

## Appendix 1. Supplemental figures

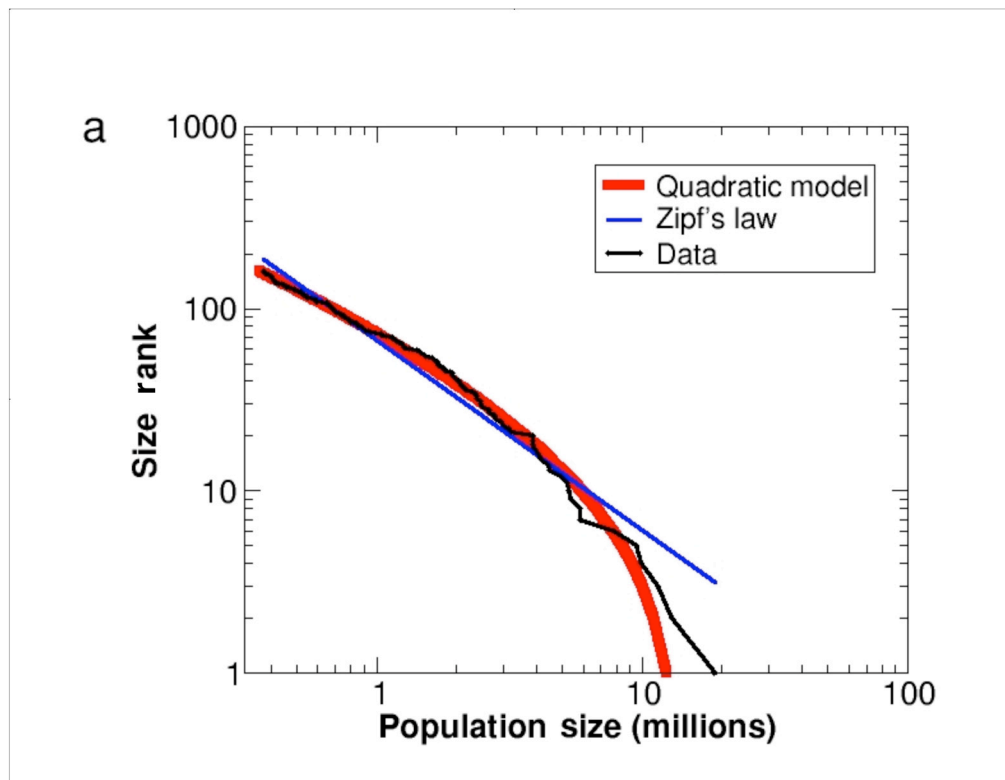


Figure A1. Zipf's law and a quadratic model illustrated for the top 160 Metropolitan Statistical Areas of the United States in 2005: plot of city population size vs the rank of the city with respect to population size. The power law exponent of the linear regression (Zipf's law) is  $-1.041$  ( $R^2 = 0.96$ ,  $p < 0.0001$ ). The best-fit parameters for the quadratic model are  $a_0 = 0.000661$ ,  $b = 0.130$  ( $R^2 = 0.996$ ,  $p < 0.0001$ ). Data are from the US Census Bureau (2007).

