SOCIAL SCIENCES

Urban growth and the emergent statistics of cities

Luis M. A. Bettencourt[1,2]

Urban theory models cities as spatial equilibria to derive their aggregate properties as functions of extensive variables, such as population size. However, this assumption seems at odds with cities' most interesting properties as engines of fast and variable processes of growth and change. Here, we build a general statistical dynamics of cities across scales, from single agents to entire urban systems. We include agents' strategic behavior to produce predictable growth rates, which requires balancing relative incomes and costs over time. We implement these dynamics using stochastic differential equations and control theory to demonstrate a number of general emergent properties of cities deriving from limit theorems applied to growth rates. This framework establishes necessary conditions for scaling to be conserved by urban dynamics and shows how exponent corrections can be calculated. These ideas are tested using stochastic simulations and a long timeseries for 382 US Metropolitan Areas over nearly five decades.

INTRODUCTION

Classical approaches to urban theory—in economic geography (1–3) and, more recently, in complex systems (4)—often treat cities as spatial equilibria, where a balance of benefits and costs is achieved out of a set of social and economic exchanges, including wages, land rents, and transportation costs (1, 3–5). While these modeling approaches have proven powerful for generating quantitative predictions in agreement with many observed properties of cities (3–5), they leave unresolved two fundamental problems: the problem of statistics and the problem of growth.

Both growth and statistics denote a broad set of phenomena that must be unpacked so that we can fully appreciate what is at stake. By statistics, we mean that in dealing with real cities, we must appreciate the wide variation between individuals and places. This variation has positive manifestations in that cities are extremely diverse in terms of the types of the people and lifestyles they support, including a broad set of coexisting cultures, professions, languages, races, and ethnicities (6–9). This interdependent functional diversity is what J. Jacobs famously called organized complexity and is at the heart of the "kind of problem a city is" (10). Negative expressions of these same heterogeneities are also familiar, such as ethnic, racial, and economic segregation (11, 12), inequality, and variable access to justice and opportunity. Moreover, it has been observed that these differences between places and people within each city are persistent over time (9, 12) and do not have the fleeting character of noisy fluctuations in statistical physics. Instead, they can pile up over time and lead to patterns of cumulative advantage and disadvantage (9, 13), which are at the root of most challenges of human development. Thus, the problem of statistics in cities deals not only with the existence of structural differences on how the same quantity is distributed across different people and places but also with the temporal persistence and amplification of these effects.

By growth, we mean that (modern) cities are characterized by fast, typically exponential change across many variables. On the one hand, modern cities tend to experience annual population growth rates between about a fraction of 1 and 3 to 4%, as we shall see below. Exceptions exist at either end at least over some periods of time, as different places experience contextually specific factors. However, the principal change in modern cities is the fast pace of their economic growth and technological transformations. Across the world today, we observe rates of urban economic growth that are typically larger than those in their corresponding populations, reaching in some cases 10% a year, with 2 to 4% being typical (14). These growth rates mean that the size of a city's economy doubles every few decades, making it possible to transition from poverty to wealth in one or two generations, as has happened in many places over the last century. With such fast growth rates at play, how is it tenable to model cities as spatial equilibria? Even more importantly, how do different resource growth rates, experienced by different households, neighborhoods, or cities, shape the heterogeneity (inequality) of outcomes for different people? Why are cities not torn apart by differential growth more often?

It turns out that these two problems, of statistics and of growth, are intimately connected and must be tackled together. This is a twist on classical statistical mechanics linking the strength of fluctuations to dissipation around equilibrium (15), which, in the context of exponential growth processes in populations, takes a character that we may call fluctuation-amplification, typical of evolutionary processes (16). The literature of complex systems applied to urban growth, especially from the perspective of geography, has demonstrated the importance of this type of stochastic nonlinear dynamics with strong feedbacks for a number of decades (17, 18).

To continue to make progress on these issues, an analytical approach is necessary that identifies and articulates the essential joint mechanics of scaling, growth, and statistics in cities. Here, we show how this synthesis can be constructed and illustrate its theoretical and empirical implications using stochastic simulations and a particularly long time series of wages and populations of U.S. metropolitan statistical areas (MSAs) covering nearly five decades. The central insight is the realization that the budget constraint used to define functional cities (1–4) is not static or homogeneous across agents but must be managed adaptively to promote growth and avoid instability. This is because what is being controlled are stochastic net resource flows over time, as incomes minus costs, not static quantities such as forces. It is precisely this fluctuating difference that accumulates to drive resource growth for each agent and for

[1]Mansueto Institute for Urban Innovation, Department of Ecology and Evolution and Department of Sociology, University of Chicago, Chicago, IL 60637, USA. [2]Santa Fe Institute, Santa Fe NM, 87501, USA. Email: bettencourt@uchicago.edu

cities as populations of agents, with implications for both aggregate growth and inequality.

The manuscript is organized following the scheme of Fig. 1. We first show how scaling analysis isolates the parameters that control urban growth, providing a number of quantitative targets for explanation and prediction. We then introduce standard theory for stochastic geometric growth and outline its properties. This allows us to establish the connection between well-known models of cities as spatial equilibria and processes of stochastic growth for individual agents. This connection disaggregates a city-wide budget condition to the level of agents and requires that they act strategically in their own self-interest to control fluctuations in net incomes, which is naturally handled via adaptive temporal averaging of expenditures. This is the central assumption of this manuscript, which leads to the expectation that net income volatilities are kept finite and small and that temporal averages of incomes and expenditures become statistically dependent. A number of results follow from standard limit theorems: The statistics of resources, incomes, and costs at the agent and group level become asymptotically lognormal, even as these quantities grow exponentially. The self-similarity of growth processes across group sizes also emerges, defining running couplings

characterizing the mean growth rate for populations and its associated volatility. Explicitly computing these quantities allows us to identify the circumstances when urban scaling is conserved by the system dynamics. Conversely, we show when these conditions are violated, creating corrections to mean-field scaling exponents when volatilities are scale dependent but small and the breakdown of scaling if they become large. These procedures are illustrated using data on wages for U.S. metropolitan areas. We finish with a discussion of the significance of the results toward a general statistical dynamics of cities and its relation to analogous dynamics of resource flows in other complex systems.

RESULTS

Scaling relations and statistical deviations

Scaling analysis provides a simple and straightforward way to characterize urban quantities (*19*) and extract average agglomeration effects at play across cities of different sizes in an urban system. This section also shows that scaling analysis provides an efficient parameterization of growth processes, different from growth accounting in economics (*20*).
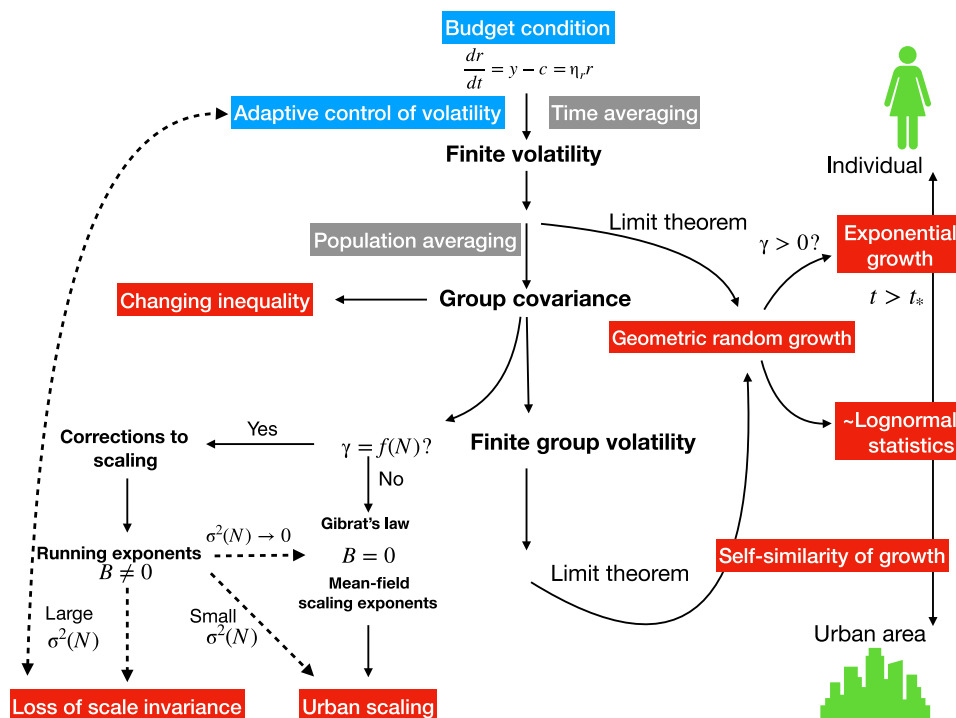

Fig. 1. Scheme of statistical theory, assumptions, and derived consequences. Basic assumptions are shown as blue boxes, and derived results are shown as red boxes; arrows indicate outcomes, while dashed lines represent alternative scenarios. The budget condition, $y - c$, is the common basic assumption for urban agents, generalizing energy conservation in simpler systems. Recognizing its dynamical, stochastic nature leads to the central assumption of the manuscript that agents must actively control its associated volatility; the simplest way to do this is through the time averaging of expenditures (consumption smoothing). Then, the resource growth rate volatility, σ^2, becomes finite and small, both at the individual and group levels. This leads to emergent stochastic geometric dynamics of resources both at the individual and population levels with exponential growth and lognormal statistics observable at long times (right). Averaging over populations derives the growth rate statistics for cities, which determines when dynamics become self-similar across scales and preserve urban scaling (left): First, if under group averaging variations of the growth rate (γ) are correlated to those in agent resources (r) inequality will change within the population. Second, if effective growth rates are independent of population size, the dynamics becomes self-similar and urban scaling is preserved over time. Alternatively, if the growth rate volatility(σ^2) is population size dependent, corrections to mean-field exponents result: They are calculable via $B \neq 0$ and are controlled by the volatility's magnitude. For large $\sigma^2(N)$, the statistics become dominated by fluctuations, and urban scaling breaks down. The existence of strong group volatilities contradicts the assumption of effective control at lower levels. This regime would be unstable, signaling the loss of control over resource flows for most of the population and entailing wide-spread crises and eventual collapse. See text for detailed notation.

