Urbanization and Structural Transformation*

Guy Michaels†
London School of Economics, CEP and CEPR
Ferdinand Rauch‡
University of Vienna and CEP
Stephen J. Redding§
London School of Economics, CEP and CEPR

February 6, 2009

Abstract

This paper presents new evidence on urbanization using sub-county data for the United States from 1880-2000 and municipality data for Brazil from 1970-2000. We show that the two central stylized features of population growth for cities – Gibrat's Law and a stable population distribution – are strongly rejected when both rural and urban areas are considered. We provide evidence that the observed U-shaped relationship between population growth and initial population density can be explained by different employment growth dynamics in the agricultural and non-agricultural sectors and the process of structural transformation away from the agricultural sector.

Keywords: Urbanization, Structural Transformation, Population Growth
JEL: E00, N10, O18, R12

---

*This paper was produced as part of the Globalization Programme of the ESRC-funded Centre for Economic Performance (CEP). We thank Yu-Hsiang Lei, Ralph Ossa, and Bart Thia for excellent research assistance. We also thank Robin Burgess, Francesco Caselli, Tim Guinnane, Alan Manning, Rachel Ngai, Steve Pischke, Chris Pissarides and seminar participants at Helsinki, LSE, Tel-Aviv University, Tinbergen Institute, University of Missouri and Zurich for helpful comments. Responsibility for any opinions, results and errors lies with the authors alone.

†Department of Economics, Houghton Street, London, WC2A 2AE, United Kingdom, *tel:* (44) 20 7852 3518, *fax:* (44) 20 7955 7595, *email:* g.michaels@lse.ac.uk.

‡Department of Economics, Hohenstaufengasse 9, A-1010, Vienna, Austria, *tel:* (43) 1 4277 37455, *fax:* (43) 1 4277 9374, *email:* ferdinand.rauch@univie.ac.at.

§Department of Economics, Houghton Street, London, WC2A 2AE, United Kingdom, *tel:* (44) 20 7955 7483, *fax:* (44) 20 7955 7595, *email:* s.j.redding@lse.ac.uk.

1 Introduction

Urbanization – the concentration of population in cities and towns – is one of the most striking features of economic development.[1] The share of the world's population living in cities grew from less than one tenth in 1300, to around one sixth in 1900 and to almost one half today.[2] While this transition from rural to urban is largely complete in developed countries such as the United States, the urbanization process continues apace in developing countries such as Brazil, China and India. In China alone, 240 million people are expected to migrate from rural to urban areas by 2025, helping to raise the share of the world's population living in cities to 60 percent by 2030.[3]

This urbanization process involves radical changes in the organization of economic activity and presents major challenges for public policy. On the one hand, dense population concentrations create new demands for resources: agricultural surpluses to feed the urban population, mass transit systems to deal with congestion, supplies of clean water and power, sanitation and waste management systems, and public health facilities. On the other hand, as millions of people choose to move from rural to urban areas, there must be some compensating advantages, including higher productivity and wages as well as a wider range of employment opportunities and goods and services.[4]

While a large literature in development economics and macroeconomics has considered the aggregate implications of structural transformation from agriculture to non-agriculture, its implications for the distribution of economic activity across space have received relatively little attention. This is perhaps surprising as there are good reasons for believing that structural transformation is central to the process of urbanization. More rapid productivity growth in agriculture than in non-agriculture can create the food surpluses to feed the urban population. Additionally, agriculture and non-agriculture differ substantially in terms of their intensity of land use and the extent to which their productivity is tied to persistent

[1] The US Census Bureau defines an urban area as territory consisting of core census blocks with a population density of at least 1,000 people per square mile and surrounding census blocks with a population density of at least 500 people per square mile (Census 2000d).

[2] The historical figures are from Bairoch (1988) and the present-day figures from United Nations (2005).

[3] The estimates for China are from Mckinsey (2008) and those for the world are from United Nations (2005).

[4] There is a large empirical literature documenting higher productivity in urban than rural areas (see for example the survey by Rosenthal and Strange 2004). Similarly, an extensive body of research examines the relationship between urbanization and income inequality (see for example Kuznets 1955 and Black and Henderson 1999).

idiosyncratic features of locations such as soil and climate.

In this paper, we provide theory and empirical evidence that structural transformation is indeed central to understanding urbanization and the broader evolution of the distribution of population across space and over time. While most previous research in this area has concentrated on cities, rural areas accounted for a large share of the population in developed countries historically, and they continue to account for large shares of the population in developing countries today. Therefore, our analysis makes use of a new source of data for the United States that encompasses both urban and rural areas. This dataset exploits information on sub-county units, which are commonly referred to as Minor Civil Divisions (MCDs), and extends for more than a century from 1880 to 2000.[5] Our data include information on both population and employment by industry and are characterized by the following six stylized facts. First, despite substantial US population growth, as reflected in an increased mass of densely-populated areas over time, the mass of sparsely-populated areas also increases over time. As a result there is an unstable population distribution, which exhibits polarization: the difference in density between densely and sparsely-populated areas increases over time (Stylized Fact 1).

Second, while our data confirm previous findings that Gibrat's Law is a reasonable approximation for densely-populated urban areas, we show that this feature of population growth is strongly rejected when we include both rural and urban areas (Stylized Fact 2). For this more comprehensive range of locations, population growth is decreasing in initial population density at low densities, and then increasing in initial population density at intermediate densities, before becoming uncorrelated with initial population density at high densities in urban areas.[6] Although a natural explanation for the decreasing relationship between population growth and initial population density at low densities is mean reversion, the explanation for the increasing relationship at intermediate densities is less immediately

[5]We exclude Alaska, Hawaii, Oklahoma, North Dakota, and South Dakota, which had not attained statehood in 1880, and therefore are either not included in the 1880 census or did not have stable county boundaries at that time. Additionally, we use county data for some states where sub-county units are not comparable over time. We discuss in further detail below the construction of our data and the robustness of our results to the sample and specification.

[6]While the existing literature on cities concentrates on the relationship between population growth and population size, we focus on the relationship between population growth and population density to control for differences in land area across sub-county units. Although our results are qualitatively the same if we instead use population size, the population density specification is more appropriate if land area varies across sub-county units and is derived directly from our theoretical model. In our data, there is a strong and approximately log linear relationship between population density and population size, which is consistent with the theoretical model developed below.

clear.

Our third stylized fact is that the correlation between population growth and initial population density is systematically related to differences in employment structure between agriculture and non-agriculture (industry and services). In particular, the share of agriculture in employment declines sharply over the range of values for initial population density for which population growth increases. Our fourth stylized fact is that there is a higher variance in the distribution of employment per square kilometer in non-agriculture than in agriculture (so non-agricultural employment is more concentrated across space). Finally, our last two stylized facts are concerned with differences in employment dynamics in the two sectors. In agriculture, employment growth is decreasing in initial population density (Stylized Fact 5). In contrast, in non-agriculture, employment growth is largely uncorrelated with initial population density (Stylized Fact 6).

As our data span a long historical time period during which the economic environment in the United States changed considerably, we undertake a wide range of robustness checks to sample and econometric specification. We estimate our baseline specification non-parametrically to allow for a flexible relationship between population growth and initial population density. We show that our results are robust to the inclusion of state fixed effects, which in this cross-section specification control for changes in institutions and other characteristics of states that can affect population growth. Among several robustness checks, we find a similar pattern of results when we restrict the sample to a subset of the former thirteen colonies that have similar organizations of administrative functions at the county and sub-county level and the most stable administrative boundaries over time. Additionally, we find a similar pattern of results if we aggregate sub-county units within the immediate vicinity of a city to construct larger economic units, suggesting that our results are not driven by suburbanization around the boundaries of existing cities. Finally, while counties are relatively coarse spatial units for examining the transition from rural to urban, we also find the same pattern of results at the county level.

Most importantly, we also replicate our entire analysis for Brazil for the period 1970-2000. Like the United States during the nineteenth and twentieth centuries, Brazil experienced rapid structural transformation away from agriculture during 1970-2000, and therefore we would expect the mechanisms emphasized in our model to apply. Even though these data are for a different country with distinct institutions and physical geography, and are collected at a different level of spatial aggregation and for a different time period, we find strikingly similar results to those we document for the United States. This similarity of the results in a quite different context reassures us that our findings are not driven by idiosyncratic features of the data or the institutional environment in the United States.

To make sense of our empirical findings, we develop a simple general equilibrium model of the distribution of population across locations that distinguishes between agriculture and non-agriculture. Workers are geographically mobile and the distribution of population across space is determined by the requirement that workers are indifferent between locations. Within each location, land is allocated endogenously to residential and commercial use depending on its relative return in the two types of activities. Land used commercially can be employed in either the agricultural or non-agricultural sector, and in equilibrium it is allocated to the sector in which it has the higher value marginal product. As idiosyncratic shocks to productivity in each sector and location occur, workers move across locations to arbitrage away real wage differences, and land endogenously switches between agricultural and non-agricultural use within locations.

As in the large macroeconomics literature concerned with unbalanced growth, structural transformation away from agriculture occurs because productivity growth is more rapid in agriculture than in non-agriculture and there is inelastic demand between the two goods.[7] In the presence of inelastic demand, more rapid agricultural productivity growth leads to a more than proportionate decline in the relative price of the agricultural good, which in turn leads to a reallocation of employment from agriculture to non-agriculture.[8] Stronger mean reversion in productivity in agriculture than in non-agriculture implies other things equal greater productivity dispersion in non-agriculture than in agriculture and hence a non-agricultural employment share that is increasing in population density at high densities. The combination of stronger mean reversion in agricultural productivity, a non-agricultural employment share that is increasing in population density at high densities, and structural transformation away from the agricultural sector accounts for the six stylized facts.

Our paper is related to a large body of work in urban economics and economic geography.

[7] See in particular Baumol (1967), Galor (2005), Galor and Weil (2000), Goodfriend and McDermott (1995), Ngai and Pissarides (2007) and Rogerson (2008).

[8] As discussed further below, the model can be extended to incorporate the other leading explanation for structural transformation in the macroeconomics literature, namely non-homothetic preferences and technological progress that raises real incomes. See among others Echevarria (1997), Gollin et al. (2002) and Matsuyama (2002).

Recent research on the relationship between population growth and size in the literature on cities includes Duranton (2007), Eeckhout (2004), Gabaix (1999) and Rossi-Hansberg and Wright (2007). While this literature typically finds that the Pareto distribution provides a reasonable approximation to the observed distribution of city sizes, Rossi-Hansberg and Wright (2007) find systematic departures from the Pareto distribution in the upper and lower tails. Evidence of departures from a Pareto distribution is also found by Eeckhout (2004), who uses data on Census Designated Places (CDPs) in the United States to provide evidence that the population distribution is log normal, as implied by Gibrat's Law of constant proportional growth.[9] While Gibrat's Law is typically found to provide a reasonable approximation to observed population growth rates in the cities literature, there is also some evidence of departures from Gibrat's Law. Using data for different countries, Black and Henderson (2003), González-Val et al. (2008) and Soo (2007) find evidence of correlations between population growth and initial population size.[10] Dividing the surface of the continental United States into a uniform grid of six-by-six mile squares, Holmes and Lee (2008) find that population growth from 1990-2000 is highest at intermediate values of initial population density. Two issues in the existing cities literature are the treatment of new cities and the minimum population size to be considered as a city. Our approach addresses both of these issues by considering the entire distribution of population across both rural and urban areas.[11]

Our focus on the reallocation of economic activity from agriculture to non-agriculture also connects with theories of new economic geography, including Fujita et al. (1999) and Krugman (1991). Although reductions in trade costs in these models can result in a polarization of population across space, they do not provide natural explanations for why Gibrat's Law is a reasonable approximation for observed city population growth (see for example the discussion in Davis and Weinstein 2002) or for why Gibrat's Law is violated when both rural and urban areas are considered. While a large literature has examined the empirical determinants of the distribution of economic activity across states and counties in the United States, much of this literature abstracts from the reallocation of economic activity from

[9] See also Eeckhout (2008) and Levy (2008).

[10] Research on the empirical determinants of city growth includes among others Glaeser et al. (1992), da Mata et al. (2007) Ioannides and Overman (2004), and is surveyed in Gabaix and Ioannides (2004). The role of industrial specialization is emphasized in Henderson (1974).

[11] Within the cities literature, Henderson and Wang (2007) and Henderson and Venables (2008) examine the emergence of new cities as a source of growth in the urban population, while Henderson (2005) and Williamson (1965) examine the relationship between the share of the urban population and economic development.

agriculture to non-agriculture.[12] Closer in spirit to our work is Caselli and Coleman (2001), who examine structural transformation and the convergence of incomes between Southern and Northern US states. Also related is Desmet and Rossi-Hansberg (2007), who examine differences in patterns of employment growth between the manufacturing and service sectors using US county data, and relate these differences to technological diffusion and the age of sectors. Neither paper examines the relationship between structural transformation and urbanization – an analysis for which our newly-constructed sub-county data are especially well suited.

In addition to the macroeconomic literature discussed above, our research is related to the development and economic history literatures. Early work documenting the importance of structural change to economic development is surveyed in Syrquin (1988), while more recent research on the interlinkages between industrial and agricultural development is reviewed in Foster and Rosenzweig (2008). Influential work on the history of urban development in the United States includes Kim (2000) and Kim and Margo (2004), although for reasons of data availability this research has again largely concentrated on cities.

The remainder of the paper is organized as follows. Section 2 discusses our main dataset for the United States, outlines our empirical strategy, presents our main empirical findings, and reports the results of a number of robustness checks. Section 3 presents the results of an additional robustness check using Brazilian data. Section 4 outlines our theoretical model and Section 5 shows that it can quantitatively account for the patterns observed in our data. Section 6 concludes.

2 US Data and Stylized Facts

2.1 Data and Samples

This section begins by introducing the US data that we use in this paper and the samples that we construct. We then document a set of stylized facts that shed light on the dynamics of urban and rural population growth from 1880-2000.[13]

In order to analyze these dynamics, we require data on land area, population, and sectoral employment for geographic units that are consistent over time. Since we are interested in

[12]See Beeson et al. (2001), Ellison and Glaeser (1999), Glaeser (2008), Kim (1995) and Rappaport and Sachs (2003) among others.

[13]For further discussion of the US data and the samples discussed below, see the web-based technical appendix.

both rural and urban areas, we also require that these geographic units partition the land area that we analyze. In other words, we want a dataset that covers the entire population and all the land - from the largest cities to the smallest farms. And since we are interested in examining rural and urban population dynamics, we prefer that our geographic units be fine enough to separate urban areas from rural ones.

While these criteria may seem natural, it is not easy to find an existing dataset that satisfies them all. The literature on urban growth in the US has often analyzed counties or Metropolitan Statistical Areas (MSAs), which are groups of counties. And although counties satisfy most of our requirements, they often pool together urban centers with their surrounding countryside. So while we include counties in our analysis, we are also interested in data that provide finer spatial aggregation. One dataset that is less aggregated than the county dataset includes incorporated places - this is the dataset used by Eeckhout (2004). But while this data is useful for studying urban growth dynamics, it does not contain information on many rural areas, where the majority of US population lived before the 20th century.

Since existing datasets are not fully satisfactory for our purposes, we construct a new dataset using minor civil divisions (MCDs). MCDs have been used to report population in parts of the US, especially in the Northeast, since the first census in 1790 (see Census 2000c). But as we discuss below, we are interested not only in population but also in sectoral employment. And since the earliest available digitized employment data for MCDs comes from the 1880 Census, we chose 1880 as the starting year for our analysis. Over time, MCDs became a standard tool for partitioning counties throughout (almost) the entire US.[14] It is this feature of MCDs that makes them so suitable for our analysis: they provide the finest level of geographical disaggregation for which we can analyze urbanization and structural transformation over more than a century.

