

Urban Evolutions: The Fast, the Slow, and the Still

By GILLES DURANTON*

With the use of French and US data, new and systematic evidence is provided about the rapid location changes of industries across cities (the fast). Cities are also slowly moving up and down the urban hierarchy (the slow), while the size distribution of cities is skewed to the right and very stable (the still). The model proposed here reproduces these three features. Small, innovation-driven shocks lead to the churning of industries across cities. Then, cities slowly grow or decline following net gains or losses of industries. These changes occur within a stable distribution. The quantitative implications of the model are also explored. (JEL R12, R32)

A fundamental but much neglected issue in urban development is the role played by the churning of industries across locations. Jane Jacobs (1970) provides early anecdotal evidence about what the key mechanisms at stake might be. In the late nineteenth century, New York City was the capital of the photographic industry, whereas Rochester, New York, was the leading city for precision instruments. George Eastman, while working at improving optical instruments in Rochester, invented an emulsion-coating machine that enabled him to mass-produce photographic dry plates. His company soon took over the market for photographic film. As a consequence, Rochester replaced New York as the main center for the industry. Rochester, 50 years later, was still the capital of the US film industry, whereas New York was that of the duplication industry. Then, Haloid Company, a firm specialized in the manufacturing of photographic papers and operating

in the shadow of Eastman Kodak, came up with a new process for making copies without the need for developing. The process, called xerography, made Rochester the new capital of the duplication industry in place of New York (again), where the previous dominant technology, the varityper, was produced. These two industries came to represent an important part of Rochester's employment.

Section I presents novel and systematic evidence that there is indeed considerable industrial churning across cities. Industries move and cities experience rapid changes in the composition of their economic activity. This stylized fact, the *fast*, is closely related to two better-known features of urban evolutions: the *slow*, where cities move slowly up and down the distribution of city sizes as they grow or decline relative to other cities; and the *still*, where the size distributions of cities tend to be stable over time and skewed to the right.

The model in Section II provides microeconomic foundations for the churning of industries across cities. Formally, it embeds an extension of Gene M. Grossman and Elhanan Helpman's (1991) quality ladder model in an urban framework. In each industry, research firms try to invent the next step up the quality ladder in order to reap monopoly profits. Research firms may be successful in their own industry or may develop a new leading quality in another industry. Local spillovers induce research firms in an industry to co-locate with production in the same industry and in most industries, successful innovators need to start

* Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: gilles.duranton@utoronto.ca; Web site: <http://individual.utoronto.ca/gilles/default.html>). This is a much revised version of a paper previously circulated as "City Size Distributions as a Consequence of the Growth Process." I am grateful to Vernon Henderson, Yannis Ioannides, Janet Kohlhase, Tomoya Mori, Henry Overman, John Parr, Diego Puga, Kwok Tong Soo, Dan Trefler, Hyoung Gun Wang, Fabrizio Zilibotti, conference and seminar participants, and four referees for helpful comments, suggestions, and encouragement. I am also grateful to Pierre-Philippe Combes for sharing his data and helping me with them. Funding from the Leverhulme Trust is gratefully acknowledged.

producing where they did their research. This implies that own-industry innovations lead to a change of monopoly but to no change of location for an industry. By contrast, cross-industry innovations imply not only a change of monopoly but also, typically, a change of location, since the old and new monopolies are not generally located in the same city. Hence, innovation-driven shocks provide the basis for the growth and decline of cities so that urban evolutions result from the aggregation of small *industry-level shocks*.

The results derived in Section III show that, over a specific time period, a city typically experiences employment gains in some sectors and losses in others. Since gains and losses partially offset each other, net employment changes in cities are smaller than gross employment flows, and cities move slowly up and down the size distribution. Furthermore, these changes occur within an approximately stable right-skewed distribution of city sizes. Hence the model can replicate the three stylized facts above.

The quantitative predictions of the model are explored in Section IV using simulations. The model is calibrated to replicate French and US city-size distributions. The simulations approximate the US distribution well. The model does better than Zipf's law according to a natural efficiency criterion.¹ For France, the performance is also good. Interestingly, the model can replicate the nonregular features of the French and US Zipf's curves like their concavity, although it exaggerates them.

These systematic deviations can be reduced or eliminated altogether by considering a natural extension of the model. In Section V, the returns to innovative activities in cities are affected by their size: positively, through (reduced-form) dynamic agglomeration economies, and negatively, through crowding. When agglomeration economies dominate crowding, the probability of innovating in a city increases more than proportionately to its size. This reduces the Zipf's coefficient in the upper tail and increases it in the lower tail. Under empirically plausible values for the trade-off between agglomeration

economies and crowding (Stuart S. Rosenthal and William C. Strange 2004), it is possible to replicate the US and French urban systems very closely.