The starting point is to write an extensive positive quantity, $Y_i(t)$, here the total wages paid in city i over a time period t, as

$$Y_i(N_i, t) = Y_0(t) N_i(t)^\beta e^{\xi_i(t)} \qquad (1)$$

where $Y_0(t)$ is time dependent but independent of city population size $N_i(t)$, which also changes over time due to standard demographic processes. The parameter β is the scaling exponent, and the quantity $\xi_i(t)$ is the time-dependent deviation from the average scaling prediction for city i. This expression is exact because any deviation from the average scaling relation, $Y_i(N_i, t) = Y_0(t)N_i(t)^\beta$, is absorbed in the residuals $\xi_i(t)$ (see section S1).

Averaging over cities can now be used to isolate a few quantities of interest. To do this, let us first take the logarithm of Eq. 1

$$\ln Y_i(N_i, t) = \ln Y_0(t) + \beta \ln N_i + \xi_i(t) \qquad (2)$$

followed by the average over cities $\langle \ln Y(t) \rangle = \ln Y_0(t) + \beta \langle \ln N_i \rangle$. This ensemble average is defined explicitly by $\langle \ln Y(t) \rangle = \frac{1}{N_c} \sum_{i=1}^{N_c} \ln$

$Y_i(N_i, t)$, $\langle \ln N(t) \rangle = \frac{1}{N_c} \sum_{i=1}^{N_c} \ln N_i(t)$, where N_c is the total number of cities in the urban system such as the United States. We will refer to the quantities $\langle \ln Y(t) \rangle$ and $\langle \ln N(t) \rangle$ as centers (21), which are collective coordinates tracking the temporal motion of the entire system of cities (yellow symbols in Fig. 2, A and B) analogous to center-of-mass coordinates in many-body physics.

By definition, the ensemble average of the deviations is zero, $\langle \xi(t) \rangle = 0$, so that the second (variance) and higher moments of the ξ become the leading quantities of interest. From these expressions, we can write the deviations $\xi_i(t)$ as

$$\xi_i(t) = \ln \frac{Y_i(N_i, t)}{Y_0(t) N_i^\beta(t)} = [\ln Y_i(N_i, t) - \langle \ln Y(t) \rangle] - \beta[\ln N_i(t) - \langle \ln N(t) \rangle] \qquad (3)$$

The first expression is the most common interpretation of the ξ_i as (multiplicative) residuals from the scaling relation, whereas the second makes their status as deviations from the collective coordinates (centers) explicit. For these reasons, the $\xi_i(t)$ provides a city size-independent
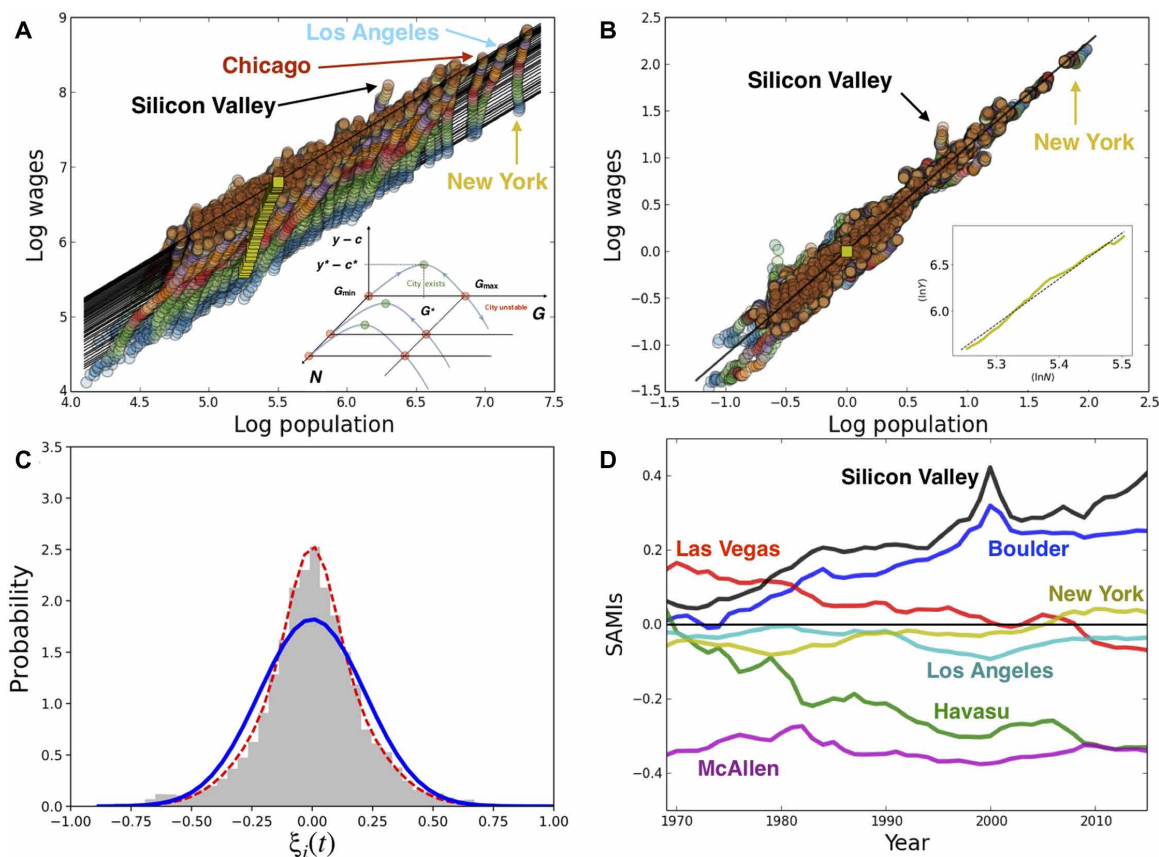

Fig. 2. Urban scaling and the dynamics of growth and deviations. (**A**) Total wages for U.S. metropolitan areas 1969–2016. Each circle is a city in a given year from blue (1969) to brown (2015). Yellow squares show the urban system's centers ($\langle \ln N \rangle$, $\langle \ln Y \rangle$), which account for collective economic and population growth (movement, upward and to the right). Urban scaling relations for each year (black lines) are derived through the consideration of a short-term spatial equilibrium (inset), which changes on a very slow time scale. (**B**) Centered data, obtained from (A) by removing the center's motion (inset). This allows the decomposition of temporal change into two separate processes: collective growth (center's motion) and deviations from scaling, $\xi_i(t)$, characteristic of each city i. We see that scaling with a common exponent (global fit $\beta = 1.114$, 95% confidence interval = [1.111, 1.117], $R^2 = 0.935$) is preserved over time, and net growth is a property of the urban system and not of individual cities. (**C**) The statistics of deviations, obtained from the residuals of the centered scaling fit of (B). While the ξ distribution is well localized and symmetrical, it is not very well fit by a normal distribution (blue line). Instead, the red dashed line, which follows from theory developed in the paper, produces a much better account of the data. (**D**) The deviations $\xi_i(t)$ of a few selected cities: Silicon Valley (San Jose–Santa Clara MSA) and Boulder, CO show two of the more exceptional trajectories in wage gains for their city sizes, whereas Las Vegas, NV and Havasu, AZ illustrate wage losses. New York City, Los Angeles, and the exceptionally poor McAllen, TX show no relative change in their positions over nearly 50 years.

measure of city performance and are also known as scale-adjusted metropolitan indicators (SAMIs) (*22*). Characterizing these three quantities, the two centers and the statistics of the ξ, gives a complete description of growth and deviations in a system of cities and separates collective effects from idiosyncratic events in each city.

Figure 2 illustrates the meaning of these quantities. Figure 2A shows the total wages, $Y_i(t)$, for U.S. MSAs between 1969 and 2016 (47 years), year by year (colors light blue to brown). The growth trajectory of some specific cities, such as New York City, Los Angeles, Chicago, or Silicon Valley (San Jose–Santa Clara MSA), is easily visualized in this way. The solid lines show the scaling relation for each year (see caption for details). We see how scaling is a good fit to the data each year, reproducing a slowly shifting spatial equilibrium in each instance (inset) (*4*). We also see how the position of the center (yellow squares) moves from year to year, reflecting the overall growth in population (shifts to the right) and especially in wages (movement upward). These results include both real and nominal growth of wages due to inflation.