The most common types of MCDs are towns and townships, but in some areas election precincts, magisterial districts, parishes, election districts, plantations, reservations, boroughs or other categories were used as MCDs. As some of these names suggest, in many states MCD boundaries coincide with those of local government bodies. In New England in particular, MCDs are actively functioning units of local government, in many cases since the 17^{th} Century. But in other states MCDs are often statistical entities with few (or no)

[14] In many Western states sub-county units were initially called MCDs but were reclassified as census county divisions (CCDs) in 1950, when the map of sub-county units in many of these states was redrawn. For simplicity, we refer to both MCDs and CCDs as MCDs (see chapter 8, Census 2000c) and discuss in further detail in the web-based technical appendix how we link MCDs over time.

other functions (see Chapter 8 Census 2000c). Given the variation in their functions, it is not surprising that the size and shape of MCDs also vary from state to state. For example, in the Midwest MCDs are often follow a chessboard patterns with squares of 6 miles per side; this design dates back to the Land Ordinance of 1785 and the Northwest Ordinance of 1787 (see Prescott 2003). As one travels West or South, the size of MCDs tends to grow, and they tend to become less regular and less stable over time. To address concerns that differences in the geographical and institutional organization of MCDs could affect population growth and employment structure, we report robustness checks where we consider states with similar geographical and institutional organizations of MCDs, and where we consider more aggregated spatial units such as counties.

To overcome changes in MCD boundaries over time, we aggregate some MCDs to create geographic units that are stable over time. This aggregation process involved considerable work using historical maps and gazetteers, and it is described in further detail in the web-based technical appendix. To give the reader a brief idea of the aggregation process, we matched the approximate centroid of each 1880 and 1940 MCD to the 2000 MCD in which it fell. We then aggregated any 2000 MCD that did not contain at least one 1880 MCD and one 1940 MCD to the nearest 2000 MCD that did. This aggregation process enables us to track the evolution of population at a fine level of spatial detail over 60-year intervals.[15] One reason for restricting ourselves to these years is that adding more years would have forced us to aggregate further. But perhaps more importantly, we only know the employment structure of MCDs for 1880 (using the individual-level census records from the North Atlantic Population Project) and for the very recent censuses, such as 2000 (using data from the US census American Factfinder tool, see Census 2000b). Since our analysis uses both population and employment data, adding more years for which we don't have employment data would have not contributed much. Finally, we used the 2000 census to calculate the land area in each geographic unit.

The extent of aggregation required to construct time-consistent units varies across states. In some states, especially in the Northeast and the Midwest, MCDs corresponded to local administrative units that were very stable over time, so little aggregation was required. We therefore divided states into samples: little aggregation was required in A states, more was

[15] While all MCDs in our baseline sample of "A and B" states have non-zero population in all three years of our sample, there are 7 MCDs in the C states that have zero population in 1880. These are dropped when we construct population growth rates.

needed in B states, and more still in C states. The geographic distribution of states across these three groups is shown in Map 1. In choosing our baseline sample, we sought to include as many states as possible while limiting the extent of aggregation, since the aggregation process might entail some imprecision. We therefore choose as our baseline sample the A and B states, for which $1-1$ matches between the 1880 and 2000 censuses involving no aggregation exceeded 70 percent.[16] But as we discuss below, we also construct alternative samples that either include more states (in some cases using county-level data) or restrict our sample to A states, where very little aggregation was required. In our baseline sample there are, on average, 13 units ("MCDs") per county. The average unit spans 115km^2, with a population of 2,400 in 1880 and 8,800 in 2000.

2.2 Empirical Strategy

We are interested in characterizing the population density distribution and the relationship between population growth and the initial population density distribution. In both cases, we adopt a nonparametric approach that imposes minimal structure on the data.

To characterize the population density distribution, we divide the range of values for log population density, x, into discrete bins of equal size δ. We index MCDs by m and bins by $\mathsf{b} \in \{1, ..., \mathsf{B}\}$. Denoting the set of MCDs with log population density in bin b by Φ_b and denoting the number of MCDs with log population densities within this set by n_b, we estimate the population density distribution, $\hat{\mathsf{g}}\left(\mathsf{x}_m\right)$, as follows:

$$\hat{\mathsf{g}}\left(\mathsf{x}_m\right) = \frac{\mathsf{n}_b}{\mathsf{n}}, \qquad \mathsf{n} = \sum_{b=1}^{B} \mathsf{n}_b, \qquad \text{for } \mathsf{x}_m \in \Phi_b. \tag{1}$$

Thus the estimated probability of observing a population density within the range of values included in bin b equals the fraction of MCDs with population densities in this range. This corresponds to a simple histogram, which yields a consistent estimate of the true underlying probability density function (Scott 1979). We choose bin sizes of $\delta = 0.1$ log points, which provide a fine discretization of the space of values for log population density, while preserving a relatively large number of MCDs within bins. Although this approach provides a simple and flexible characterization of the population density distribution, which connects closely with the other components of our analysis below, we also find similar results using related non-parametric approaches such as kernel density estimation (Silverman 1986).

[16] As in most cases our geographic units consist of a single MCD, we refer for simplicity to these units as "MCDs", even though they are sometimes aggregations of MCDs.

To characterize the relationship between population growth and the initial population density distribution, we follow a similar approach. We approximate the continuous function relating population growth to initial population density using a discrete-step function consisting of mean population growth within each initial population density bin:

$$\mathsf{y}_{mt} = \mathsf{f}\left(\mathsf{x}_{mt-T}\right) = \sum_{b=1}^{B} \varphi_b \mathsf{I}_b, \qquad \varphi_b = \frac{1}{\mathsf{n}_b} \sum_{m \in \Phi_b} \mathsf{y}_{mt}, \qquad \text{for } \mathsf{x}_m \in \Phi_b. \tag{2}$$

where m again indexes MCDs, b again indexes bins and t indexes time. In this specification, bins are defined over initial population density, x_{mt-T}; y_{mt} is average population growth from time $\mathsf{t} - \mathsf{T}$ to t; and I_b is an indicator variable equal to one if $\mathsf{x}_{mt-T} \in \Phi_b$ and zero otherwise.

This specification corresponds to a regression of population growth on a full set of fixed effects for initial population density bins. We report both mean population growth and the 95 percent confidence intervals around mean population growth for each initial population density bin. The confidence intervals are based on heteroscedasticity robust standard errors adjusted for clustering by county, and hence allow for correlated errors across MCDs within counties.[17] While this non-parametric specification allows for a flexible relationship between population growth and initial population density, we again find similar results using other related non-parametric approaches, such as locally weighted linear least squares regression (Cleveland 1979) and kernel regression (Härdle 1990). A key advantage of the specifications in (1) and (2) is that we can preserve the same discrete bins when analyzing the population density distribution, the relationship between population growth and the initial population density distribution, and the variables for the agricultural and non-agricultural sector discussed below.

As our model yields predictions for the functional form of the relationship between population growth and initial population density, we also estimate parametric versions of specification (2) of the form:

$$\mathsf{y}_{mt} = \rho \mathsf{x}_{mt-T} + \mathsf{u}_{mt}, \tag{3}$$

where ρ is a parameter to be estimated, u_{mt} is a stochastic error, and we again report standard errors clustered by county.

[17]When displaying the results of the specification (2) graphically, we remove the top and bottom one percent of the observations from the graphical representation, but not from the regressions. The bins at these extremes of the distribution contain few observations and have correspondingly large standard errors. Hence they tend to cloud rather than to illuminate the true picture.

Finally, our model highlights a relationship between population dynamics and employment dynamics in the agricultural and non-agricultural sectors. Therefore, in addition to the specifications for population in (1)-(3), we also estimate related specifications for employment in the agricultural and non-agricultural sectors.

2.3 Stylized Facts

To better understand the process of urbanization and structural transformation in the US from 1880-2000, we present a set of 6 stylized facts. These facts highlight the instability of the spatial distribution of economic activity over this time period − a pattern of results which lies in stark contrast to the stability documented within the sample of cities in the literature on urban growth. These facts also suggest that this instability is closely related to the different growth dynamics of the agricultural and non-agricultural sectors, and to structural transformation away from agriculture.

We begin by reporting a number of descriptive statistics for our baseline sample of "A and B" states in Column (1) of Table 1. Figures 1-6 then display the results of the non-parametric specifications (1) and (2) for population and for employment in the agricultural and non-agricultural sectors. Our first stylized fact is that the distribution of log population density across MCDs has become more dispersed from 1880-2000. As shown in Panel A of Column (1) in Table 1, the standard deviation of the distribution of log population density increased over this period from 0.97 to 1.56, which is both statistically significant and larger than the increase in mean log population density. Figure 1 confirms this increase in dispersion by displaying the results from specification (1). Although the US population increased substantially between 1880 and 2000, as reflected in Figure 1 in an increased mass of densely-populated areas, the figure also shows an increased mass of sparsely-populated areas. The population density distribution therefore exhibits polarization, with some low-density areas depopulating as other higher-density areas experience rapid population growth. This instability of the overall distribution of population stands in sharp contrast to the stability of the distribution of city sizes (e.g. Duranton 2007). Existing research for cities finds that the population size distribution is approximated by a Pareto distribution in the upper tail (e.g. Gabaix 1999) or a lognormal distribution for a wider range of city sizes (Eeckhout 2004).[18]

Second, Gibrat's law that population growth and population size are uncorrelated is

[18] The null hypothesis that the 1880 and 2000 distributions are drawn from the same Pareto or lognormal distribution is strongly rejected.

clearly violated. In Figure 2, we display the results from our population growth specification (2), where the dark solid line denotes mean population growth within each initial population density bin and the lighter dashed lines denote the 95 percent confidence intervals. As shown in the figure, log population density in 1880 is strongly predictive of population growth from 1880-2000. A similar relationship is found if we replace initial population density with initial population, as discussed further below. As Figure 2 shows, for low population densities, there is a negative correlation between population density in 1880 and subsequent population growth. But above the threshold of a log population density of about 2,[19] population density in 1880 is positively correlated with subsequent population growth.[20] The magnitudes of these differences are large: MCDs with log density of about 0 or 4 in 1880 experienced population growth at a rate of about 1 percent from 1880-2000. By contrast, MCDs with a log population density around 2 barely grew on average. As shown in Panel B of Column (1) in Table 1, these differences are statistically significant. We also note that at levels of log population density above 4 population density seems to be largely uncorrelated with population growth; this is the range that typically includes urbanized areas. Hence this finding is broadly consistent with the literature that finds Gibrat's law is a reasonable approximation for city population growth.[21] And yet for most of the population density distribution, and in the range that includes most of the 1880 population, we see a strong positive correlation between initial population density and subsequent growth.

Third, the share of agriculture in employment drops steeply in the range where population density in 1880 and subsequent growth are positively correlated. Figure 3 presents the results from specification (2) using the share of agriculture in employment in 1880 as the left-hand side variable rather than population growth. As shown in the figure, the agricultural employment share in 1880 drops from about 0.8 for MCDs with log density of 2 to about 0.2 for MCDs with log density of 4. Panel C of Column (1) in Table 1 shows that this difference is statistically significant. For denser MCDs the share continues to decline, but at a much

[19] Population densities in logs (levels) compare as follows: 2 (7), 4 (55) and 6 (403), where these figures are expressed as the log number (number) of people per square kilometer.

[20] While classical measurement error in 1880 population could induce a negative correlation between population growth and 1880 population density, this does not account for the positive correlation between these variables observed above a log population density of around 2, and our use of individual-level records from Census data mitigates measurement error concerns.

[21] While Gibrat's law has been confirmed in a number of studies for several countries (e.g. Gabaix 1999, Eeckout 2004, Eaton and Eckstein 1997, Ioannides and Overman 2004), other studies find some evidence of violations of Gibrat's Law even for urban population samples (e.g. Black and Henderson 2003, González-Val et al. 2008, Soo 2007).

slower rate.[22]

Fourth, the distribution of employment per square kilometer across MCDs has a lower standard deviation in agriculture than in non-agriculture in both 1880 and 2000. As shown in Panel D of Column (1) in Table 1, this difference is statistically significant at conventional critical values. Figure 4 presents the results from specification (1) for employment in agriculture and non-agriculture in 1880 and 2000. As shown in the figure, the employment density distribution in agriculture has thinner tails than its non-agricultural counterpart.[23] Therefore there are more observations with extreme low and high values of employment density for non-agriculture than for agriculture, reflecting the greater spatial concentration of non-agricultural employment. Furthermore, a comparison of Figures 1 and 4 suggests that 1880 population was distributed in a similar way to 1880 agricultural employment, while 2000 population is more spatially concentrated and distributed in a similar way to 2000 non-agricultural employment. This reflects the decline in agriculture's share of employment, which fell from 12 percent to 0.5 percent for the US as a whole and from 63 percent to 6 percent for the average MCD, with the difference between these two sets of figures reflecting the much greater spatial concentration of employment in non-agriculture than in agriculture discussed above.

Fifth, agricultural employment growth appears to follow a mean-reverting process. To document this stylized fact, we consider the subsample of MCDs for which agriculture accounted for more than 80 percent of 1880 employment. Although the share of agricultural employment in this subsample was over 88 percent in 1880, it fell to below 10 percent in 2000, and hence this subsample does not entirely capture agricultural dynamics alone. Nevertheless, since this subsample was at least initially mostly agricultural, it is likely to capture the main features of agricultural growth.[24] Figure 5 displays the results from non-parametric specification (2) for this subsample using agricultural employment growth as the left-hand

[22] The share of employment in total population was about 0.33 in 1880 and 0.48 in 2000. In both years, it was relatively stable across the population density distribution, suggesting that labor force participation is not strongly related to population density and hence that employment dynamics are a reasonable predictor of population dynamics.

[23] We also find that non-agricultural employment per square kilometer is more unequally distributed than agricultural employment in both 1880 and 2000 using standard measures of inequality such as the Gini Coefficient, the Theil Index, the difference between the 90th and 10th percentiles, and the difference between the 99th and 1st percentiles.

[24] We also find mean reverting processes when we consider population growth (rather than employment growth) for both 1880-2000 and 1880-1940 for the same agricultural subsample. During the 1880-1940 period, agriculture remained an important employer in much of the US at both the beginning and end of the period.

side variable. As apparent from the figure, densely-populated MCDs in this subsample exhibited much slower growth of agricultural employment from 1880-2000 than sparsely populated MCDs. Panel E of Column (1) in Table 1 reports the results from parametric specification (3) for this subsample, again using agricultural employment growth as the left-hand side variable. This confirms our finding of mean reversion: the coefficient on log population density in 1880 in the parametric specification is -0.006 and significant (p-value < 0.001). From the size of this coefficient, each additional log point of population density in 1880 is associated on average with just over half a percentage point lower rate of agricultural employment growth. We find very similar results if we instead relate agricultural employment growth to log agricultural employment density in 1880: the coefficient on initial log agricultural employment density is -0.006 and statistically significant.

Sixth, in contrast to the results for the agricultural sector, non-agricultural employment growth is uncorrelated with 1880 population density. To demonstrate this, we consider the subsample of MCDs for which agriculture accounted for less than 20 percent of 1880 employment. In this subsample the share of non-agricultural employment was higher than 90 percent in 1880 and higher than 98 percent in 2000. Figure 6 displays the results from non-parametric specification (2) using non-agricultural employment growth as the left-hand side variable, while Panel F of Column (1) in Table 1 reports the results from the analogous parametric specification (3). As apparent from the figure, non-agricultural employment grew at about 1.2 percent per year. This positive growth rate is very different from the (mostly) negative growth rates of agricultural employment shown in Figure 5. Moreover, in sharp contrast to the results for the agricultural sector, non-agricultural employment growth is uncorrelated with 1880 population density. As reported in Panel F of Column (1) in Table 1, the coefficient on log population density in 1880 in the parametric specification is -0.0002 and statistically insignificant (p-value $= 0.515$). We also find very similar results if we instead relate non-agricultural employment growth to log non-agricultural employment density in 1880. The coefficient on log non-agricultural employment density is -0.00021, which is more than an order of magnitude smaller than the corresponding coefficient for the agricultural sector, and statistically insignificant.