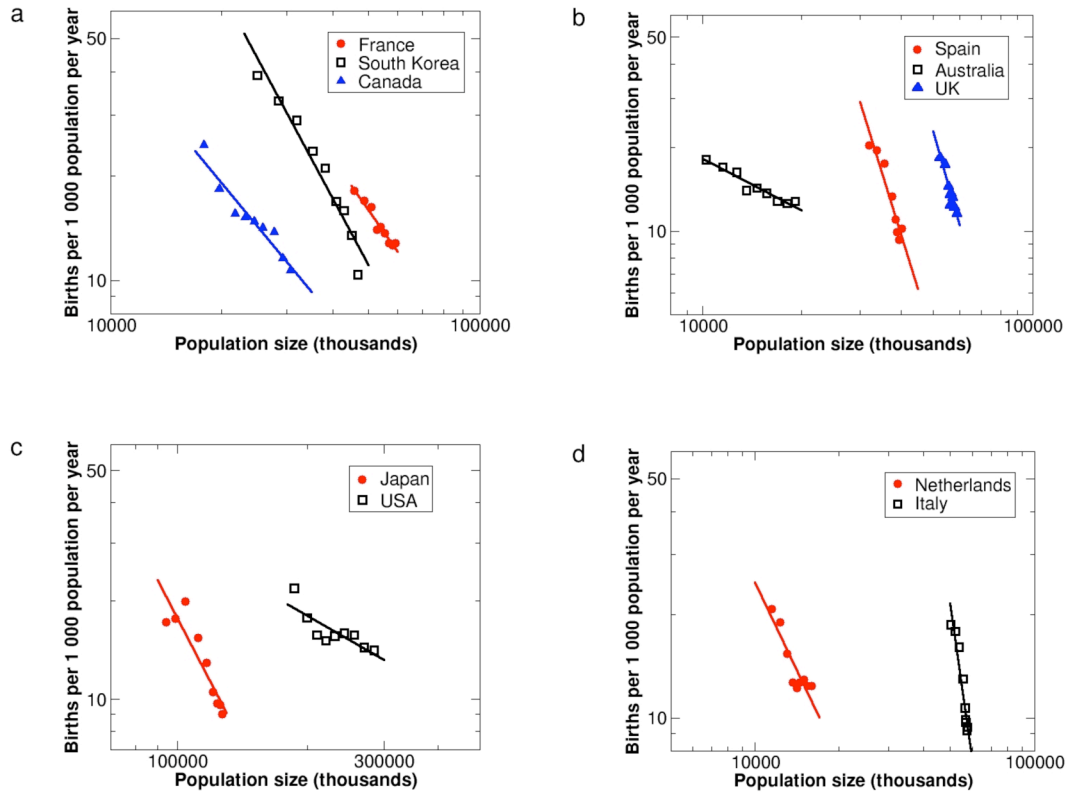


Figure A2. (a) Birth rate versus population size (thousands of individuals) for France, South Korea and Canada, every five years from 1960–2000. Power law exponents are  $-1.53$ ,  $-1.98$  and  $-1.30$  and  $R^2$  values are 0.94, 0.95 and 0.92 for France, S. Korea and Canada respectively. (b) Birth rate vs population size (thousands) for Spain, Australia and the UK, every five years from 1960–2000. Power law exponents are  $-3.83$ ,  $-0.62$  and  $-4.27$  and  $R^2$  values are 0.91, 0.93 and 0.89 for Spain, Australia and the UK respectively. (c) Birth rate vs population size (thousands) for the USA and Japan every five years from 1960–2000. Power law exponents are  $-0.76$  and  $-2.53$  and  $R^2$  values are 0.68 and 0.83 for the USA and Japan respectively. (d) Birth rate vs population size (thousands) for the Netherlands and Italy every five years from 1960–2000. Power law exponents are  $-1.71$  and  $-5.82$  and  $R^2$  values are 0.82 and 0.92 for the Netherlands and Italy respectively. Data are from US Dept of Labor (2007) and UN (2006, 2007).