Figure 2B shows the result of removing collective growth by moving all data clouds so that their centers coincide at the origin (0,0) (*21*). We see that removing the center's motion (inset) results in the reproduction of the same scaling pattern at each time, with additional small and slow moving deviations, $\xi_i(t)$, changing only slightly from year to year. Figure 2C shows the histogram of these deviations (gray) about the overall best fit scaling relation (Fig. 2B), pooled across all years. We observe that the statistical distribution is well localized and symmetrical about the origin, but that is not very well fit by a normal distribution (blue line). A better fit is provided by another model (dashed read line), which will be derived below. Last, in Fig. 2D, we see the change in $\xi_i(t)$ over time for some of the most extreme trajectories. Specifically, some cities became substantially richer in relative terms over this period (Silicon Valley and Boulder), some experienced loss in economic status (Las Vegas, NV and Havasu, AZ), and a few others, such as New York, Los Angeles and McAllen, TX (one of the worst performers, by this measure), have not changed much. These trajectories also show how slow the change in the ξ's is. Particular events that affect different cities at specific times are easily identifiable, such as the dot-com economic boom and bust around 2000, specifically for Silicon Valley and Boulder.

Stochastic growth dynamics

We now seek to connect the macroscopic statistics of entire cities to a microscopic general model of single agent's behavior. We note that the budget condition, which is taken as the starting point of an important set of urban theories (*1–4*), is a version of the fundamental law of energy conservation. Hence, it must apply to every single agent and to cities as collections of agents.

To see this, let us start by introducing a variable, $r(t)$, denoting the accumulation of the net quantity of $y(t)$ over time t. For example, if y stands for an agent's income, then r becomes its monetary wealth, but we should think of r more generally as resources that can be grown over time and used in turn (reinvested) to generate more y and so on. An important noneconomic example, which applies to premodern human societies and other biological populations, is when r is stored energy and y is an energy income per unit time. We write the dynamics of r, given y, as

$$\frac{dr(t)}{dt} = y(t) - c(t) = \eta_r(t)\, r(t) \qquad (4)$$

where η_r is the stochastic growth rate of resources. The first equality in Eq. 4 is pure accounting, stating that resources grow by the difference between income and costs, c, (i.e., net income) over some time interval, dt. Costs in cities are local and include real estate rents, transportation, and consumption, as well as others, such as health care and losses resulting from crime or poor urban services. In this sense, costs and benefits are also affected by migration decisions about where to live and work. The centerpiece of this equation is the difference between income and costs $y(t) - c(t)$, which must be balanced by all agents in their specific environments. For urban agents, this difference is the "budget condition" for the spatial equilibrium that defines a city according to the Alonso model of economic geography and urban scaling theory (*1, 3, 4*). In the original version, this difference is typically set to zero, although the meaning of incomes and costs is rather flexible and can include savings (*23*). In urban scaling theory, this difference is nonzero in general (*4*) and becomes the target of maximization through the self-consistency of infrastructural and social network properties. This implies that a positive difference between incomes and transportation costs is necessary for cities to exist (Fig. 2A, inset) and to generate exponential resource growth. It also implies that the scaling of resources, incomes, and costs has the same population size dependence characterized by a single common exponent for all these quantities, β > 1.

The second equality in Eq. 4 is a definition of the growth rate η_r implying that $\eta_r(t) \equiv \frac{y(t)}{r(t)} - \frac{c(t)}{r(t)}$. Equation 4 is not an arbitrary modeling choice: It is the standard starting point for modeling population growth and human behavior where time, effort, or resources are invested strategically. Among many examples, it is the standard model for city population growth (*24*), the standard model of financial mathematics and asset pricing (*25, 26*), and the one good wealth accumulation model, which is the basic tool in economics to analyze dynamical issues of wealth inequality (*23*). Stochastic proportional growth and resulting lognormal (and power law) distributions are associated with many forms of human behavior, including the statistics for the time to complete a task, epidemic dynamics, demography and even the statistics of marriage age [see (*27*) for a review].

Let us see how this model works in practice. Because the equation is nonlinear (the stochastic term is multiplicative), we have to be careful and use the rules of stochastic (Itô) calculus to integrate its solution in time, leading to

$$\ln\frac{r(t)}{r(0)} = \left(\bar{\eta}_r - \frac{\sigma_r^2}{2}\right)t + \Theta(t) \qquad (5)$$

where the $\bar{\eta}_r$ and σ_r^2 are the mean and variance of η_r, respectively, in the usual sense of those obtained over the probability density of η_r. Now, let us define the average effective growth rate $\gamma_r = \bar{\eta}_r - \frac{\sigma_r^2}{2}$. This quantity is fundamental in geometric random growth models and will recur in the discussion below. Keeping track of physical dimensions tells us that $\bar{\eta}_r$ and γ_r are temporal rates and have dimensions of $(\text{time})^{-1}$. Thus, the standard deviation (SD) σ_r (known as the volatility) has dimensions of $(\text{time})^{-1/2}$. The stochastic noise $\Theta(t)$ is the sum over the integration time, t, or more explicitly, $\Theta(t) = \sum_{l=1}^{t} \varepsilon_r(t_l)$, with $\varepsilon_r(t_l) = \eta_r(t_l) - \bar{\eta}_r$. This is a random variable with zero mean. Because it is the sum of stochastic variables, we expect $\Theta(t)$ to express universal behavior as a consequence of the central limit theorem (*25*). In the simplest case, where η_r is statistically independent across time with finite variance, we obtain that $\Theta = \sigma_r W(t)$ is a Wiener

process, which is a normal variable with zero mean and variance $\sigma_\Theta^2(t) = \sigma_r^2 t$. This will later define the property of ergodicity for stochastic growth, which means that for long-time averages of growth rates, the mean dominates the variance. This is not to be confused with the more general property with the same name in statistical physics, which means that all allowable spaces of a dynamical system will be sampled subject to constraints. The point of the present paper is to show that path dependence for particular cities can co-exist with simple emergent statistics for the ensemble of cities.

A number of key general results follow from this solution and associated limiting theorems. First, (i) the central limit of Θ implies that $\ln\frac{r(t)}{r(0)}$ approaches, in the same limit of long times, a Gaussian variable with time-dependent mean $\gamma_r t$ and variance $\sigma_r^2 t$. This implies, in turn, that (ii) $r(t)$ is asymptotically distributed as a lognormal variable, a result that will become important later. In turn, (iii) the temporal mean growth rate $\frac{1}{t}\ln\frac{r(t)}{r(0)} = \gamma_r + \frac{\Theta(t)}{t} \to \gamma_r$, for long times,

as a result of the behavior of $\Theta \sim t^{1/2}\sigma_r$. Last, (iv) the characteristic time, $t_* = \frac{\sigma_r^2}{\left(\bar{\eta}_r - \frac{\sigma_r^2}{2}\right)^2}$, marks the interval necessary for net exponential growth to become apparent over the (shorter-term) effect of fluctuations, which average out for longer times.

These properties are illustrated in Fig. 3 (A to C), obtained from numerical simulations of Eq. 4, with η_r taken as Gaussian white noise. The asymptotic behavior of all quantities depends on whether the effective growth rate is positive $\gamma_r > 0$ or equivalently $\bar{\eta}_r > \sigma_r^2/2$ (Fig. 3, A and C). When this condition holds, there is net growth (Fig. 3A, blue and orange trajectories). Growth of r becomes apparent on a time scale longer than t_* (Fig. 3A), which can be very short when the volatility is small. In this regime, the distribution narrows on the scale of the mean as t becomes larger, and predictable exponential growth emerges. However, when $\bar{\eta}_r < \sigma_r^2/2$ the mean growth rate is negative, and $r(t)$ decays toward zero while experiencing large
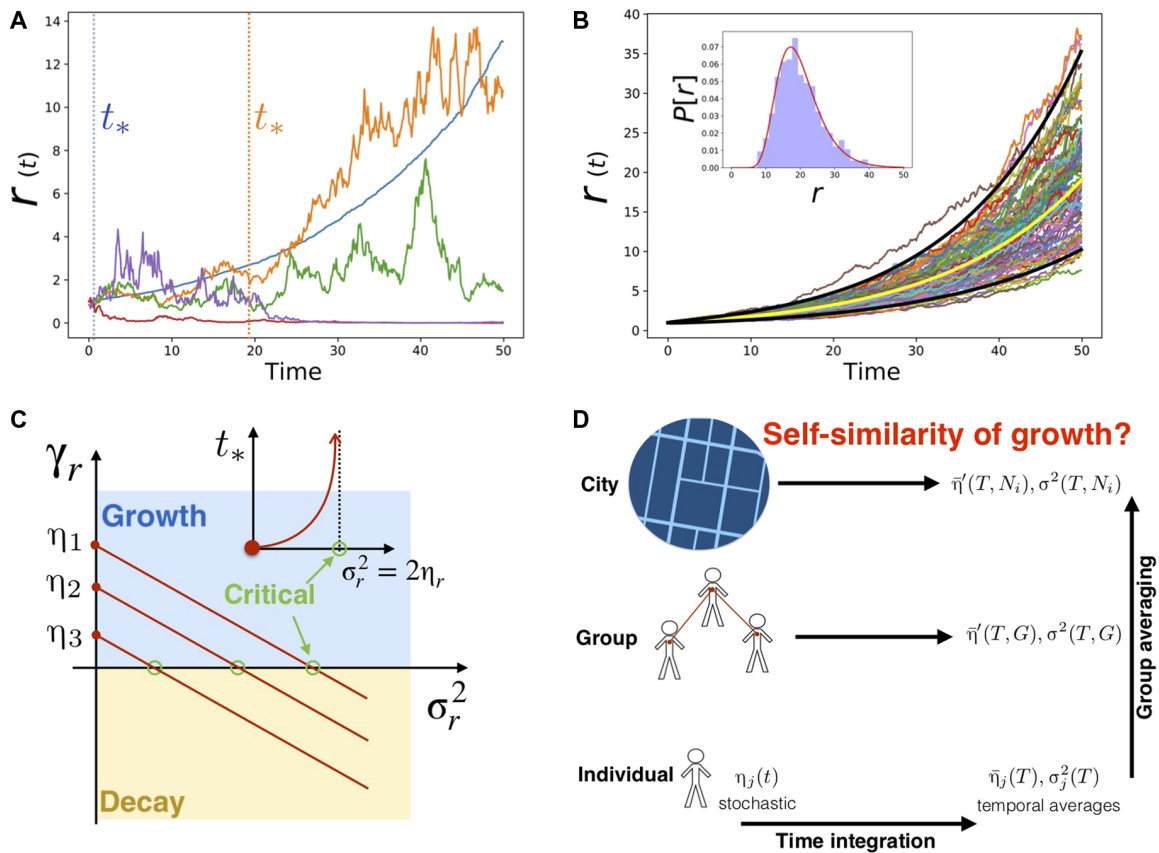