2.4 Robustness of the Stylized Facts

Having documented the 6 stylized facts for our preferred sample of MCDs, we now examine their robustness to different samples and specifications. The results of these robustness checks are summarized in Columns (2) to (8) of Table 1, while Figure 7 replicates the non-parametric population growth specification (2) displayed in Figure 2 for each of the robustness checks.

One potential concern about our preferred sample is that imperfect matching of MCDs across censuses could have affected our estimates. For example, some of the population and employment of MCDs with intermediate densities could have been assigned to MCDs with either higher or lower densities. To address this concern, the second column of Table 1 shows that all of our stylized facts remain intact when we restrict the sample to MCDs in the "A states" (to which we also refer as the restricted sample). In this restricted sample match rates are well over 90 percent, so imperfect matching is unlikely to be the cause of our finding. Figure 7 also shows non-parametrically that the U-shape we document in the second stylized fact is still strongly apparent in this sample.

Another possible concern is that we use a level of aggregation that is too fine. For example, people could live in one MCD and commute to work in another MCD, which could in turn influence the correlation between population growth and population density. As a first step to address this concern, we replicate our analysis using county-level data, since fewer people commute across county boundaries than across MCD boundaries. In the third column of Table 1, we report results using county-level data for 45 states and Washington DC.[25] As this robustness check includes a more comprehensive set of states than our baseline "A and B" sample, it ensures that our findings are not being driven by the particular geographic distribution of states in the baseline "A and B" sample. To provide a comparison, the fourth column restricts the county sample to the baseline "A and B" states. And the fifth column reports results using a hybrid sample of MCDs for states where matching was possible and counties for other states.

Our results are robust across all three specifications with two exceptions. The first stylized fact does not hold in Column (3), where the standard deviation of log population in 1880 is higher than in 2000, though the difference is not statistically significant at conventional critical values. The sixth stylized fact does not hold in Columns (3) and (5), where

[25] As noted in footnote 5, we exclude Alaska, Hawaii, Oklahoma, North Dakota and South Dakota, which had not attained statehood in 1880, and therefore are either not included in the 1880 census or did not have stable county boundaries at that time.

we find some evidence of mean reversion in both agriculture and non-agriculture, though the estimated rate of mean reversion in non-agriculture is substantially lower than that in agriculture. These exceptions are perhaps not surprising because the samples in Columns (3) and (5) include Western states that were not yet fully settled in 1880. Early settlement dynamics in these states, around the time of the "Closing of the frontier" (identified in the 1890 Census), are likely to be quite different from those elsewhere. As the Western states include areas that were largely uninhabited in 1880, they have correspondingly high standard deviations of log population in 1880, accounting for the exception to stylized fact 3.[26] Relatedly, the future settlement of areas that were largely uninhabited in 1880 provides a natural explanation for mean reversion that is unrelated to employment structure, consistent with the exception to stylized fact 6. Despite these caveats, the remainder of the stylized facts hold in these specifications, and in areas that were well-settled by 1880 all of our results are robust to aggregating MCDs up to the county level.[27]

While county-level data are consistent with our results, a further concern is that the aggregation they provide is insufficient around large cities. Some large cities (Metropolitan Statistical Areas (MSAs)) span multiple counties and may be characterized by commuting across county boundaries. Additionally, the suburbanization that occurred in the second half of the twentieth century can extend beyond county boundaries and could have influenced population dynamics in the neighborhood of large cities. To address these concerns, we undertake further aggregation. One possibility would be to aggregate counties based on 20th-century definitions of MSAs, but these definitions are themselves endogenous to population growth during our sample period. Therefore we instead aggregate MCDs based on 1880 characteristics using a flexible approach that allows us to consider various levels of aggregation. Starting with our baseline sample, we identify as "cities" MCDs that had a

[26] Consistent with this, we find that the higher standard deviation of log population in 1880 than in 2000 is driven by a tail of very sparsely populated counties in 1880. Indeed, the interquartile range of the population distribution is greater in 2000 than in 1880, so that stylized fact 3 is confirmed using measures of dispersion that are less sensitive to the tails of the distribution.

[27] To further test whether our results are affected by the US's Westward expansion, we restricted our baseline "A and B" sample to states that were part of the original 13 colonies. All the stylized facts are robust to this restriction, except part of stylized fact 3 (the downward slope of the u-shape). We do not find that population growth for log density 0 is significantly larger than for log density 2. But this finding is not surprising, since only two MCDs fall in the category of log population density 0 in this restricted sample. When we further restrict our sample to A states within the 13 colonies (New York and New England, except Maine), the remaining stylized facts all hold, except that we find no significant mean reversion in the agricultural subsample (stylized fact 5). But this is probably again due to small sample size. There are only 78 observations (in 48 counties) in the agricultural subsample for A states that were part of the original colonies (out of 4439 observations for this sample), reflecting the relatively urban character of these states.

log population per square kilometer larger than 6. To each of these cities we add the land area, population, and employment of any MCD whose geographic centroid lies within 25 kilometers of each city.[28] We label the resulting sample a suburban sample, since it pools together large urban centers with their surrounding areas. As shown in Column (6) of Table 1 and Panel D of Figure 7, all of our stylized facts hold in this suburban sample. We also experimented with other levels of aggregation, including defining "cities" as MCDs with 50,000 or 100,000 or more inhabitants in 1880 and using a distance threshold of 50 kilometers, and again found a similar pattern of results.

As a further robustness check, we examined whether the upward-sloping relationship between population growth and initial population density observed in Figure 2 for log densities in between 2 and 4 is robust to restricting the sample to MCDs with an above median distance to one of our "cities." Re-estimating our non-parametric specification (2) for this subsample, in which the distance to a "city" is greater than 170 kilometers, we continue to find a strong upward-sloping and highly statistically significant relationship between population growth and initial population density for log densities in between 2 and 4. Taking these results together, commuting and suburbanization in the neighborhood of large cities do not appear to be driving the pattern of violations of Gibrat's Law observed in our data.[29]

Although we examine the relationship between population growth and initial population density to control for variation in land area across MCDs, existing research concentrates on the relationship between population growth and initial population size. Therefore, while initial population density and size are strongly and approximately log linearly related in our data, another concern is that the violation of Gibrat's Law is driven by the use of initial population density rather than initial population size. To address this concern, Column (7) of Table 1 and Panel E of Figure 7 display results using log initial population size. Given that log population is measured in different units from log population density, we do not expect the inflection point at which the population growth relationship switches from being downward-sloping to upward-sloping relationship to occur at the same numerical values, and therefore the statistical tests based on values of 0, 2 and 4 in Table 1 do not apply to this specification and are not reported. Nonetheless, we observe the same qualitative pattern, and each of our stylized facts holds if we use initial log population size instead of initial log

[28]When two or more cities and their surrounding areas overlapped, we merged them together.

[29]While suburbanization is primarily associated with the use of the automobile as a form of mass transit, we also note that we find a very similar pattern of empirical results for the period 1880-1940, prior to the large-scale dissemination of the automobile after the end of the Second World War.

population density.

Another alternative hypothesis is that the observed relationship between population growth and initial population density could be influenced by omitted locational fundamentals, such as institutions and natural endowments. While institutions and endowments are captured in the model developed below in so far as they influence location-specific productivities in the agricultural and non-agricultural sector, the empirical concern is that locational fundamentals have a direct effect on population growth and are correlated with initial population density. To explain our results, these locational fundamentals would need to have a non-linear relationship with population growth and initial population density, to have the same non-linear relationship with the share of agricultural employment and initial population density, and to have differential effects on the correlation between employment growth and initial employment in the agricultural and non-agricultural sectors.[30]

To provide evidence that such a direct effect of locational fundamentals is not driving our results, we first regress each of our left-hand side variables (population growth, the share of agriculture in employment, and employment growth in agriculture and non-agriculture) on state fixed effects (to control for state policies and institutions) and on measures of proximity to natural endowments (rivers, lakes and coastlines, and mineral endowments). We next take the residuals from these regressions and implement our tests for Gibrat's Law (stylized fact 2), the share of agriculture in employment (stylized fact 3) and the relationship between employment growth and initial employment in agriculture and non-agriculture (stylized facts 5 and 6). As shown in Column (8) of Table 1 and Panel F of Figure 7, these four stylized facts are robust to controlling for locational fundamentals.[31]

Finally, the population of urban locations can grow through a number of channels, including migration from rural areas, international migration or differences in fertility. While the model developed below abstracts from international migration and fertility, it could be extended to include them, and the assumption of population mobility implies that agents are indifferent across locations. As a result of this indifference condition, the populations

[30] As a first robustness check to address the concern about institutional differences, we also re-estimated our baseline specification for the subset of the A states that were part of the original 13 colonies. Within this subset of the A states, MCDs are towns and townships with similar administrative functions. Once again, we find a similar pattern of results, as discussed in footnote 27 above.

[31] As the relationship between population and locational fundamentals can change over time, and as the relationship between employment and location fundamentals can differ between the agricultural and non-agricultural sectors, we do report standard deviations for log population and employment after controlling for locational fundamentals (stylized facts 1 and 4).

of all locations are linked together in the model. Although the United States has relatively high levels of population mobility, the presence of barriers to mobility could in principle break this link between locations' populations, with the result that local variation in international migration, fertility and mortality could directly affect local population. As a final robustness check, we therefore include a number of controls for initial demography, including international migration, fertility, education and race, using the same methodology as for Column (8) above. While these controls are likely to be themselves endogenous to employment structure, and are therefore not included in our baseline specification, we continue to find a similar pattern of results when they are included.[32]

Taken together, the evidence presented in this section shows that our stylized facts are robust characteristics of the US growth experience in the 20th Century. But are they also relevant for more recent experiences of structural transformation in other countries? To shed more light on this issue, we next examine urbanization and structural transformation in Brazil.

3 Brazilian Data and Stylized Facts

3.1 Data and Samples

The most populous country in the Western Hemisphere after the US is Brazil. Like the US, Brazil is divided into states. And just as US states are divided into counties, Brazilian states are divided into municipalities. Since municipality boundaries have changed over time, the Instituto de Pesquisa Econômica Aplicada (IPEA) has created "áreas mínimas comparáveis" (AMCs), geographic units that are much more stable over time. The 5,507 municipalities that existed in 1997 were pooled into 3,659 AMCs, which allow us to consistently analyze data from 1970-2000.[33] Although we could analyze Brazilian data before 1970, this would entail considerable further aggregation of municipalities, which would make it harder to distinguish urban from rural areas. Therefore we choose 1970 as the starting point for our

[32] All of the stylized facts are robust to the joint inclusion of the following four demographic control variables: the initial share of the population that is white, the share of the population aged 14-18 in education (as a measure of human capital), the share of the population that was born outside the United States (as a measure of international migration), and the share of the population aged less than six (as a measure of fertility).

[33] New municipalities were created after 2000, but the 1997 municipalities were used in the 2000 Census, the latest Census that we analyze in this paper. For further discussion of the Brazilian data and the samples discussed below, see the web-based technical appendix.

analysis. It is worth noting that agriculture's share in employment in the average AMC declined from 71 percent to 43 percent from 1970-2000, and its share in overall employment fell from 46 percent to 20 percent. Therefore the period we analyze involved considerable structural transformation.

The average Brazilian AMC spans 2,323km^2, with a population of 25,817 in 1970 and 46,421 in 2000. While AMCs are on average larger than the units that we analyze in our US sample, the difference is due in part to the fact that the interior regions of Brazil have larger and more sparsely populated AMCs. Therefore, while our baseline sample uses all of Brazil, we also demonstrate the robustness of our results to using a restricted sample that includes the Northeast, Southeast and South regions in Brazil only. In these areas, the average AMC spans 923km^2, and had a population of 26,013 in 1970 and 44,125 in 2000. These units are still substantially bigger than in the US sample, suggesting that it might be harder to separate urban from rural areas in Brazil.

3.2 Stylized Facts

Having described Brazilian AMCs, we now examine whether their population dynamics are characterized (at least qualitatively) by the same stylized facts as for US MCDs. Panel A in Figure 8 and Table 2 shows that the standard deviation of log population density across Brazilian AMCs increased from 1970-2000, confirming our first stylized fact. Additionally, Panel B in the same Figure and Table shows that low density areas and high density areas grew faster than areas of intermediate density. Therefore the U-shaped relationship between population growth and initial population density, characterized in stylized fact 2, also holds for Brazil. One quantitative difference between Brazil and the US is, however, that the increasing segment of this U-shape is not 2-4 (as in the US), but rather 4-6. This difference partly reflects differences in the relative distribution of agricultural and non-agricultural employment in Brazil and the US, as evident in Figures 4 and 8 (Panel D).

Furthermore, Panel C in Figure 8 and Table 2 shows that the increasing segment of the U-shaped population growth relationship is located in the same range of initial population densities where a sharp decline in agriculture's share of employment is observed, as in the US (stylized fact 3). This provides further corroborating evidence that the U-shape is indeed related to employment structure. Panel D in Figure 8 and Table 2 also confirms that agricultural employment has a lower standard deviation than non-agricultural employment (stylized fact 4). Finally, the last two stylized facts - that agricultural employment is mean reverting and non-agricultural employment is uncorrelated with initial density, are also confirmed for Brazil, as shown most clearly in the final two panels of Table 2 and also in Figure 8.[34]

In summary, we find a striking similarity in the relationship between population growth and employment structure in Brazil and the United States. This similarity of the results in two quite different contexts and time periods suggests that our results are unlikely to be driven by idiosyncratic features of the data or institutional environment for an individual country but rather capture more systematic features of the relationship between urbanization and structural transformation.

4 The Model

In this section we outline a simple theoretical model that generates the main stylized features of population growth found in our empirical work and shows how they can be explained by the process of structural transformation from agriculture to non-agriculture.[35] The model is a natural extension of existing research on the distribution of population across space (Eeckhout 2004) to incorporate a distinction between agriculture and non-agriculture. Productivity differences in these two sectors across locations determine both employment structure and the distribution of population. While productivity differences and agglomeration forces in non-agriculture provide forces for the concentration of population and employment, residential and commercial land use provide dispersion forces. Over time, structural transformation due to more rapid productivity growth in agriculture than in non-agriculture reallocates employment between sectors and population across locations.

4.1 Endowments, Preferences and Technology

The economy consists of a fixed number of locations $\mathsf{i} \in \{1, ..., \mathsf{I}\}$, which are grouped in our data into larger statistical units called MCDs. Each location is endowed with a quantity of land H_i, which can be used residentially or commercially. Land allocated to commercial use

[34]For Brazil, to ensure a sufficient sample size, we construct the non-agricultural subsample using AMCs that have an agricultural employment share in 1970 of less than less than 0.4 (instead of less than 0.2 for the US). Nonetheless, if we also use a threshold of less than 0.2 for Brazil, we continue to find no statistically significant relationship between non-agricultural employment growth and initial population density.

[35]A more detailed exposition of the model is contained in a web-based technical appendix available on request.

in each location can be employed in either agricultural or non-agricultural production, but cannot be simultaneously employed in both. Therefore each location specializes completely in either the agricultural or the non-agricultural good.[36] Furthermore, as the model abstracts from labor force participation, employment in a location's sector of specialization equals its population.[37] The economy as a whole is endowed with S_t workers, who are mobile across locations, and are each endowed with one unit of labor that is supplied inelastically with zero disutility.

