The Urban Mortality Transition and the Rise of Poor Mega-Cities

Remi Jedwab and Dietrich Vollrath*

November 3, 2016

Abstract

Today the world's largest cities lie in poor countries, unlike the historical norm. Also unlike the "killer cities" of history, today's poor mega-cities grow not just through in-migration but also through their own natural increase. First, we use novel historical data to document that these poor mega-cities arose at the same time as the post-war urban mortality transition. Second, we develop a general equilibrium model of location choice with heterogeneity in congestion costs and demographics across locations to account for this. In the model, the aggregate rate of population growth and the locational choice of individuals interact. Third, we calibrate the model to data from a sample of poor countries, and find that informal urban areas (e.g. slums) grew in large part because relative to other locations they were able to absorb additional population easily. We show that between 1950 and 2005 the urban mortality transition by itself would have doubled the urbanization rate as well as the size of informal urban areas in this sample. Of these effects, one-third can be attributed to the direct amenity effect of lower urban mortality rates, while the remainder is due to higher population growth disproportionately pushing people into informal urban areas. Fourth, policy analysis suggests that family planning programs may be as effective as higher productivity or urban infrastructure and institutions, and more effective than migration restrictions and slum clearance, at transforming mega-cities of the developing world into "engines of growth".

JEL: R11; R12; R13; O11; O40; O18; J10; N00

Keywords: Mega-Cities; Urbanization; Urban Mortality; Congestion; Slums; Quality of Life; Amenities; Structural Change; Demographic Regime; Growth; History

*Remi Jedwab: Associate Professor of Economics, Department of Economics, George Washington University, jedwab@gwu.edu. Dietrich Vollrath: Associate Professor of Economics, Department of Economics, University of Houston, devollra@central.uh.edu. We would like to thank Quamrul Ashraf, Hoyt Bleakley, Filipe Campante, Klaus Desmet, Gilles Duranton, Edward Glaeser, Vernon Henderson, Erik Hurst, Stelios Michalopoulos, Giacomo Ponzetto, Diego Puga, Paul Romer, Stuart Rosenthal, Esteban Rossi-Hansberg, Rodrigo Soares, Nico Voigtlaender, Hans-Joachim Voth, David Weil and seminar audiences at the CEPR Symposium on Development Economics (UPF), FGV-EESP, George Mason, Georgetown, George Washington, Harvard (Cities Conference), Los Andes, Maryland, Michigan, New South Wales, NYU, OCDE, Oxford, Pompeu Fabra-CREI, PUC-Rio, Rosario, SEA (New Orleans), SMU, Syracuse, United Nations-ECA (Addis Ababa), United Nations-ESCAP, Urban Institute, USAID, Wharton, Williams, World Bank, World Bank-George Washington University Urbanization Conference and WEHC (Kyoto) for helpful comments. We thank the Institute for International Economic Policy and the Elliott School of International Affairs (SOAR) at George Washington University for financial assistance.

1. Introduction

Urbanization has gone hand in hand with economic growth throughout history (Henderson, 2010; Duranton, 2014b). However, the post-war period witnessed "poor country urbanization" (Glaeser, 2014) that has generated poor mega-cities in developing nations. Dhaka, Karachi, Kinshasa, Lagos, and Manila are some of the largest cities on the planet today. In contrast, only six of the currently largest 30 cities (London, Los Angeles, New York, Osaka, Paris, and Tokyo) are in high income countries. The prevalence of poor mega-cities today runs counter to historical experience. In the past, the world's largest cities were in the most advanced economies. Today's mega-cities are unlike their historical peers in that their size is *not* indicative of prosperity.

We first document this puzzling phenomenon by building the largest available database on the spatial aspects of the demographic transition from antiquity to modern times, and show that the poor mega-cities of the post-war era experienced rates of natural increase - the birth rate minus the death rate - well above levels seen in historic mega-cities. This was due to the drop in urban death rates following the *urban mortality transition* of the mid-20th century. Poor mega-cities grew in absolute terms because of both in-migration *and* natural increase, setting them apart from historical "killer" cities that grew only through in-migration. Beyond their absolute pace of growth, we further show that poor mega-cities today are very densely populated, with much of this density due to a high share of population in slums. They also have low human capital, with high dependent population shares and low education completion rates.

Second, to understand the *rise of the poor mega-cities*, we construct a general equilibrium model of population allocation across locations, similar in spirit to work on structural change (Caselli and Coleman II, 2001; Gollin et al., 2002; Duarte and Restuccia, 2010; Michaels et al., 2012; Buera and Kaboski, 2012; Lagakos and Waugh, 2013). We incorporate locational choice similar to the literature on equilibrium city size (Henderson, 1974; Duranton and Puga, 2004; Desmet and Rossi-Hansberg, 2013, 2014; Desmet and Henderson, 2014; Duranton, 2014b) and combine that with elements of models of population and growth (Galor and Weil, 1999, 2000; Ashraf and Galor, 2011). Compared to Desmet and Rossi-Hansberg (2013, 2014), we distinguish between different types of urban locations, and allow aggregate population growth to depend on the distribution of population across all locations. Our model allows population to move freely among locations that differ in their demographics, productivity growth, and elasticity of welfare with respect to population. For a stable allocation of population across locations these elasticities are all negative, and their heterogeneity is consistent with locations having production functions that vary in the importance of fixed factors of production and/or in the strength of congestion and agglomeration effects.

A key insight of the model is that the aggregate population growth rate influences the distribution of population across locations. In particular, an increase in the population growth rate (arising in any location) leads to an increase in the population share of the location with the lowest elasticity of welfare with respect to population. Those locations are able to absorb large population increases without lowering the welfare of residents significantly, and thus they disproportionately grow when population growth accelerates. Note that this does not mean these locations are uncrowded. Rather, precisely because of their low elasticity of welfare with respect to population they will absorb more people, and be the most densely packed locations in the economy.

Third, to quantify the effects of the urban mortality transition (UMT), we calibrate the model to a sample of 41 poor and initially unurbanized countries observed from 1950 to 2005 with data on population growth, urbanization, and slum shares. We use three types of locations so that we can match the model to the available data. In addition to *rural* locations, we specify two types of urban locations: *formal* (e.g. business districts) and *informal* (e.g. slums). A key outcome of the calibration is that informal urban areas display the lowest elasticity of welfare of all the locations, and in turn absorbed a large fraction of the additional population growth created by the UMT, contributing to the rise of poor mega-cities and the high share of slums in those cities.

We find that the UMT may have accounted for two-fifths (9.7 percentage points of the observed 22.1 point increase) of urbanization, and two-thirds in the informal share of cities between 1950 and 2005. In addition, the UMT may have increased total urban population size by three-fourths and doubled the total informal urban population. While we want to be careful about making welfare statements involving mortality, our calibrated model indicates that welfare was slightly *lower* due to the UMT, as the increased population growth created negative congestion effects that offset the calibrated gain in welfare arising from higher urban life expectancy.

Our calibration then indicates that roughly one-third of the effect of the urban mortality transition on urbanization acted through the direct *amenity* effect of lower urban mortality rates. The remaining two-thirds were due to the relatively low *welfare elasticity* of informal urban areas, which allowed them to absorb a significant fraction of the additional population growth created by the urban mortality transition.

Fourth, using our calibration, we then forecast urban populations and welfare forward over time. Without significant declines in population growth, it appears that poor countries will continue to see rapid expansion of urban areas, and this will occur mostly through the growth of informal urban locations. Poor mega-cities may thus be in a "poverty trap" of continued expansion, congestion and informality.

Finally, the importance of demographics for the future prospects of poor mega-cities is also apparent when we examine the effects of various policies. Family planning policies that reduce birth rates appear to be just as effective at increasing welfare and reducing the size of informal urban areas as policies raising formal productivity growth or lowering congestion costs in formal urban locations via urban infrastructure investments and/or better urban institutions. In contrast, restricting migration into cities generally, or specifically into informal urban areas (e.g. slum clearance), does little but lower welfare by forcing people to remain in rural areas.

This paper adds to the literature on structural change and urbanization in several ways. First, if externalities (and particularly knowledge spillovers) are the "engines of growth" (Romer, 1986, 1990; Lucas, 1988), and if cities facilitate interactions between people (Jacobs, 1969, 1984), cities should promote economic growth (Glaeser et al., 1992; Lucas, 2004). However, the rise of over-congested mega-cities in poor countries raises questions about their origin and their potential contributions to growth. While there is some work on urbanization without growth (Fay and Opal, 2000), little attention has been given to the specific growth and characteristics of these mega-cities (see Duranton (2014b), Desmet and Henderson (2014) and Brueckner and Lall (2015) for recent surveys which reference poor mega-cities briefly). To our knowledge, only Desmet and Rossi-Hansberg (2013, 2014), Glaeser (2014) and Henderson et al. (2016b) address this important issue for the growth process. In addition, with the exception of Jedwab et al. (2015), the literature has largely ignored the role of population growth in "poor country urbanization".[1] Yet, this explanation appears to be quantitatively

[1] In a related paper, Jedwab et al. (2015) study the correlations between natural increase and urbanization for a restricted sample of developing countries from 1750 to date. They focus on the growth of the urban population as a whole relative to the total population, hence *urbanization rates*, while here we look at the absolute growth of the largest cities in the world, hence *poor mega-cities*. Their work is also

important.[2] Second, our paper connects with the literature on optimal city size (Henderson, 1977; Abdel-Rahman and Anas, 2004; Albouy, 2008; Eeckhout and Guner, 2015; Albouy et al., 2015b). However, with the exceptions of Au and Henderson (2006) and Desmet and Rossi-Hansberg (2013, 2014), this literature focuses on developed countries. Moreover, these papers only compare cities of different sizes, rather than considering the whole distribution of population, including the rural areas. The literature has also focused on studying agglomeration effects, and says little about congestion in poor mega-cities. A few exceptions are Desmet and Rossi-Hansberg (2013, 2014), Duranton (2014a, 2016) and Hanlon and Tian (2015). Third, locational choices are explained by both local productivities and local quality of life amenities (Rosen, 1979; Roback, 1982; Glaeser et al., 2001; Albouy, 2008, 2016; Albouy and Stuart, 2014; Hanlon, 2015; Ambrus et al., 2015; Chauvin et al., 2016). We find that a specific urban amenity – higher urban life expectancy – has directly contributed to urbanization in developing countries. Fourth, we document that these poor mega-cities have high child dependency ratios and are low-skilled, in contrast to the literature that emphasizes that cities help disseminate knowledge (Glaeser et al., 1992; Moretti, 2004; Gennaioli et al., 2013). Lastly, our quantitative analysis is able to deliver clear policy prescriptions.

Our work also adds to the literature on the economic effects of demography. Population growth promotes economic growth if high population densities encourage human capital accumulation or technological progress (Becker et al., 1999; Galor and Weil, 2000; Lucas, 2004; Desmet et al., 2015). But the presence of a fixed factor of production (e.g. agricultural land and urban space) implies that living standards are inversely proportional to population size. Due to congestion effects in urban areas, "Malthusian" forces need not disappear just because economies urbanize.[3]

The poor mega-city outcome is one that arises perversely because of the success of

mostly descriptive, while our quantitative analysis allows us to characterize the trajectory of these poor mega-cities, and identify the policies that may alleviate the problem of poor mega-city growth.

[2]Our explanation does not rule out other explanations, such as urban bias (Ades and Glaeser, 1995; Davis and Henderson, 2003), natural disasters (Henderson et al., 2016a), trade (Glaeser, 2014; Gollin et al., 2015; Venables, 2016) and institutional frictions (Glaeser, 2014; Henderson et al., 2016b).

[3]The literature on the relationship of population and growth is vast (see Galor (2012) for a survey). Barro and Becker (1989), Becker et al. (1990) and Manuelli and Seshadri (2009) provide models of the negative relationship of income and fertility. Unified growth models depend on a rise in the demand for human capital to induce sustained growth, which is driven by acceleration in technological change (Galor and Weil, 1999, 2000; Ashraf and Galor, 2011). This demand is also driven by the decline in benefits of child labor (Doepke and Zilibotti, 2005), health improvements (Cervellati and Sunde, 2005; Soares, 2005), trade (Galor and Mountford, 2006, 2008), culture (Tertilt, 2005), or structural change (Vollrath, 2011).

interventions that limited urban mortality rates while urban fertility remained relatively high. In this, our work is similar to others that emphasize the negative economic effect of mortality interventions and/or the positive economic impacts of mortality increases. Like Young (2005) and Voigtländer and Voth (2013b) we match a model to the data, and quantify the effects of mortality changes on welfare. Unlike Acemoglu and Johnson (2007) and Bloom et al. (2014) we do not pursue reduced form empirics.[4] We do not have sufficient data on wages, amenities, or welfare more generally, across time or locations. Nor is there a clear strategy to identify the causal effect of urban mortality relative to overall mortality. With our analysis, we can quantify the mechanisms through which the urban mortality transition drove urbanization, and examine the impact of policies.[5]

Lastly, the paper does not assume or imply spatial misallocation, unlike the literature on the urban-rural income gap (Gollin et al., 2013; Young, 2013) and the broader literature on misallocation and productivity (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013; Duranton et al., 2015). While we can allow for migration restrictions that generate a suboptimal wedge in wage growth between locations, in our analysis low welfare comes from the fact that individuals do not internalize the negative externalities of their location decisions on congestion, aggregate fertility, and living standards.

In the next section we document the rise of poor mega-cities, and the underlying demographic features of these cities. We then present and calibrate our model, and calculate the impact of the urban mortality transition on the growth of poor mega-cities. The final section discusses the policy implications of the findings.

2. The Rise of Poor Mega-Cities

Table 1 shows the largest 30 cities in the world in select years from 1700 to 2015. The mega-cities of 1700 were located in the most economically advanced countries in that period. While London and Amsterdam had wages that were relatively high then, cities such as Beijing, Istanbul, and Tokyo all had wages equivalent to those found in cities such as Paris and Naples (Özmucur and Pamuk, 2002; Allen et al., 2011).

[4]Other studies that consider the effect of differential mortality on population and growth include Weil (2007); Bleakley (2007, 2010); Ashraf et al. (2013); Voigtländer and Voth (2013a).

[5]The paper also contributes to the related economic history literature. Other studies have focused on urban mortality in England or the U.S. in the 19th century (Williamson, 1990; Costa and Kahn, 2006; Haines, 2008; Voigtländer and Voth, 2013b; Hanlon and Tian, 2015; Hanlon, 2015), whereas we track the demographic behavior of many cities, and their "surrounding" areas, for many countries and over a long period of time. For the post-war period, using the Demographic and Health Surveys and census information from IPUMS, we also provide information on measures of congestion at the city level.

By 1900, the nature of the list of largest cities changed along several dimensions. First, the absolute sizes were roughly ten times larger than in 1700. Second, the cities dominating this list were the leading cities from the richest countries. London, New York, and Paris are all found on the list. Further down, we see Boston, Manchester, and Philadelphia, all centers of industrialization. There were several large cities in poor countries: Beijing, Kolkata, and Mumbai. But none approached the size of the world leaders. Comparing 1900 to 1700, one can also see that their growth in this period was not on the same scale as the richer cities. Beijing and Istanbul increased from 700,000 to about 1 million in the same 200 years that London went from 600,000 to 6.5 million.

In 1950 the top cities remained those in advanced nations, but we see the very beginnings of mega-city growth in poor countries. Kolkata and Shanghai both had more than four million inhabitants, putting them in the top 10 cities in the world in 1950. Beijing, Cairo, Mexico, and Mumbai were all over 2 million inhabitants.

In 2015, the nature of the largest cities has changed dramatically. First, the absolute scale of cities is 3-5 times larger than in 1950. Second, the composition of the list is now dominated by developing countries. Only London, Los Angeles, New York, Osaka, Paris, and Tokyo are in rich countries. Instead we see cities such as Mexico and Mumbai. Poor countries have cities present on this list, such as Dhaka (Bangladesh), Karachi (Pakistan), Kinshasa (DRC), Lagos (Nigeria), and Manila (Philippines). These cities all have at least 11 million people, making them larger in size than nearly all cities in 1950.

The shift of the world's largest cities from rich to poor countries can be also seen in Figure 1. This plots the number of cities with more than one million inhabitants for two groups, developed countries (based on their 2015 GDP per capita) and developing countries. In 1900 and 1950 nearly all million-plus sized cities were in currently developed nations. This switched between 1950 and 2015, however, and this reversal is projected to increase well into this century (2030).[6]

Further, city size in the past was a robust indicator of city-level living standards, but that relationship has broken down over time. For the pre-1910 period, our data are for welfare ratios calculated using wages and price indices for minimal consumption baskets in different cities (see Jedwab and Vollrath (2015) for details on the sources used for each city). The 111 observations are at the city-year level, so that for several cities we have multiple observations over time. We rank all these observations based on their

[6]The patterns hold if we exclude China and India or use other cut-offs than one million inhabitants.

absolute city size. We then rank all these observations based on their welfare ratio, and plot the rank of welfare ratios against the rank of city size. Figure 2 shows there was a positive relationship historically (correlation of 0.62). This is not to say that the cities of industrializing Europe or North America had high *absolute* living standards. But city size indicated something regarding *relative* living standards at the time. For the modern period, we do not have wage measures comparable to historical data. However, we do have a city product index for a sample of 111 cities of at least 300,000 inhabitants in 2010 (UN-Habitat, 1998, 2012). For these observations we plot the rank of living standard against the rank of size, and find a correlation of only 0.28 (Figure 2).[7]

To highlight the differences over time, several cities are shown in the figure. From 1700 to 1900 both Amsterdam and London shift to the top right, indicating they were growing in size as they developed. We do not have living standard data for New York in 1700, but New York was one of the wealthiest and largest cities by 1900. From that point forward, however, these cities slipped down the rankings in size while maintaining their position in living standards. In comparison are a number of currently poor mega-cities. Kolkata in 1825, Delhi in 1875, and Jakarta in 1900 are all small and poor. Yet they have all moved up to become among the largest cities of the world by 2010. However, this has not been associated with a move up in the rankings in living standards. Other poor cities such as Dacca, Kinshasa and Lagos are among the largest cities in the world.

This change in the composition of the largest cities is only going to be exacerbated in the future. The final column of Table 1 shows in parentheses the projected growth rate from 2015–2030 for each mega-city (see $\Delta\%2015-2030$). Rich mega-cities such as Tokyo (-0.1%), New York (0.5), Paris (0.6) and London (0.7) have growth rates close to zero, suggesting that they are reaching their "steady state" size. In comparison, poor mega-cities similar in size to the rich mega-cities are still growing fast, such as Lagos (4.2%), Kinshasa (3.7), Dhaka (3.0) and Karachi (2.7). Dhaka and Karachi both have almost the same population size as New York today, at about 17-19 millions, yet both cities are predicted to reach 25-27 millions by 2030, while New York will stay at 19 million.

3. Characteristics of Poor Mega-Cities

3.1. The Urban Mortality Transition and Mega-City Growth

[7]The product index is measured using such inputs as capital investment, formal employment, trade, savings, exports and imports, and household consumption. Details are available in the UN reports.

Historically, in-migration was the dominant source of new city dwellers as the rates of natural increase were low in urban areas, typically because of high urban mortality rates. This can be seen in Figure 3, where we compare the crude birth rate (CBR) and crude death rate (CDR) of *mega-cities* across different historical eras. We used various sources described for each country-period in Table 2 – population censuses, demographic surveys, sanitary reports, and historical studies (e.g., books, theses and academic articles) – to reconstruct historical demographic data for a sample of the largest cities in the world, from antiquity to date. For the 392 city-period observations in our sample, we also obtained when possible the same demographic data for the *other cities* of the same country-period, the *rural areas*, and the entire *country*.[8]

We first focus on the 100 largest cities as of 1900, according to Chandler (1987). Panel A plots the birth and death rates for 38 of these cities in or before 1800 where we have data. The sample includes various pre-industrial cities such as ancient Rome, Teotihuacan, Renaissance Florence, London in the 17th century, and Boston and Philadelphia in 1750. As can be seen, birth rates were high on average (37.3 per 1,000 people), but death rates were also high (36.9). Mortality was high in cities due to poor water quality, inadequate waste disposal, and dense areas favoring the spread of contagious diseases (Williamson, 1990; Haines, 2008; Costa and Kahn, 2006; Voigtländer and Voth, 2013b; Ambrus et al., 2015). As a result, the cities all lie near the 45-degree line, indicating that they experienced no natural increase (0.4, i.e. 0.04% per year).[9]

Panels B-D then show that cities remained near the 45-degree line in the first and second half of the 19th century (Panel B: 1820s-1850s; Panel C: 1880s), and at the turn of the 20th century (Panel D: 1900s), when using available data for 33, 69 and 89 of the 100 largest cities in 1900 respectively. City natural increase was still low on average, at 5.0-6.1 per 1,000 people (0.5-0.6% per year). The samples include various emblematic cities of the Industrial Revolution, such as Boston, Liverpool, Manchester and Philadelphia.[10]

[8]For the oldest location-periods in our sample, we also use the work of anthropometrists who studied graves and skeletons to obtain data on average fertility (i.e. crude birth rates) and life expectancy (i.e. crude death rates). For the other location-periods in our sample, the sources used often directly report the crude birth rate and/or the crude death rate and/or the crude rate of natural increase of the location (city, urban areas, rural areas, country). The exact source (author-title-publication-date) for each location-period observation is available from the authors upon request. Lastly, for each mega-city, we use its share in the total urban population, its demographic rates, and the demographic rates of the whole urban sector, to obtain the mean demographic rates for all other cities in the same country.