Fig. 3. General properties of stochastic geometric growth and their consequences for cities. (A) Example of growth trajectories for a simple process of geometric Brownian motion (Eq. 4). The blue trajectory shows typical growth with small fluctuations and positive effective growth rate, the orange line shows a similar situation with larger fluctuations, and the green line shows a trajectory with critical $\gamma_r = 0$. The purple and red lines illustrate negative effective growth rate trajectories. The critical growth time, t_*, is shown for growing trajectories. (B) An ensemble of trajectories with stochastic growth rates similar to those of U.S. MSAs, starting with the same initial conditions. The yellow line shows the temporal trajectory of the ensemble average, and the black lines show the 95% confidence interval. Note that both the mean and the SD are time dependent (see text). The inset shows the resource distribution at a later time, which becomes asymptotically lognormal (red line). (C) The general properties of stochastic growth imply that a positive growth rate is necessary to overcome temporal decay due to rate fluctuations. If volatility increases, growth will ultimately stop, and decay will ensue. The critical point $\gamma_r = \bar{\eta}_r - \frac{\sigma_r^2}{2} = 0$ is characterized by large fluctuations with a diverging t_* so that agents will not be able to tell whether they are experiencing growth and may be unable to exert effective control (see text). (D) Under general conditions, multiplicative random growth can be self-similar across group sizes, providing a simple theory that applies at all scales, from individual agents to populations and cities (see section S3) (29). However, the key parameters of the theory "run" across scales and are in general sensitive to both group size, temporal averaging, and inequality. These dependences define urban scaling as a dynamical statistical theory beyond the mean-field approximation.

fluctuations (Fig. 3A, red and purple trajectories). When $\bar{\eta}_r \simeq \sigma_r^2/2$, there is very little growth or decay, and the dynamics appear purely random (Fig. 3A, green trajectory). For $\gamma_r \geq 0$, r displays wide fluctuations (asymptotically distributed as a lognormal; Fig. 3B, inset), which appear wider and wider for larger t (Fig. 3B). Multiplicative growth processes thus have the curious property that because drift is asymmetric, a positive mean growth rate $\bar{\eta}_r > 0$ is not sufficient to guarantee long-term growth; instead, a finite threshold $\bar{\eta}_r > \sigma_r^2/2$ must be overcome. Approaching this threshold from a regime with growth, an agent will experience wild fluctuations as t_* goes to infinity (Fig. 3C, inset) and will struggle to tell whether growth persists and estimate its time scale to plan. As a consequence, low volatility and positive average rates are necessary for sustained growth (Fig. 3C). Given these results for individual agents, it becomes critical to establish the conditions for these dynamics to apply also for populations such as for entire cities (Fig. 3D) and to determine how corresponding growth rates change across scales.

Stable growth through budget control

From the general properties of stochastic growth processes, we can conclude that any agent seeking growth must aim at a positive mean growth rate and small volatility. The conundrum is that the volatility and the mean growth rate are, to a large extent, properties of the environment, outside the agent's control. What is under the agent's control, however, are his/her own actions, which we show next can adapt to extrinsic circumstances via processes local in time so as to produce low volatility and stable growth.

It is important to realize that, besides levels of population aggregation, there is also a hierarchy of time scales involved in the process of balancing costs and benefits and observing growth (Fig. 3D). Over the very short term, there will be moments when the agent is resource flow negative, e.g., while shopping. However, judicious choices over time should result in more even positive net flow over the longer term, integrating together periods when incomes are larger than costs (at work) and vice versa (at home, socializing, etc.). This process of balancing costs and benefits over time is necessary in dissipative complex systems because there are always resources lost in any activity or exchange. Balancing costs and benefits over time creates strong correlations between y, c, and r and results on ratios, y/r and c/r, that can become independent of the level of wealth, as we show next.

Consider the basic accounting (Eq. 4) for a single agent, $y(t) - c(t) = \eta_r(t)r(t)$. As we have seen, dividing by $r(t) > 0$ gives us the definition of the growth rate $\eta_r(t)$. Defining the two resulting ratios as $b(t) \equiv y(t)/r(t)$ and $a(t) \equiv c(t)/r(t)$ and averaging over time leads to

$$\frac{1}{t}\int_0^t dt'\,[b(t') - a(t')] = \bar{b} - \bar{a} + \frac{1}{t}\int_0^t dt'(\eta_r(t') - \bar{\eta}_r) \rightarrow \bar{b} - \bar{a} = \bar{\eta}_r \tag{6}$$

This means, in general, that we can also define $\eta_r(t) = \bar{\eta}_r + \varepsilon_r(t)$, where $\varepsilon_r(t)$ is the error (or fluctuations) away from the growth rate's temporal mean such that $\frac{1}{t}\int_0^t dt'\varepsilon_r(t') \rightarrow 0$, as we have seen for $\Theta(t)$ in the previous section.

What kind of process sets the statistical properties of these fluctuations? On a short-term basis, fluctuations will be large if a, b vary strongly and independently of each other. Then, the amplitude of ε_r will be large over some period of time and, if negative, may deplete stored resources ($r \rightarrow 0$), placing the agent at risk of death or bankruptcy. Thus, it is in the vital self-interest of the agent to act so as to minimize, or at least control, fluctuations.

How is this to be achieved? The point is that the variations in expenditures, $a(t)$, should not just be seen as passive costs but rather as strategic dynamical investments under the agent's control. Conversely, the returns on this investment, $b(t)$, are stochastic and will always fluctuate because of environmental factors (Fig. 4A). Thus, $a(t)$ should be chosen to generate a target growth rate and reduce fluctuations, in other words, to achieve stable and predictable growth (Fig. 4, B and C).

To demonstrate how this can be achieved, we write the returns as $b(t) = \bar{b} + v(t)$ and the investment as $a(t) = \bar{a} + u(t)$. Here, $v(t)$ are (stochastic) variations in returns, whereas $u(t)$ will play a role of a control variable adjusted by the agent. This leads to

$$\bar{b} - \bar{a} + v(t) - u(t) = \bar{\eta}_r + \varepsilon_r(t) \rightarrow \varepsilon_r(t) = v(t) - u(t) \tag{7}$$

We must now specify how control is implemented to tame the errors. Most general practical controllers are in the Proportional-Integral-Derivative (PID) class (*28*), which specifies $u(t)$ as a function of the error, $\varepsilon_r(t)$, as

$$u(t) = k_P\varepsilon_r(t) + k_I\int_0^t \varepsilon_r(t')\,dt' + k_D\frac{d\varepsilon_r}{dt} \tag{8}$$

where k_P, k_I, and k_D are constants (in time) to be chosen by the agent. These three terms allow for different kinds of strategy to reduce fluctuations: k_P is the magnitude of an instantaneous response against the fluctuation, k_I refers to averaging of the error over time (known as smoothing, because averaged errors are smaller and converge to zero), and k_D describes a corrective reaction in the direction of the temporal change in the error. Of these, only smoothing by time integration will prove essential. Note also that $u(t)$ is a simple quantity that can be updated locally in time via the current observed error, $\varepsilon_r(t)$, and its addition and subtraction to the integral and difference, which requires remembering only two numbers. The stochastic dynamics of the errors is best captured via the derivative of Eq. 8

$$\frac{du}{dt} = k_P\frac{d\varepsilon_r}{dt} + k_I\varepsilon_r(t) + k_D\frac{d^2\varepsilon_r}{dt^2} \rightarrow k_D\frac{d^2\varepsilon_r}{dt^2} + (k_P + 1)\frac{d\varepsilon_r}{dt} + k_I\varepsilon_r = \frac{dv}{dt} \tag{9}$$

This equation for the error describes a simple driven oscillator: It is familiar from stochastic calculus when we take $\frac{dv}{dt}$ to be white noise with variance Ω^2. The solution is provided in section S2, showing that ε_r converges to a normal distribution with zero mean and variance $\sigma_r^2 = 2\Omega^2\frac{k_P+1}{k_I}$ (see Fig. 4, C and D). Making k_I larger has the double effect of accelerating the temporal convergence to a time-independent distribution and narrowing the error variance. The other parameters, k_P and k_D, can be set to zero, leading to very simple control based on the temporal averaging of the fluctuations. The effect of the environmental variance Ω^2 is simply to increase the error variance proportionally. Thus, the control process effectively filters out environmental shocks and makes the net-income variance smaller as a function of parameters chosen by the agent. This is a very simple general mechanism that allows agents to cope with environmental uncertainty and generate stable growth by adjusting their expenditures over time. Much more sophisticated strategies are possible that can maximize growth rates if more of the structure of returns, b, are known (*28*).

