Guy Michaels, London School of Economics, CEP and CEPR

Ferdinand Rauch, University of Vienna and CEP

Stephen J. Redding, London School of Economics, Yale School of Management, CEP and CEPR

Rethinking urbanisation

Urbanisation – the concentration of population in cities and towns - is one of the most striking features of economic development. 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 around one half today. 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.

Understanding the determinants and pattern of urbanization is central to a host of policy issues. On the one hand, urbanization raises major environmental and social concerns. The conversion of open countryside to built up areas involves the loss of natural habitat, the elimination of recreational space and increased demands for agricultural food surpluses to feed the urban population. At the same time, dense population concentrations create major infrastructure demands: rapid transits systems for commuting, supplies of clean water and power, sanitation and waste management systems, and public health facilities. On the other hand, millions of people have chosen to move from rural to urban areas, despite the high levels of mortality observed in cities in nineteenth century Europe and in cities in the developing world today. Therefore there must be compensating advantages, at least for those choosing to migrate to urban areas, including higher productivity and wages, a wider range of employment opportunities and goods and services, and a greater diversity along ethnic, cultural and other dimensions.

Despite the magnitude of the population reallocations from rural to urban areas that have been observed in recent history, there has been relatively little study of the process through which rural areas become urban. Partly for reasons of data availability, most existing research on population growth focuses on urban areas. This research emphasizes two key stylized facts. First, population growth appears largely uncorrelated with initial population size (Gibrat's Law). Second, there is a stable population distribution that can be approximated by Zipf's Law (the second largest location has half the population of the largest location, the third largest location has a third the population of the largest location, and so on). This exclusion of rural areas from existing research is a concern because historically in developed countries and in developing countries today these areas account for a large share of the overall population.

Our research shows that these two stylized facts for population growth are strongly rejected if both rural and urban areas are considered. Using newly constructed data based on sub-county units in the United States from 1880-2000 and based on municipalities in Brazil from 1970-2000, we find that population growth is at first decreasing in initial population density at low densities, before increasing in initial population density at intermediate densities, and only becoming largely

uncorrelated with initial population density at the high densities observed in urban areas. The resulting U-shaped relationship for population growth is shown in the figure below and results in an unstable population distribution over time. Along the upward-sloping segment of the U-shape, locations with a high initial population density grow relatively rapidly and become urban areas, while other locations with a lower initial population density grow relatively slowly and remain rural areas. As a result, even though there has been a substantial increase in the U.S. population between 1880 and 2000, we find an increase in the number of sparsely-populated locations as well as an increase in the number of densely-populated locations.

Our research explains these patterns of population growth in terms of the distinction between agriculture and non-agriculture (i.e. manufacturing and services). Two key ideas play an important role in our analysis. First, the employment growth process differs between agriculture and non-agriculture. In agriculture, employment growth is negatively related to initial employment size, because after positive shocks the relative employment of locations reverts towards a long-run mean determined by for example climate and soil. In contrast, in non-agriculture, employment growth is largely unrelated to initial employment size, because persistent factors such as climate and soil impose less constraints on employment growth. As a result, the share of non-agricultural employment in the population is increasing in initial population density, at least for sufficiently high initial population densities. Second, over time, faster productivity growth in agriculture than in non-agriculture has resulted in a reallocation of employment away from the agricultural sector.

The combination of these two ideas can explain the observed U-shaped relationship between population growth and initial population density. At low population densities, agricultural employment dominates, and mean reversion in agriculture generates a decreasing relationship between population growth and initial population density. In contrast, at high population densities, non-agricultural employment dominates, and the absence of mean reversion implies that population growth is largely uncorrelated with initial population density. In between, the positive correlation between the share of non-agriculture in employment share and initial population density, combined with faster employment growth in non-agriculture than in agriculture, generates an increasing relationship between population growth and initial population density.

Our findings have a number of policy implications. As large developing countries such as China and India industrialize, the resulting increase in the urban population is likely to create major pressures on the environment and infrastructure. Our finding that population growth is increasing in initial population density at intermediate densities, and is largely uncorrelated with initial population density at higher densities, suggests that more densely-populated areas are likely to experience larger increases in the urban population. While our analysis suggests that some locations with higher initial population density will experience rapid population growth, other locations with lower initial population density will experience declines in relative and possibly absolute population. Therefore the rural depopulation that has afflicted many small towns in the American mid-West could also become a feature of the agricultural peripheries of currently-developing nations. Taken together, the major changes in the distribution of population across space that our analysis points to will require major reorientations in public infrastructure, which are likely to provide considerable challenges for governments throughout the developing world.

Population growth and density, 1880-2000
Failure of Gibrat's Law
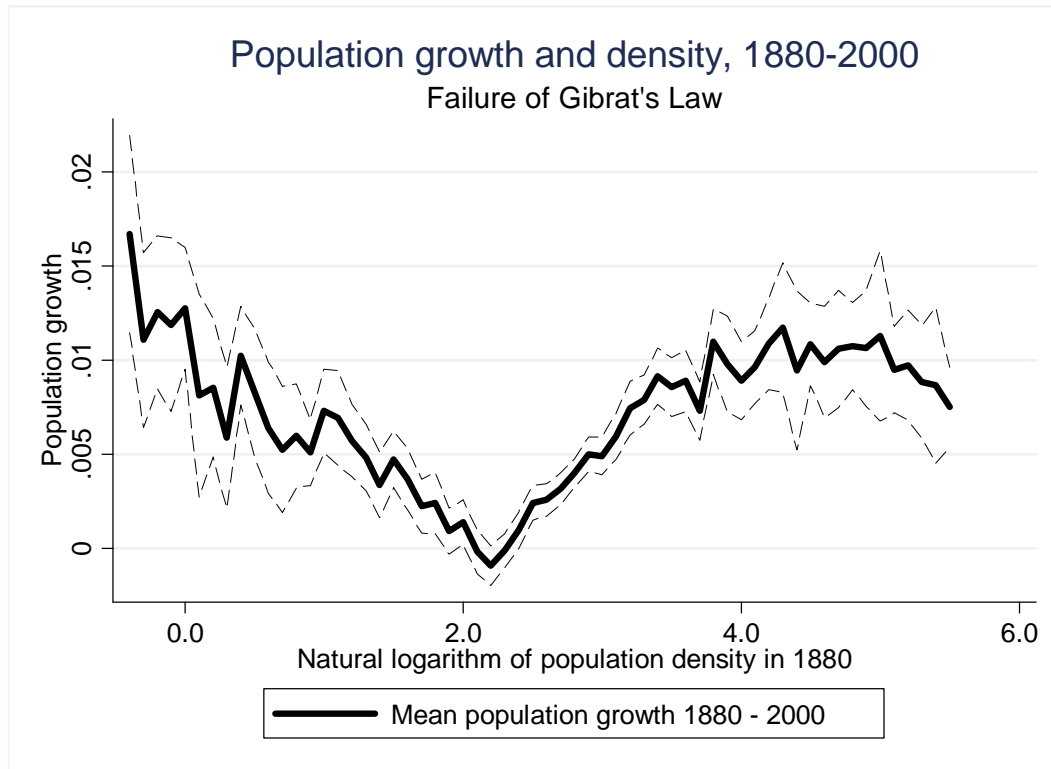
Note: The solid line shows mean population growth rate from 1880-2000 for United States Minor Civil Divisions (MCDs). The dashed lines show 95 percent confidence intervals based on robust standard errors clustered by county.