[9]All the points in Panel A represent "normal" periods, but each city was at times afflicted by severe shocks to mortality. For example, during the Black Death, cities had death rates of 250-750.

[10]We were able to find data for many industrializing cities in the 1820s and/or the 1850s and/or the

Their growth, which averaged 3% per year in the 19th century (Jedwab et al., 2015), mostly occurred through in-migration. In the rest of the world during the same period, birth rates and death rates were also high, and rates of natural increase low, as seen for Beijing, Cairo, Delhi, Kolkata, Mexico and Montevideo in Panels B-D.

In the post-war period, there was a distinct change in city demographics. In Panels E and F, we study the 100 largest cities of the future (in 2030 according to the United Nations (2014)), and show their birth and death rates when available in the 1960s (N = 63) and the 2000s (N = 100) respectively. Focus first on the relatively rich cities in the lower left of Panel E, in the 1960s. Their birth rates fell along with their death rates, and so their rate of natural increase remained small. Overall, it is apparent from Panels A to E that historically cities were "sliding down" the 45 degree line as they grew. In comparison are the nascent poor mega-cities in the upper left of Panel E, well above the 45 degree line. In the 1960s these cities differed from earlier eras in one distinct way: their death rates were very low. Dhaka, Karachi, Kinshasa and Lagos, despite being in countries with much lower levels of income per capita, all had death rates that were similar to those seen in London or Paris in the same year. Developing mega-cities in the 1960s were mainly "shifted left" compared to their historical peers. This can also be seen by comparing the respective locations of Cairo, Delhi and Mexico in Panels D and E. This led to large rates of natural increase for emerging mega-cities. For example, in the African mega-cities in the figure, crude rates of natural increase were roughly 3.5% per year. Even absent migration, these cities would have doubled in size every 20 years.

This difference continued in the 2000s (Panel F). Rich mega-cities remained in the same position as in the 1960s. The developing mega-cities shifted down to lower birth rates. However, the death rates in poor mega-cities were lower than the historical comparisons, for the most part falling below 10 per thousand. Thus in the 2000s poor mega-cities continued to have rapid rates of natural increase. A notable exception were Chinese cities (e.g. Beijing), which in the 1960s (Panel E) looked similar to other developing cities, but moved in the 2000s (Panel F) to a pattern of death and birth rates similar to rich mega-cities. Johannesburg has a high death rate due to HIV.

The deviation of the developing cities from the historical norms appears due to the

1880s, such as Amsterdam, Antwerp, Baltimore, Berlin, Birmingham, Boston, Bradford, Brussels, Chicago, Detroit, Dublin, Glasgow, Hamburg, Leeds, Liege, Liverpool, London, Lyon, Manchester, Milan, Montreal, Munich, Newcastle, New York, Osaka, Paris, Philadelphia, Pittsburgh, Sheffield, St. Louis and Tokyo.

mortality transition. Following World War II (WW2), there was a sudden improvement in health in poor countries (Stolnitz, 1955; Davis, 1956; Preston, 1975), due to: (i) the discovery of effective techniques for mass production of antibiotics such as penicillin (1942) and streptomycin (1946), which treated diseases such as cholera and dysentery, (ii) the invention of vaccines against the yellow fever (1937), poliomyelitis (1962) and measles (1963), (iii) the creation of the World Health Organization (1948), which disseminated knowledge of the new technologies to poor countries, and (iv) disease eradication campaigns, whether against smallpox or malaria. From the perspective of the poor countries of that period, this *international epidemiological transition* represented an exogenous shock to mortality (Acemoglu and Johnson, 2007).

As can be seen in Figure 4, another consequence of the mortality transition was to raise rates of natural increase in the largest cities up to the rates typically seen in rural areas. Rural natural increase was already high before the 20th century (Panel C), due to high rural fertility (Panel A) and low rural mortality (Panel B). Urban fertility was relatively low (Panel A), while urban mortality was high due to the urban penalty (Panel B). As a result, urban rates of natural increase remained low (Panel C). This changed post WW2, as the international epidemiological transition diffused to large cities and colonizers invested in urban public health by building water supply and sewerage systems and preventing disease outbreaks (Njoh and Akiwumi, 2011; Njoh, 2013; Fox, 2014; Fox and Goodfellow, 2016). This shock to mortality occurred in both the largest cities and the other cities of the same country (Panel B), which thus have similar natural growth rates (Panel C). Cities also had rates of natural increase similar to rural areas in this period. Due to the *urban mortality transition* (UMT), natural increase was now equally high across all areas.[11] Finally, note that there was also a slow rural mortality transition, as rural rates of natural increase slightly increased.

Panel C also shows for the 392 city-period observations the mean population growth rate of the entire country (see *Total*). Population growth greatly accelerated in the 20th century. When crudely decomposing the change in total natural increase between the 1900s and the 1960s into its two main sources, urban natural increase and rural natural

[11]The urban mortality transition observed in Panel B is starker if we focus on developing countries only (based on their 2015 GDP per capita, i.e. the countries whose income level is below the income of Slovakia, the last country to have become a developed country according to the International Monetary Fund). Their urban death rate was stable around 35, but 40 for the poorest, before the urban mortality transition ("pre-1800s"-"1900s"), after which it dropped to 10 and 15 respectively ("1960s"). 40 was the urban death rate of developed countries before they started developing with the Industrial Revolution.

increase, we find that half of it was explained by the former. The urban rate of natural increase dramatically increased during the same period. In addition, these countries also slightly urbanized between 1900 and 1960. Had urban natural increase remained low, this increased urbanization would have slowed population growth. But since urban natural increase was high, urbanization increased total population growth.

How much of urban natural increase comes from the informal areas versus the formal areas of cities will prove important when quantifying the effects of the UMT in the model section. To answer this question, we used similar sources as for the largest cities to obtain the crude birth rate and death rate of the slum areas of four industrializing cities in the late 19th century and eight poor mega-cities today (see Table 2).[12] Knowing the city population share of the slum areas, we reconstruct the demographic rates of the non-slum areas. In the 19th century, both slum and non-slum areas had low rates of natural increase, at 0.5% per year. Slum fertility was high, but slum mortality was also high. In the poor mega-cities of today, the rates of natural increase are higher, at 1.0% per year in the non-slum areas and 2.6% per year in the slum areas. Natural increase is thus disproportionately higher in slums, which suggests that poor mega-cities are likely to remain informal in the future. Simple decompositions show that three-fourths of the change in urban natural increase between the late 19th century and today can be explained by the higher rate of natural increase in slums.[13]

Lastly, a consequence of the decline in urban mortality was that the absolute growth of cities expanded to previously unseen levels. For each of the 30 largest cities in 2015 (Table 1), we collected information on the largest annual change in population (in thousands) experienced by the city, as well as the decade this occurred (data not shown but available upon request). We find that poor mega-cities such as Delhi (620 thousand during the 2000s), Dhaka (440; 2000s), Karachi (410; 2000s) and Lagos (350; 2000s) added population at rates well above those seen in historical mega-cities. London (90; 1890s), New York (220; 1920s) and Paris (110; 1950s) never received inflows of people at rates like those seen in the poor mega-cities of today. For poor countries in the 20th century, the UMT created a positive shock to the absolute growth of cities, and given

[12]We consider London's slums, Manchester's townships, New York's tenement-house wards and Paris' "quartiers pauvres" in the 1880s, and Manshiet in Cairo, Korail in Dhaka, Orangi in Karachi, Tondo in Manila, Neza in Mexico, Dharavi in Mumbai, Kibera in Nairobi and Rocinha in Rio in the 2000s.

[13]When comparing the poor mega-cities of the 19th century and the 20th century respectively, we find that slum fertility has remained high, at 35 per 1,000 people. But while slum mortality was high in the 19th century, at 30 per 1,000 people, it is now as low as in the non-slum areas, at 7.5 per 1,000 people.

congestion effects, this created the potential for poor mega-cities to emerge.

3.2. Congestion in Poor Mega-cities

We now document how measures of congestion relate to growth in the 100 largest cities of the world today. We examine city-level correlations with respect to natural increase, to show that mega-cities growing quickly tend to be highly congested and poor.

Figure 5.A plots the projected growth rate from 2015-2030 of the 100 projected largest cities in 2030 (source: United Nations (2014)) against their rates of natural increase in the 2000s.[14] The relationship is strongly positive, which indicates that natural increase will likely be a meaningful driver of city population growth (R-squared of 0.51). The Chinese cities (e.g. Beijing) are outliers to this relationship, with low rates of natural increase, but high expected growth rates due to in-migration (excluding them raises the R-squared to 0.69). Cities that are expected to stagnate in size (e.g. Tokyo) all have rates of natural increase of roughly zero. In Figure 5.B we find a similar correlation between past city natural increase (1960s) and past city growth (1960-2010; source: United Nations (2014)) (R-squared of 0.50), but for 63 mega-cities only.

In both Figures 5.A and 5.B, the unconditional effect of natural increase (when defined per 100 people) is about 1. Each newborn in a city appears to raise the population of that city by one in the long run. This holds even though migration costs *within* countries are low and population can reallocate across locations. Duranton (2014a) finds that the effect of the log birth rate (per 100 people) in 1993 on the change in log population between 1993 and 2010 for about 1,000 Colombian municipalities is 0.24. However, finding a coefficient of 1 *across* countries is not surprising given that (i) our analysis is mostly cross-country, and people from megacities in poor countries cannot easily migrate to rich countries;[15] and (ii) other cities and rural areas in poor countries also grow at high rates (see figure 4), and may be as congested, thus shutting down another potential "safety valve" for the residents of poor mega-cities.

Poor mega-cities that grow through natural increase differ from historical cities that grew mainly through in-migration. Figures 5.C-5.H plot several characteristics of cities against their rate of natural increase in the 2000s (we then have data for 100 megacities).

[14] The projected growth rates are based on non-linear extrapolation given the rates of growth pre-2015, estimated independently of the rates of natural increase. The relationships are thus not mechanical.

[15] Regressing total population growth (%) on aggregate natural increase (per 100 people) for a panel of 201 countries every 5 years in 1960-2010, we also find a coefficient of 1 (not shown).

Please note that the correlations in the figures do not show causal effects. Our goal is simply to characterize the conditions that characterize conditions in poor mega-cities.

First, in figure 5.C is log density, which shows a positive relationship (source: Demographia (2014)). Rich mega-cities such as New York (2,000 inh. per sq km) or Paris (4,000) are much less densely populated than poor mega-cities such as Dhaka (44,000), Karachi (23,000) or Kinshasa (17,000). While density is used as a proxy for economic development when comparing the cities within a country, it measures underdevelopment when comparing the cities across countries, consistent with the literature showing that the income elasticity of housing demand is positive (Rosenthal and Ross, 2015), and higher at lower income levels (Albouy et al., 2015a). [16]

This density is potentially indicative of congestion rather than strong agglomeration effects. Figure 5.D shows that national slum shares are much higher in countries with poor mega-cities (sources: UN-Habitat (2003b, 2012)). Roughly two-thirds of the urban residents live in slums in countries containing Dhaka, Karachi, Kinshasa, or Lagos. Figure 5.E also shows the strong negative correlation between the "city infrastructure index" designed by UN-Habitat (1998, 2012) and city natural increase (R-squared of 0.63). This is not surprising considering that poor mega-cities had annual growth rates of 3-7% in 1960-2010 (Figure 5.B), thus doubling in population size every 10-25 years.[17]

The high density of poor mega-cities does not indicate a large supply of productive workers. Figure 5.F shows the child dependency ratio (the share of those under 14 to population aged 15-64) across our sample of cities (source: reconstructed for each city using the *Demographic and Health Surveys* (DHS) and census information from IPUMS). In poor mega-cities with high rates of natural increase, this ratio reaches more than 50 percent (R-squared of 0.66). Further, the labor force that does exist in poor mega-cities is low skilled. In the poorest mega-cities, the share of college-educated workers is close to zero (Figure 5.G; sources: DHS and IPUMS). Lastly, in Figure 5.H, we show that city natural increase correlates negatively with our measure of city income,

[16]Densities are defined using both the population and the area of the agglomeration. Focusing on the central place or slums only should not affect the results. Manhattan's density is 28,000 people per sq km. However, mega-cities in poor countries contain areas with much higher densities. The slums of Mumbai, Nairobi, Dhaka and Cairo have densities of 350,000, 300,000, 200,000 and 110,000 respectively. The Lower East Side in New York was the densest slum of the developed world, reaching 140,000 in 1910. Other slums in developed cities were less dense: Les Halles in Paris (100,000) and the East End in London (90,000).

[17]The index combines information on access to water, sanitation, electricity, roads, and housing (details available in the UN reports). The results are similar if we use data on the slum share at the city level for fewer observations, the housing price-to-income ratio, or measures of commuting and pollution.

the "city product index" of UN-Habitat (1998, 2012), for the 2000s (R-squared of 0.52). It could be that mega-cities that grow through natural increase are poor because they are highly congested, whether in terms of housing, infrastructure, children or unskilled workers. Or birth rates could remain high in poor mega-cities, because their residents do not have the incentives to have and invest in fewer children.[18]

Note that one should not necessarily expect to observe the same relationships for the rural areas (or the country as a whole). First, we only find a weak correlation between rural population growth or measures of rural density and rural natural increase (figures not shown, but available upon request). One interpretation is that when rural areas become "congested" due to rural natural increase, rural residents can move to the urban areas, as if they serve as a "safety valve". The quantitative analysis below will show that when population growth is high the extra population reallocates itself into the urban areas. Additionally, the relationship between natural increase and child dependency ratios is stronger for the rural areas than for the urban areas (not shown). If rural-to-urban migrants are mostly of working-age, rural areas are disproportionately left with children.[19] Lastly, the gap between urban and rural tertiary completion rates are much smaller for countries with a high rate of natural increase (not shown), thus suggesting that poor mega-cities are not that different from their rural areas.[20]

Overall, poor mega-cities appear to be developing as "giant villages" rather than as high-productivity agglomerations. Mega-cities constitute the modern sector of their countries. Yet, poor mega-cities do not appear that "modern", suggesting that we should not automatically expect urbanization to have transformative effects.

4. A Model of Locational Choice

To understand how the urban mortality transition (UMT) could have contributed to the growth of poor mega-cities, we build a general equilibrium model of location choice

[18]These results hold if we look at primary and secondary completion rates, informal employment, unemployment, poverty and the ratio of the city product index to the log of city size. These relationships also hold if we consider the urban sector as a whole, or secondary cities only (not shown).

[19]Poor megacities today also have higher child dependency ratios than rich megacities in the past. The ratios for Bamako, Karachi and Kinshasa are 72, 63 and 70 respectively. This is higher than the same ratios for London in 1851 (45; source: 1851 census), New York in 1870 (50; source: 1870 census), and close to ratios for the rural areas of the UK and the U.S. in the same years (65 and 70 respectively).

[20]Poor megacities today also have lower literacy rates than rich megacities in the past. When using the DHS and IPUMS, the literacy rates that we obtain for Bamako, Dakar, Dhaka and Karachi are 43, 48, 67 and 62 respectively. This is lower than the literacy rate for London in 1851 (80; source: 1851 census), and New York and the rural areas of the U.S. in 1870 (95 and 80 respectively; source: 1870 census).

with heterogeneity in demographics and the effect of population growth on welfare. Heterogeneity in the welfare effect of population means that aggregate population growth influences the distribution of population across locations. With people able to move between locations, the growth in welfare will be equalized across locations. But because the welfare elasticity with respect to population varies, locations with low elasticities absorb a larger percent gain in population than locations with high elasticities. The UMT, which boosted population growth rates, accelerated the shift of population into relatively low-congestion locations. Once we have the structure of the model in place, we will then be able to more carefully quantify the effect of the UMT by calibrating the model to data from a set of developing countries.

4.1. Individual Utility

We assume that all individuals in the economy are identical, except that they reside in a specific location j from a given set J. Their utility is determined by

$$V_j = \ln w_j + \ln Q_j + \beta \ln(1/CDR_j). \tag{1}$$

w_j and Q_j refer to the net wage and amenities, respectively, in location j. By net wage, we mean the wage net of housing and commuting costs. CDR_j is the crude death rate in location j to which an individual is exposed, and hence $1/CDR_j$ can be seen as a measure of life expectancy.[21] The utility function captures the standard Rosen-Roback trade-off of net wages and amenities within locations. β measures the utility weight of life expectancy.[22] Taking as given the sets of (w_j, Q_j, CDR_j) in each location, individuals will move between locations until V_j is equalized across locations.

4.2. Agglomeration, Congestion, and the Distribution of Population

Let

$$\begin{aligned} \ln w_j &= \ln a_j^w - \epsilon_j^w \ln N_j \\ \ln Q_j &= \ln a_j^Q - \epsilon_j^Q \ln N_j \end{aligned} \tag{2}$$

where a_j^w and a_j^Q are the exogenous determinant of net wages and amenities in location j, respectively. Both terms may grow over time, and are meant to capture the determinants of wages or amenities that do not depend on population size in location

[21]We do not have the data to do age-specific death rates separately by location, so we proceed as if the death rate is constant at all ages. In this case the mapping from the death rate to life expectancy is exact.

[22]The implied utility weight on wages and amenities is identical, as in Desmet and Rossi-Hansberg (2013). However, the choice of using the same weight for wages and amenities is irrelevant to the results.

j. ϵ_j^w and ϵ_j^Q are the elasticities of wages and amenities, respectively, with respect to population size in location j, N_j.

Given the specification in (3), we can write the utility function more concisely as

$$V_j = \ln a_j - \epsilon_j \ln N_j - \beta \ln CDR_j, \tag{3}$$

where

$$\begin{aligned} \ln a_j &= \ln a_j^w + \ln a_j^Q \\ \epsilon_j &= \epsilon_j^w + \epsilon_j^Q \end{aligned} \tag{4}$$

are combined terms that capture the location-specific determinants of wages and amenities ($\ln a_j$) and the *location-specific elasticity of welfare with respect to population size* (ϵ_j). The elasticity ϵ_j captures the overall response of welfare to population size, combining a number of effects. The number of workers will affect nominal wages through the production function for output. Population size will affect housing prices and commuting costs through the production functions for housing and transportion, which in turn affects net wages. Finally, the number of people in a location affects the welfare value of amenities, depending on the production function for those amenities. ϵ_j is a reduced form effect of population on welfare, and we are not attempting to derive a micro-founded model of the production of output, housing, transportion, and amenities. Rather, our interest is in examining the implications of population growth on the distribution of population across locations, given those underlying elasticities.

Given the specification in (3) we can now describe the conditions for the equilibrium distribution of individuals across locations.

1. **Labor mobility**. Individuals move costlessly between locations, implying that $V_j = \overline{V}$ for all locations j, where $\overline{V}$ is the equilibrium level of utility.
2. **Stability**. $\epsilon_j > 0$ for all locations j.
3. **Adding up**. $\sum_{j=J} N_j = N$. This closes the model and determines $\overline{V}$.