We see how averaging expenditures over time (known as consumption smoothing in economics) gives a general mechanism
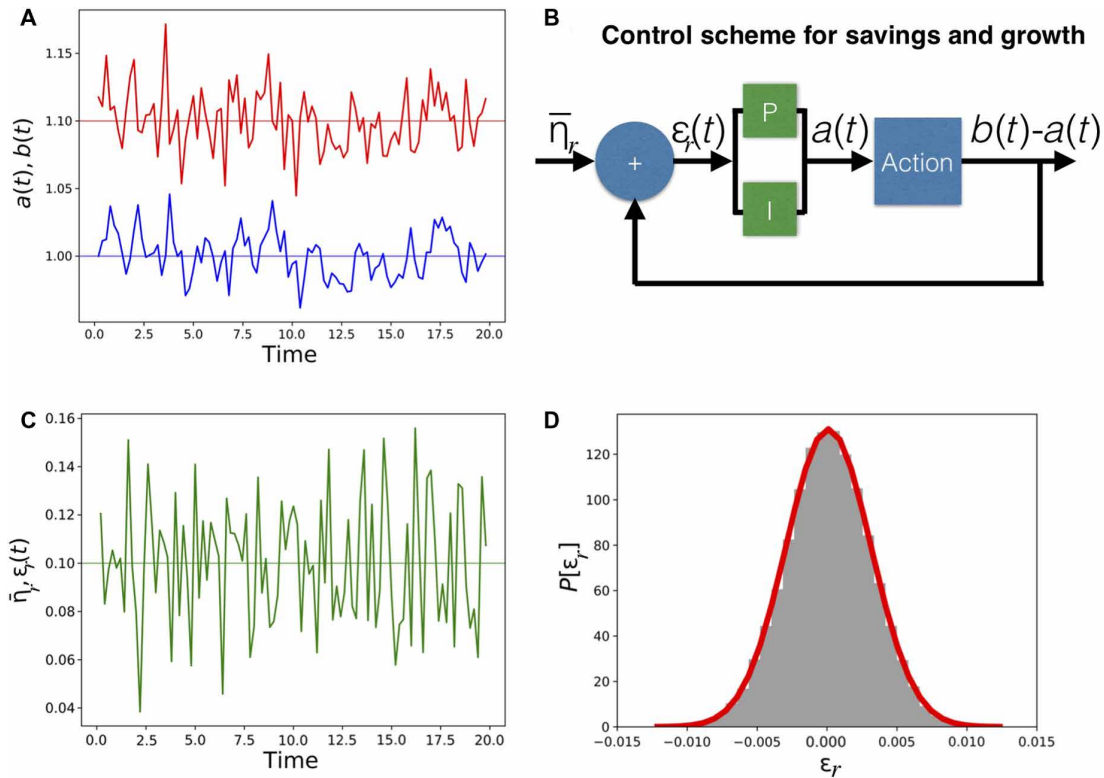
Fig. 4. Dynamically balancing income and costs via feedback control leads to simple statistics for resource growth rates. (**A**) Example trajectories for the income-to-resources and costs-to-resources ratios, $b(t)$ (red) and $a(t)$ (blue), respectively. Note that when income is larger than costs, there can be growth, but fluctuations need to be controlled. (**B**) Control scheme to deliver average growth rate and tame fluctuations $\varepsilon_r(t)$. Costs $a(t)$ become a control variable that, in part, adapts to environmental fluctuations to generate $\varepsilon_r(t)$ with small, known variance. (**C**) The dynamics of the resulting error $\varepsilon_r(t)$ is now centered around zero and (**D**) displays a Gaussian distribution (red line) with variance given by the ratio of the environmental variance to control parameters (see text). In this way, adaptive agents' behavior can lead to predictable growth in stochastic environments with a chosen variance.

whereby agents can make their average resource growth rate take on a target value, up to stochastic fluctuations with variances given by the balance between the unpredictability of the environment and the quality of their control. Effective control generates strong statistical correlations between income and costs over time, which constitute the basis for (a spatial) equilibrium. In this light, variations between agents may persist as the result of differences in their specific experienced environments and/or the quality of their control. Exposing these issues requires the consideration of averages over populations of agents as in Fig. 3D, to which we now turn.

Population dynamics and emerging inequality

To compute the growth dynamics for a city, we now define the averages over a population of size G, $r_G = \frac{1}{G}\sum_{j=1}^{G} r_j$, where r_j are individual j's resources and so on for growth rates, incomes, and costs (see section S6 for a summary of notation). To derive the corresponding dynamics, we take these averages over Eq. 4

$$\frac{dr_G}{dt} = y_G - c_G = (\eta r)_G \qquad (10)$$

where we dropped the r subscripts on the rate, for simplicity, so that in this section, $\eta_{rG} \to \eta_G$. The average of the product is

$$(\eta r)_G = \frac{1}{G}\sum_{j=1}^{G}\eta_j r_j = \eta_G r_G + \text{covar}_G(\eta, r) = [\eta_G + \text{covar}_G(\eta, r/r_G)]r_G \qquad (11)$$

The quantity $\eta_G' \equiv \eta_G + \text{covar}_G(\eta, r/r_G)$ is the effective stochastic growth rate for the group average resources, r_G. This quantity equals the simple arithmetic group average, η_G, plus a correction due to the fact that growth rate variations may not be statistically independent from variations in resources across individuals. The covariance term is familiar from evolutionary theory in the context of the Price equation (*16*) and signals selection. For example, if richer individuals experience higher growth rates across the group, then the average growth rate will be higher and vice versa. This flags the important issue that pursuing the highest possible group-level growth rates in a heterogeneous population will increase inequality. Conversely, pursuing growth such that poorer individuals enjoy higher rates leads to more equitable outcomes in distribution but subtracts from the average η_G because the covariance is negative.

To complete the derivation, we now characterize the mean and stochastic components of η_G'. We express the individual growth rate, as in the previous section, $\eta_j = \bar{\eta}_j + \varepsilon_j$, which leads to $\eta_G = \frac{1}{G}\sum_{j=1}^{G}\eta_j = \frac{1}{G}\sum_{j=1}^{G}(\bar{\eta}_j + \varepsilon_j) = \bar{\eta}_G + \varepsilon_G$, where $\bar{\eta}_G$ is the group mean of individual temporal means and ε_G is a stochastic noise term resulting from the group average of the errors for each individual. The properties of ε_G are inherited from those of each agent and their statistical correlations. The mean remains zero, while the variance is given by $\sigma_G^2 = \frac{1}{G^2}\sum_{j,k}^{G}\sigma_j\sigma_k\rho_{jk}$, where σ_j and σ_k are the volatilities for agents j and k and ρ_{jk} is the correlation matrix between them. (The correlation matrix is symmetric, with $-1 \leq \rho_{jk} \leq 1$ and with

ones along the diagonal, corresponding to each agent's squared volatilities).

In the simplest case, when errors are statistically independent across agents, $\rho_{jk} = 0$ for $k \neq j$, and if all SD are the same, $\sigma_k^2 = \sigma_r^2$, we have that $\varepsilon_G = \frac{1}{G}\sum_{j=1}^{G}\varepsilon_j \to \sigma_G^2 = \frac{1}{G}\sigma_r^2$. Then, the magnitude of fluctuations is reduced by group size and vanishes in the infinite G limit. Thus, if errors are independent across individuals, both long times and large-population pooling leads to a convergence to the behavior set by the temporal means. This, curiously, implies that the group average grows faster than the agent's temporal average in general and provides a strong quantitative argument for pooling resources either via government action or risk management instruments, such as insurance (26).

The case of nonindependent variables is interesting because the treatment of the last section suggests that it would follow from different agents either experiencing correlated fluctuations and/or generating coordinated "institutional" control responses, which is likely in many circumstances. When all variables are fully correlated $\rho_{jk} = 1$ and $\sigma_G^2 = \sigma_r^2$, the volatility associated with r_G becomes independent of group size. In urban settings, we may expect some correlation between agents as they experience a common spatial and socioeconomic environment of the city. For U.S. MSAs, σ_G^2 is approximately constant in G (see fig. S3).

The covariance term between individual growth rates and resources adds additional corrections

$$
\begin{aligned}
\mathrm{covar}_G\left(\eta, \frac{r}{r_G}\right) &= \left[\frac{1}{G}\sum_{j=1}^{G}\left(\frac{\bar{\eta}_j}{\bar{\eta}_G}-1\right)\left(\frac{r_j}{r_G}-1\right)\right]\bar{\eta}_G + \left[\frac{1}{G}\sum_{j=1}^{G}\left(\frac{\varepsilon_j}{\varepsilon_G}-1\right)\left(\frac{r_j}{r_G}-1\right)\right]\varepsilon_G \\
&= \mathrm{covar}_G\left(\frac{\bar{\eta}}{\bar{\eta}_G}, \frac{r}{r_G}\right)\bar{\eta}_G + \mathrm{covar}_G\left(\frac{\varepsilon}{\varepsilon_G}, \frac{r}{r_G}\right)\varepsilon_G
\end{aligned}
\tag{12}
$$

With these results in hand, we can now write the time evolution of average group resources as

$$
\begin{aligned}
\frac{dr_G}{dt} &= \eta'_G r_G, \text{ with } \eta'_G = \bar{\eta}'_G + \varepsilon'_G, \bar{\eta}'_G = \left[1 + \mathrm{covar}_G\left(\frac{\bar{\eta}}{\bar{\eta}_G}, \frac{r}{r_G}\right)\right]\bar{\eta}_G, \\
\varepsilon'_G &= \left[1 + \mathrm{covar}_G\left(\frac{\varepsilon}{\varepsilon_G}, \frac{r}{r_G}\right)\right]\varepsilon_G
\end{aligned}
\tag{13}
$$

We see that the statistical behavior of r_G is set by the dependencies of these quantities (see section S3 for discussion). When $\bar{\eta}'_G$ and $\sigma_G'^2$ are independent of r_G (but may depend on G and t) and ε'_G obeys the conditions of the central limit theorem, the population average resources r_G will follow a multiplicative random growth process (Fig. 3D). This process, similar to Eq. 4, will then integrate to give

$$
\ln\frac{r_G(t)}{r_G(0)} = \left(\bar{\eta}'_G - \frac{\sigma_G'^2}{2}\right)t + \sigma'_G W(t)
\tag{14}
$$

showing that if $W(t)$ converges to a normal variable as the result of the central limit theorem, then the statistics of $r_G(t)$ become lognormal at long times (see section S3 for an example and further discussion of necessary conditions, exceptions, and related results) (29). It is important to stress that growth rates and volatilities now "run" (i.e., change) with group sizes, G, and time, t, depending on the correlations captured by the several covariance terms (see Fig. 3D).