Each worker has the same Cobb-Douglas preferences and allocates a constant share of expenditure (α) to a consumption index of tradeable goods and the remaining share $(1-\alpha)$ to the consumption of residential land.[38] The tradeable goods consumption index (C_{it}) is defined over consumption of agriculture (c_{Ait}) and non-agriculture (c_{Nit}) and is assumed to take the constant elasticity of substitution (CES) form:

$$C_{it} = \left[\psi_{At} c_{Ait}^{\rho} + \psi_{Nt} c_{Nit}^{\rho}\right]^{1/\rho}, \qquad 0 < \kappa = \frac{1}{1-\rho} < 1, \quad \psi_{At}, \psi_{Nt} > 0, \tag{4}$$

where ψ_{At} and ψ_{Nt} are preference parameters that capture the relative strength of consumer preferences for the agricultural and non-agricultural goods. Consistent with empirical evidence and a large literature in macroeconomics, we assume that agricultural and non-agricultural consumption are complements, so that the elasticity of substitution between the two goods (κ) is strictly less than one.[39]

The non-agricultural and agricultural goods are produced under conditions of perfect competition and are costlessly tradeable across locations. Output in each sector (Y_{jit}) depends on labor input (L_{jit}), land input (H_{jit}), a productivity parameter (θ_{jit}) and a local externality in the size of the sector ($S_{jit}^{\eta_j}$):

$$Y_{jit} = S_{it}^{\eta_j} \theta_{jit} L_{jit}^{\mu_j} H_{jit}^{1-\mu_j}, \qquad 0 < \mu_j < 1, \ 0 \leq \eta_j < 1. \tag{5}$$

where $j \in \{A, N\}$ indexes agriculture (A) and non-agriculture (N). While we allow for positive externalities in non-agriculture ($0 < \eta_N < 1$), we assume for simplicity that there

[36] The assumption that locations are completely specialized in agriculture or non-agriculture simplifies the characterization of the model's dynamics. MCDs are in general incompletely specialized, as they can contain both agricultural and non-agricultural locations.

[37] The model's abstraction from labor force participation is motivated by the empirical finding noted above that labor force participation is not strongly related to population density in our data.

[38] For empirical evidence using US data in support of the constant housing expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magne (2008).

[39] The assumption of an elasticity of substitution between agriculture and non-agriculture of less than one is consistent with empirical findings of larger changes over time in nominal consumption shares than in real consumption shares (see for example Kravis et al. 1983).

are no externalities in agriculture ($\eta_A = 0$), although all we require is that externalities in agriculture are less strong than those in non-agriculture, which is consistent with the much greater spatial concentration of employment in non-agriculture discussed above.

Productivity in each sector is assumed to have a secular component (Γ_{jt}), which is common across locations but changes over time, and an idiosyncratic component (σ_{jit}), which varies across locations and over time:

$$\theta_{jit} = \Gamma_{jt} \left(1 + \sigma_{jit}\right) \theta_{jit-1}^{\nu_j}, \qquad 0 < \nu_j \leq 1, \tag{6}$$

where ν_j captures the degree of mean reversion in productivity over time, and the idiosyncratic component of productivity is assumed to be independently and identically distributed with mean zero, and bounded support satisfying $1 + \sigma_{jit} > 0$.

4.2 Equilibrium Land Use and Population

After observing the vector of agricultural and non-agricultural productivity shocks in each period, each worker chooses location, consumption of the agricultural good, consumption of the non-agricultural good, and residential land use to maximize their utility taking the population distribution as given. Since relocation is costless, the worker's optimization problem reduces to choosing these variables to maximize their instantaneous flow of utility. The distribution of population across locations is therefore determined by the requirement that real wages are equalized across all locations populated in equilibrium.

With perfectly competitive goods and factor markets, labor and land are paid their value marginal product. Equilibrium commercial land use in each location is determined by whichever sector offers the higher value marginal product for land. In general, the equilibrium rental rate for land varies across locations and is determined by the requirement that residential and commercial land use sum to the location's endowment of land. As each worker allocates a constant share of expenditure to tradeable goods consumption and residential land use, and as the production technology in each sector is also Cobb-Douglas, a constant equilibrium fraction of land in each location is allocated to residential and commercial use, with this fraction depending on which good is produced.

Combining real wage equalization and equilibrium land use, the equilibrium population density in each location can be determined as a function of its productivity in its sector of specialization and its land endowment:

$$\frac{S_{jit}}{H_i} = \Lambda_{jt}^{\xi_j} \theta_{jit}^{\xi_j} H_i^{\eta_j \xi_j}, \qquad \xi_j \equiv \frac{1}{(1 - \mu_j) + \frac{1-\alpha}{\alpha} - \eta_j} > 0, \tag{7}$$

where Λ_{jt} is constant across locations specialized in the same good j at a given point in time t and is defined in the web-based technical appendix.

Combining equilibrium population density (7) and productivity dynamics (6), we obtain the following relationship between population growth and initial population density for locations that remain specialized in the same sector over time:

$$\ln\left(\frac{S_{jit}}{S_{jit-1}}\right) = \vartheta_{jt} + \xi_j \ln(1 + \sigma_{jit}) - (1 - \nu_j) \ln\left(\frac{S_{jit-1}}{H_i}\right), \tag{8}$$

where ϑ_{jt} is constant across locations that specialize in the same good j in both t and $t-1$ and is defined in the web-based technical appendix.

Therefore, for locations that remain specialized in the same sector over time, the correlation between population growth and initial population density in the model depends on the extent of mean reversion in productivity shocks over time. As we find empirically that non-agricultural employment growth is largely uncorrelated with initial population density, we assume $\nu_N = 1$, which implies constant proportional growth in non-agricultural productivity and the population of non-agricultural locations (Gibrat's Law). Similarly, as we find empirically that agricultural employment growth is negatively correlated with population density, we assume $0 < \nu_A < 1$, which implies mean reversion in agricultural productivity and the population of agricultural locations. This assumption of mean reversion in agricultural productivity but not in non-agricultural productivity is not only consistent with observed patterns of employment growth in our data, but is also consistent with the idea that the relative productivity of locations is more strongly influenced by locational fundamentals, such as soil and climate, in agriculture than in non-agriculture. While idiosyncratic shocks to agricultural productivity occur, the re-assertion of these locational fundamentals over time could be responsible for greater mean reversion in productivity in agriculture than in non-agriculture.[40]

With inelastic demand between the two tradeable consumption goods in (4), more rapid technological progress in the agricultural sector than in the non-agricultural sectors leads

[40] For empirical evidence of stronger mean reversion in agricultural productivity than in non-agricultural productivity, see for example Martin and Mitra (2001).

to a more than proportionate fall in the relative price of the agricultural good and a reallocation of employment from agriculture to non-agriculture over time.[41] As this change in employment structure proceeds, population is reallocated away from locations with relatively high productivity in agriculture towards locations with relatively high productivity in non-agriculture. Furthermore, the more than proportionate fall in the relative price of the agricultural good reduces the value marginal product of land in agriculture relative to that in non-agriculture, which results in endogenous switches in land use from agriculture to non-agriculture. These endogenous switches in land use are in general associated with a violation of Gibrat's Law, as discussed further in the web-based technical appendix.

4.3 Structural Transformation and the Six Stylized Facts

The model provides a natural explanation for the six stylized facts and explains how they are related. This explanation is based on differences in productivity dynamics in agriculture and non-agriculture and structural transformation away from agriculture. On the one hand, constant proportional growth in non-agricultural productivity ($\nu_N = 1$) generates the lack of correlation between non-agricultural employment growth and initial population density (stylized fact 6). On the other hand, mean reversion in agricultural productivity ($0 < \nu_A < 1$) gives rise to the decreasing relationship between agricultural employment growth and initial population density (stylized fact 5).

Since productivity growth in non-agriculture exhibits constant proportional growth, non-agricultural productivity is unbounded from above. In contrast, as agricultural productivity growth exhibits mean reversion, agricultural productivity is bounded from above. Together these properties of productivity growth in the two sectors provide a natural explanation for a higher standard deviation of employment in non-agriculture than in agriculture (stylized fact 4) and a share of employment in agriculture that is negatively correlated with population density at high densities (stylized fact 3).

The combination of mean reversion in agriculture, an agricultural employment share that is decreasing in population density at high densities and more rapid employment growth in

[41] See the web-based technical appendix for further discussion. While there is substantial empirical evidence of more rapid technological progress in agriculture than in agriculture (see again Martin and Mitra 2001) and inelastic demand between these broad categories of goods (see for example the discussion in Ngai and Pissarides 2007), structural transformation away from the agricultural sector could be also generated by labor-augmenting technological change and complementarity between labor and land in agriculture. Similarly, common technological progress in both sectors combined with non-homothetic preferences can also generate structural transformation, as discussed further in the web-based technical appendix.

non-agriculture than in agriculture explains the U-shaped relationship between population growth and initial population density (stylized fact 2). Finally, the upward-sloping segment of this U-shaped relationship accounts for the increased dispersion of population (stylized fact 1), since along the upward-slope more densely-populated locations experience more rapid population growth than less-densely populated locations.

5 Counterfactuals

In this section, we use relationships from the model to provide evidence that structural transformation can account not only qualitatively but also quantitatively for the relationship between population growth and population density observed in our data. The quantitative analysis builds on four key components of the model. First, as MCDs comprise multiple locations that specialize completely in either agriculture or non-agriculture, MCD population growth can be written as a weighted average of employment growth in agriculture and non-agriculture.[42] Second, the share of agricultural employment in the population is negatively correlated with population density at high densities. Third, the relationship between employment growth and population density differs between agriculture and non-agriculture. Fourth, the relationship between employment growth and initial population density depends on whether a location continues to specialize in the same sector in both time periods or whether it endogenously switches between sectors.

To illustrate the explanatory power of each of these components, we generate a sequence of counterfactual predictions for MCD population growth, each of which uses progressively more components. We next compare the predicted relationship between population growth and initial population density from each of these counterfactuals to the actual relationship observed in the data. We undertake this comparison in two ways. First, we estimate our non-parametric specification (2) and display the results for predicted and actual population growth graphically in Figure 9. Second, to provide further evidence on the predictive power of the model, we regress actual population growth on predicted population growth and include a number of control variables. We first undertake the analysis using our US data before examining whether the model can also quantitatively account for our results using the Brazilian data. For brevity, we concentrate on results for the US data with our baseline

[42] Consistent with the model's abstraction from labor force participation, we predict population growth using employment data, and compare the results to observed population growth.

sample of "A and B" states. However, we find a qualitatively similar pattern with the other samples, as expected from the robustness checks above, and as discussed further below.

As a first step, Counterfactual 1 uses the property that MCD population growth is a weighted average of employment growth in agriculture and non-agriculture and makes the following assumptions: (a) a common rate of employment growth within each sector across all MCDs, (b) the same share of agricultural employment in the population across all MCDs, and (c) no switching between agriculture and non-agriculture. To measure employment growth in locations that produce the agricultural good in both periods, we use average agricultural employment growth in the agricultural sample from Table 1. Similarly, to measure employment growth in locations that produce the non-agricultural good in both periods, we use average non-agricultural employment growth in the non-agricultural sample from Table 1.[43] As Counterfactual 1 assumes the same employment growth rates within each sector and the same agricultural employment share for all MCDs, it predicts the same rate of population growth for all MCDs, as shown in Figure 9.[44]

Counterfactual 2 is the same as Counterfactual 1, except that it allows the agricultural employment share to vary across MCDs by using the 1880 value of this variable for each MCD. Therefore, in this second counterfactual, the cross-section variation in population growth is predicted solely from the cross-section variation in the initial agricultural employment shares combined with common values of average employment growth within each sector for all MCDs. As evident from Figure 9, the employment share of an MCD in agriculture and non-agriculture in 1880 goes a good way towards explaining its population growth from 1880-2000, providing strong evidence for the importance of structural transformation in shaping observed population dynamics.

Counterfactual 3 is the same as Counterfa agriculture with initial population density. The results of the regressions for agriculture and non-agriculture are reported in Columns (1) and (2) of Table 3. As shown in Figure 9, enriching the model in this way makes the downward-sloping relationship between population growth and initial population density observed at low densities more pronounced.[45]

In Counterfactuals 1-3, we measured the common value of employment growth within each sector using employment growth in the most and least agricultural MCDs, which contain locations least likely to switch between sectors. In contrast, Counterfactual 4 takes into account the possibility of switching between sectors by allowing for a more flexible relationship between population growth and initial patterns of specialization in agriculture and non-agriculture.

Specifically, in Counterfactual 4, we regress total employment growth in each MCD on the 1880 agricultural employment share, the 1880 log population density, and the interaction term between these two variables. The inclusion of the initial agricultural employment share captures the role of structural transformation in shaping population growth, while the inclusion of initial log population density allows for the possibility of mean reversion in non-agriculture, and the inclusion of the interaction term between the two variables captures the extent to which mean reversion in agriculture differs from that in non-agriculture.

As column (3) of Table 4 shows, the agricultural employment share in 1880 is negatively correlated with subsequent population growth, reflecting structural transformation away from agriculture. Additionally, from the negative coefficient on the interaction term, the share of agriculture in 1880 employment has an even more negative effect on subsequent population growth in areas that were initially denser, reflecting mean reversion in agriculture. After controlling for these two terms, 1880 log population density is not significant, consistent with an absence of mean reversion in non-agriculture. We therefore use the coefficients from Column (4), which excludes initial log population density, to construct Counterfactual 4 shown in Figure 9.

As apparent from the figure, actual population growth rates are substantially more variable than predicted population growth rates and the actual data exhibit a sharper change

[45] As a robustness check, we also augmented the non-agricultural employment growth regression with initial population density, which although not shown in Figure 9 had no visible effect, since from Table 1 employment growth is largely uncorrelated with initial population density in non-agriculture. Finally, we experimented with allowing for richer forms of scale dependence within each sector by introducing polynomials in initial population density, which also had little effect on the relationship between predicted population growth and initial population density.

in slope than the predicted values from the counterfactual. Nonetheless, predicted population growth closely replicates the observed pattern of violations of Gibrat's Law: the downward sloping relationship between population growth and initial population density at low densities, the upward sloping relationship at intermediate densities, and the largely flat relationship at high densities. The mean reversion in population growth rates at low initial population densities evident in Counterfactual 3 is further enhanced in Counterfactual 4, consistent with the idea that some of the mean reversion is the result of switches from agriculture to non-agriculture. Additionally, mean predicted population growth for Counterfactual 4 is closer to mean actual population growth, because some of the higher employment growth in non-agriculture is associated with these switches in land use, which are allowed for in Counterfactual 4.

To provide further evidence on the ability of structural transformation to explain observed patterns of population growth, and to compare the predictions of the counterfactuals against alternatives, Table 4 reports regressions of actual against predicted population growth from the model using our preferred Counterfactual 4. Whereas the non-parametric estimates that were displayed in Figure 9 are means for population size bins, the regressions exploit variation across individual MCDs. To provide a benchmark, we begin in Column (1) by regressing actual population growth rates on a constant. In Column (2), we augment that regression with the predicted population growth rates from the counterfactuals. Clearly there are many idiosyncratic factors affecting the population growth of individual MCDs that are not captured by our model, which results in a much larger variance of actual than of predicted population growth rates, as reflected in the regression R^2. Nonetheless, the coefficient on predicted population growth is positive, highly statistically significant and statistically indistinguishable from one.[46] Therefore, despite the much greater variance in statistically significant. Columns (6) to (8) show that the same pattern of results holds for the more restrictive sample of A states, the county sample and the suburban sample.