The condition regarding stability states that for each location the effect of population on welfare is negative. With this assumption, the equilibrium is stable to any perturbations. If one person were to randomly move away from location j to location k, then welfare would rise in location j and fall in location k. Given labor mobility, this would incent someone to move back from k to j, restoring the original equilibrium. In contrast, if $\epsilon_j < 0$ and hence welfare is *increasing* in population size, then if someone leaves location j for location k, welfare in j actually *falls*, but *rises* in location k. This would incent more migration away from j and into k, meaning the original equilibrium was not stable. With the assumption of free mobility and all $\epsilon_j > 0$, this implies that the economy is at a stable location equilibrium. The exact allocation across locations will change over time due to changes in demographic conditions, changes in the level of wages/amenities, and as we will see below, changes in the aggregate population.

This does not exclude the possibility of, or assume away, agglomeration effects. There may be agglomeration effects at work, but as in standard models of urban areas (Henderson, 1974; Duranton and Puga, 2004) the allocation of population across locations will only be stable if agglomeration effects are more than offset by congestion effects or diminishing marginal returns to labor. One could imagine that agglomeration effects dominate congestion effects and the impact of fixed fators in some locations within low-population countries (Becker et al., 1999; Desmet et al., 2015). While this hypothesis may apply to prehistorical, antique and medieval societies that were characterized by low population densities, they do not characterize well poor countries in the 20th century, which were already densely populated by the time of the UMT.

Denote the growth rate of any variable X by using a "hat", so that $\hat{X} = \dot{X}/X$. Then we can write the growth rate of welfare in location j as

$$\hat{V}_j = G_j - \epsilon_j \hat{N}_j - \beta C\hat{D}R_j \tag{5}$$

where

$$G_j = \hat{a}^w_j + \hat{a}^Q_j. \tag{6}$$

G_j captures changes in welfare unrelated to the population growth rate of location j.[23]

4.3. The Dynamics of Location Size and Welfare

At this point, we can establish several results regarding population, location, and welfare, and how they relate to changes in location-specific mortality rates.

The effect of aggregate population growth on welfare. Define the share of population in location j as $s_j = N_j/N$. By definition, the growth rate of aggregate population, N, can be written as the weighted sum

[23]When calibrating the model, we will not distinguish wages and amenities due to a lack of data. However, fast population growth crowds out both productive amenities and consumption amenities. Therefore, the effects in terms of amenities should reinforce the effects in terms of wages. Not being able to distinguish wages and amenities is thus not essential, unlike in studies of the Rosen-Roback model.

$$\hat{N} = \sum_{j=1}^{J} s_j \hat{N}_j. \tag{7}$$

If we rearrange (5) we have that $\hat{N}_j = (G_j - \beta C\hat{D}R_j - \hat{V}_j)/\epsilon_j$ for each location. Inserting this into (7) yields the following relationship

$$\hat{N} = \sum_{j=1}^{J} s_j \frac{G_j - \beta C\hat{D}R_j - \hat{V}_j}{\epsilon_j}. \tag{8}$$

To proceed further, note that because individuals are fully mobile over locations, it must be the case that $\hat{V}_j = \hat{V}$ for any location. Using that fact, we can solve for

$$\hat{V} = \overline{\epsilon} \sum_{j=1}^{J} \frac{s_j}{\epsilon_j} \left(G_j - \beta C\hat{D}R_j \right) - \overline{\epsilon}\hat{N} \tag{9}$$

where

$$\overline{\epsilon} = \frac{1}{\sum_{j=1}^{J} \frac{s_j}{\epsilon_j}} \tag{10}$$

is the harmonic mean of the elasticities across all locations. Note that by the assumption that $\epsilon_j > 0$ for all locations, it must be that $\overline{\epsilon} > 0$ as well.

Aggregate welfare growth is similar to welfare growth in any given location. It depends on a weighted sum of the location-specific rates G_j and crude death rates $C\hat{D}R_j$, where the weights depend both on their share in population, s_j, and their elasticities, ϵ_j. A higher elasticity ϵ_j mutes the effect of any given G_j, because the high elasticity means that the growth of the population in that location cannot be very large without lowering welfare. On the presumption that $\beta > 0$, declines in the crude death rate, $C\hat{D}R_j < 0$, also contribute positively to welfare.

Finally, *aggregate* population growth, $\hat{N}$, detracts from welfare growth in all locations, with the size of the effect depending on $\overline{\epsilon}$.

The effect of aggregate population growth on the distribution of population. Using the aggregate welfare growth in (9), we can combine that with (5) to solve for the location-specific growth rate of population in location i as

$$\hat{N}_i = \frac{\overline{\epsilon}}{\epsilon_i} \left(\hat{N} + (G_i - \beta C\hat{D}R_i) - \sum_{j=1}^{J} \frac{s_j}{\epsilon_j} \left(G_j - \beta C\hat{D}R_j \right) \right). \tag{11}$$

Aggregate population growth, $\hat{N}$, is positively related to population growth in location i. Regardless of where new people originate, the additional population is spread across locations to ensure that welfare is equalized. The strength of $\hat{N}$ on $\hat{N}_i$ depends on the ratio $\bar{\epsilon}/\epsilon_i$, the aggregate elasticity relative to the location-specific elasticity. When $\epsilon_i < \bar{\epsilon}$, location i will tend to grow more quickly than the aggregate population, and hence its share of the population will increase. ϵ_i is the percent loss of welfare for a percent gain in population, so the lower is ϵ_i, the higher the percent gain in population a location can absorb while still maintaining similar welfare growth to other locations. The logic for a location with $\epsilon_i > \bar{\epsilon}$ is just the opposite, and these locations will grow more slowly than aggregate population growth.

In both cases, however, these direct effects of population growth are potentially offset or exaggerated by differentials in welfare growth from G_i, and changes in crude death rates $C\hat{D}R_j$. The second two terms in (11) capture the difference between these individual effects and their aggregates across locations. If $(G_i - \beta C\hat{D}R_i) > \sum_j (G_j - \beta C\hat{D}R_j)s_j/\epsilon_j$, then location i will grow relatively quickly as people move there to take advantage of its higher productivity, better amenities, and/or lower mortality.

Regardless, the speed of aggregate population growth will always influence the growth rate of locations. To be more concrete, the share of population in each location, s_i, evolves according to the following simple relationship

$$\hat{s}_i = \hat{N}_i - \hat{N}. \tag{12}$$

The derivative of the growth rate of population share, $\hat{s}_i$, with respect to aggregate population growth, $\hat{N}$, can be found using (11),

$$\frac{\partial \hat{s}_i}{\partial \hat{N}} \begin{cases} > 0, & \text{if } \epsilon_i < \bar{\epsilon} \\ < 0, & \text{if } \epsilon_i > \bar{\epsilon} \\ = 0, & \text{if } \epsilon_i = \bar{\epsilon} \end{cases} \tag{13}$$

Unless all locations have identical values of ϵ, then there must exist at least one location with $\epsilon_i > \bar{\epsilon}$ and one with $\epsilon_i < \bar{\epsilon}$. Heterogeneity in elasticities across locations implies that aggregate population growth influences the distribution of population.

Amenity effects of mortality changes on the distribution of population. As can be seen in (11), if the crude death rate in location j is falling, $C\hat{D}R_j < 0$, then this will increase the growth rate of population in location j as individuals move there to take advantage of the lower death rates. The strength of this amenity effect depends on the ratio of $\bar{\epsilon}/\epsilon_j$, not only on the size of β, the utility weight on life expectancy.

Effects of population growth on future population growth. The full consequences of population growth depend on how the changing distribution of population across locations influences population growth itself. Each location has an exogenous crude birth rate, denoted CBR_j. We will explore extensions of this to allow for endogenous changes in birth rates after demonstrating the baseline model and calibrations. Aggregate population growth is dictated by these birth rates combined with the death rates, CDR_j,

$$\hat{N} = \sum_{j=1}^{J} s_j \left(CBR_j - CDR_j \right). \tag{14}$$

The shares s_j are the way in which aggregate population growth is endogenous in the baseline model. As population shares change, this may lower or raise $\hat{N}$ itself depending on the pattern of child costs or death rates across locations. If higher population growth leads to a shift into locations with lower child costs or death rates, then it will accelerate further population growth, which would imply even further shifts of population into that location. The possibility exists that population growth begets more population growth and leads to a concentration of population in a low-elasticity location.[24]

5. The Role of the Urban Mortality Transition

To account for the rise of the poor mega-cities we documented earlier in the paper we will use the general model constructed in the prior section, and calibrate it to match the observed experience of poor countries so that we can perform counterfactuals.

To begin, we have to be more specific about what precisely we mean by locations, so that we can match them up to available data. We will work with three locations: *rural, urban formal, and urban informal*, which we think of as differing along several dimensions: production, housing, transportation, and amenities.[25]

The rural location (denoted by subscript r below) involves production of not only agriculture, but also service or manfuacturing activities of rural residents. Housing

[24]Ways to endogenize population growth would work through the parameters related to child costs or the death rate itself. Following Becker (1960) - through income effects - or Galor (2011) - through the value of human capital - we could introduce changes to child costs related to the composition of population across locations. Changes in mortality rates (Kalemli-Ozcan, 2002; Soares, 2005), by changing the expected value of children, would be an alternative way of modeling endogenous population growth changes. We explore some of these possibilities after the baseline calibration.

[25]One could easily allow for a large number of smaller locations that fall under each category; there could be multiple rural locations, for example. We are assuming that all of those smaller locations of a given type have a common exogenous wage/amenity growth rate, G_i, and a common elasticity, ϵ_i. With that being the case, examining (9) and (11) it is clear that tracking the individual locations will not offer any additional information beyond that found in tracking the three major categories.

consists mainly of individual houses, with few space constraints. There are few organized transportation networks or amenities. In comparison, the formal urban location (denoted by subscript f below) operates a "modern" sector production technology. Manufacturing and professional services would be examples of what we have in mind. Housing involves multi-unit buildings, and faces both physical and institutional (i.e. zoning) constraints. Transportation and amenities are organized centrally to some extent. Finally, the informal urban location (denoted by subscript l below) is located in an urban area and involves a production technology that involves low-level personal services. Housing faces some space constraints, but is not subject to strict institutional limits, and consists mainly of individual houses (e.g., shacks). Transportation and amenities are either absent or organized unofficially. Slums are an example of an informal location, although we think the idea of informal locations may encompass more than those places strictly defined as slums.[26]

5.1. Calibration for a Sample of Poor Countries, 1950-2005

Our model is general enough to give rise to different potential evolutions of the distribution of population across the three locations. Depending on the parameters, the model could predict that the economy remains rural, or urbanizes through formal urban areas (i.e. rich mega-cities) or informal urban areas (i.e. poor mega-cities).

To evaluate the importance of the UMT for the rise of poor mega-cities, we would like to compare actual outcomes to a counter-factual case where the UMT did not occur. To do this, we need to know the values of the parameters in poor countries, and to find those we calibrate our model using information both from our data on poor mega-cities, as well as aggregate data from a sample of 41 poor countries in the period from 1950-2005. Formally, we selected only countries with (a) at least 1 million inhabitants in 1950, (b) an urbanization rate below 20% in 1950, (c) data available on the slum share in 2005, and (d) slum shares of at least 30% in 2005. The choice of 2005 is dictated by the availability of widespread slum data in this year (UN-Habitat, 2003b, 2012).[27]

[26]While the characteristics of locations differ, we assume that all individuals are homogenous, and do not have location-specific preferences or skills. For example, we do not assume that people born in informal urban locations have preferences for or are more capable of working in formal urban locations than people born in rural areas. We do not have the data to track individuals by their location of origin, or a way of measuring location-specific skills for individuals.

[27]The sample contains 28 countries from Africa, 11 from Asia including India and China, 1 from the Middle East, and 2 from Latin America. The calibration returns similar results if we drop China and India.

Initial urban share and informal shares. We use 1950 as the initial period for our calibration. For our sample, the average urbanization rate in 1950 is 8.9%. Explicit information on the informal share of either the housing or labor market in 1950, or a proxy such as the slum share, does not exist. We thus assumed a value of 50%. Historical evidence suggests that mega-cities in the developing world were relatively more formal during the pre-1950 era than today (UN-Habitat, 2003a; Njoh and Akiwumi, 2011; Njoh, 2013; Fox, 2014; Fox and Goodfellow, 2016). The average slum share for the 41 countries is 64.2% in 2005, so 50% implies an increase from 1950 to 2005. Results are similar if we assume that cities were mostly formal (i.e. a formal share of 60%) or informal (a formal share of 40%) in 1950 (not shown, but available upon request).[28]

Crude death rates. The initial CDR in 1950 for both urban locations is set to 40 (per thousand), the average in our mega-city data for poor countries prior to the UMT (see section 3.1., and footnote 11 in particular). After the UMT is completed, the CDR for both urban formal and informal locations is set to 15. To capture the onset of the UMT we let the urban CDRs decay exponentially from 1950 forwards, with a half-life of 3 years. This implies that by 1959, urban CDRs are only 18.25. The results are not sensitive to assuming that the half-life of the UMT ranges between 2-10 years (not shown). For the rural location, we set the CDR to be 20, in line with our data, and that is held constant over the entire period from 1950 to 2005. Our calibration thus takes the *rural* mortality transition as a given, so that we can focus on the effects of the UMT.[29]

Crude birth rates. We use information from the location-level demographic data presented earlier, along with the crude rates of natural increase in our sample of 41 poor countries over time, to set the time path of crude birth rates. For 1950, we set the initial rural crude birth rate (CBR) to 43 (per thousand), the informal CBR to 43, and the formal CBR to 38. The urban location values are consistent with what we observe in our sample of mega-cities. Additionally, these initial rates combined with our initial crude

[28] UN-Habitat (2003a) describes how poor mega-cities such as Ahmedabad, Cairo, Ibadan, Karachi, Kolkata, Mexico, Nairobi, Rio de Janeiro, and Sao Paulo saw a boom in their slum areas from the 1950s, and especially in the 1960s-1970s. For example, the "well-planned town [of Ibadan, the 3rd largest city of Nigeria] turned into a slum. In 1963, half of the city's core area consisted of slum dwellings, growing to 70 per cent of the town's total number of derelict housing in 1985." In Nairobi between 1971 and 1995, "the share of informal-settlement village inhabitants rose from one third to an estimated 60 per cent." In Mexico City, "by 1960, only 54 percent of all dwellings [...] had toilets with running water". However, Mexico City's "irregular settlements now [...] house more than 60 per cent of the city's population."

[29] We have calibrated the model incorporating an explicit rural mortality transition, with the rural CDR falling from 35 to 20 at the same pace as urban mortality rates fell. This exacerbates the effect of the UMT because it implies larger changes in population growth rates (not shown, but available upon request).

death rates mean that our model matches the observed crude rate of natural increase in 1950 in our sample of countries, which was roughly 21 (per thousand).

For the 41 countries, there was a demonstrable rise in the aggregate crude rate of natural increase from 1950 to roughly 1985, and then it began to decline. The crude rate of natural increase was 22.2 in 1960, rose to 28.7 in 1985, and then fell to 25.2 in 2005. This fluctuation was largely due to changes in crude birth rates. To capture this, we parametrically set the changes in crude birth rates in the model to match the observed behavior in our sample. Specifically, we use the following quadratic

$$\Delta CBR_{t+1} = \phi_1 \times t + \phi_2 \times t^2. \tag{15}$$

Fitting this to the observed sample data, we set $\phi_1 = 0.286$ and $\phi_2 = -0.0041$. We will discuss below the nature of these changes in the birth rate, as they are likely related to the changes in the death rates from the UMT. For the purposes of calibrating the model, however, we require the actual time path of the birth rates to match the data.

Productivity and amenity growth. The exogenous rate G_i in a location includes the growth of net wages and amenities unrelated to population size. Separate data on the growth of net wages and amenities does not exist. For our purposes, assigning the *relative* values of G_i across locations is most important, as these will dictate the movement of people across locations, and allow us to determine the *change* in welfare growth between different counterfactual scenarios. The absolute size of the the G_i terms will dictate the absolute growth in welfare over this period, something we cannot measure and so the choice of the absolute levels of G_i will necessarily be arbitrary.

In relative terms, we assume that G_i in rural and informal locations is identical. Fuglie (2010) shows that agricultural TFP growth was below 1% for nearly all developing regions in the 1960s-1970s. However, since that time TFP growth rates have increased across all regions, reaching 2-3% per year in some cases. Overall, he reports developing countries with agricultural TFP growth of 1.4% from 1961-2007. Block (2010) reports agricultural TFP growth rates for Africa over different time periods as well. Growth rates are negative through much of the 1960s and 1970s, but have been as high as 2.8% in the most recent decade, and were 2% from the 1980s into the 1990s. For informal locations, they are dominated by workers in personal services and small-scale retail trade, so they are likely to have low TFP growth (Duarte and Restuccia, 2010; McMillan et al., 2014). Duarte and Restuccia indicate that productivity growth in services among the poorest countries in their sample was around 1% per year. We set $G_l = G_r = 0.025$ to account for both productivity growth and any unobserved amenity growth in these locations.

We assume that $G_f = 0.05$, twice as large as in the other locations. This choice implies that absent any population growth or changes in death rates, workers would be moving from rural to formal urban areas to take advantage of the higher productivity and/or amenities of formal locations. Thus we are allowing for a typical process of structural change over time, and even in the absence of the UMT there will be urbanization. The justification for assuming that G_f is larger than in informal or rural areas is based, from the productivity perspective, on the types of industries we associate with formal locations. Manufacturing experiences unconditional convergence (Rodrik, 2013), and hence growth rates of productivity are rapid relative to informal locations. Duarte and Restuccia (2010) and McMillan et al. (2014) both find evidence that productivity growth in sectors we associate with formal urban locations - manufacturing, finance, utilities - all had substantially faster productivity growth than other sectors. It seems also valid to assert that amenity growth in formal locations was as least as rapid as in other locations over this period.[30]

Preferences. The utility weight on life expectancy (β) will determine how important changes in the CDR's across locations are to changing the distribution of population. We follow Becker et al. (2005) and Weil (2010) who compare the value of changes in life expectancy with changes in income per capita. Specifically, Becker et al. (2005) convert the changes in life expectancy to equivalent changes in annual income. Using this information, we can back out an implied value of β such that the observed growth in life expectancy is equivalent to their equivalent change in annual income. We use their estimates for Africa, as the majority of our sample comes from that region. Becker et al. (2005) calculate that the increase in life expectancy in this region between 1960 and 2000 (a 12% increase from 41 to 46 years) was equivalent to an increase of 72 intl. dollars (a 4.9% increase on a base of 1,470 dollars). This implies a value of β of 0.41.[31]

[30]While it seems natural to assume that the growth of productivity and amenities may be higher in formal urban areas, our findings below on the relative importance of the UMT on urbanization are not driven entirely by this assumption. Changing G_f to be lower, or equal to G_l and G_r, still shows that the UMT had a meaningful impact on urbanization and the size of informal locations. The absolute quantitative effect becomes smaller as G_f is set lower, but the effect of the UMT remains substantial.

[31]The implied value from South Asia is $\beta = 1.70$ and from Latin America is $\beta = 1.56$. The lower value from Africa, implying less utility weight on life expectancy, is consistent with recent work by Jones (2016) on how people's preference for consumption relative to health falls as they become richer. Regardless, using higher values of β would only increase the implied effects of the UMT on urbanization

Rural welfare elasticity. For the rural location elasticity ϵ_r, sources point to a high negative elasticity of output per capita with respect to population size in predominantly rural countries. Lee (1987) finds an elasticity of -1.6 and Weir (1991) finds -1.2, both using an agricultural production function with a low degree of substitutability between land and population, and pre-Industrial Revolution data. Lee (1997) updated his own estimate to -1.0 more recently. Acemoglu and Johnson (2007) have work on the effect of the international epidemiological transition that implies an elasticity of -1.2 in a sample of low and middle-income countries in 1940-1980, all of which had very low urbanization rates about 10-15% in 1950.[32] This elasticity is derived from the production technology alone. However, we feel this is a good approximation for rural areas, as housing faces few constraints in terms of space, and due to their low densities, the existing amenities and transportation networks in rural areas are less prone to congestion. Hence we use -1.2 as our preferred estimate of ϵ_r.[33]

Urban welfare elasticities. The urban elasticities, ϵ_f and ϵ_l, are estimated by targeting specific moments from the sample of 41 poor countries at the onset of the UMT. We target the average urbanization rate in 2005 and the informal share in 2005. The informal share is set based on the slum share in urban areas (United Nations, 2014). We calibrate these values because there is no clear empirical analog for these values we can draw on. Recall that these elasticities combine the elasticity of wages and amenities with respect to population. While choosing a specific production function for a location might provide the elasticity of wages with respect to population, it does not indicate the elasticity of amenities. Hence we calibrate these values so that they ensure the model matches the observed growth in both urbanization and informal shares over time.