From stochastic growth to statistical scaling relations

We are now ready to express the quantities in scaling relations as functions of stochastic growth rates. This will provide us with a

statistical theory that derives urban scaling beyond mean-field calculations (4). To keep the notation simple, the index i denotes cities. We take each city to be a group with $G = N_i$ and write the simplified notation $\gamma_i = \gamma'_{N_i}$, $\sigma_i = \sigma'_{N_i}$, and so on. We will also write the averages of these quantities over the ensemble of cities as $\gamma = \langle\gamma_i\rangle$ (see section S6 for a summary of notation).

Running scaling exponents and the emergence of scale invariance

Let us see when a power law scaling relation is a conserved quantity of the stochastic growth dynamics. We start with the integral trajectory for total resources, $R_i(t)$, $\ln\frac{R_i(t)}{R_i(0)} = \gamma_i t + \sigma_i W(t)$. This equation is ergodic in the sense of stochastic population dynamics because long-time averages coincide with ensemble averages (30)

$$
\left(\frac{1}{t}\ln\frac{R_i(t)}{R_i(0)} - \gamma_i\right)^2 = \frac{\sigma_i^2 W^2(t)}{t^2} \to \frac{\sigma_i^2}{t} \to 0
\tag{15}
$$

This property of long-time means specifies necessary conditions for scaling to hold over time. To see this, define B_i

$$
B_i(\ln N_i) \equiv \frac{d\gamma_i}{d\ln N_i} \to \gamma_i(t) = \int B_i\, d\ln N_i \simeq \gamma^{(0)} + \bar{B}_i \ln N_i + O[(\ln N_i)^2]
\tag{16}
$$

where $\gamma^{(0)}$ is independent of time and scale. B_i varies slowly with $\ln N_i$ so that $\bar{B}_i$ is also independent of scale but could depend on time. B_i is analogous to a beta function expressing the change ("running") of a coupling with scale in statistical physics (15). Replacing it into Eq. 15 obtains

$$
R_i(t) \to R_i(0)\, e^{\gamma_i t} \simeq Y_0(t)\, N_i(t)^{\beta + \bar{B}_i t}
\tag{17}
$$

which shows that if $\bar{B}_i$ is nonzero, then the scaling exponent $\beta(t) \simeq \beta + \bar{B}_i t$ becomes time dependent in general and is not conserved by the dynamics of growth. Scaling relations will then vary over time, becoming steeper (larger exponent), if $\bar{B}_i > 0$, or shallower, if $\bar{B}_i < 0$. It is also possible that the integral Eq. 16 yields a more complicated function of $\ln N_i$ and time. Under time averaging and control, it is natural for $\bar{B}_i \sim 1/t$ as we have seen, resulting in a time-independent change of scaling exponent.

To see this, consider that the volatility $\sim \frac{\sigma_i^2}{2}$ in the effective growth rate is, in general, both time and population scale dependent, while the mean $\bar{\eta}_i$ is independent of both. This means that, in most circumstances, $B_i(\ln N_i) = -\frac{1}{2}\frac{d\sigma_i^2}{d\ln N_i}$, which should be small because of the agent's control over fluctuations. Consider the example $\sigma_i^2(N_i) = \frac{\sigma_r^2}{t N_i^\alpha}$, $\to B_i = \frac{\alpha}{2}\sigma_i^2(N_i)$, which leads to the exact result, $\beta \to \beta - \frac{\sigma_r^2}{2N^\alpha \ln N}$. This shows that the scaling exponent β, while time independent, increases with city size, N. In this case, only at sufficiently large $N >> (\sigma_r^2/2\beta)^{1/\alpha}$ will the value of β coincide with that predicted by mean-field scaling theory (4). This is not an issue if σ_r^2 is small. Otherwise, for smaller cities, β may become measurably smaller than for larger ones. Because the magnitude of variations away from scaling is urban system and quantity dependent, this may help account for some variations of observed scaling exponents in different nations and for different urban properties (31, 32). It also implies correlations between the behavior of the prefactor, $Y_0(t)$, the ξ variance, and the scaling exponent β, as noted recently in (33).

These results show that strict scaling invariance is predicated on $B_i = \frac{d\gamma_i}{d\ln N_i} \to 0$, which is analogous to a renormalization group fixed point in statistical mechanics (34) applied to the population

growth rate. Away from this fixed point, we have now shown how to compute corrections to scaling exponents, which are the result of the scale-dependent statistics of growth rates. Last, note that the scale independence of growth rates for cities is a standard assumption known as Gibrat's law (or law of proportional growth) (24). This assumption is necessary to derive Zipf's law for the statistics of city sizes. Figure S3 illustrates this general analysis with the growth rates and variances for wages in U.S. metropolitan areas since 1969, showing that the effective growth rates are city size independent to an excellent approximation, justifying the observed persistence of scaling with a time-invariant exponent.

Equations of motion for prefactors and scaling residuals
We now translate stochastic growth into equations of motion for both scaling prefactors and residuals; details of the derivations are given in section S4. For the prefactors, we obtain

$$\frac{d \ln R_0}{dt} = \frac{d\langle \ln R \rangle}{dt} - \beta \frac{d\langle \ln N \rangle}{dt} \tag{18}$$

which is a function of only the centers' dynamics. Because the centers are averages over all cities, no higher order statistics plays out in these quantities. This dynamics of the scaling prefactor is important because it measures the urban system (nation) wide per capita baseline growth, a form of endogenous intensive economic growth.

For the residuals, we obtain

$$\frac{d\xi_i^r}{dt} = (\gamma_i - \gamma) - \beta \frac{d}{dt}(\ln N_i - \langle \ln N \rangle) + (\varepsilon_i - \varepsilon) = \gamma_i - \gamma - \beta(\gamma_{N_i} - \gamma_N) + (\varepsilon_i - \varepsilon) \tag{19}$$

where $\gamma_{N_i} = \frac{d}{dt} \ln N_i$ and $\gamma_N = \frac{d}{dt}\langle \ln N \rangle$. This equation has a number of interesting properties: The most important is that it essentially describes a random walk driven by the ε terms, which set the variance, $\langle (\xi^r)^2 \rangle$. The two other terms enforce the convergence to the population averages in terms of growth rates of resources and population and guarantee that $\langle \xi^r \rangle = 0$ is preserved by the urban system's growth dynamics.

The emergent statistics of urban indicators
The statistics of resources follows from integrating Eq. 19, leading to the general expectation that the statistics of the ξ^r become normal at long times (see section S4). This means, in turn, that cumulative urban indicators (stocks) are expected by the same argument to be lognormal, as we saw more directly above. Flow quantities, such as income or costs, are often more accessible empirically (Fig. 2). Their statistics follow from the analysis of the previous sections, where we wrote $Y_i = b_i R_i = (\bar{b}_i + v_i) R_i \rightarrow \ln Y_i = \ln R_i + \ln (\bar{b}_i + v_i)$. Substituting the scaling relations for R_i, Y_i, this implies that $\xi_i(t) = \xi_i^r(t) + \ln R_0 - \ln Y_0 + \ln b_i$. Taking averages over cities obtains the constraint $\ln Y_0 - \ln R_0 = \langle \ln b \rangle$, which allows us to write

$$\xi_i(t) = \xi_i^r(t) + \ln b_i - \langle \ln b \rangle \tag{20}$$

This shows that the statistics of income are set by two different processes, the first resulting from the statistics of associated resources and the second due to stochastic returns. The first piece is characterized by the accumulation of variations over time, which entails time averaging and is expected to become approximately normal. The second term is instantaneous and consequently not subject to limit theorems. Hence, it can have more arbitrary statistics.

To see this, we return to the analysis of stochastic returns b_i under agents' adaptive control to obtain the explicit time evolution equation

$$d\xi_i(t) \simeq d\xi_i^r(t) + (\ln \bar{b}_i - \langle \ln \bar{b} \rangle) dt + \left[\frac{\Omega_i}{\bar{b}_i} dW_i(t) - \langle \frac{\Omega}{\bar{b}} \rangle dW(t) \right] \tag{21}$$

where the force dv_i/dt was taken here to be white noise dW_i (the differential of the Wigner process W_t) with variance Ω_i^2, as above. If dv_i/dt has nonrandom components, the expression is similar but more complicated. Here, dW is the average stochastic force over cities, and we assumed that fluctuations are uncorrelated to population variations in Ω and $\bar{b}$. This also implies that the quantity $\Delta_i = \Delta_i^r - (\ln \bar{b}_i - \langle \ln \bar{b} \rangle)$ inherits the property of ergodicity from Δ_i^r. Figure S4 shows the income growth rates for U.S. metropolitan areas over time, including its noise-driven equation of motion and the property for wages where fluctuations away from the mean trajectory of growth fall over time (roughly as $1/t$, inset) to become negligible for long times.