As the regressions in Columns (1) through (8) of Table 4 are estimated across MCDs, they exploit in part variation across the initial population density bins used in our non-parametric specification (2) as displayed in Figure 9. As a final step, we examine whether our estimates can explain variation in population growth not only across but also within population density bins. In Column (9) of Table 4, we therefore augment the baseline specification from Column (2) with a full set of fixed effects for initial population density bins. Even focusing solely of variation within initial population density bins, we continue to find a positive coefficient on predicted population growth that is large in magnitude and statistically significant. Column (10) of Table 4 shows that we continue to find the same pattern of results if we further augment this specification with our measures of proximity to natural endowments, and county fixed effects.

As an robustness check, the remainder of this section shows that we also find a very similar pattern of results for Brazil. The four counterfactuals are constructed in the same way for Brazil as for the United States.[47] The employment growth regressions used in these counterfactuals for Brazil are reported in Table 5 (analogous to Table 3 for the US). Having constructed the four counterfactuals, Figure 10 displays the results of estimating our non-parametric specification (2) for Brazil using actual and predicted population growth. As for the US, controlling for the initial agricultural employment share has considerable predictive power for patterns of population growth (Counterfactual 2). Controlling for mean reversion in agriculture generates the downward-sloping relationship between population growth and initial population density at low densities (Counterfactual 3). Finally, a more flexible relationship between population growth and initial patterns of specialization to allow for switches from agriculture to non-agriculture again enhances the explanatory power of the model (Counterfactual 4).

Following the same structure as for the US, Table 6 reports the results of regressions of actual against predicted population growth from the model using our preferred Counterfactual 4. While actual population growth again has a much higher variance than predictede□acb8(resso tically significant and statistically indistinguishable from one.[48] Therefore we again find a close correspondence between actual and predicted population growth. While Columns (1)-(4) include all AMCs, we find a similar pattern of results in Column (5), where we restrict attention to AMCs in the Northeast, Southeast and South of Brazil, which are smaller in geographic scope and are therefore likely to permit a finer discrimination between rural and urban areas. In Columns (6) and (7), we show that the model has explanatory power within as well as across population density bins by including a full set of fixed effects for population density bins.

Overall, there is considerable evidence that structural transformation can account for the quantitative as well as the qualitative patterns of observed population growth. Given the many differences between the United States and Brazil, and between the time periods considered, the similarity of the results in these two different contexts provides strong evidence in support of an explanation based on structural transformation.

6 Conclusion

While as recently as the nineteenth century less than one tenth of the world's population lived in cities, urban residents now account for a growing majority of the world's population. Arguably few other economic changes have involved as dramatic a transformation in the organization of society. In this paper, we present new evidence of six stylized facts about urbanization and develop a simple theoretical model that accounts for these facts both qualitatively and quantitatively.

Making use of a new source of sub-county data for the United States from 1880-2000, we find an unstable population distribution that exhibits polarization. This polarization reflects a population growth rate that is at first decreasing in initial population density at low densities, before increasing in initial population density at intermediate densities, and finally becoming largely uncorrelated with initial population density at high densities characteristic of urban areas.

Our model explains these systematic departures from Gibrat's Law of constant proportional growth in terms of differences in productivity dynamics between agriculture and non-agriculture. While agricultural productivity is mean reverting, non-agricultural productivity

[48] Again the standard errors are adjusted for predicted population growth being generated in a prior regression (Pagan 1984).

exhibits constant proportional growth. Over time, faster productivity growth in agriculture than in non-agriculture and inelastic demand between the two goods leads to structural transformation and a decline in the share of agriculture in employment.

At low population densities, where agricultural employment dominates, mean reversion in agriculture generates the observed decreasing relationship between population growth and initial population density. In contrast, at high population densities, where non-agricultural employment dominates, population growth is largely uncorrelated with initial population density. In between, faster employment growth in non-agriculture than in agriculture combined with a positive correlation between the non-agricultural employment share and population density leads to the observed increasing relationship between population growth and initial population density.

This pattern of empirical results is robust across a wide range of specifications and samples. Moreover, we find a strikingly similar pattern of results for Brazil from 1970-2000 as for the United States from 1880-2000. The ability of structural transformation to account for our findings in these two quite different contexts provides strong evidence in its support. Our results suggest that structural transformation is not only central to macroeconomic issues, such as growth and employment, but also has important microeconomic implications for the organization of economic activity within countries. As the reallocation of population from rural to urban areas has wide-ranging implications for public policy, urbanization is likely to remain a central policy issue as large developing countries such as Brazil, China and India experience rapid structural change.

References

Bairoch, P (1988) "Cities and Economic Development: From the Dawn of History to the Present", University of Chicago Press.

Baumol, William J. (1967) “Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crises,” American Economic Review, 57(3), 415-26.

Beeson, Patricia E., David N. DeJong and Werner Troesken (2001) “Population growth in U.S. counties, 1840–1990,” Regional Science and Urban Economics, 31(6), 669-699.

Black, Duncan and Vernon Henderson (1999) “A Theory of Urban Growth,” Journal of Political Economy, 107(2), 252-284.

Black, Duncan and Vernon Henderson (2003) “Urban Evolution in the USA,” Journal of Economic Geography, 3, 343-372.

Brazil Census (1970, 2000), Integrated Public Use Microdata Series — International (IPUMS International)

Caselli, Francesco and Wilbor John Coleman II (2001) “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation,” Journal of Political Economy, 109(3), 584-616.

Census (1940), US Census Bureau, Sixteenth US Census of Population and Housing, www.census.gov/prod/www/abs/decennial/1940.htm

Census (2000a), US Census Bureau, US Census of Population and Housing, Shapefile downloaded from
www.census.gov/main/www/cen2000.html

Census (2000b), US Census Bureau, American Factfinder, available at
factfinder.census.gov

Census (2000c), US Census Bureau, Bureau’s Geographic Areas Reference Manual, downloaded from
http://www.census.gov/geo/www/garm.html.

Census (2000d), US Census Bureau, "Census 2000 Basics", Chapter 4: Geographic Areas, http://www.census.gov/mso/www/c2000basics/chapter4.htm.

Cleveland, William S. (1979) “Robust Locally Weighted Regression and Smoothing Scatterplots,” Journal of the American Statistical Association, 74(368), 829-836.

Davis, Morris A. and François Ortalo-Magne (2008) “Household Expenditures, Wages, Rents,” University of Wisconsin-Madison, mimeograph.

Davis, Donald R. and David Weinstein (2002) "Bones, Bombs, and Break Points: The Geography of Economic Activity ", American Economic Review, 92(5), 1269-1289.

Desmet, Klaus and Esteban Rossi-Hansberg (2007) " Spatial Growth and Industry Age," NBER Working Paper, 13302.

Duranton, Gilles (2007) "Urban Evolutions: The Fast, the Slow, and the Still," American Economic Review, 97(1), 197-221.

Eaton, Jonathan and Zvi Eckstein (1997) "Cities and growth: Theory and evidence from France and Japan", Regional Science and Urban Economics, 27, 443-374

Echevarria, Cristina (1997) "Changes in Sectoral Composition Associated with Economic Growth," International Economic Review, 38(2), 431-452.

Eeckhout, Jan (2004) "Gibrat's Law for (All) Cities," American Economic Review, 94(5), 1429-1451.

Eeckhout, Jan (2008) "Gibrat's Law for (All) Cities: Reply," American Economic Review, forthcoming.

Ellison, Glenn and Edward L. Glaeser (1999) "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?" American Economic Review, 89(2), 311-316.

ESRI (1999), Arc View Database Access (Version 2.1a), Environmental Systems Research Institute, Inc.

Foster, Andrew D. and Mark R. Rosenzweig (2008) "Economic Development and the Decline of Agricultural Employment," Chapter 47 in (eds) T. Schultz and John Strauss, Handbook of Development Economics, Volume 4(5).

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables (1999) "The Spatial Economy: Cities, Regions and International Trade," Cambridge, MA: MIT Press.

Gabaix, Xavier (1999) "Zipf's Law for Cities: An Explanation," Quarterly Journal of Economics, 114(3), 739-767.

Gabaix, Xavier and Yannis Ioannides (2004) "The Evolution of City Size Distributions," Handbook of Regional and Urban Economics, Volume 4, in (eds) Vernon Henderson and Jacques Thisse, North-Holland.

Galor, Oded and David Weil (2000) "Population, Technology, and Growth: From the Malthusian Regime to the Demographic Transition and Beyond," American Economic Review, 90(4), 806-828.

Galor, Oded (2005) “From Stagnation to Growth: Unified Growth Theory,” in (eds) Philippe Aghion and Steven Durlauf, Handbook of Economic Growth, Volume 1A.

GIS (2003), Global GIS DVD-Global Coverage DVD 1st ed. U.S. Geological Data Series DDS-62 A-H Data Collections and Custom Programming the United States Geological Survey Published by the American Geological Institute, copyright 2003.

Goodfriend, Marvin and John McDermott (1995) “Early Development,” American Economic Review, 85(1), 116-133.

Glaeser, Edward L. (2008) “The Rise of the Sunbelt,” Southern Economic Journal, 74(3), 610-643.

Glaeser, Edward L., Hedi D. Kallal, Jose A. Scheinkman and Andrei Shleifer (1992) “Growth in Cities,” Journal of Political Economy, 100(6), 1126-1152.

Gollin, Douglas, Stephen Parente and Richard Rogerson (2002) “The Role of Agriculture in Development,” American Economic Review, 92(2), 160-164.

González-Val, Rafael, Luis Lanaspa and Fernando Sanz (2008) “New Evidence on Gibrat’s Law for Cities,” MPRA Working Paper, 10411.

Härdle, Wolfgang (1990) "Applied Nonparametric Regression", Econometric Society Monographs, Cambridge University Press: Cambridge.

Henderson, J. Vernon (1974) “The Sizes and Types of Cities,” American Economic Review, 64(4), 640-56.

Henderson, J. Vernon (2005) “Urbanization and Growth,” in (eds) Philippe Aghion and Steven Durlauf, Handbook of Economic Growth, North Holland.

Henderson, J. Vernon and Anthony J. Venables (2008) “The Dynamics of City Formation,” NBER Working Paper, 13769.

Henderson, J. Vernon and Hyoung Gun Wang (2007) “Urbanization and City Growth: the Role of Institutions,” Regional Science and Urban Economics, 37(3), 283-313.

Holmes, Thomas J. and Sanghoon Lee (2008) “Cities as Six-by-Six Mile Squares: Zipf’s Law,” University of Minnesota, mimeograph.

Ioannides, Yannis and Henry G. Overman (2004) “Spatial Evolution of the U.S. Urban System,” Journal of Economic Geography, 4(2), 131–156.

IPEA (2008) - Instituto de Pesquisa Econômica Aplicada, Data Section, available at http://www.ipeadata.gov.br

Kim, Sukkoo (1995) “Expansion of Markets and the Geographic Distribution of Economic

Activities: The trends in U.S.," Quarterly Journal of Economics, 110(4), 881-908.

Kim, Sukkoo (2000) "Urban Development in the United States, 1690-1990," Southern Economic Journal, 66(4), 855-880.

Kim, Sukkoo and Robert A. Margo (2004) "Historical Perspectives on U.S. Economic Geography" in (eds) Vernon Henderson and Jacques Thisse, Handbook of Regional and Urban Economics, Volume 4, North-Holland.

Kravis, Irving B., Alan W. Heston and Robert Summers (1983) "The Share of Services in Economic Growth," in F. Gerard Adams and Bert G. Hickman, eds., Global Econometrics: Essays in Honor of Lawrence R. Klein, Cambridge: MIT Press, 188-218.

Krugman, Paul (1991) "Increasing Returns and Economic Geography," Journal of Political Economy, 99(3), 483-499.

Kuznets, Simon (1955) "Economic Growth and Income Inequality," American Economic Review, 45, 1-28.

Levy, Moshe (2008) "Gibrat's Law for (All) Cities: A Comment," American Economic Review, forthcoming.

Martin, Will and Devashish Mitra (2001) "Productivity Growth and Convergence in Agriculture and Manufacturing," Economic Development and Cultural Change, 49(2), 403-422.

da Mata, Daniel, Uwe Deichmann, J. Vernon Henderson, Somik V. Lall and Hyoung G. Wang (2007) "Determinants of City Growth in Brazil," Journal of Urban Economics, 62(2), 252-272.

Matthews, Robert C., Charles H. Feinstein and John C. Odling-Smee (1982) "British Economic Growth 1856-1973", Stanford: Stanford University Press

Matsuyama, Kiminori (2002) "The Rise of Mass Consumption Societies," Journal of Political Economy, 110, 1035-1070.

Mckinsey (2008) "Preparing for China's Urban Billion", Mckinsey Global Institute: New York.

Ngai, Rachel and Chris Pissarides (2007) "Structural Change in a Multisector Model of Growth," American Economic Review, 97(1), 429-443.

NAPP 2006, North Atlantic Population Project and Minnesota Population Center. NAPP: Complete Count Microdata. NAPP Version 1.0 [computer files]. Minneapolis, MN: Minnesota Population Center [distributor], 2006. [www.nappdata.org]

Pagan, Adrian (1984) "Econometric Issues in the Analysis of Regressions with Generated

Regressors," International Economic Review, 25(1), 221-247.

Prescott, Samuel T. (2003) "Federal Land Management - Current Issues and Background", Nova Publishing

Rappaport, Jordan and Jeffrey Sachs (2003) "The United States as a Coastal Nation," Journal of Economic Growth, 8(1), 5-46.

Rogerson, Richard (2008) "Structural Transformation and the Deterioration of European Labor Market Outcomes", Journal of Political Economy, 116(2), 235-259.

Rosen, Kenneth T. and Mitchel Resnick (1980) "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy", Journal of Urban Economics, 8(2), 165-186.

Rosenthal, Stuart S. and William C. Strange (2004) "Evidence on the Nature and Sources of Agglomeration Economies," Handbook of Urban and Regional Economics, (eds) J. Vernon Henderson and Jacques-Francois Thisse.

Rossi-Hansberg, Esteban and Mark L. Wright (2007) "Urban Structure and Growth," Review of Economic Studies, 74(2), 597-624.

Scott, David W. (1979) "On Optimal and Data-based Histograms," Biometrika, 66(3), 605-610.

Silverman, B. W. (1986) "Density Estimation for Statistics and Data Analysis", Chapman and Hall: London.

Simon, Herbert A. (1955) "On a Class of Skew Distribution Functions," Biometrika, 42(3/4), 425–440.

Soo, Kwok Tong (2005) "Zipf's Law for Cities: A Cross-country Investigation", Regional Science and Urban Economics, 35(3), 239-263.

Soo, Kwok Tong (2007) "Zipf's Law and Urban Growth in Malaysia," Urban Studies, 44(1), 1-14.

Syrquin, Moshe (1988) "Patterns of Structural Change," Chapter 7 in (eds) H. Chenery and T. N. Srinivasan, Handbook of Development Economics, 1, 205-273.

United Nations (2005) "Word Urbanization Prospects - The 2005 Revision," published October 2006, accessible via www.un.org/esa/population/publications/WUP2005/2005wup.htm.

Williamson, Jeffrey G. (1965) "Regional Inequality and the Process of National Development: A Description of the Patterns," Economic Development and Cultural Change, 13(1), 3-45.