Numerically, solving for the elasticities(ϵ_f and ϵ_l) is not computationally intense, as the model is a series of static allocation problems. The only other information necessary, aside from what has been discussed, is that we normalize the size of the aggregate population to be one in 1950. We use a discrete-time approximation of the

and informalization, as individuals would have more incentive to move to low-mortality urban areas.

[32]-1.2 comes from dividing their estimate of the effect of life expectancy on GDP per capita, -2.43, by the effect of life expectancy on population, 2.04 (see column (3) of their Tables 8 and 9).

[33]A value of this elasticity of one, or more, is still consistent with a constant returns rural production function in which the elasticity of output with respect to labor is less than one. Adding a new non-working resident to a rural area lowers output *per capita* without raising output, with an elasticity of negative one, similar to our preferred value. Further, starting with an assumed value of the $0 > \epsilon_r > -1$, as would be implied by a constant returns production function, alters the calibrated values of the other elasticities, but does not change the effects of the UMT by a material amount.

model, setting one period equal to one year. The solution is found by using a non-linear least squares routine to solve for the values of the parameters that force the model to match the two moments we target: the urbanization rate in 2005 (31.0%), and the informal share in 2005 (64.2%). Table 3 shows the results of the calibration, along with the assumed initial values. We find that $\epsilon_f = 1.32$ and $\epsilon_l = 0.60$. The data suggests that in our specific sample of poor countries formal urban locations have elasticities slightly higher than rural ones, while informal locations are implied to have elasticities about half of the rural value. In other words, their rural and formal sectors are not be able to absorb new residents as easily as their informal sector.[34]

Additional clarifications. First, it may not be immediately clear why welfare in the formal urban locations of these countries would be as sensitive to population size as rural areas, or larger than in informal locations. Formal urban locations presumably enjoy some sort of agglomeration effects, and physical land would appear to be less important in production. However, these developing countries are likely to use backward technologies in production, housing, and transportation that are more susceptible to congestion. The institutional structure may mean that production, housing and transportation cannot, or are not allowed to, expand at a pace equal to that of population. In contrast, informal urban locations can adapt relatively easily to population inflows because their production (e.g. street vending), housing (e.g. shacks), and transportation (e.g. rickshaws) have low fixed costs and are not subject to the same institutional constraints as in formal areas. The difference in formal and informal urban elasticities is entirely consistent with the data we presented in Figure 5(c) comparing mega-cities at different income levels, as well as footnote 16 discussing why slums are so densely populated despite consisting mostly of low-rise structures. High-elasticity built-up (formal) locations in rich mega-cities have low population densities compared to low-elasticity slum (informal) locations in poor mega-cities, even though formal locations have technologies (e.g. skyscrapers) that relieve some space constraints.

Second, the elasticities we calibrate are intrinsically different from the "net agglomeration effects" (i.e. agglomeration benefits net of congestion costs) that have been studied in the literature, where one regresses measures of productivity and pecuniary and non-pecuniary living costs on population for a sample of cities belonging

[34]We can separately calibrate the urban elasticities for each country in our sample. If we do so, the median values of ϵ_f and ϵ_l are essentially identical to those calibrated using the sample averages.

to the same country. Econometrically, the effects that studies estimate are the *relative* effects of moving one person from one city to another larger city within a same country. The effects that we are calibrating are the *absolute* effects of adding an extra person to the urban sector of one country. In that case, the entire urban sector may become more congested and urban welfare may decrease. To our knowledge, this parameter has not been estimated in the literature, hence our need to rely on calibration.[35]

Finally, we consider two broad urban locations, as opposed to tracking individual cities that could possibly vary in their elasticities or inherent wage/amenity growth rates. If there were substantial variance in the elasticities or wage/amenity growth across cities, then this should show up as changes in the distribution of city sizes within economies over time. However, for our sample, there is little evidence of substantial changes in this distribution. Crude rates of natural increase were effectively identical in mega-cities as secondary cities throughout our period of study (see Figure 4). In addition, the average primacy rate (the share of the largest city in urban population) for our sample of 41 countries in 1950 was 31.3%, and 32.8% in 2005, for essentially no change. This suggests that we may not be missing variation across cities when we track outcomes in different urban locations at the aggregate level. The poor mega-cities we discussed earlier in the paper are the most visible examples of the rapid growth experienced by all cities in these countries, but the phenomenon is common to all.

5.2. Implications of Calibration for the Sample of Poor Countries

Before examining quantitative exercises, we can examine some general implications of our calibrated parameters for the UMT and the growth of mega-cities in our sample of poor countries. The acceleration of population growth should have had several effects:

1. The economy urbanized more quickly. The UMT had a direct effect on urbanization by making urban locations more desirable to live in given individual's preference for lower crude death rates. In addition, there was an indirect effect of the UMT on urbanization given the differences in elasticities between locations. The values of the elasticities, along with the initial shares of population indicate that $\epsilon_r > \bar{\epsilon}$. Hence in response to the increase in population growth from the UMT, the growth rate of the

[35] In addition, with the notable exception of Duranton (2014a), there are almost no studies of these *relative* congestion effects in developing countries. Further, there are no almost studies of these *relative* agglomeration effects for most of the 41 poor countries (incl. 28 African countries) in our sample. Chauvin et al. (2016) focus their analysis on Brazil, China and India, three middle-income countries.

rural population share, $\hat{s}_r$, fell, and hence urbanization occurred faster.

2. Urbanization occurred primarily in informal locations. With $\epsilon_l < \bar{\epsilon}$ the growth rate of the informal share, $\hat{s}_l$, rose in response to the UMT. At the same time, with $\epsilon_f > \bar{\epsilon}$, the growth rate of the formal share was pulled down by faster population growth in the UMT. The UMT not only urbanized the economy, but did so by pushing people into the location with the lowest elasticity of welfare - informal areas.

3. Population growth sped up. The immediate effect of the UMT is to raise $\hat{N}$ directly. But in addition, by urbanizing the economy primarily through informal locations, the UMT led to higher aggregate $CRNI$ as population moved into the relatively high-growth informal locations. The UMT thus eliminated the natural limitation on population growth that historical economies experienced as they urbanized.

4. Ambiguous welfare effects. By raising $\hat{N}$, welfare growth was pushed down directly due to negative effects of population growth in all locations.. But in addition, because informal locations have relatively low values of G_l, the increased share of population in informal locations meant that aggregate welfare growth became slower after the UMT. On the other hand, because the crude death rate fell, there was an increase in welfare given that people value lower mortality. On net, there is no obvious conclusion to draw analytically, and this remains a quantitative question. It is important to note with respect to welfare that the shift into informal location is the welfare-maximizing response to the UMT, and there is no spatial misallocation. There is a negative externality, as individuals do not take into account their effect on congestion, but individuals are not constrained from moving between locations.

5.3. Effects of the UMT, 1950-2005

In Table 4 we work through several scenarios to quantify the role of the UMT. Row 1 shows observed data on the urbanization rate, both in 1950 (8.9%) and 2005 (31.0%), for our sample of 41 poor countries. In addition, it shows the relative size of urban populations in 2005, and urban populations grew by a factor of 15.2 from 1950 to 2005 in our sample. We do not have data on slum shares in 1950, which we use to measure the informal share. But in 2005, this share was 64.2%. The relative size of informal locations was 18.6 in 2005 compared to 1950, based on our presumed slum share of 50% in 1950.[36]

[36]The relative size of the total population in 2005 was roughly 5 times that of 1950. Productivity growth (or lower international transport costs for food as in Glaeser (2014)) in the economy must have been sufficient to either produce, or import, enough food to support this increase in population while allowing

Row 2 presents outcomes from our baseline calibrated model. We take the initial urbanization and informalization rates as given, and target the final urbanization and informalization rates of 31.0% and 64.2%, and so our calibration delivers these values exactly. For urban size, the calibration finds that urban areas are larger by a factor 14.8, compared to the observed value of 15.2. For the relative size of informal areas, the calibration delivers a value of 19.0, compared to the actual value of 18.6. As a normalization, we set the value of welfare in 2005 to be 1.00 in our calibration.

Removing the UMT. In row 3, we rerun the model from 1950 to 2005, but this time we remove the UMT. The formal and informal CDRs stay at 40 throughout. In this case, the urbanization rate would only have been 21.3% in 2005. Compared to the observed data, this indicates that the UMT accounted for 9.7 percentage points of urbanization between 1950 and 2005 (31.0 minus 21.3), or roughly 45% of the increase in the urbanization rate in this period. Urban areas in 2005 were roughly 75% larger (14.8 versus 8.4) because of the UMT. The informal share without the UMT would have been 54.8%, compared to 64.2% in the observed data, and thus 9.4 percentage points of the informal rate in 2005 can be attributed to the UMT. This is roughly 70% of the increase in the informal rate from 1950 to 2005. Without the UMT informal areas in 2005 would have been 9.2 times larger than in 1950, compared with the actual ratio of 18.6. Hence informal areas are twice as large today due to the UMT. From a welfare perspective, there are two conflicting effects at work. First, without the UMT mortality is higher and hence welfare is lower. Second, slower population growth reduces the drag on welfare due to crowding effects on wages/amenities. In our calibration, the net result is that welfare in 2005 could have been 8% *higher* without the UMT, suggesting that the latter effect of less crowding might have outweighed the former effect of higher mortality.

We now attempt to quantify the effects by which the UMT create higher urbanization and informal shares. There is the direct effect of lower mortality drawing people into cities because of their preferences for lower mortality. Then there is also the differential elasticity effects, where population move into locations with low congestion elasticities.

Direct congestion effects. In row 4, we rerun the model with the UMT occurring, but set the preference parameter on life expectancy, β, to zero. Thus the UMT does not provide any direct incentive to move to urban areas. Urbanization in this case is driven

for higher urbanization rates. Our wage/amenity growth rates implicitly take this into account, and this is why we have not explicitly accounted for food demand or subsistence constraints in the model.

only by differentials in the underlying wage/amenity growth rates, G_i, and differentials in elasticities, ϵ_i. Here, the results show that urbanization in 2005 would have been 22.7%, or 1.4 percentage points higher than without the UMT. Similarly, an informal share of 59.1, compared to 54.8 without the UMT, indicates that these effects added 4.3 percentage points to the informal share of urban areas. Welfare in this scenario is reported as 88% of the baseline welfare *with* the UMT, but this is attributable to removing the benefits of lower death rates from the welfare calculation.

Life expectancy effects. Row 5 looks at the effects of the UMT arising purely from preferences for lower death rates. We rerun the model including the UMT, but we set the values of the elasticities, ϵ_i, in all locations to be equal.[37] The urbanization rate is 25.0, which is 3.7 percentage points higher than without the UMT. For the informal share, the implied size is only 21.2%, meaning that informal areas would have shrunk as a share of urban areas. When we set the informal location elasticity equal to the other locations, the advantages of this location dissipate. Compared to formal locations, informal areas have $G_l < G_f$, and so there is little reason to move to informal areas.

Life expectancy vs. direct and indirect congestion effects. Roughly one-third (3.7/9.7) of the urbanization due to the UMT can be accounted for directly by the preference for lower mortality rates. The remaining two-thirds is due both to the direct effect of differential elasticities (row 4) as well as the interaction of the mortality preference with these differential elasticities. The low elasticity in informal areas means that when individuals move to these locations out of a desire for lower mortality, the wages/amenities in informal areas do not fall much, and hence even more individuals can move into the urban areas in search of lower mortality rates. The two effects reinforce each other. Without the lower informal elasticity, urbanization would have been much lower following the UMT. The role of differential elasticities is also present in the rise of informal shares. Preferences for lower mortality (row 5) alone would have *shrunk* the informal sector as a share of urban population, as people would have moved almost exclusively to formal locations if the elasticities are equal across locations. But in row 4, we see that informal areas would have grown to 59.1% of urban population due solely to the lower elasticities in informal areas. The combination of the two, with the low elasticity amplifying the effect of mortality preferences, is necessary to account for

[37]We use $\epsilon_i = 1.014$, the harmonic mean of the three sectoral elasticities from the baseline model.

the 9.4 percentage point effect of the UMT on the informal share (64.2% versus 54.8%). The lower elasticity in informal areas accounts for more than 100% of the growth in the informal share from the UMT. While one cannot take a percentage greater than 100% too seriously, it indicates how important the differential elasticity effects were.[38]

5.4. Endogenous Fertility and Effects of the UMT, 1950-2005

So far, we held the fertility behavior equal to that observed in the data, and this allowed us to make comparisons between the scenarios. Here, we show that the implied impact of the UMT on urbanization and informalization may be *understated* by those results.

Fertility bulge. Under almost any reasonable setting, endogenous responses of fertility to the UMT will act to make crude birth rates *higher*. This occurs through several channels. First, the drop in crude death rates contributes to higher crude birth rates by ensuring cohorts survive longer through child-bearing ages. This population momentum effect is larger, the faster is the decline in mortality (Heuveline, 1999; Guillot, 2005). Second, while total fertility rates may fall after declines in mortality due to risk-aversion or the value of human capital (Kalemli-Ozcan, 2002; Soares, 2005), the net fertility rate will not necessarily fall, and will likely increase (Doepke, 2005). Last, if fertility is linked to the opportunity cost of parents time (Becker, 1960), then lower wages induced by congestion following the UMT would act to raise fertility rates.

Regardless of the exact mechanism, the three mechanisms explained above all indicate that in the *absence* of the UMT, the crude birth rate would have been *lower*. That is, our counter-factual simulation in row 3 of Table 4 likely understates the effects of the UMT on urbanization because it holds constant the observed time path of crude birth rates, which includes their rise from 1950 to 1985. If we incorporate endogenous fertility, this will result in lower fertility rates, and through the mechanisms present in the model, result in less urbanization, smaller informal shares, and higher welfare.

Constant fertility. As a crude means of establishing the possible consequences of endogenous fertility, we first show in row 6 of Table 4 a counter-factual in which we remove the UMT, and also hold crude birth rates constant at their 1950 level. The results

[38] The overall analysis and decomposition could be different if we were able to account separately for child and adult mortality changes from the UMT. Drops in urban child mortality may create stronger amenity effect drawing people to cities, which may imply an even bigger role for the UMT. Drops in urban adult mortality may have less of an amenity effect, but could lead to bigger congestion effects if adults take more "space" than children. Using national-level data, we find that adult and child mortality each account for about half of the overall decline in crude death rates during the UMT, so our decomposition is not necessarily over- or under-estimating the amenity versus congestion effects.

show that urbanization and informalization rates would have been lower, urban and informal sizes smaller, and welfare higher, than in our first counter-factual (row 3).

Endogenous fertility. To be more sophisticated we introduce endogenous fertility into the model. We cannot incorporate age-specific effects related to population momentum due to the lack of location-specific data on age distributions and demographic behavior in this time period. We can, however, build in responses based on wages and death rates (details available in the appendix). The crucial elements are as follows. First, the fertility decision is assumed to depend on both the wage, w_j, and death rate, CDR_j, in location j. Second, in line with literature, birth rates are negatively related to wages and death rates (note that this does not imply that total fertility rates are necessarily negatively related to death rates). Specifically, the elasticity of fertility with respect to wages is set to -0.30, and the elasticity of fertility with respect to the death rate is set to -0.30, based on outside sources. The utility weight on fertility is set to be equal to that on life expectancy, also consistent with outside evidence.[39] As can be seen in row 7, the results with endogenous fertility are similar to those where we hold birth rates at their pre-UMT levels in 1950. The higher wages induced by lower population growth in turn push down fertility, which in turn helps keep wages high. Hence the overall population sizes are lower with endogenous fertility than with our prior results with a fixed fertility regime. Overall welfare is higher due to smaller population size.

5.5. Long-run Effects of the UMT, 1950-2100

With the UMT. We simulate outcomes 100 and 150 years from our start date (i.e. to 2050 and 2100). Rows 1-2 of Table 5 show the results under the assumption the UMT takes place starting in 1950, and persists over time, with the CDR in urban locations asymptoting towards 15. In row 1, we allow the birth rate to continue to follow the hump-shaped pattern discussed above, but put a floor on the CBR in each location of the value it had in 1950. Thus the fertility bulge that followed the UMT is temporary, but birth rates remain high. In row 2, fertility is determined endogenously. Regardless of the fertility option, one can see that by 2050 the urban share and the slum share would be around 50% and 65-70%, respectively. Urban areas and informal areas would also be much larger than in 1950. Using the first scenario as a baseline to normalize utility,

[39] It is not necessary that the elasticities be exactly the same value (-0.30), this was simply a coincidence. We examined outcomes with values ranging between -0.05 and -0.75 for the elasticity of fertility with respect to the crude death rate and wages, and the general results are consistent (not shown).

we see that under the endogenous fertility regime the increased size of the population lowers welfare to 78% of the baseline. Population growth pushes down wages, and this in turn pushes up fertility, which exacerbates population growth, and so on.

After 150 years, in 2100, the calibration suggests that urbanization would be 69-78%, with informal areas still making up around two-thirds of urban locations. The relative sizes of urban areas, and informal areas, are massive, although the numbers are obviously speculative. Welfare is far lower in the endogenous fertility case, emphasizing the possibility of a downward spiral in welfare. With lower death rates following the UMT, this leads to higher population size and hence lower wages (and amenities) across locations. This in turn generates higher fertility, which only exacerbates the problem.

Removing the UMT. Rows 3 through 5 perform a similar analysis, but remove the UMT, and leave urban death rates at their 1950 levels throughout. They differ in their assumed fertility regimes as well. Row 3, similar to row 1, allows birth rates to follow the hump-shaped pattern seen in the data, with a floor of the 1950 birth rates imposed. Row 4 holds birth rates constant at their 1950 levels in all periods, eliminating the hump-shaped response of fertility to the UMT. Finally, row 5 allows for fertility to be determined endogenously. In each case, one can see that without the UMT, urbanization rates after 100 years are predicted to be only around 33%, about 20 percentage points lower than with the UMT. Urban sizes range between 15-30 times larger than in 1950, only about a third of the size with the UMT. The informal share without the UMT is predicted to be under one-half. Welfare, despite the lack of the UMT keeping mortality high, is higher than in the baseline with the UMT, due to the slower population growth.

When these simulations are extended out in time until 2100, urbanization rises to roughly one-half, and the informal shares continue falling to between 12% and 36%. Consistent with this, welfare in each simulation is predicted to be higher than in our baseline with the UMT. While we do not want to make too much of these numbers, note that welfare is highest under the endogenous fertility regime, where the lack of a UMT now allows for a virtuous spiral to occur, where population growth is low, wages are high, and hence fertility continues to fall, which limits population growth further, and so on.

Table 5 suggests that demographic conditions can have large long-run effects on the "nature" of urbanization. Note that in rows 3-5, without the UMT, the relative size of urban areas do become large. By 2100 urban areas are predicted to be between 19.3 (row 5) and 85.1 (row 3) times larger than in 1950. Urbanizing without the UMT does not preclude mega-cities from forming, but it does allow them to develop into "rich" mega-cities. This long, slow, urbanization process with an urban CDR of 40 is consistent with the path followed by rich countries. In 1800, both Europe and the U.S. had an urbanization rate of about 10%, similar to the starting position of our sample. At the eve of WW2, roughly 150 years later, their urbanization rate was about 55%, close to our simulated values that range from 46-51%. The dominance of formal locations, and the associated improvement in welfare, are qualitatively consistent with the development of these countries over this period. The model predicts informal shares of 12.5-36.5%. We believe this is within range of the slum share of rich country urban areas around WW2.[40] Relative urban size varied greatly when comparing the U.S. and Europe between 1800 and WW2. In the U.S., urban areas became 150-250 times larger over this period, depending on the dates chosen. Our simulations do not indicate such a massive gain, but our model is of a closed economy while the U.S. in this period accepted massive amounts of immigrants. For the United Kingdom, the urban population was 16 times larger in 1950 than in 1800, roughly in range with our endogenous fertility scenario.

5.6. Simulating Policy Effects, 2005-2055

Table 6 shows the effects of various policy experiments that might be used to alleviate the conditions in, or slow the growth of, poor mega-cities.