Note that in the limit of strong control, at the individual level and/or as an emergent average within cities when the $\Omega_i/\bar{b}_i \ll 1$, the stochastic terms will be small, and the statistics of income will approximate that of resources as a normal distribution for the ξ_i. In addition, this derivation leads to a set of quantitative expectations that can be checked against the data: Figure 5A shows that the quantity $t^2 \langle \Delta^2 \rangle \sim \sigma^2 t$ behaves approximately like the displacement of a one-dimensional (1D) random walk. This is well described by the straight line in time with slope given by the variance, although empirically we also observe shorter periods of acceleration or deceleration relative to the main trend. Figure 5B shows an analogous picture depicting each SAMI, ξ_i, trajectory, starting all cities at $\xi_i = 0$ in 1969. This demonstrates the spread of the SAMIs over time according to the behavior of a 1D random walk (red line, the same as straight line in Fig. 5A). Figure 5C shows the volatility and mean growth rates for all cities over the 47 years and corresponding estimates from measurements of dispersion over time (Fig. 5, A and B) and over the ensemble of cities: The observed statistical agreement of these two strategies for measuring the square volatility demonstrate the ergodicity of the statistics of $\xi_i(t)$ once drift has been removed. Last, Fig. 5 (A and D) shows that the income residuals' variance is actually time dependent, spreading very slowly over time as predicted by the derived lognormal part of the distribution. The overall distribution is better described, however, by the sum of two Gaussians, one broad and one narrow, corresponding to the two terms in Eq. 21 (red dashed line in Fig. 2C). It is only because the annual volatilities are so small that this temporal pooling and a deduction of a pure lognormal behavior appeared reasonable for flow variables in earlier work (22, 35).

DISCUSSION
We showed how quantitative urban theory can be taken beyond a stationary approach based on an average budget constraint, characteristic of spatial equilibrium. In its place, we proposed the primacy of stochastic growth processes and agents' strategic behavior as the dynamical statistical theory from which more particular results follow (Fig. 1). This provides a common foundation for nonequilibrium modeling of cities across scales (17, 18) and shows how these processes are associated with urban scaling and agglomeration effects.
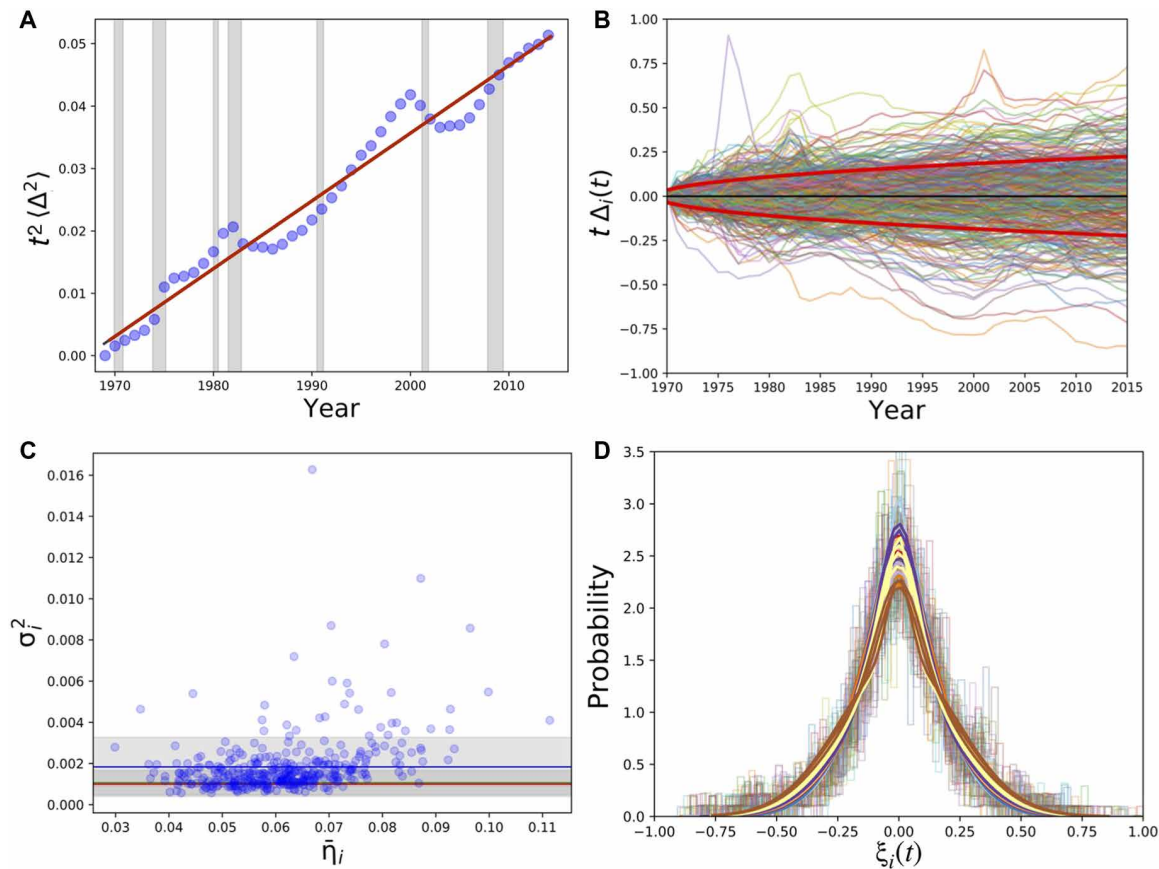
Fig. 5. Effective diffusive growth of deviations $\xi_i(t)$ **and the emerging statistics of cities.** (**A**) On the average over cities, the displacement from their initial deviations in 1969 grows linearly (red line) (gradient = 0.00108, 95% confidence interval = [0.00102, 0.00115]; intercept = −2.13279, 95% confidence interval = [−2.25885, −2.00672], $R^2 = 0.93$), as expected from pure random diffusion of the growth rates. Note that this is a mean temporal behavior and that there are periods when deviations grow faster or slower. Periods of economic recession are shown in gray. (**B**) The trajectory of deviations for all cities (different colors) but having set all deviations in 1969 to zero so that all trajectories depart from a common origin. The red line indicates the diffusive behavior, same as in (A), clearly showing that deviations tend to increase in magnitude over time. (**C**) The prediction of the wage growth volatility for U.S. MSAs by three methods: the fit of (A) and (B) and the averages over time and sets of cities, demonstrating the ergodic character of the statistical dynamics. Shaded areas show the overlapping 95% intervals in these estimates. (**D**) The distribution of deviations, year by year, using the same color scheme as in Fig. 2 (A and B). We see that, unlike our first approach in Fig. 2C, the width of the distributions is increasing slowly over time (brown most recent) and that the data for wages (a flow) should be fit by a distribution that is well described as the sum of two Gaussians: a universal broad distribution due to resource compounding and a contingent short-term narrow distribution (Eq. 20), which depends on most recent environmental shocks.

From this point of view, we see how the budget condition of spatial equilibrium models becomes the emergent property of a much more fundamental process, whereby agents subject to stochastic resource flows (incomes and costs) must develop adaptive strategies to reduce potentially fatal volatility. This point of departure is both necessary for dissipative complex systems and is very general so that it offers a number of connections with the statistical dynamics of other natural and engineered systems (28, 36, 37).

The key advantage of this bottom-up stochastic approach is that it naturally unifies processes of resource flow management ("equilibrium"), growth, and statistics. Hence, the framework emphasizes the critical role played by growth rate variations in a number of important urban phenomena. Specifically, we showed how the properties of the growth rate volatility are implicated in the (non-) preservation of urban scale invariance and set the boundary between growth and decay regimes, including the time scale for exponential growth to become manifest as fluctuations average out. In particular, we demonstrated how the general property of ergo-

dicity in population dynamics and formal demography (30) is intimately connected, together with a renormalization fixed point condition on the growth rates, to the emergence of "mean-field" scaling relations (4).

There are a number of important consequences for urban theory that these results clarify and unify. First, they show how spatial equilibrium is, after all, consistent with observed exponential growth in cities both economic and demographic, which has been an assumption in previous models. Second, they show how to derive macroscopic statistical behavior for cities and urban systems from microscopic strategic choices at the agents' level and provide expressions for how to aggregate growth rates over time and populations. Third, this process exposes issues of inequality of wealth and income and how they are compounded over time (23). Specifically, the quantities discussed here show how policies aimed at maximizing aggregate economic growth may naturally deemphasize the relative growth of poorest sections of the population. Last, and in many ways the central motivation of the paper, the results derived

here demonstrate that the statistics of most urban indicators are not universal in a simple sense. Rather, they are emergent as the consequence of limit theorems under stochastic (exponential) growth. In particular, the statistics of urban indicators that account for incomes and costs are the result of a mixture of a more universal component, inherited from their association to accumulating quantities and corresponding limit theorems, and a nonuniversal part, arising from the short term "hustle" (accidents and the quality of the agent's control) in variable stochastic environments. Thus, statistical tests to evaluate the Gaussianity of urban (log-)quantities (22, 35, 38), to be meaningful, must be performed with care and explicitly acknowledge the distinct distributions of different urban indicators.