Table 1: US – Robustness of stylized facts

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Baseline: A and B states	Only A states	Counties, 45 states and DC[1]	Counties, A and B sample	Hybrid Sample, 45 states and DC[2]	Suburban A and B states[3]	Log pop, not log density	Baseline, geo controls[4]
Panel A	Standard deviation of log population density in 1880 (σ_1)	0.967	1.025	1.757	0.963	1.272	0.904	0.833	.
	Standard deviation of log population density in 2000 (σ_2)	1.556	1.631	1.450	1.303	1.687	1.436	1.475	.
	H_0: $\sigma_1 = \sigma_2$, vs. H_1: $\sigma_1 < \sigma_2$, p-value	<0.001	<0.001	1.000	<0.001	<0.001	<0.001	<0.001	.
	Stylized Fact 1: Distribution of log population density across geographic units became more dispersed from 1880-2000 (population became more concentrated)	Yes	Yes	No[5]	Yes	Yes	Yes	Yes	.[4]
Panel B	Mean population growth at log population density 0 ($\beta_g(0)$)	0.013	0.012	0.016	0.019	0.010	0.013	.	0.013
	Mean population growth at log population density 2 ($\beta_g(2)$)	0.001	-0.001	0.007	0.007	0.002	0.001	.	0.005
	Mean population growth at log population density 4 ($\beta_g(4)$)	0.009	0.010	0.014	0.014	0.011	0.008	.	0.011
	H_0: $\beta_g(0) = \beta_g(2)$, H_1: $\beta_g(0) > \beta_g(2)$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	.	<0.001
	H_0: $\beta_g(2) = \beta_g(4)$, H_1: $\beta_g(2) < \beta_g(4)$, p-value	<0.001	<0.001	0.001	0.011	<0.001	<0.001	.	<0.001
	Stylized Fact 2: U-shaped relationship between population growth from 1880-2000 and log population density in 1880	Yes	Yes	Yes	Yes	Yes	Yes	.[6]	Yes
Panel C	Percent of agricultural in total employment at log population density 2 ($\beta_{sa}(2)$)	0.767	0.762	0.691	0.618	0.738	0.769	.	0.743
	Percent of agricultural in total employment at log population density 4 ($\beta_{sa}(4)$)	0.227	0.189	0.195	0.185	0.228	0.221	.	0.235
	H_0: $\beta_{sa}(2) = \beta_{sa}(4)$, H_1: $\beta_{sa}(2) > \beta_{sa}(4)$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	.	<0.001
	Stylized Fact 3: Share of agriculture in employment falls in the range where population density distribution in 1880 is positively correlated with population growth 1880-2000	Yes	Yes	Yes	Yes	Yes	Yes	.[6]	Yes
Panel D	Standard deviation of agricultural employment in 1880 (σ_{1a})	0.820	0.722	1.677	0.810	1.084	0.820	0.820	.
	Standard deviation of non-agricultural employment in 1880 (σ_{1na})	1.520	1.631	1.784	1.272	1.779	1.440	1.520	.
	H_0: $\sigma_{1a} = \sigma_{1na}$, vs. H_1: $\sigma_{1a} < \sigma_{1na}$, p-value	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	<0.001	.
	Standard deviation of agricultural employment in 2000 (σ_{2a})	0.858	0.853	0.806	0.617	0.936	0.851	0.858	.
	Standard deviation of non-agricultural employment in 2000 (σ_{2na})	1.623	1.689	1.530	1.359	1.767	1.503	1.623	.
	H_0: $\sigma_{2a} = \sigma_{2na}$, vs. H_1: $\sigma_{2a} < \sigma_{2na}$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	.
	Stylized Fact 4: Standard deviation of non-agricultural employment is larger than standard deviation of agricultural employment in both years	Yes	Yes	Yes	Yes	Yes	Yes	Yes	.[4]
Panel E	Regress agricultural employment growth on log population density and intercept in subsample of units with agricultural employment share > 0.8 in 1880, report slope coefficient (β_a)	-0.0060	-0.0077	-0.0067	-0.0054	-0.0066	-0.0060	-0.0056	-0.0055
	H_0: $\beta_a = 0$, H_1: $\beta_a \neq 0$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Stylized Fact 5: Agricultural employment does not follow Gibrat's law (employment growth depends on population density)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel F	Regress non agricultural employment growth on log population density and intercept in subsample of units with non-agricultural employment share < 0.2 in 1880, report slope coefficient (β_{na})	-0.0002	-0.0006	-0.0016	-0.0006	-0.0010	-0.0001	-0.0002	-0.0005
	H_0: $\beta_{na} = 0$, H_1: $\beta_{na} \neq 0$, p-value	0.515	0.287	<0.001	0.096	<0.001	0.745	0.515	0.0991
	Stylized Fact 6: Non-agricultural employment follows Gibrat's law (employment growth does not depend on population density)	Yes	Yes	No[7]	Yes	No[7]	Yes	Yes	Yes

Note: This table reports robustness tests of the 6 stylized facts using US data. All the regressions and tests reported in the table use robust standard errors clustered by county.

[1] The county sample includes all US states except Alaska, Hawaii, North Dakota, Oklahoma, and South Dakota, which had not attained statehood in 1880 and did not have stable county boundaries at that time.

[2] The hybrid sample uses the smallest geographical units available for each state. We use MCDs for the states in samples A, B, and C, and counties elsewhere. This sample excludes Alaska, Hawaii, North Dakota, Oklahoma, and South Dakota, as explained in the footnote above.

[3] In the Suburban Sample we merge any MCD with more than 100,000 inhabitants in 1880 to all the MCDs whose centroids lie within 25 kilometers of its centroid.

[4] The geographic control variables are state fixed effects, an indicator for the presence of coal, and indicators for the unit bordering on the ocean and for its centroid being within 50 kilometers from a lake or a river. As these specifications include controls, we do not test stylized facts 1 and 4.

[5] Since this sample includes many states that were not fully settled in 1880, many near-empty areas increase the standard deviation of the population density distribution in that year. When we restrict the analysis to counties in states A and B only, the stylized fact does hold (see column 4). This is reassuring, since our model is concerned with long-run equilibria, which is likely to a better characterization of the longer-settled A and B states.

[6] In this sample we do not expect the turning point of the U and the fall of the agriculture share at coefficient 2, and hence do not report these coefficients. The figures qualitatively show that there is a U-shape whose minimum coincides with the drop in agricultural employment.

[7] Since this sample includes many states that were not fully settled in 1880, many near-empty areas increase the standard deviation of the population density distribution in that year. The future settlement of areas that were near empty in 1880 is also likely to cause mean reversion that is unrelated to employment structure.

Table 2: Brazil – Robustness of stylized facts

		(1) All of Brazil (AMCs)	(2) As (1) but with state fixed effects	(3) As (1) but with geo controls[1]	(4) As (2) but with geo controls	(5) Brazil sub-sample[2]	(6) As (5) but with state fixed effects	(7) As (5) but with geo controls	(8) As (6) but with geo controls
Panel A	Standard deviation of log population density in 1970 (σ_1)	1.222	.[3]	1.222	.[3]	1.009	.[3]	1.009	.[3]
	Standard deviation of log population density in 2000 (σ_2)	1.323	.	1.323	.	1.197	.	1.197	.
	H_0: $\sigma_1 = \sigma_2$, vs. H_1: $\sigma_1 < \sigma_2$, p-value	<0.001	.	<0.001	.	<0.001	.	<0.001	.
	Stylized Fact 1: Distribution of log population density across geographic units became more dispersed from 1970-2000 (population became more concentrated)	Yes	.	Yes	.	Yes	.	Yes	.
Panel B	Mean population growth at log population density 0 ($\beta_g(0)$)	0.0239	0.0239	0.0239	0.0239	0.0146	0.0146	0.0146	0.0146
	Mean population growth at log population density 4 ($\beta_g(4)$)	0.0079	0.0134	0.0116	0.0146	0.0079	0.0053	0.0090	0.0100
	Mean population growth at log population density 6 ($\beta_g(6)$)	0.0214	0.0271	0.0265	0.0305	0.0214	0.0190	0.0240	0.0258
	H_0: $\beta_g(0) = \beta_g(4)$, H_1: $\beta_g(0) > \beta_g(4)$, p-value	<0.001	0.015	0.002	0.016	<0.001	<0.001	0.001	0.006
	H_0: $\beta_g(4) = \beta_g(6)$, H_1: $\beta_g(4) < \beta_g(6)$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Stylized Fact 2: U-shaped relationship between population growth from 1970-2000 and log population density in 1970	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel C	Percent of agricultural in total employment in 1970 at log population density 4 ($\beta_{sa}(4)$)	0.6710	0.6710	0.6710	0.6710	0.6710	0.6710	0.6710	0.6710
	Percent of agricultural in total employment in 1970 at log population density 6 ($\beta_{sa}(6)$)	0.1677	0.1933	0.1459	0.1689	0.1677	0.1933	0.1447	0.1686
	H_0: $\beta_{sa}(4) = \beta_{sa}(6)$, H_1: $\beta_{sa}(4) > \beta_{sa}(6)$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Stylized Fact 3: Share of agriculture in employment falls in the range where population density distribution in 1970 is positively correlated with population growth 1970-2000	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel D	Standard deviation of agricultural employment in 1970 (σ_{1a})	0.8933	.[3]	0.8933	.[3]	0.8869	.[3]	0.8869	.[3]
	Standard deviation of non-agricultural employment in 1970 (σ_{1na})	1.4157	.	1.4157	.	1.4287	.	1.4287	.
	H_0: $\sigma_{1a} = \sigma_{1na}$, vs. H_1: $\sigma_{1a} < \sigma_{1na}$, p-value	<0.001	.	<0.001	.	<0.001	.	<0.001	.
	Standard deviation of agricultural employment in 2000 (σ_{2a})	1.0176	.	1.0176	.	0.9954	.	0.9954	.
	Standard deviation of non-agricultural employment in 2000 (σ_{2na})	1.3754	.	1.3754	.	1.3642	.	1.3642	.
	H_0: $\sigma_{2a} = \sigma_{2na}$, vs. H_1: $\sigma_{2a} < \sigma_{2na}$, p-value	<0.001	.	<0.001	.	<0.001	.	<0.001	.
	Stylized Fact 4: Standard deviation of non-agricultural employment is larger than standard deviation of agricultural employment in both years	Yes	.	Yes	.	Yes	.	Yes	.
Panel E	Regress agricultural employment growth on log population density and intercept in subsample of units with agricultural employment share > 0.8 in 1970, report slope coefficient (β_a)	-0.0038	-0.0036	-0.0022	-0.0037	-0.0042	-0.0031	-0.0028	-0.0031
	H_0: $\beta_a = 0$, H_1: $\beta_a \neq 0$, p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Stylized Fact 5: Agricultural employment does not follow Gibrat's law (employment growth depends on population density)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel F	Regress non agricultural employment growth on log population density and intercept in subsample of units with agricultural employment share < 0.4 in 1970, report slope coefficient (ββ0θεpp4.82es07.on l120.63c≤p<.on l7 3773-07.on l3427(3784.7202.T3 17.on l10.7(376on l6).T3 17.on l6747(3773-0).T3 17.on l7297(3773-								

Note: This table reports robustness tests of the 6 stylized facts using data on Brazilian municipalities (Áreas Mínimas Comparáveis (AMCs)). All the regressions and tests reported in the table use robust standard errors.

[1] The geographic controls are twelve dummy variables indicating the presence of oil, nickel, manganese, iron, gold, copper, cobalt, and aluminum, whether the AMC borders the ocean, lies within 50 kilometers of a river, has its centroid covered with tropical or subtropical moist broadleaf forest, or is contained in the Amazonas area.

[2] This subsample uses only AMCs in the states of the Northeast, Southeast, and South official regions of Brazil, since AMCs in these regions are relatively small, allowing a clearer distinction between rural and urban areas. The three regions in this subsample cover about 90 percent of Brazil's AMCs, 36 percent of its land area and 91 percent of its population in 1970.

[3] As these specifications include controls, we do not test stylized facts 1 and 4, which involve measuring standard deviations.

Table 3: US – Constructing the counterfactuals

Employment growth rate, 1880-2000	(1) For counterfactual 3 Non-agric.	(2) For counterfactual 3 Agric.	(3) For counterfactual 4 Total	(4) For counterfactual 4 Total
Constant	0.011 (0.001)	-0.005 (0.001)	0.014 (0.001)	0.014 (0.001)
Log population density in 1880		-0.006 (0.000)	-0.0002 (0.0003)	
Share of agriculture 1880			-0.008 (0.002)	-0.007 (0.001)
(Share of agriculture in 1880) x (log population density in 1880)			-0.0010 (0.0005)	-0.0013 (0.0004)
Number of Observations	755	3,074	10,856	10,856
R^2	0	0.31	0.063	0.063
Sample:	A and B, non-agric	A and B, agric	A and B	A and B

Note: This table reports the regressions used to construct counterfactuals 3 and 4 for the US data. We construct counterfactual 3 using the predicted values of sectoral employment growth from the regressions reported in columns (1) and (2), as described in the text of the paper. We construct counterfactual 4 using the predicted values of employment growth from the regression reported in column (4), as described in the text. The non-agricultural subsample used in column (1) includes MCDs from our baseline A and B Sample for which agriculture's share of 1880 employment was less than 0.2. The agricultural subsample used in column (2) includes MCDs from our baseline A and B Sample for which agriculture's share of 1880 employment exceeded 0.8. Robust standard errors in parentheses are clustered by county.

Table 4: US – Quantifying the predictive power of counterfactual 4

	(1) Inter-cept only	(2) As (1) but with predicted growth	(3) As (2) but with geo controls[1]	(4) As (3) with state fixed effects	(5) As (4) but with county fixed effects	(6) As (2) with A Sample only	(7) As (2) but with county sample	(8) As (2) but with suburban sample	(9) As (2) with log pop density bins[2]	(10) As (5) with log pop density bins[2]
Actual population growth regression										
Predicted population growth		1.041 (0.06)	0.798 (0.055)	0.629 (0.067)	0.72 (0.052)	1.221 (0.078)	1.011 (0.057)	1.057 (0.062)	0.648 (0.079)	0.468 (0.046)
Intercept	0.475 (0.034)	-0.026 (0.045)								
R^2	0	0.098	0.183	0.303	0.345	0.173	0.433	0.098	0.151	0.284
Number of observations	10,864	10,864	10,864	10,864	10,864	4,439	2,496	10,159	10,864	10,864
Regression used to generate predicted population growth										
Share of agriculture in 1880		-1.05 (0.017)	-1.039 (0.017)	-0.982 (0.016)	-0.871 (0.014)	-1.075 (0.021)	-0.428 (0.051)	-0.74 (0.016)	-1.217 (0.046)	-0.913 (0.023)
Share of agriculture 1880 x population density 1880		-0.162 (0.006)	-0.157 (0.006)	-0.147 (0.006)	-0.151 (0.005)	-0.171 (0.008)	-0.797 (0.019)	-0.245 (0.006)	-0.077 (0.017)	-0.119 (0.009)
F – statistic[3]		5243	4876	6657	7211	4548	1194	5886	3268	8061

Note: This table shows the predictive power of counterfactual 4 for various specifications using US data. The upper panel of the table reports the regressions of actual population growth on predicted population growth. The lower panel of the table reports the regression whose fitted values are used for predicted population growth. The left-hand side variable in the lower panel of the table is total employment growth. Robust standard errors clustered by county are in parentheses. The standard errors in the upper panel of the table have been adjusted for the fact that predicted population growth is generated using a prior regression (Pagan 1984).

[1] The geographic control variables are state fixed effects, an indicator for the presence of coal, and indicators for observations bordering on the ocean and for observations whose centroid lies within 50 kilometers of a lake or a river.

[2] The log population density bin fixed effects included in these regressions are a full set of dummy variables for MCDs having population densities within intervals of 0.1 log points. For example, all MCDs with log population density from 0.1 to 0.2 are grouped together in bin 0.1.

[3] The F-value reported is for an F-test that the coefficients on the share of agriculture and the interaction term are jointly equal to zero in the prior regression used to generate predicted population growth.