Baseline. Row 1 shows our starting point in 2005, with urbanization of 31.0% and an informal share of 64.2%. We normalize the size of the urban sector and of the informal sector to one in 2005. The second panel presents the results of the simulations, for which we use the parameters of Table 3, but update the demographic rates to match observed 2005 rates.[41] Row 2 shows our baseline outcomes 50 years later in 2055, i.e. an urbanization rate of 50.8%, and an urban size 5.2 times that in 2005. As a comparison, United Nations (2014) projects an urbanization rate of 54%, and a relative size of 3.2 for the countries in our sample. Our baseline overstates the size, due primarily to the fact

[40]For example, the 1950 U.S. housing census reports that 21.2% of the U.S. urban population lived in "dilapidated" housing units and/or units without running water, which are characteristic of slums. Another example is London where according to the 1921 census, 17.7% of people lived in "overcrowded" dwellings, and 25% and 15% of dwellings were occupied by at least 2 and 3 families, respectively. One last example is Paris, where according to *Annuaires Statistiques de la Ville de Paris* the share of "insalubrious" housing units was about 35% in the 1930s, half of which were considered "dangerous" to live in.

[41]We use the sample averages in the location-level data for the 19 poor countries (minus China, due to the one-child policy) in the 2000s. Birth rates in informal locations and rural locations are set to 35, and formal locations to 20. Death rates are set to 10 in rural locations, and 7.5 in urban locations. We find similar results if we replicate the entire analysis in Table 6 using endogenous fertility (not shown).

that our baseline holds rural demographic rates constant at relatively high rates over the entire period. The informal share of urban areas remains roughly the same as in 2005, at 63.6%. Welfare in 2055 is normalized to one for comparison to the alternatives below.[42]

Productivity growth. One strategy for alleviating the conditions in poor mega-cities is to foster rapid productivity growth, either through industrialization in formal locations or "green revolutions" in rural ones. To see these effects, we separately raise productivity/amenity growth by one percentage point in each sector. Raising rural productivity/amenity growth to $G_r = 0.035$ (row 3) reduces urbanization to 38.3%, while welfare is 24% higher in this case. Compared to this, raising informal productivity/amenity growth to 3.5% (row 4) increases welfare by 30%, even though urbanization would be 61.2% and the informal share 75.5% in this scenario. Welfare is higher in the informal case because we are raising wage/amenity growth in a location with a low congestion elasticity, and this allows a large group of people to move into that location to take advantage of the higher level of G_l. Raising G_f to 6% (row 5) for formal locations increases urbanization rates slightly, but also lowers the informal share. Welfare is 10% higher in this scenario, lower than in the informal scenario (row 4) because the congestion elasticity in formal locations is much higher. Rows 6 and 7 increase the formal rate of wage/amenity growth to 8% and 10%, respectively, to explore the effects of rapid industrialization. In both cases, the urbanization rate rises compared to the baseline economy, as far as 73.9% in row 7. The informal share falls dramatically, down to 11% in row 7. The welfare gains appreciate very quickly as G_f rises.

Lower formal elasticity. An alternative policy is trying to lower the cost of population growth in locations with high wage/amenity growth (i.e. formal locations) by improving urban planning, relaxing land-use regulations, building additional infrastructure in existing megacities, facilitating the growth of less congestable secondary cities, and/or creating well-planned large cities ex nihilo. There are reasons to believe that the formal location elasticity is lower in rich countries, since they use better urban production, housing and transportation technologies, and have better urban institutions. However, we cannot estimate their elasticities using the same computation exercise as for poor countries, since this would require us having historical century-old data on slum shares for enough rich countries. But we can test how the poor mega-cities would evolve if

[42]The pattern of results is similar if we simulate the policies for 100 or 150 years forward.

we were to lower their formal elasticity, i.e. the "absorption capacity" of their formal sector. In row 8, we examine the effect of lowering the formal sector elasticity of 1.32 by about 10% to the value we use for rural locations (1.20). This raises the urbanization rate slightly to 52%, while lowering the informal share slightly to 58.9%, and welfare is 4% higher. A more dramatic change is seen in row 9, where we lower the formal sector elasticity by about 50% to equal the informal sector elasticity (0.60). In this case, the urbanization rate rises to 67.3%, as more workers are able to move into formal sectors with little effect on wages/amenities. This shows up as well in the informal share falling to 19.4%, and this implies a welfare gain of 82%. Row 10 drops the formal elasticity even further to $\epsilon_f = 0.44$, which is about two-third lower than the calibrated value of 1.32. Here, the same pattern of results holds, with urbanization much higher at 74.8%, an informal share of only 9.9%, and welfare 162% higher. The welfare effect is then similar to the effect of raising formal location wage/amenity growth to 10% per year.[43]

Higher informal congestion. Instead of improving formal locations, policies may be aimed explicitly at reducing the attraction of informal locations. This could take the form of enforcing property rights within these locations (or regulations in their sectors), or physically destroying these locations, i.e. clearing slums as Haussmann did in Paris in 1853-1870. In row 11 we simulate these policies by raising the informal elasticity to the level of the formal elasticity, $\epsilon_l = 1.32$. This lowers the informal share to 41.6%, and the size of the informal population is smaller, at only 2.5 times the 2005 starting value. However, welfare is lower, at only 79% of the baseline. Lower welfare is a result of raising the average elasticity of welfare with respect to population in the economy. Informal areas, while they have low productivity growth, provide an outlet for population growth that alleviates congestion effects in the aggregate. Making them unable to absorb that population reduces their size at the cost of lower welfare for all individuals.

Migration restrictions. Another policy approach, such as with China's hukou system,

[43]A formal elasticity half of the value we calibrated for poor countries could make sense for rich countries, although this is entirely speculative. The fact that poor megacities are more highly congested than rich megacities, whether in terms of housing, infrastructure or unskilled workers, is consistent with this difference in elasticities. First, while the price elasticity of housing supply ranges from 0.5-2.0 in developed countries (Caldera and Johansson, 2013), it ranges from 0.1-0.5 in developing countries (see Malpezzi and Mayo (1997) for Malaysia and Lall et al. (2007) for Brazil). Second, various sources indicate that an one-way commute is about 40 min in rich megacities and 80 mins in poor megacities. Struyk and Giddings (2009) writes: "One-way average commute times in Jakarta, Kinshasa, Lagos, and Manila are over 75 minutes." Lastly, another example of congestion is that various sources show rich megacities have pupil-teacher ratios in primary education around 15-20, while in poor megacities these ratios range from 30-70 (e.g. 69 in Dar es Salaam according to the World Bank's *Service Delivery Indicators Report*).

is to limit the growth of poor mega-cities by limiting in-migration to cities. We can evaluate the effect of such policies by adding a migration restriction to our model that generates a wedge in welfare growth between locations. Let the value λ_{ur} measure the wedge between welfare growth in rural and urban areas. If $\lambda_{ur} = -0.01$, for example, this means that migration restrictions keeping rural residents from leaving lead to 1% slower growth in welfare in rural locations compared to urban ones (either formal or informal). For our simulation, we set $\lambda_{ur} = -0.01$ in row 12. These kinds of migration restrictions limit urbanization to 38.3%, while city size and informal size are also smaller. However, welfare is only 90% of the baseline. The migration restrictions lower welfare by forcing a large portion of the population to stay in relatively high-congestion rural locations.[44]

In row 13, we flip the migration restrictions, imposing limits on the movement of urban people into rural areas. This may reflect either an explicit urban-biased policy, or the fact that with the UMT an increasing share of urban residents are urban-born, and urban-born individuals may acquire from birth a strong preference for urban living. Setting $\lambda_{ur} = 0.01$, which implies faster welfare growth in rural areas, this raises urbanization rates to 62.8%, well above the baseline, and informal areas would be larger. In order to achieve higher rural welfare growth, even more people must enter cities to keep population low in the rural areas. However, in this scenario welfare is found to be 11% *higher* than in the baseline. The reason is that rural workers are, for much of the simulation, a majority of the workers (i.e. the initial urbanization rate is only 31%), and their higher welfare growth due to the restrictions implies a higher weighted average.

Family planning. Finally, we consider policies regarding population growth, which are not generally thought of as urban policies, but which we show can have a significant impact on poor mega-city growth and welfare. In row 14 we impose zero population growth by setting the CBR and CDR to be equal in each location. This results in urbanization of only 39.8% by 2055, and urban size is only 1.3 times the 2005 value. The informal share of urban areas drops to 38.1%, and the size of these informal areas falls to only 80% of their size in 2005. Welfare is 213% higher than in the baseline. Lowering population growth slows down the growth of the low-congestion informal locations, resulting in lower urbanization and informal shares, and higher welfare.

[44] We obtain similar results if we extend the model to see how slum-constraining policies impact poor mega-city growth when there are cross-location congestion effects (not shown). We could indeed imagine that informal areas in cities create congestion effects in the formal areas of the same cities.

Zero population growth is an extreme policy, and in row 15 we explore a policy of family planning, making all crude birth rates equal to 20, similar to those achieved in China, Singapore or South Korea during the late 1970's (Bangladesh, Ethiopia and Indonesia have been touted as family planning success stories but they still have birth rates above 20). 20 is also the average birth rate that we obtained for the formal areas of 8 poor mega-cities today (see section 3.1.). Urbanization would only be 44.2% in 2055, and urban size only about 2.4, while the informal share would fall to 50.3%, and informal areas would only be 1.9 times as large. Welfare would be 82% higher. In rows 16 and 17, we show separately the effects of family planning in informal and rural areas, respectively. Informal family planning lowers urbanization and informal shares relative to the baseline scenario, and raises welfare by 19%. Rural family planning has larger effects, lowering urbanization to 45.8% and informalization to 54%, while raising welfare by 55%. The larger effect of rural family planning is due to the relatively large size of the rural population, and hence lowering birth rates in these locations would have a larger impact on aggregate population growth, and hence on welfare. However, rural family planning may be costlier to implement than urban family planning due to the rural population being much more scattered spatially.

5.7. Discussion of Policy Implications

Here we note briefly the possible costs and feasibility of these policies. The most direct way of improving urban conditions in poor countries is to raise the formal wage growth rate. Raising this rate from 5% to 10% per year would more than double welfare and reduce the size of informal areas within cities. The open question here is how exactly one accomplishes this increase in the growth rate. There is not an obvious set of policies that would generate an industrial revolution of the type experienced by China recently, especially as many countries in our sample do not appear to have a comparative advantage in manufacturing (Gollin et al., 2015; Venables, 2016).

Lowering the formal elasticity is promising. Creating new urban locations by "better building elsewhere" is one strategy. First, it may have massive costs. For example, we find that the cost of creating new large cities may be between 3 and 25% of one year of GDP (mean of 12.5%), based on information from Brasilia in Brazil (1956-1960; 11-13%), Islamabad in Pakistan (1960s; 6%), Abuja in Nigeria (1978-1991; 23-25%), 50 ghost cities in China (2000s-2010s; 7-24%), Kilamba outside Luanda in Angola (2008-2012;

3%), and the future potential new capital city of Egypt (2020s; 13%). Second, it does not appear that countries that have created new megacities in our sample period have shown any significant reduction in slum shares or infrastructure congestion in their other megacities.[45] Third, these cities may be motivated more by a desire to isolate elites within a country rather than a desire to alleviate congestion (Campante et al., 2013).

"Building back better" in existing cities is another strategy. Improving both urban governments and markets (e.g. forward-looking urban planning, documenting property rights, lowering the costs of land transactions, easing of urban regulations, reducing corruption in local government) would then certainly help poor mega-cities (Glaeser, 2014; Brueckner and Lall, 2015; Henderson et al., 2016b; Ashraf et al., 2016). It may not be feasible, as building back better may be socially costly. It requires clearing slums, which involves a massive displacement of population. Further, there is a question of whether poor countries will try to address their own government and market failures, as politicians may benefit from the status quo (e.g. some politicians are slum lords).

Compared to lowering formal elasticities and raising formal growth rates of productivity, Table 6 suggests that population policies may be just as effective in managing urban growth in the future. While these are not typically thought of as urban policies per se, population growth itself is a significant contributor to the pace of urbanization and the type of urban locations it occurs in. Today, most African countries and the poorest Asian and Latin American still have birth rates of around 40 per thousand. Family planning programs aimed at lowering the CBR to 20 limit the size of slums and raise welfare by up to 82% if extended across the economy. Hong Kong, Singapore, and South Korea all implemented similar "two is enough" campaigns in the 1960's and 1970's, bringing down their birth rates dramatically.[46] The WHO estimates costs for family planning programs that would lower birth rates to 20 of 1.2% of one year of GDP in Sub-Saharan Africa (Cleland et al., 2006). These policies may thus be more cost-effective than rebuilding existing cities or building new cities. The other advantage

[45] Ten of the poor megacities in our sample are in countries that have massively invested in the creation of a new capital city: Karachi and Lahore in Pakistan (Islamabad became the capital city in 1960), Belo Horizonte, Rio and Sao Paulo in Brazil (Brasilia; 1960), Dar es Salaam in Tanzania (Dodoma; 1974), Abidjan in Ivory Coast (Yamoussoukro; 1983), and Ibadan, Kano and Lagos in Nigeria (Abuja; 1991).

[46] The World Bank (2007) documents that family planning policy is not based on income levels. China, India, Indonesia, South Africa and South Korea all adopted strong family planning policies at very low income levels, while many African, Central and South American, and Western Asian countries had weak family planning policies given their income level for most of the post-WWII period.

of such policies is that they are usually implemented at the national level, which could allow national governments to bypass the problem of weak urban institutions.

Lastly, similar to the conclusions of Au and Henderson (2006), Feler and Henderson (2011) and Desmet and Rossi-Hansberg (2014), migration restrictions are counter-productive. They limit both the size and informal share of cities, but at the cost of decreasing aggregate welfare as they do not allow population to flow into the low elasticity locations.[47] Policies meant to eliminate slums *lower* welfare because they eliminate a valuable outlet for rapid population growth. Ultimately, our analysis shows that it was a surge in population growth that helped create these poor mega-cities, and reversing that surge could be an efficient way to improve living standards in mega-cities.

6. Conclusion

Poor mega-cities are a relatively new phenomena, appearing in the post-war era across the developing world. Using a novel dataset of location-level demographics over time, we document that poor mega-cities arose at the same time as the urban mortality transition (UMT) following WW2, which lowered their urban mortality rates to rich-country levels. To assess the quantitative importance of the UMT for poor mega-city development, we develop a general equilibrium model of location choice that has heterogeneity in congestion costs and demographics across locations. The model shows that population growth has a direct effect on the distribution of population across locations. Calibrating our model to a sample of poor countries, we find that the UMT may have doubled their urbanization rate as well as the size of their slums between 1950 and 2015. One-third of these effects can be traced to direct preferences for living in locations with lower mortality rates, while the remaining two-thirds are accounted for by faster population growth pushing people into informal locations. The calibrated model allows us to assess various policies, and we find that family planning programs may be as effective as higher productivity or urban infrastructure and institutions, and more effective than migration restrictions and slum clearance, at transforming the mega-cities of the developing world into "engines of growth".

[47] Au and Henderson (2006) find that migration restrictions have resulted in many undersized cities, with net output per worker 17% lower at the 50th percentile of cities, suggesting welfare costs to restrictions even higher than those we estimate (a 10% welfare loss) if agglomeration effects are this strong in China. Desmet and Rossi-Hansberg (2014) find that reducing spatial differences in efficiency and amenities in China to the levels seen in the U.S. would improve welfare by 17.7% and 22.6% respectively.

REFERENCES

Abdel-Rahman, Hesham M. and Alex Anas, "Theories of systems of cities," in J. V. Henderson and J. F. Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4 of *Handbook of Regional and Urban Economics*, Elsevier, 2004, chapter 52, pp. 2293–2339.

Acemoglu, Daron and Simon Johnson, "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, December 2007, *115* (6), 925–985.

Ades, Alberto F and Edward L Glaeser, "Trade and Circuses: Explaining Urban Giants," *The Quarterly Journal of Economics*, February 1995, *110* (1), 195–227.

Albouy, David, "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas," NBER Working Papers 14472, National Bureau of Economic Research November 2008.

—, "What Are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities," *The Review of Economics and Statistics*, 2016, *98* (3), 477–487.

— and Bryan Stuart, "Urban Population and Amenities: The Neoclassical Model of Location," NBER Working Papers 19919, National Bureau of Economic Research, Inc February 2014.

—, Gabriel Ehrlich, and Yingyi Liu, "Housing Demand and Expenditures: How Rising Rent Levels Affect Behavior and Cost-of-Living over Space and Time," Mimeo 2015.

—, Kristian Behrens, Frederic Robert-Nicoud, and Nathan Seegert, "Are Cities Too Big? Optimal City Size and the Henry George Theorem Revisited," Mimeo 2015.

Allen, Robert C., Jean-Pascal Bassino, Debin Ma, Christine Mollâ-Murata, and Jan Luiten Van Zanden, "Wages, Prices, and Living Standards in China, 1738-1925: in Comparison with Europe, Japan, and India," *Economic History Review*, 2011, *64* (s1), 8–38.

Ambrus, Attila, Erica Field, and Robert Gonzalez, "Loss in the Time of Cholera: Long-run Impact of a Disease Epidemic on the Urban Landscape," 2015.

Ashraf, Nava, Edward L. Glaeser, and Giacomo A. M. Ponzetto, "Infrastructure, Incentives, and Institutions," *American Economic Review*, May 2016, *106* (5), 77–82.

Ashraf, Quamrul and Oded Galor, "Dynamics and Stagnation in the Malthusian Epoch," *American Economic Review*, August 2011, *101* (5), 2003–41.

—, David Weil, and Joshua Wilde, "The Effect of Fertility Reduction on Economic Growth," *Population and Development Review*, 2013, *39* (1), 97–130.

Au, Chun-Chung and J. Vernon Henderson, "Are Chinese Cities Too Small?," *Review of Economic Studies*, 2006, *73* (3), 549–576.

Barro, Robert J and Gary S Becker, "Fertility Choice in a Model of Economic Growth," *Econometrica*, March 1989, *57* (2), 481–501.

Becker, Gary S., "An Economic Analysis of Fertility," in "Demographic and Economic Change in Developed Countries" NBER Chapters December 1960, pp. 209–240.

—, Edward L. Glaeser, and Kevin M. Murphy, "Population and Economic Growth," *American Economic Review*, September 1999, *89* (2), 145–149.

Becker, Gary S, Kevin M Murphy, and Robert Tamura, "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy*, October 1990, *98* (5), S12–37.

Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares, "The Quantity and Quality of Life and the Evolution of World Inequality," *American Economic Review*, March 2005, *95* (1), 277–291.

Bleakley, Hoyt, "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics*, 02 2007, *122* (1), 73–117.

—, "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure," *American Economic Journal: Applied Economics*, April 2010, *2* (2), 1–45.

Block, Steven, "The Decline and Rise of Agricultural Productivity in Sub-Saharan Africa Since 1961," NBER Working Papers 16481, National Bureau of Economic Research 2010.

Bloom, David E., David Canning, and Gunther Fink, "Disease and Development Revisited," *Journal of Political Economy*, 2014, *122* (6), 1355 – 1366.

Brueckner, Jan K. and Somik V. Lall, "Cities in Developing Countries," in "in," Vol. 5, Elsevier, 2015, chapter Chapter 21, pp. 1399–1455.

Buera, Francisco J. and Joseph P. Kaboski, "The Rise of the Service Economy," *American Economic Review*, October 2012, *102* (6), 2540–69.

Caldera, Aida and sa Johansson, "The price responsiveness of housing supply in OECD countries," *Journal of Housing Economics*, 2013, *22* (3), 231–249.

Campante, Filipe R., Quoc-Anh Do, and Bernardo V. Guimaraes, "Isolated Capital Cities and Misgovernance: Theory and Evidence," NBER Working Papers 19028 May 2013.