There is a large urban growth literature that builds on the trade-off between agglomeration economies and crowding (see Marcus Berliant and Ping Wang 2004, or J. Vernon Henderson 2006, for surveys). Like this paper, this literature often views purposeful innovative activity as a key engine of urban growth. Its main emphasis, however, is on how the growth process depends on urban agglomeration, and vice versa, an issue that is intentionally left aside here. On the other hand, the spatial churning of industries, the slow mobility of cities in the urban hierarchy, and their heterogeneity in population size are usually neglected in this literature.²

A second important strand of literature, which dates back to Herbert Simon (1955), has sought to generate distributions of city sizes that would obey Zipf's law. This literature, recently revived by the work of Xavier Gabaix (1999), is successful at generating realistic and stable distributions of city sizes with some mobility within them. It has two main limitations. First, it ignores the churning of industries across cities, the main focus of this paper. Second, and more subtly, urban evolutions in this literature are driven by ad hoc exogenous shocks: amenity shocks in Gabaix (1999), preference shocks in Juan-Carlos Córdoba (2003), and productivity shocks in Jan Eeckhout (2004) and Esteban Rossi-Hansberg and Mark L. J. Wright (forthcoming). By contrast, the model presented here offers detailed microeconomic foundations for technology shocks.³ Such microeconomic foundations are important because these shocks are

² A possible exception is Duranton and Diego Puga (2001), who examine the movements of firms from diversified to specialized cities in steady state. Their focus is nonetheless different, since churning occurs across cities that keep the same population and sectoral structure.

³ In a companion paper (Duranton 2006), a Zipf's law model with detailed microeconomic foundations is proposed. This companion paper builds on the product proliferation model of endogenous growth (Paul M. Romer 1990) instead of its quality ladder version. This companion model is unable to deal with the churning of industries—a crucial focus of the present paper—since existing industries never move. Furthermore, this Zipf's law model does not examine the higher moments of the city size distribution, nor does it examine quantitatively the deviations from Zipf's law.

¹ Zipf's law refers to the well-known empirical claim that the size distribution of cities follows a Pareto distribution with unitary exponent. See below for more information.

the fundamental drivers of the distribution of city sizes in steady state. Furthermore, sound foundations for these shocks should provide important guidance for future empirical work by telling us where and how to look for the ultimate sources of urban growth.

The rest of the paper is organized as follows. Section I provides systematic evidence about the three stylized facts mentioned above. Section II proposes a simple benchmark model. It is solved in Section III, where its main qualitative features are discussed. Simulations results are discussed in Section IV. The benchmark model is then enriched in Section V, where more realistic urban assumptions are introduced. Finally, the last section contains some conclusions.

I. Three Stylized Facts about City Sizes

A. *The Fast*

Detailed evidence about the fast, the slow, and the still of urban evolutions is provided for French and US cities. For the United States, I use employment data from the US Census Bureau County Business Patterns for 272 metropolitan areas between 1977 and 1997 at the level of 70 two-digit sectors, and population data for the same units from the US Census.⁴ For France, I used the Enquête Structure des Emplois from the French Institute for National Statistics. French employment data are collected at the level of 96 two-digit sectors for plants with more than 20 employees and are available for 217 French metropolitan areas from 1985 to 1993.⁵

⁴ County-level data are aggregated into Metropolitan Statistical Areas/Consolidated Metropolitan Statistical Areas outside New England and into New England County Metropolitan Areas in New England using year 2000 definitions. To look at churning, one needs sector-level data by location, which are publicly available only at the county level. Hence documenting the *fast* one is restricted to the largest units, that is, metropolitan areas, which are counties or aggregates of counties. However, to explore some implications of the model regarding city size distribution, it is better to use a more complete sample of cities. This is what I do below, where sector-level information is not needed. Consistent definitions are used throughout the paper.

⁵ The possible (minor) problems caused by the employment cutoff are discussed in Pierre-Philippe Combes, Thierry Magnac, and Jean-Marc Robin (2004), who also provide a more detailed description of these data. There were some data problems in 1988 in 15 cities. I excluded

It is useful to start with a comparison of the location patterns of population and employment in selected industries in the largest US cities between 1977 and 1997. The first two columns of Table 1 rank the ten largest US metropolitan areas in 1977 and indicate in brackets their changes in ranking with 1997. Changes over this 20-year period were small, with only the growth of the San Francisco area and the decline of Cleveland to be noted. This can be contrasted with the much larger changes in employment rankings for three large industries in the next three columns. In apparels (a low-tech industry), most cities changed rank. Sometimes the movements were quite large, with Washington losing 7 places while San Francisco gained 12. In transportation equipment (a mid-tech industry), there was some stability at the very top (though big changes in absolute employment), but New York and Cleveland dramatically moved from third to twenty-third and from sixth to twenty-first, respectively. Finally, instruments (a high-tech industry) also exhibits substantial changes over the 1977–1997 period. Overall, Table 1 suggests much more variability for sectoral employment than for aggregate population. This is shown formally by the last row of the table, which sums all rank variations.

To examine the spatial churning of industries in greater detail, an index of gross employment reallocation across sectors within each city c can be defined following the literature on industry dynamics (Steven J. Davis and John C. Haltiwanger 1998). For each US city, this index averages yearly relative employment gains and losses over all sectors and years:

(1) $Churn_c$

$$= \frac{1}{20 \times 70} \sum_{t=1977}^{1996} \sum_{z=1}^{70} \frac{|e_c(z, t+1) - e_c(z, t)|}{e_c(t)},$$

where $e_c(z, t)$ (>0) is employment in metropolitan area c and sector z for year t . It is also useful to compute an index of employment re-

them here (but will use their population data below when looking at the distribution of city sizes in France). Finally, note that only a short time period is available because of data problems in 1984 and a change of industrial classification in 1994.