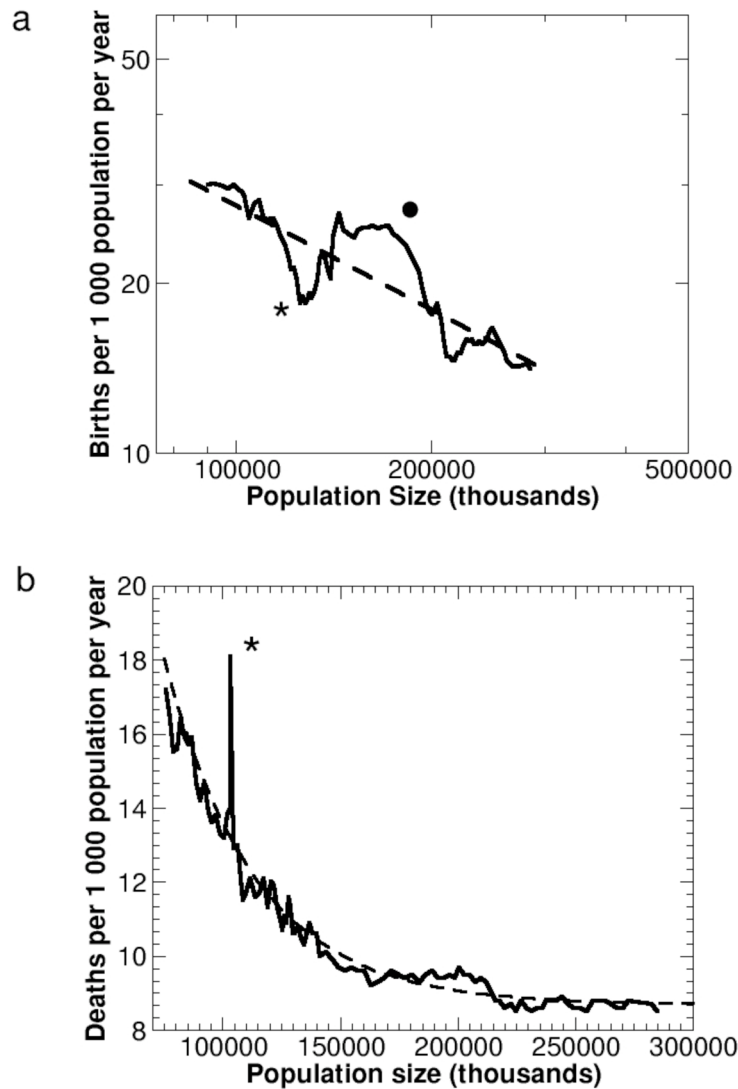


Figure A3. (a) Birth rate versus population size in the United States every year from 1900 to 2001, data (solid line) and fitted power law function (dashed line). The best fit parameter for the power law exponent is  $-0.61$ , for which  $R^2 = 0.73$ . The asterisk denotes the Great Depression and the closed circle denotes the Baby Boom. Data are from US Dept of Labor 2007. (b) Death rate vs population size in the United States every year from 1900 to 2001, data (solid line) and fitted exponential model Eq. 14 (dashed line). The best fit parameters for Eq. 14 are  $a = 8.7$ ,  $b = 64.6$ ,  $c = 2.58 \times 10^{-8}$ , for which  $R^2 = 0.97$ . The spike in death rates denoted by the asterisk was caused by the 'Spanish Flu' pandemic of 1918–1919. Data are from the US Census Bureau (2007).

## References

- Office of Productivity and Technology, Bureau of Labor Statistics, US Dept of Labor 2007. Comparative real gross domestic product per capita and per employed person: sixteen countries, 1960–2006. <[ftp://ftp.bls.gov/pub/special.requests/foreignlabor/flsgdp.txt](http://ftp.bls.gov/pub/special.requests/foreignlabor/flsgdp.txt)> (accessed 20 December 2007).
- Population Division, US Census Bureau 2007. Table 1: Annual estimates of the population of metropolitan and micropolitan statistical areas: April 1, 2000 to July 1, 2005 (cbsa-est2005-01) (accessed August 2007).
- Population Division of the Dept of Economic and Social Affairs of the United Nations Secretariat. 2006. World population prospects: the 2006 revision. <<http://esa.un.org/unpp>> and <<http://unstats.un.org>> (accessed 25 December 2007).
- Population Division of the Dept of Economic and Social Affairs of the United Nations Secretariat 2005. World urbanization prospects: the 2005 revision. <<http://esa.un.org/unpp>> and <<http://unstats.un.org>> (accessed 25 December 2007).
- United States Census Bureau 2007. The 2007 statistical abstract of the United States, historical statistics. <<http://www.census.gov/compendia/statab/hist stats.html>> (accessed 20 December 2007).

## Appendix 2. Model derivation from Cobb-Douglas production function

Here, it is illustrated how Eq. 11 of the main text can be derived starting from somewhat different principles. Rather than relying upon empirical data from Zipf's law and demographic data relating city size to GDP in order to determine the GDP per capita as a function of population size, the classic Cobb-Douglas production function is used (Douglas 1976). A country's gross domestic product  $G$  can be related to three factors of production—physical capital  $K$ , labor supply  $L$ , and technology  $T$ —through the Cobb-Douglas production function (Douglas 1976)

$$G = TL^\alpha K^{1-\alpha} \quad (16)$$

where  $\alpha$  is a constant and  $0 < \alpha < 1$  (Barro et al, 2004). This function remains one of the most widely used production functions in economics and has been verified against national economic data (Douglas 1976), although its use is not uncontroversial (Antràs 2004). As long as the variables  $T$ ,  $L$  and  $K$  in Eq. 16 are sufficiently independent, they can in turn be related to population size  $N$ , hence making  $G$  a function of  $N$ . (Note that an empirical test would actually be necessary to determine whether  $T$ ,  $L$  and  $K$  are sufficiently independent from one another.) Although some simplifying assumptions will be made, it is important to emphasize that any function of the form  $g(N) = G(N)/N = N^\theta$ , where  $\theta > 0$ , will yield results that are consistent with our conclusions. Under the assumption that the proportion of people who work and the hours they work are constant, one obtains

$$L = \lambda N \quad (17)$$

where  $\lambda$  is a proportionality constant. In the United States, for example,  $\lambda$  actually increased from 0.40 in 1910 to 0.66 in 2007 (Johnston and Williamson 2004, US Bureau of Labor Statistics 2007), but this increase was small compared to the absolute increase in population size over that time. Moreover, a function  $\lambda(t)$  increasing with time would not change the qualitative results. Capital  $K$  can also be expressed as a function of  $N$ . The development of centres of large population density is historically associated with increased productivity. As population grows over time, capital accumulates relatively more quickly due to the economics of scale and economics of agglomeration (Kuznets 1968). A larger population enables sharing of infrastructure, larger and more diverse markets, more trade, and greater labor specialization. Hence, if all other factors (such as level of technological development or economic policies) are the same for a population, its productivity is greater:

$$K = \kappa N^\theta \quad (18)$$

where  $\kappa$  is a proportionality constant and  $\theta > 1$  due to economics of scale and agglomeration. A similar assumption has been made previously at the level of urban centres and therefore should be equally valid at the country level (Glaeser et al 1992, Drennan 2002, Bettencourt et al. 2007). Note that Eq. 18 need apply only to a population in the later stages of industrialization. Combining Eq. 16–18 yields

$$g = G/N = AN^\theta \quad (19)$$

where  $\theta = (\delta - 1)(1 - \alpha) > 0$  and  $A = T\lambda^\alpha \kappa^{1-\alpha}$ .

Note there are two situations where Eq. 18 may not hold. One situation is when short-term fluctuations in population size oc-

cur, due to an influx or exodus of labor for example. In this case, population size will change but capital will not immediately respond. The second situation is when a population is open to the flow of capital or knowledge from other populations. In this situation, a country can accumulate capital faster than suggested by the population size alone (which is why small island countries can be wealthy). Hence, Eq. 18 applies only to closed populations over long timescales where the population size changes due to gradually evolving domestic birth and death rates. Also note that economic policies or perhaps even cultural differences will modulate these basic trends; this is indicated by the greater variance in the plot of birth rates versus GDP across countries ( $R^2 = 0.64$ ) as compared to within a given country ( $R^2 = 0.91$ ) (Fig. 1b in main text).

Technological innovation is thought to explain a significant portion of perpetually rising GDP per capita (Solow 1957). Here technology  $T$  has been treated as a constant, but it is possible that  $T$  also increases with  $N$  since larger populations can support more researchers. This assumption would also result in a function  $g(N)$  of the form  $g(N) \propto N^\theta$  and would yield results consistent with the remainder of our analysis. Also note that different functional relations between factors of production  $T$ ,  $L$ ,  $K$  and population size  $N$  would yield consistent results as long as the combined exponent of  $T(N)L(N)^\alpha K(N)^{1-\alpha}$  is greater than one.

Population dynamics are described by Eq. 10 of the main text, as before. Combining Eq. 9, 10 and 19 yields

$$\frac{dN}{dt} = N \left( \frac{b_0}{A^\sigma} \frac{1}{N^{\theta\sigma}} - d \right) \quad (20)$$

which is solved by

$$N(t) = \left\{ \frac{b_0}{A^\sigma d} + e^{-d\theta\sigma t} \left[ N(0)^{\theta\sigma} - \frac{b_0}{A^\sigma d} \right] \right\}^{\frac{1}{\theta\sigma}} \quad (21)$$

Equation 20 and 21 are formally identically to Eq. 11 and 12 of the main text, with

$$\theta = \beta - 1 \quad (22)$$

$$A = \frac{\gamma A}{\Omega^\beta} \quad (23)$$

Hence, one arrives at the same model from two different approaches. The second approach described in this subsection ignored city-level dynamics and used theoretical arguments based on production functions, rather than empirical data.

## References

- Antràs, P. 2004. Is the U.S. aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution. – *Contrib. Macroecon.* 4(1).
- Barro, R. and Sala i Maring, X. 2004. *Economic growth*. – MIT Press.
- Bettencourt, L. et al. 2007. Growth, innovation, scaling, and the pace of life in cities. – *Proc. Natl Acad. Sci. USA* 104: 7301–7306.
- Drennan, M. 2002. *The information economy and American cities*. – Johns Hopkins Univ. Press.
- Douglas, P. 1976. The Cobb-Douglas production function once again: its history, its testing, and some new empirical values. – *J. Political Econ.* 84: 903–16.
- Glaeser, E. et al. 1992. Growth in cities. – *J. Political Econ.* 100: 1126–1152.
- Johnston, L. and Williamson, S. 2004. The annual real and nominal GDP for the United States, 1790-present. *Economic History Service*. – URL <<http://www.eh.net/hmit/gdp>>.
- Kuznets, S. 1968. *Toward a theory of economic growth*. – W.W. Norton, New York.
- Solow, R. 1957 Technical change and the aggregate production function. – *Rev. Econ. Stat.* 39: 312–320.

US Bureau of Labor Statistics 2007. Employment situation summary,  
<<http://www.bls.gov/news.release/empstat.nr0.htm>> (accessed 25  
May 2007).