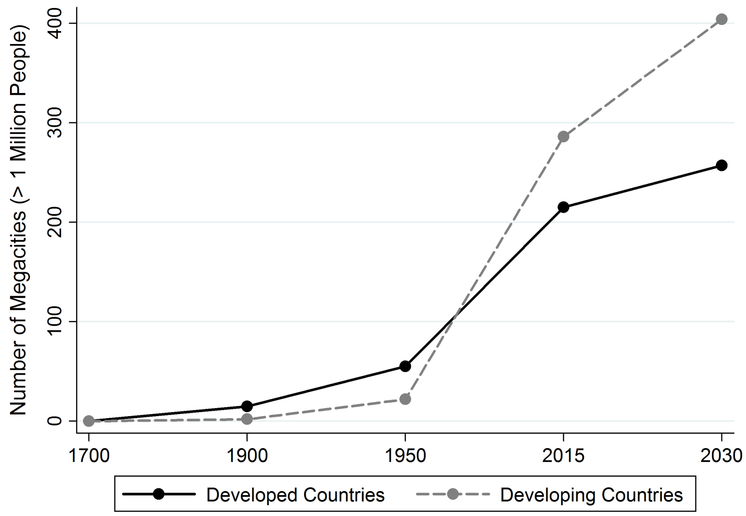
Caselli, Francesco and Wilbur John Coleman II, "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, June 2001, *109* (3),

584–616.
Cervellati, Matteo and Uwe Sunde, "Human Capital Formation, Life Expectancy, and the Process of Development," *American Economic Review*, 2005, *95* (5), 1653–1672.
Chandler, Tertius, *Four Thousand Years of Urban Growth: An Historical Census* 1987.
Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio, "What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States," *Journal of Urban Economics*, 2016, pp. –.
Cleland, John, Stan Bernstein, Alex Ezeh, Anibel Faundes, Anna Glasier, and Jolene Innis, "Family Planning: the Unfinished Agenda," *The Lancet*, 2006, *368* (9549).
Costa, Dora and Matthew E. Kahn, "Public Health and Mortality: What Can We. Learn from the Past?," in David Card Alan Auerbach and John Quigley, eds., *Poverty, the Distribution of Income, and Public Policy*, Russell Sage volume in honor of Eugene Smolensky, 2006.
Davis, James C. and J. Vernon Henderson, "Evidence on the Political Economy of the Urbanization Process," *Journal of Urban Economics*, January 2003, *53* (1), 98–125.
Davis, Kingsley, "The Amazing Decline of Mortality in Underdeveloped Areas," *American Economic Review Papers and Proceedings*, May 1956, *46*, 305–18.
Demographia, *World Urban Areas (500,000+): Population, Density.* 2014.
Desmet, Klaus and Esteban Rossi-Hansberg, "Urban Accounting and Welfare," *American Economic Review*, October 2013, *103* (6), 2296–2327.
_ and _ , "Analyzing Urban Systems : Have Megacities Become too Large?," Policy Research Working Paper Series 6872, The World Bank May 2014.
_ and J Vernon Henderson, "The Geography of Development within Countries," CEPR Discussion Papers 10150, C.E.P.R. Discussion Papers September 2014.
_ , Avner Greif, and Stephen Parente, "Spatial Competition, Innovation and Institutions: The Industrial Revolution and the Great Divergence," Mimeo 2015.
Doepke, Matthias, "Child mortality and fertility decline: Does the Barro-Becker model fit the facts?," *Journal of Population Economics*, 06 2005, *18* (2), 337–366.
_ and Fabrizio Zilibotti, "The Macroeconomics of Child Labor Regulation," *American Economic Review*, December 2005, *95* (5), 1492–1524.
Duarte, Margarida and Diego Restuccia, "The Role of the Structural Transformation in Aggregate Productivity," *The Quarterly Journal of Economics*, 2010, *125* (1), 129–173.
Duranton, Gilles, "The Colombian Urban System: Agglomeration, Urban Costs, and Growth," 2014. Working paper.
_ , "Growing through cities in developing countries," Policy Research Working Paper Series 6818, The World Bank March 2014.
_ , "Estimating the Costs of Traffic Congestion in Bogota," Mimeo 2016.
_ and Diego Puga, "Micro-foundations of urban agglomeration economies," in J. V. Henderson and J. F. Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4, Elsevier, 2004, chapter 48, pp. 2063–2117.
_ , Syed Ejaz Ghani, Arti Grover Goswami, and William Robert Kerr, "The misallocation of land and other factors of production in India," Working Paper 7221, The World Bank March 2015.
Eeckhout, Jan and Nezih Guner, "Optimal Spatial Taxation: Are Big Cities Too Small?," IZA Discussion Papers 8781, Institute for the Study of Labor (IZA) January 2015.
Fay, Marianne and Charlotte Opal, "Urbanization without growth : a not-so-uncommon phenomenon," Policy Research Working Paper 2412, The World Bank August 2000.
Feler, Leo and J. Vernon Henderson, "Exclusionary Policies in Urban Development: Under-servicing migrant households in Brazilian cities," *Journal of Urban Economics*, May 2011, *69* (3), 253–272.
Fox, Sean, "The Political Economy of Slums: Theory and Evidence from Sub-Saharan Africa," *World Development*, 2014, *54* (C), 191–203.
_ and Tom Goodfellow, *Cities and Development* 2016.
Fuglie, Keith O., "Total Factor Productivity in the Global Agricultural Economy: Evidence from FAO Data," in "Shifting Patterns of Agricultural Production and Productivity Worldwide" 2010.
Galor, Oded, *Unified Growth Theory*, Princeton, NJ: Princeton University Press, 2011.
_ , "The Demographic Transition: Causes and Consequences," *Cliometrica, Journal of Historical Economics and Econometric History*, January 2012, *6* (1), 1–28.
_ and Andrew Mountford, "Trade and the Great Divergence: The Family Connection," *American Economic Review*, May 2006, *96* (2), 299–303.
_ and _ , "Trading Population for Productivity: Theory and Evidence," *Review of Economic Studies*, 2008, *75* (4), 1143–1179.
_ and David Weil, "From Malthusian Stagnation to Modern Growth," *American Economic*

Review, May 1999, *89* (2), 150–154.
— and — , "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 2000, *90* (4), 806–828.
Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Human Capital and Regional Development," *The Quarterly Journal of Economics*, 2013, *128* (1).
Glaeser, Edward L., "A World of Cities: The Causes and Consequences of Urbanization in Poorer Countries," *Journal of the European Economic Association*, 2014, *12* (5), 1154–1199.
— , Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer, "Growth in Cities," *Journal of Political Economy*, December 1992, *100* (6), 1126–52.
— , Jed Kolko, and Albert Saiz, "Consumer city," *Journal of Economic Geography*, January 2001, *1* (1), 27–50.
Gollin, Douglas, David Lagakos, and Michael E. Waugh, "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*, 2013.
— , Remi Jedwab, and Dietrich Vollrath, "Urbanization with and without industrialization," *Journal of Economic Growth*, 2015, *21* (1), 35–70.
— , Stephen L. Parente, and Richard Rogerson, "The Role of Agriculture in Development," *American Economic Review*, May 2002, *92* (2), 160–164.
Guillot, Michel, "The Momentum of Mortality Change," *Population Studies*, 2005, *59* (3).
Haines, Michael R., *The Population of the United States, 1790-1920* 2008.
Hanlon, W. Walker, "Pollution and Mortality in the 19th Century," NBER Working Papers 21647, National Bureau of Economic Research, Inc October 2015.
— and Yuan Tian, "Killer Cities: Past and Present," *American Economic Review*, May 2015, *105* (5), 570–75.
Henderson, J. Vernon, "The Size and Types of Cities," *American Economic Review*, September 1974, *64* (4), 640–656.
— , "Cities and Development," *Journal of Regional Science*, 2010, *50* (1), 515–540.
— , Adam Storeygard, and Uwe Deichmann, "Has climate change driven urbanization in Africa?," *Journal of Development Economics*, 2016, pp. –.
— , Tanner Regan, and Anthony J. Venables, "Building the City: Sunk Capital, Sequencing and Institutional Frictions," SERC Discussion Papers 0196 April 2016.
Henderson, Vernon, *Economic Theory and The Cities* 1977.
Heuveline, Patrick, "The Global and Regional Impact of Mortality and Fertility Transitions, 1950-2000," *Population and Development Review*, 1999, *25* (4), 681–702.
Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.
Jacobs, Jane, *The Economy of Cities* 1969.
— , *Cities and the Wealth of Nations: Principles of Economic Life* 1984.
Jedwab, Remi and Dietrich Vollrath, "Urbanization without growth in historical perspective," *Explorations in Economic History*, 2015, *58* (C), 1–21.
— , Luc Christiaensen, and Marina Gindelsky, "Demography, urbanization and development: Rural push, urban pull and urban push?," *Journal of Urban Economics*, 2015, pp. –.
Jones, Charles I., "Life and Growth," *Journal of Political Economy*, 2016, *124* (2), 539–578.
Kalemli-Ozcan, Sebnem, " Does the Mortality Decline Promote Economic Growth?," *Journal of Economic Growth*, December 2002, *7* (4), 411–39.
Lagakos, David and Michael E. Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, April 2013, *103* (2), 948–80.
Lall, Somik V., Hyoung Gun Wang, and Daniel Da Mata, "Do Urban Land Regulations Influence Slum Formation? Evidence From Brazilian Cities," Technical Report 2007.
Lee, Ronald, "Population Dynamics of Humans and Other Animals," *Demography*, November 1987, *24* (4), 443–465.
— , "Population dynamics: Equilibrium, disequilibrium, and consequences of fluctuations," in "Handbook of Population and Family Economics," Vol. 1, Elsevier, 1997.
Lucas, Robert E., "On the mechanics of economic development," *Journal of Monetary Economics*, July 1988, *22* (1), 3–42.
— , "Life Earnings and Rural-Urban Migration," *Journal of Political Economy*, February 2004, *112* (S1), S29–S59.
Malpezzi, Stephen and Stephen K. Mayo, "Getting Housing Incentives Right: A Case Study of the Effects of Regulation, Taxes, and Subsidies on Housing Supply in Malaysia," *Land Economics*, 1997, *73* (3), 372–391.
Manuelli, Rodolfo E. and Ananth Seshadri, "Explaining International Fertility Differences," *The Quarterly Journal of Economics*, May 2009, *124* (2), 771–807.

McMillan, Margaret, Dani Rodrik, and Inigo Verduzco-Gallo, “Globalization, Structural Change, and Productivity Growth, with an Update on Africa,” *World Development*, 2014, *63*, 11 – 32.
Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding, “Urbanization and Structural Transformation,” *The Quarterly Journal of Economics*, 2012, *127* (2), 535–586.
Moretti, Enrico, “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions,” *American Economic Review*, June 2004, *94* (3), 656–690.
Njoh, Ambe, “Colonization and Sanitation in Urban Africa: A Logistics Analysis of the Availability of Central Sewerage Systems,” *Habitat International*, 2013, *38*, 207 – 213.
_ and Fenda Akiwumi, “The Impact of Colonization on Access to Improved Water and Sanitation Facilities in African cities,” *Cities*, 2011, *28* (5), 452 – 460.
Özmucur, Süleyman and Şevket Pamuk, “Real Wages and Standards of Living in the Ottoman Empire, 1489-1914,” *Journal of Economic History*, June 2002, *62* (2), 293–321.
Preston, Samuel H., “The Changing Relation between Mortality and Level of Economic Development,” *Population Studies*, July 1975, *29*, 231–48.
Restuccia, Diego and Richard Rogerson, “Misallocation and productivity,” *Review of Economic Dynamics*, January 2013, *16* (1), 1–10.
Roback, Jennifer, “Wages, Rents, and the Quality of Life,” *Journal of Political Economy*, December 1982, *90* (6), 1257–78.
Rodrik, Dani, “Unconditional Convergence in Manufacturing,” *The Quarterly Journal of Economics*, 2013, *128* (1), 165–204.
Romer, Paul M, “Increasing Returns and Long-run Growth,” *Journal of Political Economy*, October 1986, *94* (5), 1002–37.
_ , “Endogenous Technological Change,” *Journal of Political Economy*, 1990, *98* (5), S71–102.
Rosen, Sherwin, “Wages-Based Indexes of Urban Quality of Life,” in Peter Miezkowski and Mahlon Straszheim, eds., *Current Issues in Urban Economics*, Johns Hopkins, 1979.
Rosenthal, Stuart S. and Stephen L. Ross, *Change and Persistence in the Economic Status of Neighborhoods and Cities*, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier,
Soares, Rodrigo R., “Mortality Reductions, Educational Attainment, and Fertility Choice,” *American Economic Review*, June 2005, *95* (3), 580–601.
Stolnitz, George J., “A Century of International Mortality Trends: I.,” *Population Studies*, July 1955, *9*, 24–55.
Struyk, R.J. and S. Giddings, *The Challenge of an Urban World: An Opportunity for U.S. Foreign Assistance* 2009.
Tertilt, Michele, “Polygyny, Fertility, and Savings,” *Journal of Political Economy*, December 2005, *113* (6), 1341–1370.
UN-Habitat, *Global Urban Indicators* 1998.
_ , *The Challenge of Slums* 2003.
_ , *Slums of the World: The face of urban poverty in the new millennium?* 2003.
_ , *State of the World’s Cities Report* 2012.
United Nations, *World Urbanization Prospects* 2014.
Venables, Anthony J, “Breaking into Tradables: urban form and urban function in a developing city,” CEPR Discussion Papers 11212, C.E.P.R. Discussion Papers April 2016.
Voigtländer, Nico and Hans-Joachim Voth, “How the West ‘Invented’ Fertility Restriction,” *American Economic Review*, October 2013, *103* (6), 2227–64.
_ and _ , “The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe,” *Review of Economic Studies*, 2013, *80* (2), 774–811.
Vollrath, Dietrich, “The Agricultural Basis of Comparative Development,” *Journal of Economic Growth*, December 2011, *16* (4), 343–370.
Weil, David N., “Accounting for The Effect of Health on Economic Growth,” *The Quarterly Journal of Economics*, 08 2007, *122* (3), 1265–1306.
_ , “Endemic Diseases and African Economic Growth: Challenges and Policy Responses,” *Journal of African Economies*, November 2010, *19*, 81–109.
Weir, David, “A Historical Perspective on the Economic Consequences of Population Growth.,” in “Consequences of rapid population growth in developing countries,” Taylor & Francis, 1991.
Williamson, Jeffrey, *Coping with City Growth During the British Industrial Revolution* 1990.
World Bank, *The Global Family Planning Revolution* 2007.
Young, Alwyn, “The Gift of the Dying: The Tragedy of Aids and the Welfare of Future African Generations,” *The Quarterly Journal of Economics*, 2005, *120* (2), 423–466.
_ , “Inequality, the Urban-Rural Gap, and Migration,” *The Quarterly Journal of Economics*, 2013, *128* (4), 1727–1785.

Figure 1: **NUMBER OF MILLION-PLUS MEGACITIES, 1700-2030**

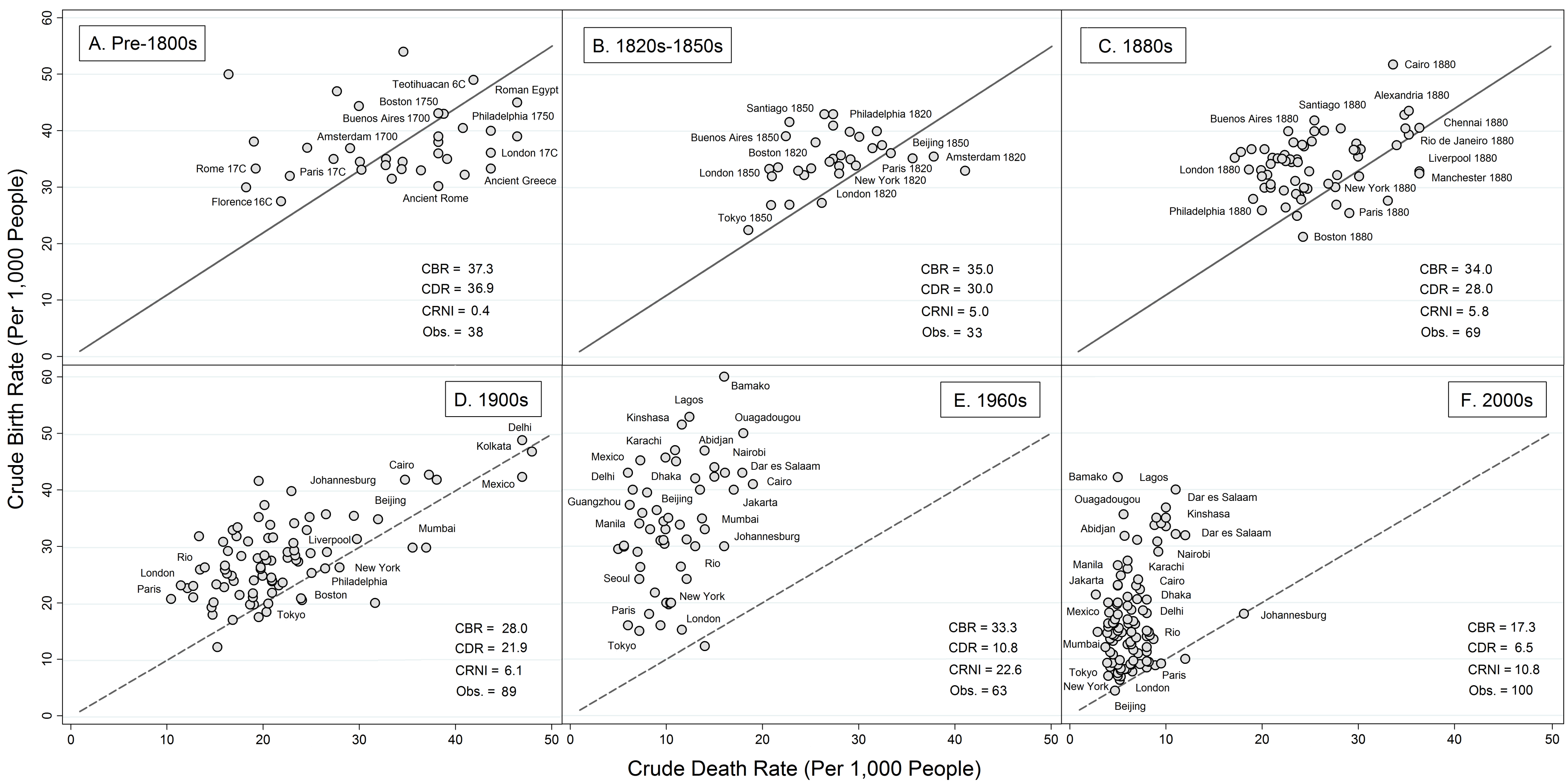


Notes: This figure shows the number of urban agglomerations of at least 1 million inhabitants in developed countries and developing countries in 1700, 1900, 1950, 2015 and 2030. Developed countries are countries whose GDP per capita (PPP, constant 2011 international $) is above $12,476 (the income level of Slovakia) in 2015. See *Web Appendix* for data sources.