The models for the budget constraint, the growth of resources, and associated control strategies introduced here are standard starting points in demography, geography, economics, and finance (23–25). They can clearly be made more complex and, where necessary, also more realistic. The concept of resources and incomes is not 1D. Issues of energy, monetary wealth, knowledge, and social capital all contribute to resource growth in human populations. The extent to which these quantities, which can all be accumulated, interact with each other is critical for a general understanding of human development. Models for the dynamics of volatilities and more sophisticated control of fluctuations and maximization of growth rates may also become important. Some inspiration should be derived from population biology and mathematical finance (16, 25), where such models are more developed. Last, data on detailed expenditures, wealth, and other financial and social characteristics are becoming increasingly available for households at finer temporal resolutions (14) and will be critical to test and improve the ideas introduced here and to identify systematic heterogeneities in agents' behavior, e.g., associated with conditions of poverty and uncertainty.

The approach developed here can be applied to other contexts beyond the contemporary United States but requires appropriate contextualization. In all societies, household adaptive management of resources is likely to remain important. However, in more collectivist societies or in those with stronger top-down governance, the aggregate management of benefits and costs will replace, to a larger extent, bottom-up agency. As shown, managing aggregate costs at the societal level can achieve considerable benefits because this strategy minimizes some risk. Its success hinges, however, on the effective investment of resources that generate society-wide benefits and their redistribution and related inequality. In circumstances of low growth, such as in most preindustrial societies, adaptive control of resources and the associated dynamics of volatility provide us with an important window into their (in)stability. This allows us to connect more proximate explanations of collapse, e.g., related to environmental stresses or violence, to the broader collective and political dynamics of societies, expressed as the capacity to manage shocks or disintegrate instead.

Empirically, the U.S. urban system, at least in terms of changes in total wages in MSAs, turns out to be very well-behaved: Its growth volatilities are almost always very small, fluctuations converge to limiting statistics quickly, and scaling relations are conserved over time. However, our theoretical results show that these properties pertain only to quantities and systems of cities with small, population size–independent growth rate volatilities. In the United States, over the past nearly 50 years, despite a number of notable events, observed average square volatilities associated with

wages and population growth are about one order of magnitude smaller than average growth rates, making their effects almost negligible. It will be interesting in the future to investigate other urban systems and quantities characterized by larger volatility, such as crime or innovation (22, 33, 35), for which the present framework makes a number of testable predictions.

The flip side of the observed constancy and stability of growth rates in American cities is that extant wage disparities become very slow to reverse. The typical square displacement in ξ over nearly five decades (Fig. 5A; $\sim\sigma^2 t$) is just 0.054. Assuming a similar rate of change in the past means that the observed variance in deviations from scaling at the beginning of our dataset (in 1969, about 0.043) would have been the product of the previous 40 years, taking us back to the time of the roaring 1920s and the subsequent great depression. Thus, the answer to the question at the beginning of this paper about predicting the magnitude of deviations from scaling in any given year is now recast not so much in terms of parameters of stationary statistics (33). Rather, this variance is the result of accounting for the accumulation of much smaller accidents and variations that make up the stochastic history of cities, which compound short-term noisy growth under partial control of heterogeneous agents over entire urban areas and long-time periods of decades (22). This is the quantitative sense in which history matters for cities, and their development becomes path dependent (18, 39, 40).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/34/eaat8812/DC1

REFERENCES AND NOTES

1. W. Alonso, *Location and Land Use: Toward a General Theory of Land Rent* (Harvard Univ. Press, 1977).
2. M. Fujita, P. Krugman, A. Venables, *The Spatial Economy: Cities, Regions, and International Trade* (The MIT Press, 2001).
3. E. L. Glaeser, *Cities, Agglomeration, and Spatial Equilibrium, Lindahl Lectures* (Oxford Univ. Press, 2008).
4. L. M. A. Bettencourt, The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
5. A. Anas, R. Arnott, K. Small, Urban spatial structure. *J. Econ. Lit.* **36**, 1426–1464 (1998).
6. C. S. Fischer, The subcultural theory of urbanism: A twentieth-year assessment. *Am. J. Sociol.* **101**, 543–577 (1995).
7. L. M. A. Bettencourt, H. Samaniego, H. Youn, Professional diversity and the productivity of cities. *Sci. Rep.* **4**, 5393 (2014).
8. H. Youn, L. M. A. Bettencourt, J. Lobo, D. Strumsky, H. Samaniego, G. B. West, Scaling and universality in urban economic diversification. *J. R. Soc. Interface* **13**, 20150937 (2016).
9. R. J. Sampson, *Great American City: Chicago and the Enduring Neighborhood Effect* (Univ. of Chicago Press, 2012).
10. J. Jacobs, *The Death and Life of Great American Cities* (Modern Library, 2011).
11. E. E. Bruch, R. D. Mare, Neighborhood choice and neighborhood change. *Am. J. Sociol.* **112**, 667–709 (2006).
12. S. F. Reardon, K. Bischoff, Income inequality and income segregation. *Am. J. Sociol.* **116**, 1092–1153 (2011).
13. L. Hannon, Poverty, delinquency, and educational attainment: Cumulative disadvantage or disadvantage saturation? *Sociol. Inq.* **73**, 575–594 (2003).
14. R. Chetty, D. Grusky, M. Hell, N. Hendren, R. Manduca, J. Narang, The fading American dream: Trends in absolute income mobility since 1940. *Science* **356**, 398–406 (2017).
15. J. Zinn-Justin, *Phase Transitions and Renormalisation Group, Oxford Graduate Texts* (Oxford Univ. Press, 2013).
16. S. A. Frank, Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. *J. Evol. Biol.* **25**, 2377–2396 (2012).
17. P. M. Allen, The importance of complexity for the research agenda in the built environment. *Architect. Eng. Des. Manag.* **4**, 5–14 (2008).
18. J. Johnson, *Non-Equilibrium Social Science and Policy: Introduction and Essays on New and Changing Paradigms in Socio-Economic Thinking* (Springer Berlin Heidelberg, 2016).
19. L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, G. B. West, Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7301–7306 (2007).

20. R. J. Barro, X. Sala-i-Martin, *Economic Growth* (The MIT Press, 2003).

21. L. M. A. Bettencourt, J. Lobo, Urban scaling in Europe. *J. R. Soc. Interface* **13**, 20160005 (2016).

22. L. M. A. Bettencourt, J. Lobo, D. Strumsky, G. B. West, Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLOS ONE* **5**, e13541 (2010).

23. T. Piketty, G. Zucman, Capital is back: Wealth-income ratios in rich countries 1700–2010. *Q. J. Econ.* **129**, 1255–1310 (2014).

24. X. Gabaix, Zipf's law for cities: An explanation. *Q. J. Econ.* **114**, 739–767 (1999).

25. W. Paul, J. Baschnagel, *Stochastic Processes: From Physics to Finance* (Springer, 2013).

26. O. Peters, A. Adamou, Insurance makes wealth grow faster. arXiv:1507.04655 [q-fin.RM] (16 July 2015).

27. E. L. Crow, K. Shimizu, *Lognormal Distributions: Theory and Applications*, (M. Dekker, 1988), vol. 88.

28. K. J. Aström, R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers* (Princeton Univ. Press, 2008).

29. G. B. Arous, L. V. Bogachev, S. A. Molchanov, Limit theorems for sums of random exponentials. *Probab. Theory Relat. Fields* **132**, 579–612 (2005).

30. S. Tuljapurkar, *Population Dynamics in Variable Environments* (No. 85 in Lecture notes in biomathematics, Springer, 1990).

31. E. Arcaute, E. Hatna, P. Ferguson, H. Youn, A. Johansson, M. Batty, Constructing cities, deconstructing scaling laws. *J. R. Soc. Interface* **12**, 20140745 (2014).

32. C. Cottineau, E. Hatna, E. Arcaute, M. Batty, Paradoxical interpretations of urban scaling laws. arXiv:1507.07878 [physics.soc-ph] (28 July 2015).

33. A. Gomez-Lievano, O. Patterson-Lomba, R. Hausmann, Explaining the prevalence, scaling and variance of urban phenomena. *Nat. Hum. Behav.* **1**, 0012 (2016).

34. J. Zinn-Justin, *Phase Transitions and Renormalization Group, Oxford Graduate Texts* (Oxford Univ. Press, 2007).

35. A. Gomez-Lievano, H. J. Youn, L. M. A. Bettencourt, The statistics of urban scaling and their connection to Zipf's law. *PLOS ONE* **7**, e40393 (2012).

36. D. W. Stephens, J. R. Krebs, *Foraging theory* (*Monographs in Behavior and Ecology*) (Princeton Univ. Press, 1986).

37. P. Guarnieri, *Decision Models in Engineering and Management* (Springer, Heidelberg, Germany, 2015).

38. J. C. Leitão, J. M. Miotto, M. Gerlach, E. G. Altmann, Is this scaling nonlinear? *R. Soc. Open Sci.* **3**, 150649 (2016).

39. W. B. Arthur, *Increasing Returns and Path Dependence in the Economy, Economics, Cognition, and Society Series* (University of Michigan Press, 1994).

40. J. Depersin, M. Barthelemy, From global scaling to the dynamics of individual cities. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2317–2322 (2018).

Citation: L. M. A. Bettencourt, Urban growth and the emergent statistics of cities. *Sci. Adv.* **6**, eaat8812 (2020).