture spending shares in China are determined by the equation (noting that the preference parameter γ_{is} is time invariant)

$$\begin{aligned} & \log \omega_{is}^a - \epsilon_a \log(1 - \omega_{is}^a) = \\ (1 - \eta) \log \left[\frac{\hat{P}_{ia}}{\hat{P}_{in}} \right] & + (\epsilon_a - 1)(1 - \eta) \log \left[\frac{\hat{\phi}_{is}}{\hat{P}_{in}} \right] + \log \omega_{is,2000}^a - \epsilon_a \log(1 - \omega_{is,2000}^a) \end{aligned} \quad (33)$$

where the international spending share remains exogenous at world value added. Migration flows, as a fraction of total hukou registrants in (j, k) in 2005, can be described as follows (in hat algebra form),

$$m_{jkis} = \frac{(\hat{W}_{is}/\hat{f}_{jkis})^\kappa}{\sum_{i' \in \mathbf{P} \setminus \{RoW\}, s' \in \{a, n\}} (\hat{W}_{i's'}/\hat{f}_{i's'is})^\kappa m_{i's'is,2000}} \quad (34)$$

where $m_{isjk,2000}$ are migration flows in the data. The welfare index \hat{W}_{is} satisfies

$$\hat{W}_{is} = \frac{\hat{\phi}_{is}}{\left[\sum_{k \in \{a, n\}} \omega_{is,2000}^k \hat{P}_{ik}^{1-\eta} \right]^{\frac{\nu}{1-\eta}} \hat{r}_i^{1-\nu}} \quad (35)$$

labour market clearing in 2005 implies

$$\hat{L}_{is} L_{is,2000} = \sum_{i' \in \mathbf{P} \setminus \{RoW\}, s' \in \{a, n\}} m_{i's'is} M_{i's',2005} \quad (36)$$

where $M_{is,2005}$ is the inferred stock of individuals with the corresponding hukou in 2005, which is exogenous. Finally, \hat{r}_i is determined by the unique land allocation \hat{H}_{is} and \hat{H}_{is}^c that equates marginal products with the price of land in consumption.

An equilibrium allocation is defined as a solution to each equation in the endogenous variables as a function of \hat{K}_{is} , \hat{a}_{is} , $\hat{f}_{is,jk}$ and $\tau_{ij,s}$.

Computation

Derivation of Welfare Index

I define CV_{is} as the fraction of income deducted in 2005 in order for a worker in (i, s) to be as well off as in 2000. Let $E_{is}(P, V)$ denote the minimum expenditure required to achieve consumption level V at prices P for workers in (i, s) . This is derived from the utility function in Equation (10) with preference parameters γ_{is} allowed to vary. Then, \tilde{C}_{is} is defined by the equation

$$(1 - CV_{is}) \frac{\phi_{is,2005}}{1 - \nu} = E(P_{is,2005}, V_{2000}) \quad (37)$$

where V_{2000} is the utility achieved in 2000 under time-invariant preferences, $\phi_{is,2005}$ value added per worker in 2005 (in this case proportional to nominal income because land spending is rebated and preferences over housing are Cobb-Douglas) and $P_{is,2005}$ are the prices faced by workers in 2005. Define $\mathbf{P}_{is}(P, V) = \frac{E_{is}(P, V)}{V}$ as the *price index* associated with the level of utility V and prices P . Also note by duality, $\phi_{is,2005} = E(P_{is,2005}, V_{2005})$. Then (37) can be re-written as

$$\frac{1}{1 - CV_{is}} = \frac{\hat{\phi}_{i,s}}{\hat{\mathbf{P}}_{is}(V_{2000})} \quad (38)$$

Where $\hat{\mathbf{P}}_{is}(V_{2000})$ is the proportional change in the price index *holding utility constant at 2000 levels*. If preferences are homothetic, this change in price index is constant in V and so (38) reduces to the proportional change in utility \hat{V} (Samuelson and Swamy, 1974). However, since preferences are not homothetic, the price index will depend on the initial level of utility V_{2000} which breaks the connection with \hat{V} .

I now derive the price index $\hat{\mathbf{P}}_{is}$. To do this, I need to define a second price index associated only with goods consumption, with the measure of goods consumption coming from the non-homothetic CES in Equation (4). Call this index $\mathbb{P}_{is}(P, U)$ associated with prices P and goods consumption level U . The connection with \mathbf{P}_{is} and \mathbb{P}_{is} is derived, noting that preferences over goods and housing is Cobb Douglas,

$$\mathbf{P}_{is} = \mathbb{P}_{is}(P, U)^\nu r_i^{1-\nu} \quad (39)$$

and so

$$\mathbf{P}_{is}(V_{2000}) = \mathbb{P}_{is}^{\hat{}}(U_{2000})^{\nu} r_i^{\hat{}}^{1-\nu} \quad (40)$$

It remains to solve for $\mathbb{P}_{is}^{\hat{}}(U_{2000})$, which is the change in the goods price index holding the measure of goods consumption U at 2000 levels. This can be written directly as (via the solution to the consumers problem)

$$\mathbb{P}_{is}^{\hat{}}(U_{2000}) = \frac{\left[\sum_{k \in \{a,n\}} U_{2000}^{(\epsilon_k - 1)(1-\eta)} P_{ik,2005}^{1-\eta} \right]^{\frac{1}{1-\eta}}}{\left[\sum_{k \in \{a,n\}} U_{2000}^{(\epsilon_k - 1)(1-\eta)} P_{ik,2000}^{1-\eta} \right]^{\frac{1}{1-\eta}}} \quad (41)$$

However, solving the consumers problem also implies $\omega_{is,2000}^k = \frac{U_{2000}^{\epsilon_k(1-\eta)} P_{ik,2000}^{1-\eta}}{\left[\frac{\phi_{is,2005}}{1-\nu} \right]^{\frac{1}{1-\eta}}}$,

which coupled with the definition $\left[\sum_{k \in \{a,n\}} U_{2000}^{(\epsilon_k - 1)(1-\eta)} P_{ik,2000}^{1-\eta} \right]^{\frac{1}{1-\eta}} = \frac{\phi_{is,2000}}{U_{2000}(1-\nu)}$ and subbing both into Equation (41) yields the index measure

$$\mathbb{P}_{is}^{\hat{}}(U_{2000}) = \left[\sum_{k \in \{a,n\}} \omega_{is,2000}^k \hat{P}_{ik}^{1-\eta} \right]^{\frac{1}{1-\eta}} \quad (42)$$

Substituting (42) into (40) and then into (38) yeilds the welfare index formula used in (35).