Figure 2: **CITY LIVING STANDARD VERSUS CITY SIZE RANK, HISTORICALLY (PRE-1910) AND IN 2010**

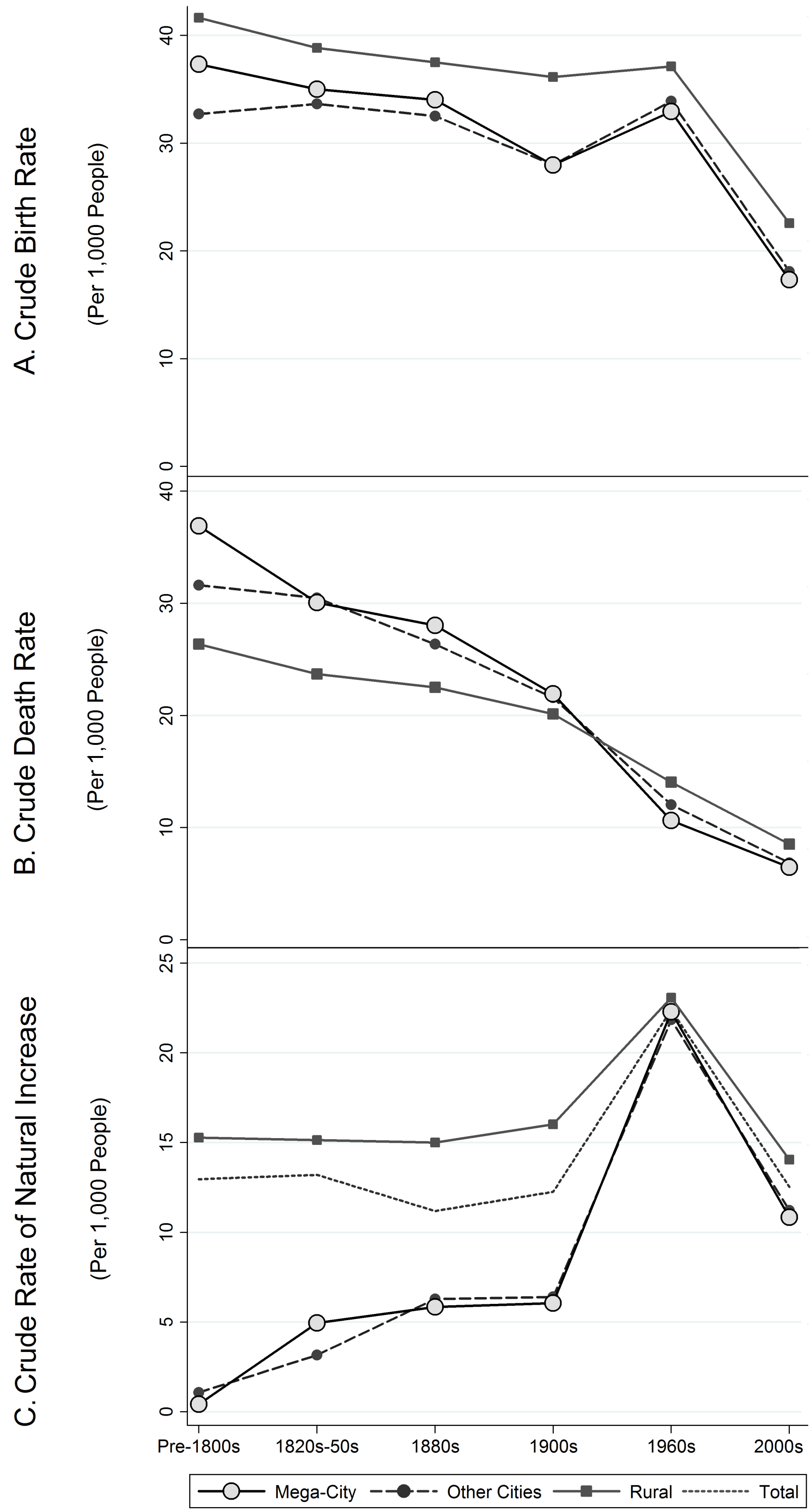
Notes: This graph displays the relationship between city living standards and city size for 111 cities of more than 300,000 inhabitants in 2010 and 111 city-year observations of more than 100,000 inhabitants pre-1910 (multiple observations for a same city). For each period, we rank the cities by living standards and city size and show the correlation between the two (linear fit estimated using as weights the population of each city-year). City living standards are proxied by the city product indexes of United Nations Habitat in 2000-2010 and welfare ratios (estimated for a "bare bones" consumption basket) before 1910. See *Web Appendix* for data sources.

Figure 3: **MEGA-CITY CRUDE BIRTH AND DEATH RATES, FROM ANTIQUITY TO DATE**



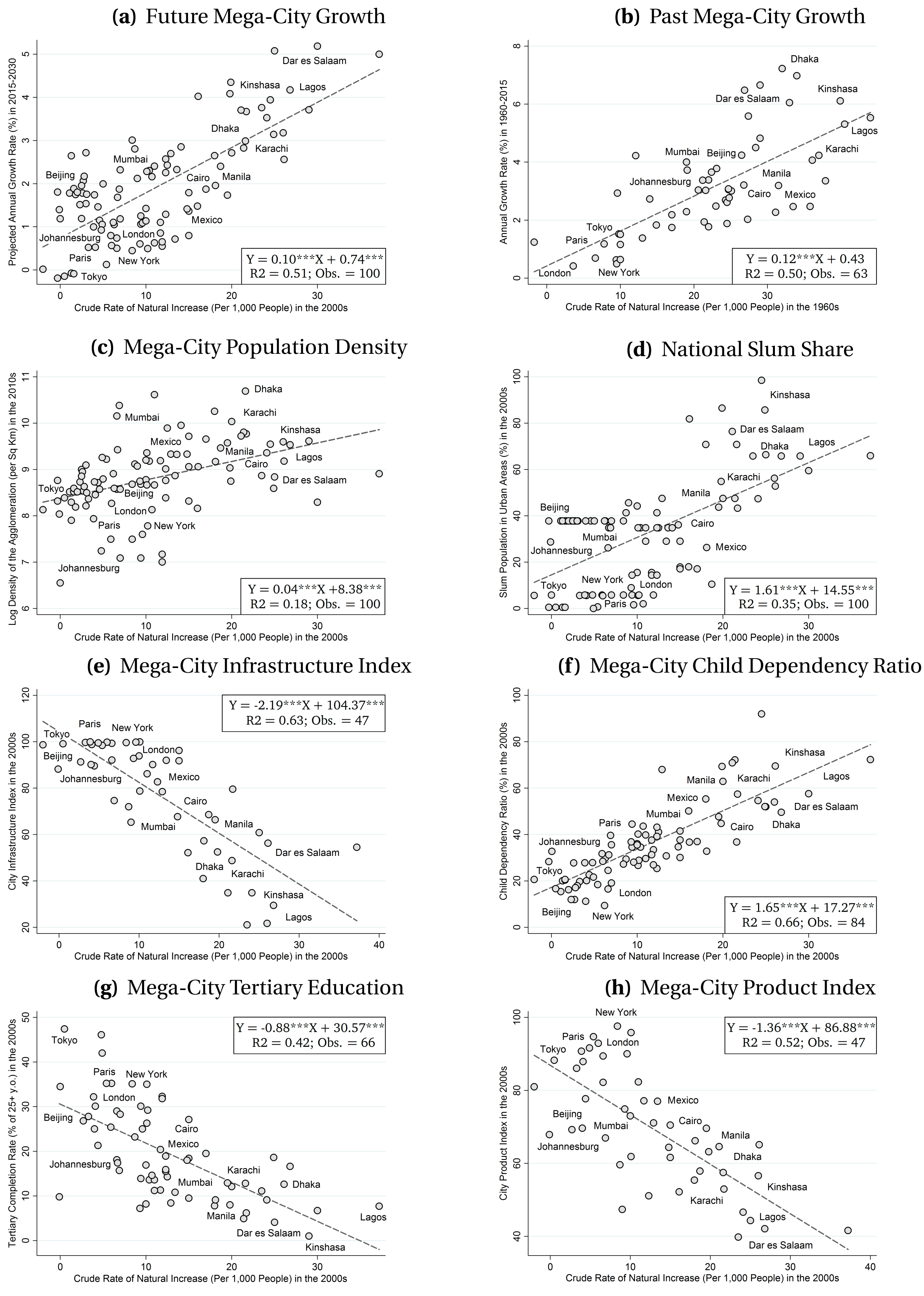
Notes: This figure shows the crude birth rates and the crude death rates for 392 megacity-period observations: 38 cities in the 1800s or before (Panel A), 33 cities in the 1820s or the 1850s (Panel B), 69 cities in the 1880s (Panel C), 89 cities in the 1900s (Panel D), 63 cities in the 1960s (Panel E), and 100 cities in the 2000s (Panel F). The cities in the 1900s and before (Panels A-D) were selected because they were among the 100 top cities in 1900. The cities in the 1960s-2000s (Panels E-F) were selected because they will be among the 100 top cities in 2030 according to United Nations (2014). The number of observations (Obs.), mean crude birth rates (CBR), death rates (CDR) and rates of natural increase (CRNI) are shown. See Table 2 for data sources.

Figure 4: **HISTORICAL DEMOGRAPHY OF MEGA-CITIES VS. OTHER AREAS**



Notes: This figure shows the unweighted mean crude rates of birth (Panel A), death (Panel B) and natural increase (Panel C) for the 392 mega-cities (*Mega-City*) as well as for the other cities (*Other Cities*), the rural areas (*Rural*), and the whole area (*Total*) of the same countries as the mega-cities in the 1800s or before (N = 38), the 1820s-1850s (33), the 1880s (69), the 1900s (89), the 1960s (63), and the 2000s (100). The sample is the same as in Figure 3. See Table 2 for data sources.

Figure 5: **MEGA-CITY NATURAL INCREASE AND CHARACTERISTICS**



Notes: This figure describes the 100 largest cities in 2030 according to United Nations (2014). (a) shows the correlation between city projected growth in 2015-2030 and city natural increase in the 2000s. (b) shows the correlation between city growth in 1960-2015 and city natural increase in the 1960s. (c)-(h) show the correlations between measures of city congestion in the 2000s and city natural increase in the 2000s: the city density (c), the national slum share (d), the city infrastructure index (e), the city child dependency ratio (f), the city tertiary education completion rate (g), and the city product index (h). See text for data sources.

Table 1: **WORLD'S LARGEST MEGACITIES (MILLIONS), 1700-2015**

Rank	1700		1900		1950		2015 (△% 2015-2030)		
1	Istanbul	0.7	London	6.5	New York	12.3	Tokyo	38.0	(-0.1)
2	Tokyo	0.7	New York	4.2	Tokyo	11.3	Delhi	25.7	(2.3)
3	Beijing	0.7	Paris	3.3	London	8.4	Shanghai	23.7	(1.8)
4	London	0.6	Berlin	2.7	Paris	6.3	Sao Paulo	21.1	(0.7)
5	Paris	0.5	Chicago	1.7	Moscow	5.4	Mumbai	21.0	(1.9)
6	Ahmedabad	0.4	Vienna	1.7	Buenos Aires	5.1	Mexico	21.0	(0.9)
7	Osaka	0.4	Tokyo	1.5	Chicago	5	Beijing	20.4	(2.1)
8	Isfahan	0.4	St. Petersburg	1.4	Kolkata	4.5	Osaka	20.2	(-0.1)
9	Kyoto	0.4	Manchester	1.4	Shanghai	4.3	Cairo	18.8	(1.8)
10	Hangzhou	0.3	Philadelphia	1.4	Osaka	4.1	New York	18.6	(0.5)
11	Amsterdam	0.2	Birmingham	1.2	Los Angeles	4	Dhaka	17.6	(3.0)
12	Naples	0.2	Moscow	1.1	Berlin	3.3	Karachi	16.6	(2.7)
13	Guangzhou	0.2	Beijing	1.1	Philadelphia	3.1	Buenos Aires	15.2	(0.7)
14	Aurangabad	0.2	Kolkata	1.1	Rio	3	Kolkata	14.9	(1.7)
15	Lisbon	0.2	Boston	1.1	StPetersburg	2.9	Istanbul	14.2	(1.1)
16	Cairo	0.2	Glasgow	1	Mexico	2.9	Chongqing	13.3	(1.8)
17	Xian	0.2	Osaka	1	Mumbai	2.9	Lagos	13.1	(4.2)
18	Seoul	0.2	Liverpool	0.9	Detroit	2.8	Manila	12.9	(1.8)
19	Dacca	0.2	Istanbul	0.9	Boston	2.6	Rio	12.9	(0.6)
20	Ayutthaya	0.2	Hamburg	0.9	Cairo	2.5	Guangzhou	12.5	(2.3)
21	Venice	0.1	Buenos Aires	0.8	Tianjin	2.5	Los Angeles	12.3	(0.5)
22	Suzhou	0.1	Budapest	0.8	Manchester	2.4	Moscow	12.2	(0.0)
23	Nanking	0.1	Mumbai	0.8	Sao Paulo	2.3	Kinshasa	11.6	(3.7)
24	Rome	0.1	Ruhr	0.8	Birmingham	2.2	Tianjin	11.2	(1.8)
25	Smyrna	0.1	Rio	0.7	Shenyang	2.1	Paris	10.8	(0.6)
26	Srinagar	0.1	Warsaw	0.7	Roma	1.9	Shenzhen	10.7	(1.1)
27	Palermo	0.1	Tientsin	0.7	Milano	1.9	Jakarta	10.3	(2.0)
28	Moscow	0.1	Shanghai	0.6	San Francisco	1.9	London	10.3	(0.7)
29	Milan	0.1	Newcastle	0.6	Barcelona	1.8	Bangalore	10.1	(2.6)
30	Madrid	0.1	St.Louis	0.6	Glasgow	1.8	Lima	9.9	(1.4)

Notes: The table shows the population (millions) of the world's largest urban agglomerations in 1700, 1900, 1950 and 2015. An urban agglomeration comprises the city proper and also the suburban fringe or thickly settled territory lying outside. (△% 2015-30) is the projected annual growth rate (%) of the urban agglomeration between 2015 and 2030 according to United Nations (2014). These growth rates are based on non-linear extrapolation given the rates of growth pre-2015. The main sources of the data are Chandler (1987) and United Nations (2014). See text for more details on data sources.

Table 2: **NATURAL INCREASE SOURCE INFORMATION BY COUNTRY 1/3**

Panel A: Largest Mega-Cities of the Future (2030)

Rank	City in 2030 [*Slum*]	Data Years	Main Sources (See Excel File for Details)
1	Tokyo	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
2	Delhi	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
3	Shanghai	1960s, 2000s	Historical Studies. Population Censuses.
4	Mumbai [*Dharavi*]	1960s, 2000s	CICRED Monograph. Population Censuses.
5	Beijing	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
6	Dhaka [*Korail*]	1960s, 2000s	CICRED Monograph. Population Censuses.
7	Karachi [*Orangi*]	1960s, 2000s	CICRED Monograph. Population Censuses.
8	Cairo [*Manshiet*]	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
9	Lagos	1960s, 2000s	CICRED Monograph. DHS.
10	Mexico [*Neza*]	1960s, 2000s	CICRED Monograph. Population Censuses.
11	Sao Paulo	1960s, 2000s	CICRED Monograph. Population Censuses.
12	Kinshasa	1960s, 2000s	Demographic Study. DHS.
13	Osaka	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
14	New York	1960s, 2000s	Historical Studies. Population Censuses.
15	Kolkata	1960s, 2000s	CICRED Monograph. Population Censuses.
16	Guangzhou	1960s, 2000s	Historical Studies. Population Censuses.
17	Chongqing	2000s	Population Census.
18	Buenos Aires	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
19	Manila [*Tondo*]	1960s, 2000s	CICRED Monograph. Population Censuses.
20	Istanbul	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
21	Bangalore	1960s, 2000s	CICRED Monograph. Population Censuses.
22	Tianjin	1960s, 2000s	Population Census.
23	Rio de Janeiro [*Rocinha*]	1960s, 2000s	CICRED Monograph. Population Censuses.
24	Chennai	1960s, 2000s	CICRED Monograph. Population Censuses.
25	Jakarta	1960s, 2000s	CICRED Monograph. Population Censuses.
26	Los Angeles	1960s, 2000s	Population Censuses.
27	Lahore	2000s	Population Census.
28	Hyderabad	1960s, 2000s	CICRED Monograph. Population Censuses.
29	Shenzhen	2000s	Population Census.
30	Lima	1960s, 2000s	CICRED Monograph. Population Censuses.
31	Moscow	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
32	Bogota	1960s, 2000s	CICRED Monograph. Population Censuses.
33	Paris	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
34	Johannesburg	1960s, 2000s	DHS. Historical Studies.
35	Bangkok	1960s, 2000s	CICRED Monograph. Population Censuses.
36	London	1960s, 2000s	Historical Studies. CICRED Monograph. Population Censuses.
37	Dar es Salaam	1960s, 2000s	CICRED Monograph. Population Censuses. DHS.
38	Ahmadabad	1960s, 2000s	CICRED Monograph. Population Censuses.
39	Luanda	2000s	DHS. Population Census.
40	Ho Chi Minh	2000s	DHS. Population Census.
41	Chengdu	2000s	Population Census.
42	Tehran	1960s, 2000s	CICRED Monograph. Population Censuses.
43	Seoul	1960s, 2000s	CICRED Monograph. Population Censuses.
44	Nanjing	1960s, 2000s	Population Census.
45	Baghdad	2000s	DHS. Population Census.
46	Chicago	1960s, 2000s	History Books. CICRED Monograph. Population Censuses.
47	Wuhan	2000s	Population Census.
48	Kuala Lumpur	1960s, 2000s	CICRED Monograph. Population Censuses.
49	Nagoya	2000s	Population Census.
50	Hangzhou	2000s	Population Census.
51	Dongguan	2000s	Population Census.
52	Surat	2000s	Population Census.
53	Foshan	2000s	Population Census.
54	Kabul	2000s	DHS. Population Census.
55	Khartoum	2000s	DHS. Population Census.
56	Suzhou	2000s	Population Census.
57	Pune	2000s	Population Census.
58	Riyadh	2000s	Demographic Study.
59	Shenyang	2000s	Population Census.
60	Xian	2000s	Population Census.
61	Hong Kong	1960s, 2000s	Wikipedia.
62	Abidjan	1960s, 2000s	Population Censuses. DHS.
63	Nairobi [*Kibera*]	1960s, 2000s	CICRED. Population Censuses. DHS.
64	Santiago	1960s, 2000s	CICRED. Population Censuses.
65	Toronto	2000s	Population Censuses.
66	Xiamen	2000s	Population Census.
67	Haerbin	2000s	Population Census.
68	Houston	1960s, 2000s	Population Censuses.
69	Chittagong	2000s	Population Census.
...	...	...	...

TABLE CONTINUED ON THE NEXT TWO PAGES (TABLE NOTES ON THE PAGE AFTER THE NEXT PAGE).

Table 2: **NATURAL INCREASE SOURCE INFORMATION BY COUNTRY 2/3**

Rank	City in 2030 [*Slum*]	Data Years	Main Sources (See Excel File for Details)
...	...	...	...
70	Madrid	1960s, 2000s	Population Censuses.
71	Dallas	1960s, 2000s	Population Censuses.
72	Yangon	1960s, 2000s	Historical Studies. Population Censuses. DHS.
73	Singapore	1960s, 2000s	Wikipedia.
74	Miami	1960s, 2000s	Population Censuses.
75	Belo Horizonte	2000s	Population Censuses.
76	Alexandria	2000s	Population Censuses.
77	Kano	2000s	DHS. Population Census.
78	Philadelphia	1960s, 2000s	Historical Studies. Population Censuses.
79	Atlanta	1960s, 2000s	Population Censuses.
80	Dakar	1960s, 2000s	Population Censuses. DHS.
81	Qingdao	2000s	Population Census.
82	Zhengzhou	2000s	Population Census.
83	Ankara	2000s	Population Census.
84	Ouagadougou	1960s, 2000s	Population Censuses. DHS.
85	Addis Ababa	1960s, 2000s	Population Censuses. DHS.
86	Dalian	2000s	Population Census.
87	Guadalajara	2000s	Population Census.
88	Washington	1960s, 2000s	Population Censuses.
89	Barcelona	1960s, 2000s	Population Censuses.
90	Zhongshan	2000s	Population Census.
91	Ibadan	2000s	DHS. Population Census.
92	Hanoi	2000s	DHS. Population Census.
93	Monterrey	1960s, 2000s	CICRED Monograph. Population Censuses.
94	Faisalabad	2000s	Population Census.
95	Kitakyushu	2000s	Population Census.
96	Sydney	1960s, 2000s	Population Census.
97	Jinan	2000s	Population Census.
98	Bamako	1960s, 2000s	Population Censuses. DHS.
99	Yaounde	1960s, 2000s	Population Censuses. DHS.
100	Aleppo	2000s	DHS. Population Census.

Panel B: Largest Mega-Cities of the Past (1900)

Rank	City in 1900 [*Slum*]	Data Years	Main Sources (See Excel File)
1	London [*Slums*]	17C, 1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Population Censuses. Historical Studies.
2	New York [*Tenement wards*]	1820s, 1850s, 1880s, 1900s	Population Censuses. Historical Studies.
3	Paris [*Quartiers pauvres*]	17C, 1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Population Censuses. Historical Studies.
4	Berlin	1800s, 1820s, 1850s, 1880s, 1900s	Historical Studies.
5	Chicago	1880s, 1900s	Population Censuses. Historical Studies.
6	Vienna	1880s, 1900s	Historical Studies.
7	Tokyo	19C, 1880s, 1900s	Historical Studies.
8	St. Petersburg	1880s, 1900s	Historical Studies.
9	Manchester [*Townships*]	1880s, 1900s	Population Censuses. Historical Studies.
10	Philadelphia	1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Population Censuses. Historical Studies.
11	Birmingham	1880s, 1900s	Population Censuses. Historical Studies.
12	Moscow	1880s, 1900s	Historical Studies.
13	Beijing	19C, 1900s	Historical Studies.
14	Kolkata	1900s	Historical Studies.
15	Boston	1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Population Censuses. Historical Studies.
16	Glasgow	1880s, 1900s	Historical Studies.
17	Osaka	1900s	Historical Studies.
18	Liverpool	1880s, 1900s	Population Censuses. Historical Studies.
19	Istanbul	1900s	Historical Studies.
20	Hamburg	1880s, 1900s	Historical Studies.
21	Buenos Aires	17C, 1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Historical Studies.
22	Budapest	1880s, 1900s	Historical Studies.
23	Mumbai	1880s, 1900s	Historical Studies.
25	Rio de Janeiro	1880s, 1900s	Historical Studies.
26	Warsaw	1900s	Historical Studies.
29	Newcastle	1880s, 1900s	Population Censuses. Historical Studies.
30	St. Louis	1880s, 1900s	Population Censuses. Historical Studies.
31	Cairo	1880s, 1900s	Historical Studies.
33	Naples	1880s, 1900s	Historical Studies.
...	...	...	...