TABLE 1—RANKINGS AND CHANGES FOR THE TEN LARGEST US METROPOLITAN AREAS, 1977 AND 1997

Rank in 1977 (and change in rank between 1977 and 1997)	Total population	Apparel	Transportation equipment	Instruments
New York	1 (0)	1 (+1)	3 (+20)	2 (+2)
Los Angeles	2 (0)	2 (-1)	2 (0)	5 (-3)
Chicago	3 (0)	11 (-3)	8 (+4)	4 (+5)
Washington	4 (0)	13 (+7)	29 (+8)	16 (-6)
Philadelphia	5 (+1)	3 (+3)	13 (-3)	7 (+4)
Boston	6 (+1)	4 (0)	17 (+24)	3 (-2)
Detroit	7 (+1)	10 (0)	1 (0)	20 (-1)
San Francisco	8 (-3)	15 (-12)	9 (0)	6 (-3)
Cleveland	9 (+5)	18 (+26)	6 (+15)	12 (+6)
Dallas	10 (-1)	5 (+4)	7 (-2)	13 (-5)
Total rank variation	(12)	(57)	(76)	(37)

Notes: The first number in each column is the 1977 ranking (among 272 metropolitan areas). It is followed (in parentheses) by the change in ranking between 1977 and 1997 (where an increase is a decline). The final row sums the absolute value of all the changes above. All calculations use the year 2000 definition of US metropolitan areas. Columns 3 to 5 give the ranking (and the 1977–1997 change) for total employment in SIC23 (apparel and other textile products), SIC37 (transportation equipment), and SIC38 (instruments and related products).

Source: US Census, County Business Patterns, and author's calculations.

allocation for each US metropolitan area over the 1977–1997 period:

$$(2) \quad \Delta Emp_c = \frac{1}{20} \sum_{t=1977}^{1996} \frac{|e_c(t+1) - e_c(t)|}{e_c(t)},$$

where $e_c(t)$ is total employment in city c for year t . Similar indices can be defined for French cities between 1985 and 1993 by changing the number of sectors and years.

By definition, $Churn_c$ is expected to be larger than ΔEmp_c . The interesting question is by how much. The difference $Churn_c - \Delta Emp_c$ is called “excess sector churning.” For the index of gross city employment reallocation across sectors, $Churn_c$, the averages are 8.26 percent for US cities and 11.40 percent for French cities. They are more than twice as large as the averages of ΔEmp_c , which are 4.10 percent for US cities and 5.20 percent for French cities. There is, thus, a large amount of excess sector churning and in this sense the movement of industries across cities is *fast* compared to the *slow* movement of cities.

When repeating the same exercise for only manufacturing sectors, I found an average of 9.81 percent for $Churn_c$ and 4.82 percent for ΔEmp_c in the United States. The figures for France are 12.24 and 6.62 percent, respectively. This points at more sectoral mobility when only manufacturing sectors are considered.

All of this churning cannot be the result of sectoral change. To see this, an index of changes in sectoral employment at the national level can be computed:

$$(3) \quad \Delta SecEmp$$

$$= \frac{1}{20 \times 70} \sum_{t=1977}^{1996} \sum_{z=1}^{70} \frac{|e(z, t+1) - e(z, t)|}{e(t)},$$

where $e(z, t)$ is employment in sector z for all metropolitan areas in year t . Calculations show that $\Delta SecEmp$ is very close to 5 percent for both France and the United States.⁶

To complement this analysis, I used Markov transition matrices for employment across areas in each sector. For each US metro area, I used sectoral employment in 1977, 1982, 1987, 1992, and 1997. For these years, the distribution of employment for each sector was divided into four quartiles and the maximum likelihood probability of moving between them was computed. A convenient way to summarize the

⁶ More disaggregated studies typically indicate that sectoral change accounts for only a tiny part of the observed employment churning (Davis and Haltiwanger 1998). In this respect, and in absence of more detailed sectoral divisions, sectoral churning will be underestimated since many changes of employment within finely defined sectors will wash out at a more aggregated level.

information is to compute the mean first passage time from the lowest quartile to the highest. This corresponds to the expected time a city in the lowest quartile takes to reach the highest quartile of employment in the sector (see Duncan Black and Henderson 2003 for details). In the United States, the cross-sector average of mean first passage time is 277 years. Despite relatively short periods of five years, the transition matrices are stationary at a 5-percent confidence level for 66 of 70 sectors (including all manufacturing sectors).

This mean first passage time of 277 years for sectors should be compared to a mean first passage time of 1,428 years for the population of US metropolitan areas (see below for details). Viewed differently, over a five-year period, 97.1 percent of metropolitan areas in the top quartile remain in the same cell, whereas 90.1 percent do so in the bottom quartile. For industries in cities, the corresponding figures are much lower at, on average, 89.3 percent and 75.1 percent, respectively.⁷ This suggests that the mobility of employment in sectors is much faster than that of city population.

A second interesting feature is that downward mobility is faster than upward mobility. The mean first passage time from the top cell to the bottom is only 176 years, showing that within a sector it takes longer to reach the top than to fall from it. Interestingly, this faster downward mobility is driven by a number of factors consistent with the model below. First, sectors can literally fall from grace in some cities. Every five years, 1.5 percent of city sectors in the top decile fall to the bottom, whereas less than 1 percent do the opposite. Second, more mobility at the bottom of the distribution implies that any progression is easily reversed. Finally (and consistent with the model below), note that cities with very little employment in some industries tend to remain with very little employment.