Table 5: Brazil – Constructing the counterfactuals

	(1)	(2)	(3)	(4)
	For counterfactual 3		For counterfactual 4	
Employment growth rate 1970-2000	Non-agric.	Agric.	Total	Total
Constant	0.039	0.00216	0.045	0.043
	(0.001)	(0.00111)	(0.004)	(0.001)
Log population density in 1970		-0.0038	-0.0005	
		(0.0004)	(0.0008)	
Share of agriculture 1970			-0.0317	-0.0291
			(0.0044)	(0.0016)
(Share of agriculture in 1970) x (log population density in 1970)			-0.0037	-0.0043
			(0.0009)	(0.0004)
Number of Observations	384	1,651	3,659	3,659
R^2	0	0.059	0.262	0.262
Sample:	AMCs non-agric.	AMCs agric.	AMCs	AMCs

Note: This table reports the regressions we used to construct counterfactuals 3 and 4 for the Brazilian municipalities (Áreas Mínimas Comparáveis (AMCs)) data. We construct counterfactual 3 using the predicted values of sectoral employment growth from the regressions reported in columns (1) and (2), as described in the text of the paper. We construct counterfactual 4 using the predicted values of employment growth from the regression reported in column (4), as described in the text of the paper. The non-agricultural subsample used in column (1) includes AMCs for which agriculture's share of 1970 employment was less than 0.4 due to the small sample size using a threshold of 0.2 (but results are similar using a 0.2 threshold). The agricultural subsample used in column (2) includes AMCs for which agriculture's share of 1970 employment exceeded 0.8. Robust standard errors are in parentheses.

Table 6: Brazil – Quantifying the predictive power of counterfactual 4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Intercept only	As (1) but with predicted growth	As (2) but with geo controls[1]	As (3) but with state fixed effects	As (4) but with subsample[4] only	As (2) but with log pop density bins[2]	As (4) but with log pop density bins
Actual population growth							
Predicted population growth		1.024	0.968	1.112	1.122	0.909	0.915
		(0.035)	(0.035)	(0.040)	(0.042)	(0.036)	(0.036)
Intercept	0.269	0.010					
	(0.009)	(0.010)					
R^2	0	0.196	0.315	0.378	0.350	0.287	0.385
Number of observations	3,659	3,659	3,659	3,659	3,659	3,659	3,659
Regression used to generate predicted population growth							
Share of agriculture in 1970		-0.810	-0.821	-0.755	-0.885	-0.693	-0.708
		(0.013)	(0.015)	(0.015)	(0.017)	(0.044)	(0.045)
Share of agriculture 1970 x population density 1970		-0.122	-0.125	-0.129	-0.073	-0.158	-0.159
		(0.003)	(0.003)	(0.004)	(0.005)	(0.012)	(0.012)
F – statistic[3]		5460	5651	4728	4088	4042	4212

Note: This table shows the predictive power of counterfactual 4 for various specifications using the Brazilian municipalities (Áreas Mínimas Comparáveis (AMCs)) data. The upper panel of the table reports the regression of actual population growth on predicted population growth. The lower panel of the table reports the regression whose fitted values are used for predicted population growth. The left-hand side variable in the lower panel of the table is total employment growth. Robust standard errors are in parentheses. The standard errors in the upper panel of the table have been adjusted for the fact that predicted population growth is generated using a prior regression (Pagan 1984).

[1] The geographic controls are twelve dummy variables indicating the presence of oil, nickel, manganese, iron, gold, copper, cobalt, and aluminum, whether the AMC borders the ocean, lies within 50 kilometers of a river, has its centroid covered with tropical or subtropical moist broadleaf forest, or is contained in the Amazonas area.

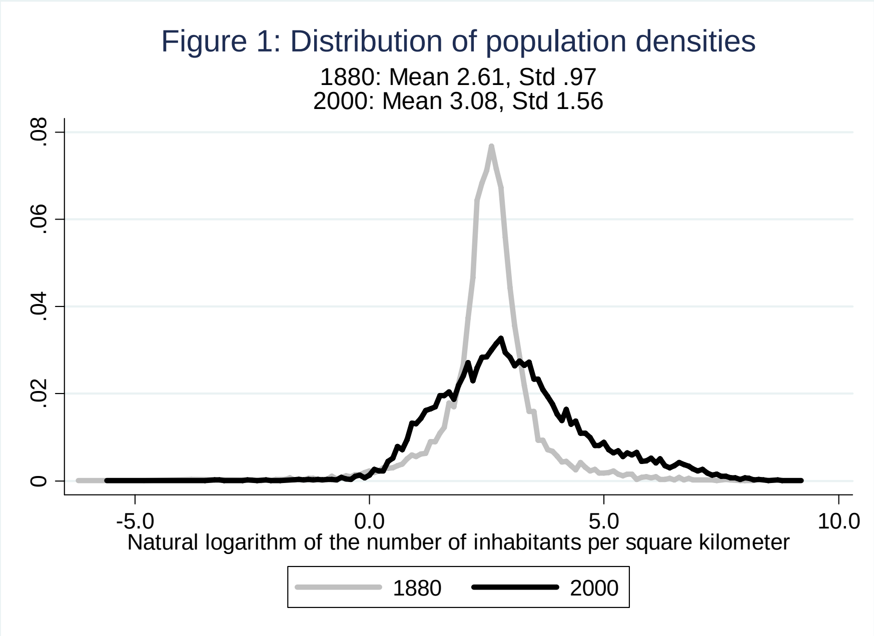
[2] The log population density bin fixed effects included in these regressions are a full set of dummy variables for MCDs having population densities within intervals of 0.1 log points. For example, all AMCs with log population density from 0.1 to 0.2 are grouped together in bin 0.1.

[3] The F-value reported is for an F-test that the coefficients on the share of agriculture and the interaction term are jointly equal to zero in the prior regression used to generate predicted population growth.

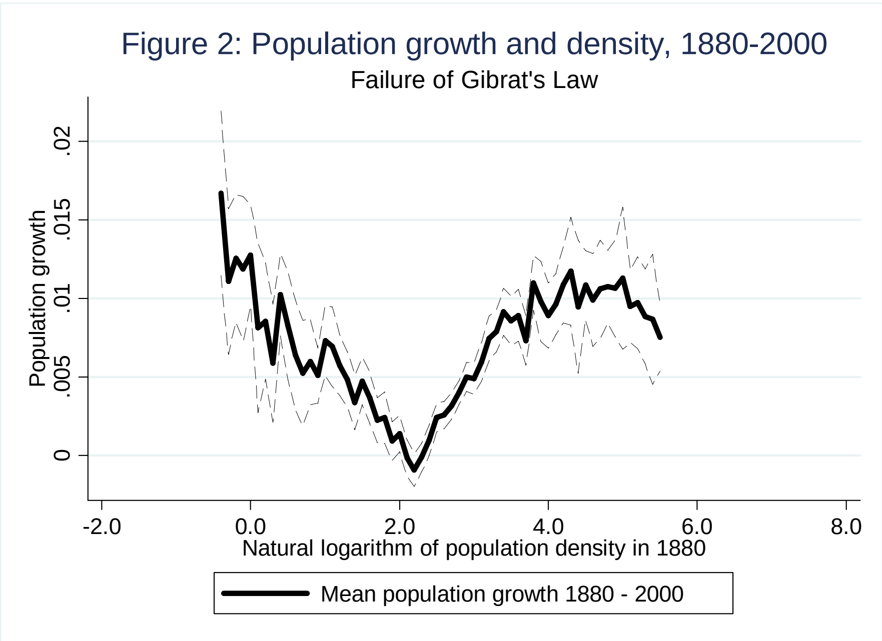
[4] This subsample uses only AMCs in the states of Northeast, Southeast, and South official regions of Brazil, since AMCs in these regions are relatively small, allowing a clearer distinction between rural and urban areas. The three regions in this subsample cover about 90 percent of Brazil's AMCs, 36 percent of its land area and 91 percent of its population in 1970.

Map 1: US MCD data by state and county

Note: This map shows the geographical distribution of states for the various samples. Our baseline sample consists of A and B states. The classification A, B and C corresponds to the quality of the match rate between 1880 and 2000 MCDs. In states classified as A (Connecticut, DC, Indiana, Iowa, Massachusetts, New Hampshire, New York, Rhode Island, Vermont), the 1-1 match rate between 1880 and 2000 MCDs is larger than 0.9. In states classified as B (Illinois, Maine, Maryland, Michigan, Missouri, North Carolina, Ohio), the match rate is larger than 0.7. In states classified as C (Arkansas, California, Delaware, Georgia, Kansas, Minnesota, Nebraska, New Jersey, Pennsylvania, South Carolina, Utah, Virginia, West Virginia, Wisconsin), 1880 MCD data are available but the match rate is lower than 0.7. For states in the counties sample (Alabama, Arizona, Colorado, Florida, Idaho, Kentucky, Louisiana, Mississippi, Montana, Nevada, New Mexico, Oregon, Tennessee Texas, Washington, Wyoming), 1880 MCD data are not available. We exclude Alaska, Hawaii, Oklahoma, North Dakota, and South Dakota, which had not attained statehood in 1880, and therefore are either not included in the 1880 census or did not have stable county boundaries at that time.

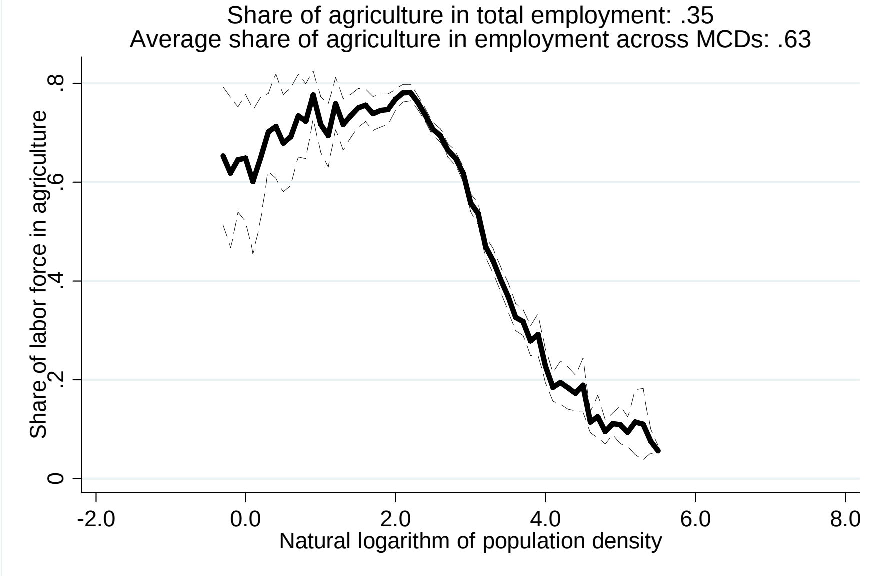


Note: This figure shows the distribution of log population per square kilometer in 1880 and 2000 estimated using non-parametric specification (1) for the sample of "A and B" states. Population density bins are defined by rounding down log population density for each MCD to the nearest single digit after the decimal point. For example, all MCDs with log population density from 0.1 to 0.2 are grouped together in bin 0.1. See the web-based technical appendix for further details on data.



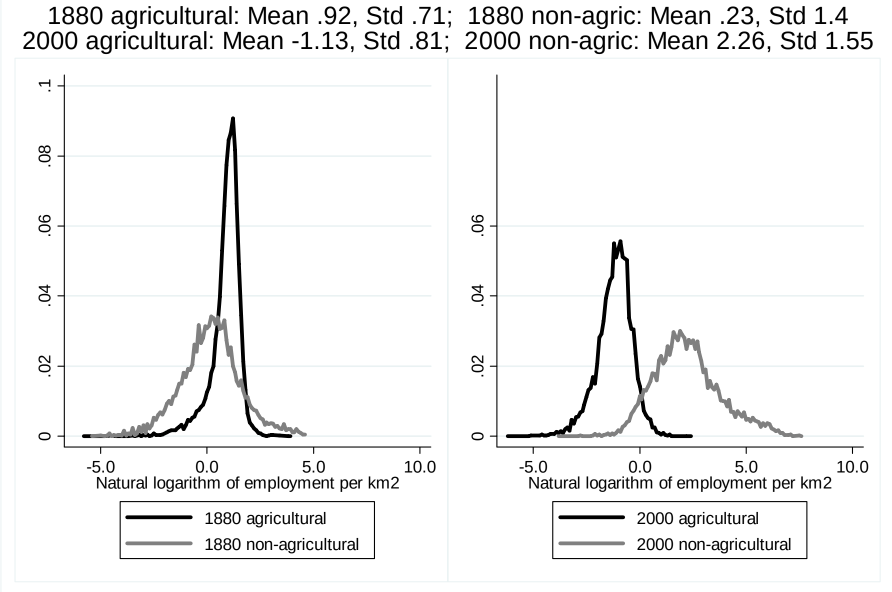
Note: The solid line shows mean population growth rate from 1880-2000 within each population density bin based on estimating non-parametric specification (2) for the sample of "A and B" states. Population density bins are defined by rounding down log population density for each MCD to the nearest single digit after the decimal point. The dashed lines show 95 percent confidence intervals based on robust standard errors clustered by county. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880. See the web-based technical appendix for further details on data.

Figure 3: Share of agriculture and population density in 1880



Note: The solid line shows the mean share of agriculture in 1880 employment within each population density bin based on estimating non-parametric specification (2) for the sample of "A and B" states. Population density bins are defined by rounding down log population density for each MCD to the nearest single digit after the decimal point. The dashed lines show 95 percent confidence intervals based on robust standard errors clustered by county. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880. See the web-based technical appendix for further details on data.

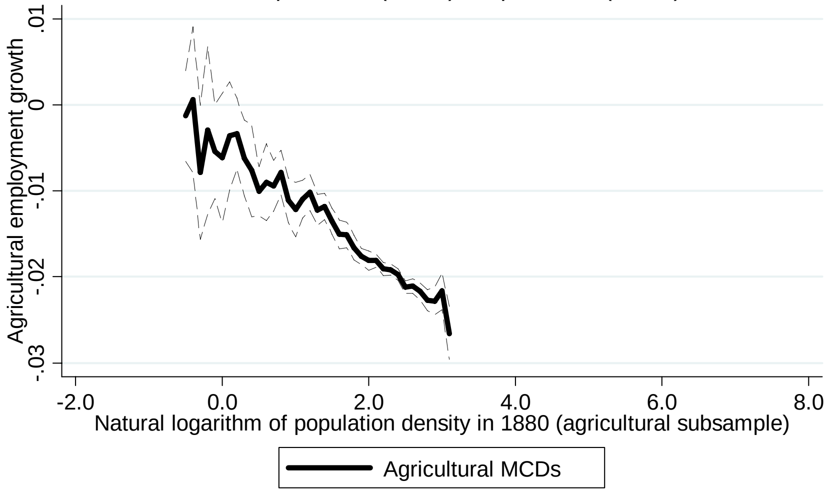
Figure 4: Density of employment



Note: This figure shows the distribution of log agricultural employment and log non-agricultural employment (employment in industry and services) per square kilometer in 1880 and 2000 estimated using non-parametric specification (1) for the sample of "A and B" states. Employment density bins are defined by rounding down log employment density for each MCD to the nearest single digit after the decimal point. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880. See the web-based technical appendix for further details on data.