TABLE CONTINUED ON THE NEXT PAGE (TABLE NOTES ON THE NEXT PAGE).

Table 2: NATURAL INCREASE SOURCE INFORMATION BY COUNTRY 3/3

Rank	City in 1900	Data Years	Main Sources (See Excel File)
...	...	...	...
34	Pittsburgh	1880s, 1900s	Population Censuses. Historical Studies.
35	Brussels	1850s, 1880s, 1900s	Population Censuses. Historical Studies.
36	Barcelona	1880s, 1900s	Historical Studies.
37	Dresden	1880s, 1900s	Historical Studies.
38	Madrid	1880s, 1900s	Historical Studies.
39	Leipzig	1900s	Historical Studies.
40	Amsterdam	1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Historical Studies.
41	Lyon	1880s, 1900s	Historical Studies.
42	Baltimore	1880s, 1900s	Population Censuses. Historical Studies.
43	Chennai	1880s, 1900s	Historical Studies.
44	Munich	1850s, 1880s, 1900s	Historical Studies.
45	Milan	1900s	Historical Studies.
46	Melbourne	1880s, 1900s	Historical Studies.
47	Sydney	1880s, 1900s	Historical Studies.
48	Prague	1880s, 1900s	Historical Studies.
49	Copenhagen	17C, 1700s, 1750s, 1880s, 1900s	Historical Studies.
51	Odessa	1900s	Historical Studies.
52	Hyderabad	1880s	Historical Studies.
53	San Francisco	1880s, 1900s	Population Censuses. Historical Studies.
54	Rome	17C, 1700s, 1880s, 1900s	Historical Studies.
55	Cologne	1900s	Historical Studies.
56	Leeds	1880s, 1900s	Population Censuses. Historical Studies.
57	Wroclaw	1880s, 1900s	Historical Studies.
58	Cincinnati	1880s, 1900s	Population Censuses. Historical Studies.
59	Marseille	1880s, 1900s	Historical Studies.
60	Sheffield	1880s, 1900s	Population Censuses. Historical Studies.
61	Edinburgh	1880s, 1900s	Historical Studies.
63	Cleveland	1880s, 1900s	Population Censuses. Historical Studies.
64	Dublin	1850s, 1880s, 1900s	Historical Studies.
65	Mexico	1900s	Historical Studies.
66	Rotterdam	1880s, 1900s	Historical Studies.
68	Minneapolis	1900s	Population Censuses. Historical Studies.
69	Lisbon	1880s, 1900s	Historical Studies.
70	Kyoto	1900s	Historical Studies.
71	Antwerp	1800s, 1820s, 1850s, 1880s, 1900s	Population Censuses. Historical Studies.
72	Buffalo	1900s	Population Censuses. Historical Studies.
76	Belfast	1900s	Historical Studies.
77	Turin	1880s, 1900s	Historical Studies.
78	Montreal	1850s, 1880s, 1900s	Historical Studies.
79	Bristol	1880s, 1900s	Population Censuses. Historical Studies.
80	Alexandria	1880s, 1900s	Historical Studies.
81	Bordeaux	1880s, 1900s	Historical Studies.
82	Bradford	1880s, 1900s	Population Censuses. Historical Studies.
83	Stockholm	1700s, 1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Historical Studies.
86	Wuppertal	1850s, 1880s, 1900s	Historical Studies.
87	Detroit	1900s	Population Censuses. Historical Studies.
88	Riga	1880s, 1900s	Historical Studies.
88	Liege	1850s, 1880s, 1900s	Historical Studies.
90	New Orleans	1880s, 1900s	Population Censuses. Historical Studies.
91	Santiago	1850s, 1880s, 1900s	Historical Studies.
92	Lille	1880s, 1900s	Historical Studies.
93	Hanover	1850s, 1880s, 1900s	Historical Studies.
94	Milwaukee	1900s	Population Censuses. Historical Studies.
95	Bucharest	1880s, 1900s	Historical Studies.
97	Washington	1880s, 1900s	Population Censuses. Historical Studies.
99	Genoa	1900s	Historical Studies.
100	Montevideo	1750s, 1800s, 1820s, 1850s, 1880s, 1900s	Historical Studies.
_	Florence 16C	16C	Historical Studies.
_	Teotihuacan	6C	Historical Studies.
_	Roman Egypt	Antiquity	Historical Studies.
_	Ancient Rome	Antiquity	Historical Studies.
_	Ancient Greece	Antiquity	Historical Studies.

Notes: This table shows the 100 largest megacities and their population in 2030 according to United Nations (2014) (*Panel A*), 86 of the 100 largest megacities and their population in 1900 according to Chandler (1900) (*Panel B*), and the main sources of information for the crude rates of birth, death and natural increase for each city-period. We also collect data for a number of cities that were among the richest cities before 1900 (below the dashed line). For most city-periods, we also obtained using the same sources the same demographic data for the other cities of the same country-period, the rural areas, and the entire country. The slums for which we also obtained using the same sources the same demographic data are shown into brackets after the name of the city. The sources are: *CICRED Monograph* = 1974 Country monograph from CICRED. *DHS* = Report(s) from the Demographic and Health Survey(s) of USAID. *Historical Studies*. Various academic or policy articles and books giving specific estimates of the demographic rates for the city-period observation. *Population Census(es)* = Summary report(s) of the population census(es). *Wikipedia* = Webpage "Demographics of X" on Wikipedia. See the excel file megacity_urban_rural_demographics.xlsx for details on each demographic rate and each source.

Table 3: **CALIBRATION PARAMETER VALUES**

Parameter	Value	Source
Set externally:		
Urbanization rate (%) in 1950	8.9	Sample average
Informal share of urban areas (%) in 1950	50.0	See text for details
Initial population size in 1950	1.0	Normalization
Pre-UMT formal CDR (per 000)	40	Location-level data, 1900s
Pre-UMT informal CDR (per 000)	40	Location-level data, 1900s
Post-UMT formal CDR (per 000)	15	Location-level data, 2000s
Post-UMT informal CDR (per 000)	15	Location-level data, 2000s
UMT half-life	3	See text for details
Initial rural CDR (per 000)	20	Location-level data, 1960s-2000s
Initial rural CBR (per 000)	43	Location-level data, 1960s
Initial informal CBR (per 000)	43	Location-level data, 1960s
Initial formal CBR (per 000)	38	Location-level data, 1960s
Informal prod./amenity growth (G_l)	0.025	See text for details
Rural prod./amenity growth (G_r)	0.025	See text for details
Formal prod./amenity growth (G_f)	0.05	See text for details
Preference parameters (β)	0.41	From Becker et al (2005)
Rural congestion elasticity (ϵ_r)	1.20	See text for details
Targeted:		
Formal congestion elasticity (ϵ_f)	1.32	Urbanization rate in 2005 (mean 31.0%)
Informal congestion elasticity (ϵ_l)	0.60	Informal share in 2005 (mean 64.2%)

Notes: The table shows the parameter values used in the baseline simulation of the model. *Location-level data* refers to the data presented in section 3.1.. *Sample average* refers to the sample of 41 poor countries.

Table 4: **CALIBRATED OUTCOMES AT DIFFERENT DEATH RATES**

	Urbanization Rate (%)		Urban Size		Informal Share (%)		Informal Size		Welfare
	1950	2005	1950	2005	1950	2005	1950	2005	2005
1. Observed data:	8.9	31.0	1.0	15.2	n/a	64.2	1.0	18.6	n/a
Calibrated model:									
2. With UMT	8.9	31.0	1.0	14.8	50.0	64.2	1.0	19.0	1.00
3. Without UMT	8.9	21.3	1.0	8.4	50.0	54.8	1.0	9.2	1.08
4. With UMT, $\beta = 0$	8.9	22.7	1.0	10.7	50.0	59.1	1.0	12.7	0.88
5. With UMT, $\epsilon_i = \bar{\epsilon}$	8.9	25.0	1.0	11.7	50.0	21.2	1.0	5.0	1.17
6. W/o UMT, constant fertility	8.9	19.9	1.0	6.5	50.0	50.1	1.0	6.5	1.33
7. W/o UMT, endogenous fertility	8.9	19.4	1.0	5.9	50.0	48.4	1.0	5.7	1.44

Notes: Row 2: Baseline model. The urban death rate falls from 40 (per thousand) to 15 exponentially, with a half-life of 3 years. The birth rates follow the observed path in the data. Row 3: We hold the urban death rate constant at 40 over the entire period, and also have birth rates follow the observed path. Row 4: We allow the UMT to occur, but shut down the preference of individuals for lower death rates. Row 5: We allow the UMT to occur, and individuals to have a preference for lower mortality, but set the congestion elasticity in each location to be equal to 1.014. Row 6: We hold the urban death rate at 40, but also hold the birth rate in each location at the 1950 level. Row 7: We hold urban death rates at 40, but have the birth rate determined endogenously by changes in the death rates and wages (see Web Appendix for details of this model).

Table 5: **LONG-RUN OUTCOMES USING THE CALIBRATED MODEL**

From 1950 forward	100 years (in 2050):					150 years (in 2100):				
	Urb. Rate	Urb. Size	Inf. Share	Inf. Size	Welf.	Urb. Rate	Urb. Size	Inf. Share	Inf. Size	Welf.
1. With UMT, CBR floor	49.8	72.1	64.7	93.4	1.00	68.9	341.3	62.1	423.8	1.00
2. With UMT, endog fertility	53.0	99.6	69.5	138.5	0.78	77.6	996.5	77.0	1533.7	0.45
3. Without UMT, CBR floor	34.9	30.0	49.7	29.9	1.39	51.4	85.1	36.5	62.1	2.22
4. Without UMT, no chg in CBR	33.2	23.2	44.8	20.8	1.74	50.1	68.3	32.3	44.1	2.74
5. Without UMT, endog fertility	30.7	14.9	36.2	10.8	2.58	46.2	19.3	12.5	4.8	10.17

Notes: Row 1: Long-run outcomes with parameters set according to Table 3 and with the UMT occurring, meaning urban death rates decline exponentially from 40 (per thousand) to 15 with a half-life of 3 years. Birth rates match the observed data through 2005, and then are allowed to decline to their 1950 levels (which occurs in 2020). Row 2: Outcomes with the UMT, but allowing birth rates to be endogenous. Row 3: Outcomes if urban death rates are held at 40, and birth rates follow the observed path to 2005 and then are allowed to decline to their 1950 levels. Row 4: We hold urban death rates at 40 and birth rates at their 1950 levels. Row 5: We hold urban death rates at 40, but allows fertility to evolve endogenously.

Table 6: **LONG-RUN OUTCOMES UNDER DIFFERENT POLICIES**

	Urb. Rate	Urb. Size	Inf. Rate	Inf. Size	Welf.
1. Initial values (in 2005):	31.0	1.0	64.2	1.0	n/a
After 50 years (2055):					
2. Baseline (CDR$_r$ = 10; CDR$_l$ = CDR$_f$ = 7.5; CBR$_r$ = CBR$_l$ = 35, CBR$_f$ = 20; G_r = G_l = 0.025; G_f = 0.05; ϵ_r = 1.20; ϵ_l = 0.60; ϵ_f = 1.32)	50.8	5.2	63.6	5.1	1.00
Productivity changes:					
3. Higher rural growth ($G_r = 0.035$)	38.3	3.9	58.9	3.6	1.24
4. Higher informal growth ($G_l = 0.035$)	61.2	6.4	75.5	7.5	1.30
5. Higher formal growth ($G_f = 0.06$)	53.6	5.4	52.7	4.4	1.10
6. Higher formal growth ($G_f = 0.08$)	62.1	6.0	29.1	2.7	1.47
7. Higher formal growth ($G_f = 0.10$)	73.9	6.7	11.0	1.1	2.42
Congestion costs:					
8. Lower formal congestion ($\epsilon_f = \epsilon_r = 1.20$)	52.0	5.3	58.9	4.8	1.04
9. Lower formal congestion ($\epsilon_f = \epsilon_l = 0.60$)	67.3	6.2	19.4	1.9	1.82
10. Lower formal congestion ($\epsilon_f = 1.32 \div 3 = 0.44$)	74.8	6.6	9.9	1.0	2.62
11. Higher informal congestion ($\epsilon_l = \epsilon_f = 1.32$)	38.7	3.9	41.6	2.5	0.79
Migration costs:					
12. Rural-to-urban mig restriction ($\lambda_{cr} = -0.01$)	38.3	3.9	58.9	3.6	0.90
13. Urban-to-rural mig restriction ($\lambda_{cr} = 0.01$)	62.8	6.4	66.7	6.7	1.11
Population growth changes:					
14. Zero population growth (CDR = CBR)	39.8	1.3	38.1	0.8	3.13
15. Family planning (CBR = 20)	44.2	2.4	50.3	1.9	1.82
16. Informal Family planning (CBR$_l$ = 20)	48.7	4.2	59.9	3.9	1.19
17. Rural Family planning (CBR$_r$ = 20)	45.8	3.0	54.0	2.5	1.55

Notes: Simulated outcomes in 2055 using different policy interventions. Each simulation uses the parameters of Table 3. Death rates and birth rates are set to observed rates in 2005. Welfare is the equivalent variation in net wage necessary for the baseline economy (row 1) to match welfare in the given scenario.

NOT FOR PUBLICATION - Appendix: Endogenous Fertility

To incorporate endogenous fertility into the model, we modify the utility function to be as follows

$$V_j = \ln w_j + \ln Q_j + \beta \ln(1/CDR_j) + \gamma \ln n(w_j, CDR_j, \tau_j) \tag{16}$$

where $n(w_j, CDR_j, \tau_j)$ is the number of births as a function of the wage, w_j, and the crude death rate, CDR_j. τ_j is a location-specific cost of having children. This function is the optimal outcome of a choice problem facing individuals who take the wage and crude death rate in a location as given, deciding how many children to have.

Putting this in terms of growth rates, we have the following

$$\hat{V}_j = \hat{w}_j - \beta C\hat{D}R_j + \gamma\phi^n_w \hat{w}_j + \gamma\phi^n_{CDR} C\hat{D}R_j. \tag{17}$$

The term ϕ^n_w is the elasticity of fertility with respect to wages, and ϕ^n_{CDR} is the elasticity of fertility with respect to the death rate. There are several assumptions located in this expression. First, we assume that $\hat{Q}_j = 0$, or that amenities do not grow. We do this because the fertility decision is directly related to wages, and we have no way of separately tracking amenity from wage growth in the model. Hence, to solve the model we need a way of pinning down wage growth, and assuming that amenities do not grow is the most direct way of doing this. Second, we have assumed that τ_j does not change over time, although it drives level differences between locations in crude birth rates.

Re-arranging this expression for the growth rate of welfare, we have

$$\hat{V}_j = (1 + \gamma\phi^n_w)\hat{w}_j - (\beta - \gamma\phi^n_{CDR}) C\hat{D}R_j, \tag{18}$$

which shows that welfare growth is a combination of wage growth and changes in mortality, similar to the baseline model. Here, the effect of wage growth is modified by the effect that it has on optimal fertility behavior (the $\gamma\phi^n_w$ term) and the effect of mortality is modified by the effect it has on fertility (the $\gamma\phi^n_{CDR}$ term).

Based on parameter values from the literature noted below, $\phi^n_w < 0$, meaning that fertility is negatively related to wages. Thus the effect of wage growth on welfare growth is *smaller* with endogenous fertility. An increase in wages raises welfare, but also lowers optimal fertility, and as people value fertility ($\gamma > 0$), this lowers welfare.

For mortality, the literature also indicates that $\phi^n_{CDR} < 0$, or lower mortality raises net fertility. Combined with the assumptions that $\gamma > 0$, this implies that endogenous fertility makes the effect of changes in mortality even larger in absolute size. A decline in mortality directly raises welfare, but through its effect of raising fertility has an additional positive effect on welfare. Endogenous fertility thus tends to put greater weight on mortality rates and less on wage growth in determining welfare.

With wage growth determined by $\hat{w}_j = \hat{a}_j^w - \epsilon_j^w \hat{N}_j$, as in the original model, we have that

$$\hat{V}_j = G_j - \epsilon_j \hat{N}_j - \theta_j C\hat{D}R_j \tag{19}$$

In terms of our prior notation, this implies that now

$$\begin{aligned} G_j &= (1 + \gamma\phi_w^n)\hat{a}_j^w \\ \epsilon_j &= (1 + \gamma\phi_w^n)\epsilon_j^w \\ \theta_j &= (\beta - \gamma\phi_{CDR}^n) \end{aligned} \tag{20}$$

Note that the structure of utility growth is identical to what we have in the baseline model, as it depends on welfare growth not related to population, G_j, population growth, $\hat{N}_j$, and changes in the crude death rate, $C\hat{D}R_j$. The interpretation of G_j, ϵ_j, and the term θ_j are now different, as they incorporate terms involving fertility, but solving the model from here is identical to our baseline model.

What does change is the exact specification for growth in population. Rather than our set-up from before, now we have that

$$\hat{N} = \sum_{j=1}^{J} s_j \left(n(w_j, CDR_j, \tau_j) - CDR_j \right). \tag{21}$$

Aggregate population growth will change over time both because of changes in the allocation of individuals across locations, the s_j terms, but also because of changes in fertility due to wages and crude death rates across locations.

To simulate the model, we require several additional parameters. Work by Vogl (2016) suggests that the utility weight on fertility is roughly equal to that on consumption, which in terms of this model implies that γ should be roughly equal to one.

For simulation, we need the elasticity of the crude birth rate with respect to both wages and crude death rates. Jones et al. (2010) provide elasticities of fertility with respect to income for the U.S. from 1826-1960, and find that they generally decline from between -0.30 and -0.40 in the 1800's to about -0.20 in the 1950's. Young (2005) estimates an elasticity of fertility with respect to wages of -0.35 using household survey data from South Africa. Total fertility rates and crude birth rates may deviate because of age structure. As a check on these elasticities, we use cross-country evidence from the United Nations (2012), and regressed the log crude birth rate on log GDP per capita using a panel of developing countries between 1950-2010, finding an elasticity of -0.20. We use a value of $\phi_w^n = -0.3$ in our calibration.

The elasticity of crude birth rates with respect to crude death rates, ϕ_{CDR}^n is much harder to pin down. In the end, the exact value does not turn out to be important, and the calibration is not sensitive to the exact choice of this elasticity. We use a value of $\phi_{CDR}^n = -0.30$ as a baseline, but modifying this between 0 and -1 does not create material differences in our results. The value of θ_j follows directly from the values of β, γ, and ϕ_{CDR}^n.

Finally, we require that the crude birth rates in each location in the initial period of the calibration are identical to those obseved in the data from 1950, $n(w_j, CDR_j, \tau_j) = CBR_{j,1950}$, which would implicitly set the values of τ_j. As we only require the changes in fertility over time, solving for these is not necessary.

To keep the simulations with endogenous fertility comparable, we keep the values of the G_j and ϵ_j terms identical to the baseline calibration. From (21) we know that these terms have a slightly different interpretation in the model with endogenous fertility. An alternative would be to entirely recalibrate the model in the endogenous fertility setting, and then replicate all aspects of the baseline analysis. We have done this, and the overall results are not demonstrably different from our baseline model. So for comparison purposes, we present the endogenous fertility results in the paper holding the values of the G_j and ϵ_j terms equal to their values in the baseline.

Appendix References

Jones, Larry E., Alice Schoonbroodt, and Michele Tertilt, "Fertility Theories: Can They Explain the Negative Fertility-Income Relationship?," in "Demography and the Economy", National Bureau of Economic Research, December 2010, pp. 43-100.

United Nations, *World Population Prospects* 2012.

Vogl, Tom S., "Differential Fertility, Human Capital, and Development," *The Review of Economic Studies*, 2016, 83 (1), 365401.

Young, Alwyn, "The Gift of the Dying: The Tragedy of Aids and the Welfare of Future African Generations," *The Quarterly Journal of Economics*, 2005, 120 (2), 423466.