The same exercise can be conducted for France, using 1985, 1989, and 1993. Given that

the population threshold for French cities in the data is very low and that there is a threshold of 20 employees per plant, there are too many zeroes for the analysis to be conducted meaningfully with the entire set of French data. The sample is thus restricted to the largest 100 cities and to the 32 industries for which less than 25 percent of these cities have zero employment. The mean first passage time is 207 years from bottom to top, and stationarity is not rejected in 29 industries. This is much faster than the mean first passage time for city employment (446 years). As in the United States, downward mobility is also faster than upward mobility (a mean first passage time of 154 years instead of 207).

These findings are consistent with previous work by Black and Henderson (1999). They compute mean first passage times for three high-tech and three capital goods sectors across US metropolitan areas (using a different data source). They find that the mean first passage from the bottom to the top cell is on average 90 years for high-tech sectors and 150 years for capital goods sectors. In a more systematic analysis of industry mobility across cities, Mohammad Arzaghi and James C. Davis (2005) also obtain results that are consistent with those computed here.⁸ In detailed industry studies, Mark Beardsell and Henderson (1999) and Nancy E. Wallace and Donald W. Walls (2004) examine the location patterns of the US information technology industry over time. Despite different time frames and different definitions for this industry, both studies find substantial turbulence. According to Beardsell and Henderson (1999), mean first passage time from the lowest to the highest cell (of four) for a typical US metropolitan area is only 39 years. In their detailed examination of employment changes by location between 1989 and 2002, Wallace and Walls (2004) find both big winners and big losers. While employment grew nationally by 20 percent in the industry over that period, it

⁷ Obviously these figures depend on the cutoffs. Taking a more restrictive set of cutoffs (say the fifth, fifteenth, and fortieth centile instead of the twenty-fifth, fiftieth, and seventy-fifth) yields longer mean first passage times (427 years for sectors and 4,027 years for city population). However, the much faster relative mobility of city sectors is robust to the choice of cutoff. If anything, taking smaller cells at the top magnifies the results reported above.

⁸ Using a radically different approach for US states between 1972 and 1992, Guy Dumais, Glenn Ellison, and Edward L. Glaeser (2002) find very small changes in the geographic concentration of most industries, but very significant changes in where they concentrate. This again suggests that industries, even the more concentrated ones, are quite mobile over time. These results are confirmed by Salvador Barrios et al. (2005) for two European countries.

grew by nearly half or more in Washington, DC, San Francisco, and Atlanta, despite their high initial bases. Conversely, Dutchess, New York, and Colorado Springs (twenty-first and twenty-third in 1989) lost roughly half of their employment in this sector over the same period. All this suggests, again, substantial industry mobility.

B. The Slow

These high levels of industry churning must be contrasted with much slower changes in city-wide employment. In their analysis of 39 French cities between 1876 and 1990, Jonathan Eaton and Zvi Eckstein (1997) use Markov transition matrices. They assess mobility directly by looking at the off diagonal terms. They find that about 20 percent of cities change cell (among six) every 20 years or so—arguably a small proportion. Black and Henderson (2003) also find fairly low levels of mobility for 282 US metropolitan areas between 1900 and 1990, since only about 15 percent of them change cell (among five) every ten years.

To complete this evidence and to provide a benchmark for the figures about industry mobility reported above, the transition matrix analysis can be repeated for city-level data. The mean first passage time is 1,428 years for the United States and 446 years for France.⁹ In both cases, this implies far less mobility than for sectors. Despite slower mobility, some aggregate employment churning is nonetheless observed. To evidence this, the average index of employment changes for US metropolitan areas (see equation (2)) is compared to an index of employment change for all metro areas:

$$(4) \quad \Delta UrbEmp = \frac{1}{20} \sum_{t=1977}^{1996} \frac{|e(t+1) - e(t)|}{e(t)},$$

where $e(t)$ is the employment of all metropolitan areas in year t . A similar index can be defined for France between 1985 and 1993. $\Delta UrbEmp$ is equal to 2.77 percent for the United States and 3.32 percent for France. The difference between (2) and (4) can be called “excess em-

ployment churning” and is equal to 1.34 percent for the United States and 1.88 percent for France. Put differently, employment changes in cities are about 50 percent larger than they need to be to accommodate the changes in aggregate metropolitan employment.

C. The Still

To examine the stability of size distribution of French urban areas, Eaton and Eckstein (1997) use three different approaches. After normalizing the overall French urban population, they first compute the Lorenz curves of the spatial Gini coefficients for French urban areas for each census year. The Lorenz curves are remarkably close to each other. Second, for each year of data they estimate

(5)

$$\log \text{Rank}_c = \text{Constant} - \zeta \log \text{Size}_c + \text{error}_c,$$

where c is a subindex for cities. They find that ζ , the *Zipf's coefficient*, remains approximately constant over time. Finally, they assume that cities follow a first-order Markov transition process, estimate the transition probabilities between different size cells, and compute the ergodic probability distribution. This ergodic distribution is very close to the current distribution.