Figure 5: Agricultural growth and population density

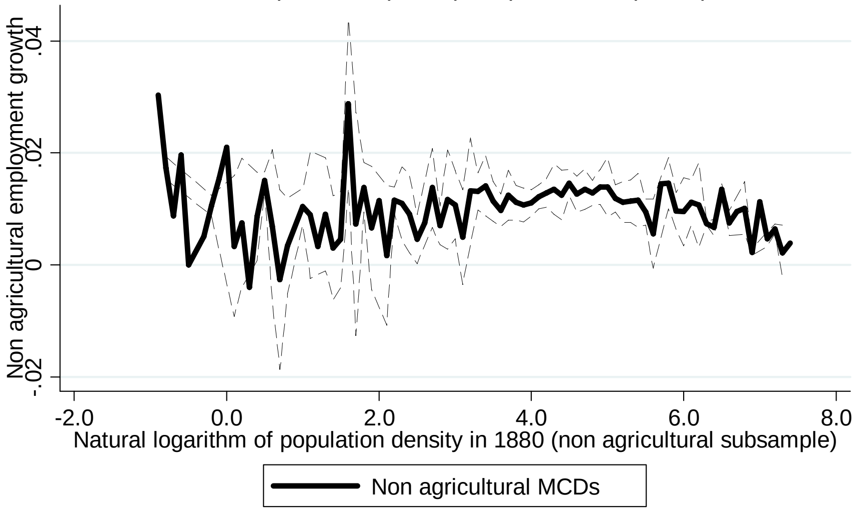
Regression of agricultural employment growth on log population density:
intercept -.0053 (.0009), slope: -.006 (.0004)



Note: The solid line shows the mean growth rate of agricultural employment from 1880-2000 within each population density bin based on estimating non-parametric specification (2) for the agricultural subsample (an agricultural share in 1880 employment of greater than 0.8) within "A and B" states. Population density bins are defined by rounding down log population density for each MCD to the nearest single digit after the decimal point. The dashed lines show 95 percent confidence intervals based on robust standard errors clustered by county. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880. See the web-based technical appendix for further details on data.

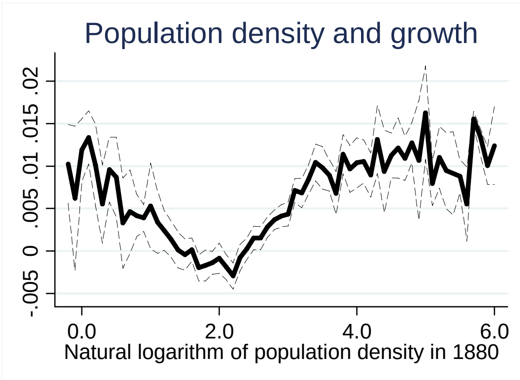
Figure 6: Non agricultural growth and population density

Regression of non agricultural employment growth on log population density:
intercept: .0123 (.0016), slope: -.0002 (.0003)

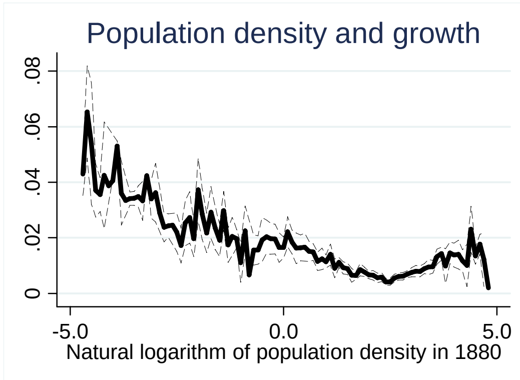


Note: The solid line shows the mean growth rate of non-agricultural employment (employment in industry and services) from 1880-2000 within each population density bin based on estimating non-parametric specification (2) for the non-agricultural subsample (an agricultural share in 1880 employment of less than 0.2) within "A and B" states. Population density bins are defined by rounding down log population density for each MCD to the nearest single digit after the decimal point. The dashed lines show 95 percent confidence intervals based on robust standard errors clustered by county. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880. See the web-based technical appendix for further details on data.

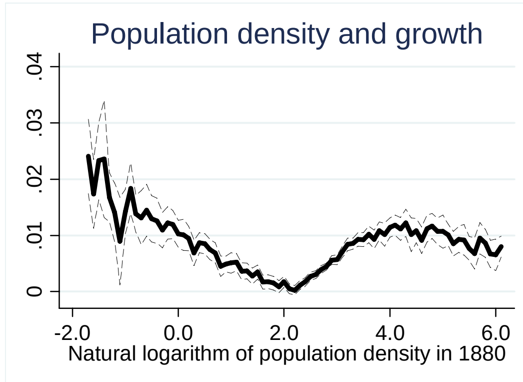
Figure 7: Robustness of failure of Gibrat's Law



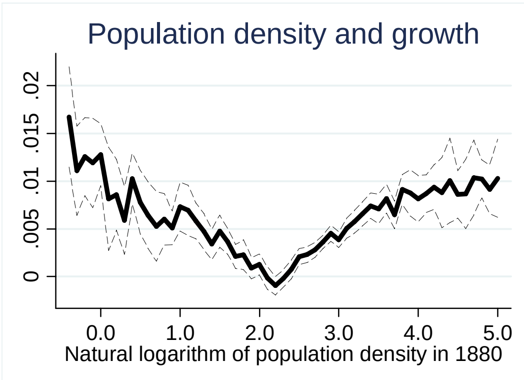
Panel A: A states sample



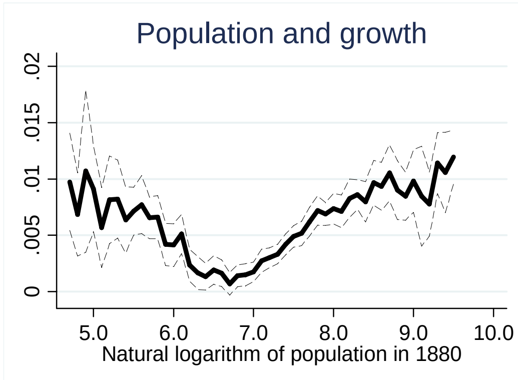
Panel B: Counties sample



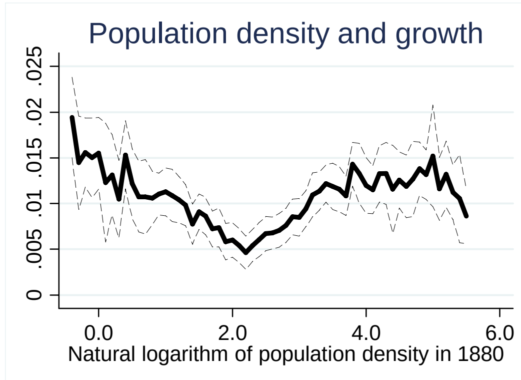
Panel C: Hybrid sample



Panel D: Suburban sample



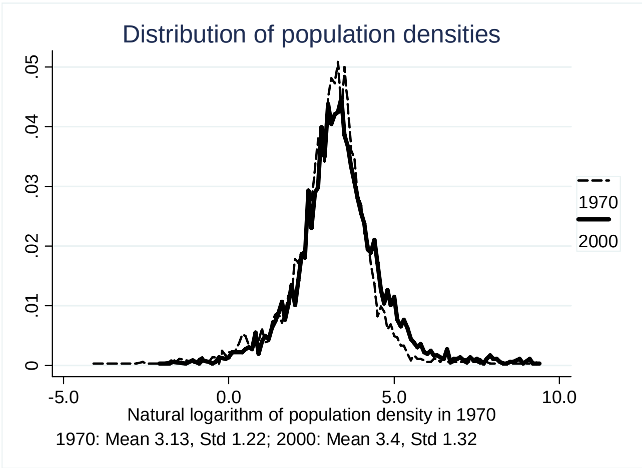
Panel E: A and B, population not densities



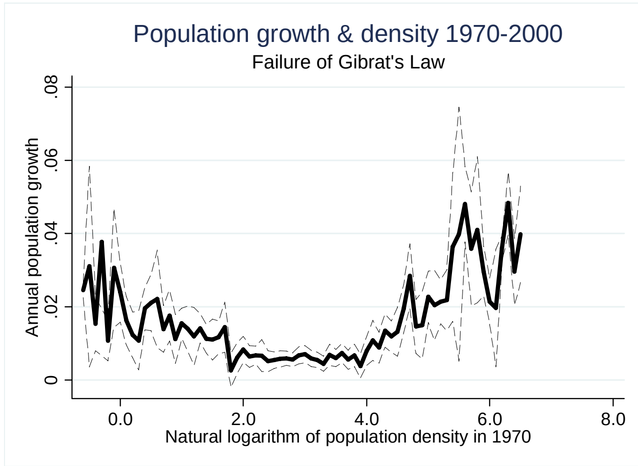
Panel F: A and B sample, state controls

Note: This figure shows the robustness of the failure of Gibrat's Law (Figure 2) by reproducing it for other samples. The various samples used here are described in the web-based technical appendix. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880.

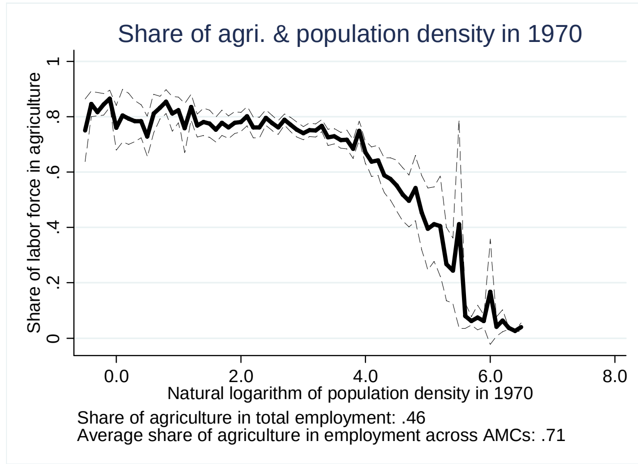
Figure 8: Brazilian Results



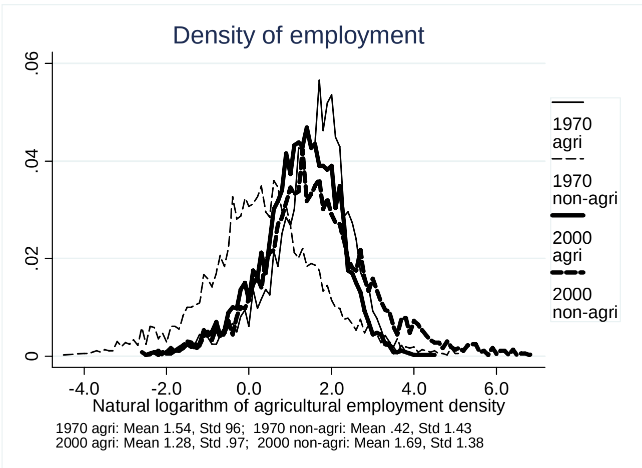
Panel A



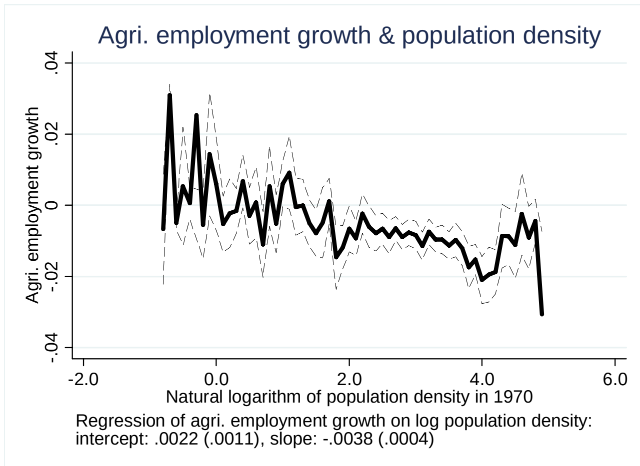
Panel B



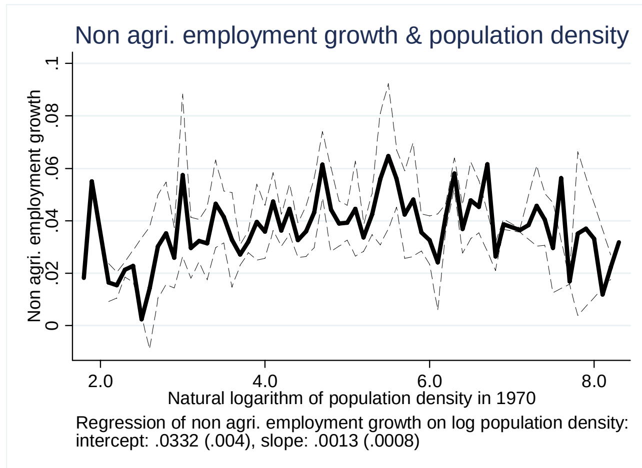
Panel C



Panel D



Panel E

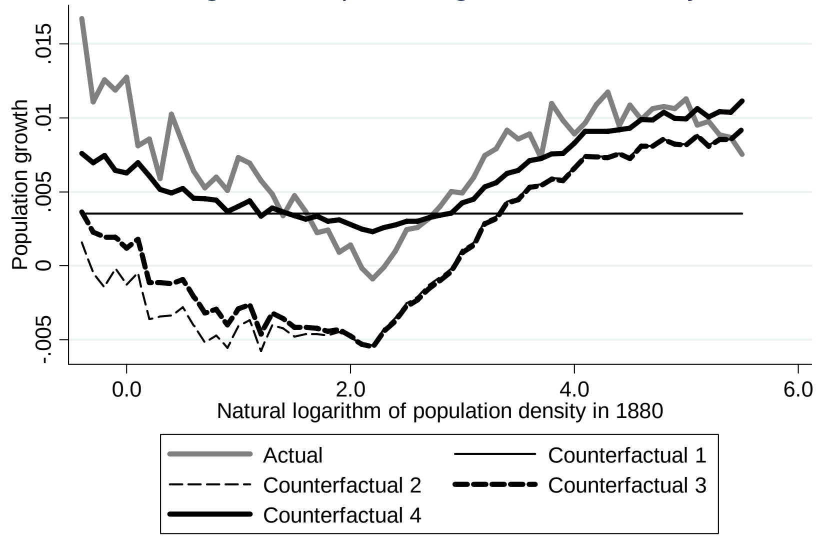


Panel F

Note: This figure reproduces Figures 1 to 6 but uses Brazilian instead of US data.

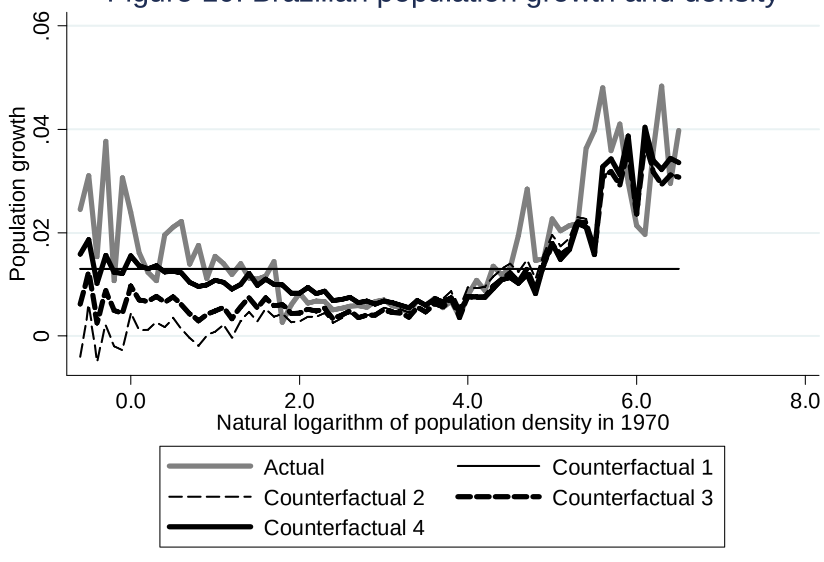
.

Figure 9: Population growth and density



Note: This figure shows mean actual and predicted population growth from 1880-2000 within each population density bin based on estimating non-parametric specification (2) for the sample of "A and B" states. Population density bins are defined by rounding down log population density for each MCD to the nearest single digit after the decimal point. Since population density bins at the extreme ends of the distribution typically contain at most one observation, the figure (but not the estimation) omits the 1 percent most and least dense MCDs in 1880. Counterfactuals 1-4 use progressively more components of the model to generate predicted population growth as discussed in the paper and web-based technical appendix.

Figure 10: Brazilian population growth and density



Note: This figure reproduces Figure 9 but uses Brazilian instead of US data.