For US metropolitan areas, Black and Henderson (2003) find that their probability distribution function in 1990 is remarkably close to that in 1900. When estimating equation (5) for each decade, the Zipf's coefficient is close to constant over time. Black and Henderson (2003) estimate transitions probabilities for a Markov process as well. Like Eaton and Eckstein (1997) on French data, they find that the 1990 distribution is very close to the ergodic distribution. These findings broadly confirm those of Linda Harris Dobkins and Yannis M. Ioannides (2000), who use different US data.

D. Some Links

The link between the last two stylized facts, the still and the slow, is straightforward. Cities can be thought of as experiencing population shocks within a steady-state distribution. The link between industry mobility and city population changes deserves closer scrutiny.

⁹ These two figures cannot be compared directly because the size thresholds differ widely for the two countries. Note also that, in both cases, stationarity is not rejected.

Population changes in cities could well be the consequence of local industry shocks, but this need not be the case. Causality could in principle run in the opposite direction. Anecdotal evidence suggests nonetheless that technology shocks in industries constitute an important channel for urban growth and decline. Extreme examples include, among others, the demise of the steel industry in Pittsburgh and the rise of Internet-related industries in San Jose.

Beyond anecdotal evidence, Edward N. Coulson (1999) and Gerald A. Carlino, Robert H. DeFina, and Keith Sill (2001) provide more systematic evidence on the role of industry shocks to explain US metropolitan employment fluctuations. Using a time series methodology to disentangle common shocks and sector-level shocks locally, both papers attribute an overwhelming share of local fluctuations to the latter. These findings complement those of Olivier Jean Blanchard and Lawrence F. Katz (1992), showing that in- and out-migration is the main channel of adjustment to local shocks in US states.

II. The Model

The model is a variation of Grossman and Helpman’s (1991) quality ladder model of growth which is embedded below in an urban framework. Consider an economy with a large (discrete) number of industries, n , each of which produces one good that can potentially be supplied in an infinite number of qualities.¹⁰ Quality j of good z is given by $q_j(z) = \delta^j$, with $\delta > 1$. At time $t = 0$, the quality of all goods is normalized to unity so that any good must be improved j times to reach quality j . Quality improvements stem from research investments, which are described below.

A. Preferences

Consider a population of long-lived households whose mass is normalized to one and whose instantaneous utility is given by

¹⁰ Here, there is a small difference with Grossman and Helpman (1991): a discrete set of industries instead of a continuum. This prevents the law of large numbers from applying at the level of individual cities.

$$(6) \quad u(t) \equiv \sum_{z=1}^n \frac{1}{n} \log \left[\sum_{j=1}^{\bar{j}(z,t)} q_j(z) d_j(z, t) \right],$$

where $d_j(z, t)$ is the consumption of quality j of good z at time t , and $\bar{j}(z, t)$ its highest available quality at time t . For reasons made clear below, location indices can be ignored for the time being. Total expenditure is

$$(7) \quad E(t) \equiv \sum_{z=1}^n \sum_{j=1}^{\bar{j}(z,t)} p_j(z, t) d_j(z, t),$$

where $p_j(z, t)$ is the price of quality j of good z at time t . The objective of consumers is to maximize the discounted sum of their future instantaneous utilities

$$(8) \quad U \equiv \int_0^{\infty} u(\tau) e^{-\rho\tau} d\tau,$$

subject to the intertemporal budget constraint

$$(9) \quad \int_0^{\infty} E(\tau) e^{-R(\tau)} d\tau \leq W(0),$$

where $R(\tau)$ is the cumulative interest factor between 0 and τ , and $W(0)$ is the net present value of the stream of income plus the initial asset holdings at $t = 0$.

The maximization of the consumers’ program can be performed in two stages. First, allocate instantaneous expenditure, $E(t)$, to maximize $u(t)$ and then choose the intertemporal allocation of expenditure. The maximization of instantaneous utility (6) for any positive level of consumption expenditure implies equal shares of expenditure across industries. Then, to solve for the allocation of expenditure within industries, define $J(z, t) \equiv \text{Argmin}_{j \leq \bar{j}(z,t)} (p_j(z, t)/q_j(z))$, the quality of good z for which the ratio of price to quality is the lowest. When $J(z, t)$ is unique (and it is so in equilibrium), demand in industry z is then given by

$$(10) \quad d_j(z, t) = \begin{cases} \frac{E(t)}{np_j(z, t)} & \text{for } j = J(z, t), \\ 0 & \text{otherwise.} \end{cases}$$

Inserting these demands into (6) yields

$$(11) \quad u(t) = \frac{1}{n} \sum_{z=1}^n [\log E(t) - \log n + \log q_j(z, t) - \log p_j(z, t)],$$

where $p_j(z, t)$ and $q_j(z, t)$ are the price and quality of $J(z, t)$, respectively. Equation (11) can now be used to solve the optimal consumption path whose solution is characterized by

$$(12) \quad \dot{E}/E = \dot{R} - \rho,$$

together with the budget constraint and a transversality condition. After normalizing total expenditure $E(t)$ to n through the choice of numéraire, equation (12) implies $\dot{R} = \rho$, that is, the nominal interest rate is always equal to the subjective discount rate.

B. Technology

As is standard in quality ladder models of growth, there is competition in each industry to innovate and occupy the next step up the quality ladder. Research is performed by research firms (which produce no output in equilibrium). Any successful innovator is rewarded with a patent giving it a monopoly right over the production of this quality, $\bar{j}(z, \cdot)$. This patent cannot be licensed, and it expires only when yet another successful innovator manages to develop the following quality step. Thus, there is free entry among price-setting oligopolists for qualities below $\bar{j}(z, t)$. In all industries and irrespective of quality, producers need one unit of labor to produce one unit of good.

Free entry and unit marginal costs imply that for any nonleading quality, $j < \bar{j}(z, t)$, the price, $p_j(z, t)$, is equal to the wage rate, $w(t)$. Together with (10), this implies that any quality leader in industry z has a revenue $p_{\bar{j}}(z, t)d\bar{j}(z, t) = E(t)/n = 1$ when $p_{\bar{j}}(z, t)q_{\bar{j}}(z, t) \leq p_{\bar{j}-1}(z, t)q_{\bar{j}-1}(z, t)$, that is, when $p_{\bar{j}}(z, t) \leq \delta w(t)$. Prices above $\delta w(t)$ imply zero demand for the industry's leading quality. Hence, with unit elastic demand, any industry leader maximizes its profits by selling its quality at the limit price $p = \delta w$. Since the assumptions about product development ensure that in every industry there is a

unique quality leader, this leader is also the only active firm in the industry: $\bar{j}(z, t) = J(z, t)$. Using $p = \delta w$ and $pd = 1$, the profit of any industry leader is thus

$$(13) \quad \pi = (p - w)d = 1 - \frac{1}{\delta}.$$

There is free entry in the race to be the next leader in each industry. A research firm k in industry z , by investing $\lambda^k(z)$ units of research labor for a time interval of length dt to work on the highest existing quality, $\bar{j}(z, t)$, succeeds in inventing the next step up the quality ladder in this industry, $\bar{j}(z, t) + 1$, with probability $\beta\lambda^k dt$. Thus, as in Grossman and Helpman (1991), research firms use the state-of-the-art technology $\bar{j}(z, t)$ in an industry as a base to invent the next step up the quality ladder in the same industry.¹¹

There is, however, a slight difference with Grossman and Helpman's (1991) framework regarding the research technology. A research effort targeted at improving industry z may be successful, not only in this particular industry (as just described), but also in any other industry because of serendipity in the research process. Frederic M. Scherer (1984) provides very strong empirical support regarding the pervasiveness of such cross-industry innovations.¹² Formally, a research firm k in industry z , by investing $\lambda^k(z)$ units of research labor over dt , succeeds in inventing the next step up the quality ladder in industry $z' \neq z$ with probability $\gamma\lambda^k dt$ with $\gamma < \beta$.

¹¹ As in most endogenous growth models, innovations have both a private good dimension (patenting) and a public good aspect (increase of own-industry stock of knowledge). Note, also, that without any cost advantage in research, industry leaders do not attempt to innovate since, in case of success, the incremental profit would be less than that of a new entrant. Thus, in equilibrium, research is performed only by would-be entrants.

¹² So far, after the choice of a discrete number of industries as opposed to a continuum, this is only the second difference with Grossman and Helpman (1991). In their setting, none of these differences would change their results in any meaningful way. In a spatial setting, however, the possibility of cross-industry innovations is crucial to allow a research firm located in a given city to "capture" an industry located in another city, and thus provides a reason for city size to change.

In total, a research firm k in industry z , which invests $\lambda^k(z)$ over dt , expects to invent the next step up the quality ladder in industry z with probability $\beta\lambda^k(z)dt$ and in each of the other industries with probability $\gamma\lambda^k(z)dt$. After denoting $\lambda(z) \equiv \int \lambda^k(z) dk$, the sum of all research investments made by research firms in industry z , the probability of an innovation taking place in industry z over dt is $\iota(z)dt$, where

$$(14) \quad \iota(z) \equiv \beta\lambda(z) + \gamma \sum_{z' \neq z} \lambda(z').$$

Note that this probability increases in both the industry and the aggregate research efforts.¹³

For the sake of clarity, the assumptions presented here stick as closely as possible to the canonical model of Schumpeterian growth developed by Grossman and Helpman (1991): a multi-industry model where firms compete and invest in research in order to reap the monopoly profits associated with the highest quality. Self-sustaining and nonexplosive growth is possible since new innovations are neither more difficult nor easier than past ones. This well-known model can now be embedded in a very simple urban setting.

C. Cities

Consider m cities across which final goods are freely tradable. There are many more industries than cities: $n \gg m$. Workers are freely mobile, and there are neither advantages nor costs to city size so that any city can accommodate any number of workers at zero cost. This last assumption is relaxed in Section V.

Regarding research, recall that a quality improvement in an industry requires the knowledge associated with the leading quality. In turn, this knowledge is available only to research firms located in the same city as the industry leader. A possible justification is that one may

learn about leading technologies only by observing how industry leaders produce, through small talk with workers involved in production, or by being involved indirectly in production as supplier. All this requires physical proximity. This assumption of local knowledge spillovers has received ample empirical support. (See Rosenthal and Strange (2004) for a review of the literature.)

Turning to the location of production, note that a mechanism is needed to allow industries to change location. The possible mobility of all industries may, however, lead cities to lose all their industries and thus fall into a zero population trap. The simplest way to satisfy these modeling requirements is to assume that each industry is of one of two types: *first-nature* or *second-nature*.

First-nature industries are immobile. They provide both a first-nature justification for the existence of cities and a way by which to identify them. For simplicity, each city hosts one such industry. Any successful innovator in a first-nature industry, if located in a different city, must then relocate at no cost to implement its innovation. One may think of some natural advantages, like a primary resource, that tie these industries to some particular cities. For instance, any improvement in coal extraction can be implemented only close to coal fields.

The remaining $n - m$ industries are labeled second-nature, in the sense that production must take place where the last quality innovation occurred. A possible justification for this assumption is that, in this type of industries, the production of the highest quality depends on the many workers who took part in the innovation. Although they are individually freely mobile, coordinating the relocation of these workers to a particular city may be difficult.¹⁴ In this respect, note that in many industries where quality innovations are rather complex, such as in the electronics industry, the highest-quality products are nearly always manufactured close to the research centers where they were developed (see Masahisa Fujita and Ryoichi Ishii 1998 for evidence regarding Japanese electronics firms).

In summary, first-nature industries are “anchors” that prevent cities from disappearing,

¹³ On a more technical note, we consider only the case of a single innovation taking place between t and $t + dt$. Following a classical reasoning in the literature, the case of two or more innovations taking place between t and $t + dt$ can be neglected. Formally, the probability of exactly k innovation happening over dt is given by $(\iota dt)^k e^{-\iota dt}/k!$. The time interval dt can be made arbitrarily small. Consequently, the probability of two or more innovations taking place between t and $t + dt$ can be neglected since it is a function of $(dt)^2$ and terms of higher order.

¹⁴ Alternatively, assume that state-of-the-art knowledge is too complex to be codified and exported to another city.

whereas second-nature industries provide a reason for the growth and decline of cities.¹⁵

III. Steady-State Growth and City Size Distribution

A. Steady-State Economic Growth

In absence of costs or benefits to city size, profits are independent of location. The model can thus be solved for research and growth independently from the urban structure. Denote by v the stock market value of an industry leader. If this stock market value is the same across industries (which is the case at the symmetric equilibrium), then firm k , at a cost of $w\lambda^k(z)dt$, can expect to win $(\beta + (n - 1)\gamma) \times \lambda^k(z) \times vdt$. Profit maximization by research firms implies that, in equilibrium,

$$(15) \quad v = \frac{w}{\beta + (n - 1)\gamma}.$$

Turning to the stock market valuation of firms, industry leaders pay a dividend πdt over the period dt since their profits are not reinvested in research. The value of an industry leader appreciates by $\dot{v}dt$ when no research firm succeeds in inventing the next step up the quality ladder. This occurs with probability $1 - \iota(z)dt$. In the opposite case, when an innovation takes place in the industry, the value of its leader goes to zero. This loss of v occurs with probability $\iota(z)dt$. Summing these terms (and neglecting the terms in $(dt)^2$) implies that the (instantaneous) expected yield for a shareholder is $\pi + \dot{v} - \iota(z)v$. Investing in any firm is risky, but perfect diversification is possible since there is always one leader, so that profit is constant in each industry. Hence, firms are valued so that their expected stock market return is equal to

¹⁵ Although this dichotomy between first- and second-nature industries may seem a bit extreme, it captures in a stylized way the existing opposition between knowledge-driven and more footloose industries, on the one hand, and those bound by natural resources, on the other. Several other mechanisms could have generated results, similar to those below, albeit in a more complicated way. For instance, the existence of a durable stock of housing with the possibility of luring away a few industries (at a small cost) would also prevent cities from disappearing.

the safe interest rate, \dot{R} , which is itself equal to the subjective discount rate, ρ . Consequently

$$(16) \quad \frac{\pi + \dot{v}}{v} - \iota(z) = \rho.$$

Equations (13), (15), and (16) imply the following no-arbitrage equation:

$$(17) \quad \frac{\dot{w}}{w} + \left(1 - \frac{1}{\delta}\right) \frac{\beta + (n - 1)\gamma}{w} = \iota(z) + \rho.$$

In steady state, equations (14)–(16) imply a symmetric research effort in all industries: $\lambda(z) = \lambda$. The aggregate probability of an innovation taking place in an industry is then $\iota(z) = \iota = (\beta + (n - 1)\gamma)\lambda$. After inserting this in equation (17), rearranging yields the following differential equation guiding (nominal) wages:

$$(18) \quad \lambda = \frac{1}{\beta + (n - 1)\gamma} \frac{\dot{w}}{w} + \left(1 - \frac{1}{\delta}\right) \frac{1}{w} - \frac{\rho}{\beta + (n - 1)\gamma}.$$

This no-arbitrage condition is such that higher wages, which make research more expensive, have a negative effect on employment in research. Furthermore, a higher efficiency of research (β or γ) raises returns to this activity and thus employment therein. Finally, an increase in the discount rate, ρ , is equivalent to an increase in interest rate. This lowers the net present value of future profits and thus reduces employment in research.

The model is closed by equating labor market demand and supply. Recall that each monopoly employs $1/(\delta w)$ units of labor and that research employment is the same in all industries. Consequently, aggregate labor demand is equal to $n[1/(\delta w) + \lambda]$. Since aggregate labor supply is inelastic and equal to one, labor market clearing implies

$$(19) \quad \lambda = \frac{1}{n} - \frac{1}{\delta w}.$$

The evolution of the economy is depicted in Figure 1. The (NN) locus is the no-arbitrage

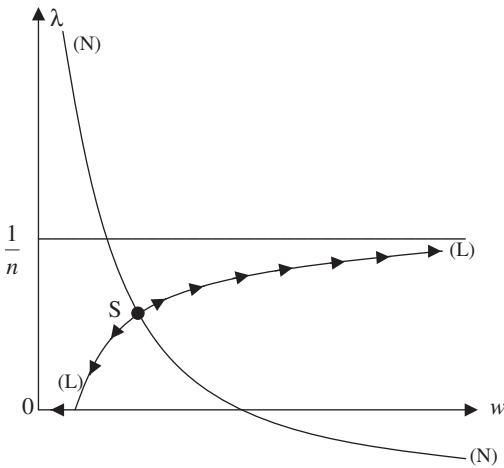


FIGURE 1. DETERMINATION OF THE EQUILIBRIUM

condition (18) under $\dot{w} = 0$, and the (LL) locus is the labor market clearing condition (19). The economy must always lie along (LL) for the labor market to clear. For values of w below (NN) , wages and research employment would fall to zero. Since expenditure is constant, expected profits must rise above research costs, a contradiction with free entry in research. A similar argument applies for values of w above (NN) . The economy must thus always be in steady state at point S .

The steady-state values of w and λ solve equations (18) and (19) with $\dot{w} = 0$:

$$(20) \quad w = \frac{1}{\frac{1}{n} + \frac{\rho}{\beta + (n - 1)\gamma}},$$

$$(21) \quad \lambda = \frac{1}{n} \left(1 - \frac{1}{\delta} \right) - \frac{\rho}{[\beta + (n - 1)\gamma]\delta}.$$

Note first that the steady state is characterized by constant employment in each industry for both production and research, as well as constant nominal wages: with aggregate expenditure normalized to n , consumers benefit from growth through lower quality-adjusted prices. As usual in the endogenous growth literature, the subjective discount rate must be low enough, i.e., $\rho \leq (\delta - 1)[\beta + (n - 1)\gamma]/n$, for research employment to be positive.

In steady state, the instantaneous probability of an innovation taking place in the economy over a period dt is $undt$. Using equation (14), this is equal to $(\beta + (n - 1)\lambda)ndt$, with λ given by (21). Each innovation then increases the aggregate quality-adjusted output by $(\delta - 1)/n$. After simplification, the expected growth rate, g , of quality-adjusted output in the economy over dt is

$$(22) \quad g_{\text{output}} = un \times \frac{\delta - 1}{n} = \left(1 - \frac{1}{\delta} \right) \left[(\delta - 1) \left(\gamma + \frac{\beta - \gamma}{n} \right) - \rho \right].$$

This expected growth rate is increasing with the size of the quality improvements, δ . There is a direct effect caused by larger quality improvements and an indirect effect, whereby larger improvements imply higher profits and thus stronger incentives to do research. The growth rate also increases with β and γ , the two efficiency parameters of research. It is also obviously decreasing with the rate of time preference, ρ . More interestingly, growth decreases with the number of industries (keeping population constant). This is a dilution effect: with more industries, a research investment has a probability of success, β , to yield a monopoly over a smaller part of the economy, $1/n$, and a probability of success, γ (lower than β), over a larger part of the economy, $(n - 1)/n$.

B. Steady State of the Probability Distribution of City Sizes

Turning to city sizes, recall that each firm with a monopoly over the highest quality in an industry is the sole producer in this industry. Furthermore, because of local spillovers, research is geographically tied to production. Consequently, symmetry across industries and free worker mobility imply that the population of a city is $1/n$ times its number of monopolies. Thus, the latter quantity is a sufficient statistic to describe a city. As a shorthand, the number of active industries in a city is referred to as its size.

To explore the properties of the model with respect to the size distribution of cities, I first solve analytically for the *steady state of the probability distribution*. This is a situation

where the number of cities of any size is expected to stay constant. More formally, it is such that

$$(23) \quad E[m_i(t + dt)] - m_i(t) = 0,$$

where $m_i(t)$ is the number of cities with i industries at time t . Note that this corresponds to a limit case for which cities would be perfectly divisible. In particular, condition (23) imposes point-wise stationarity, which cannot generally be satisfied with a finite discrete number of cities. This limit case is nonetheless a very useful approximation to understand the working of the model and to exhibit its main properties. A number of other properties, such as convergence toward steady state and the fluctuations around it when cities are not divisible, are explored by means of simulations in Section IV.

Note first that in absence of cross-industry innovation between t and $t + dt$, the urban structure is left unchanged. Changes in the size distribution of cities happen only when a second-nature industry is improved by a research firm located in a different city. In this case, a city loses an industry while another gains one.

Conditional on the occurrence of a cross-industry innovation between t and $t + dt$, the number of cities of size 1, m_1 , increases by one unit when a second-nature industry located in a city of size 2 is successfully improved by research in an industry located in another city of size 2 or above. All $n(n - 1)$ possible cross-industry innovations occur with the same probability. Since there are m_2 second-nature industries in cities of size 2, each of which can each be improved by one of $n - m_1 - 2$ other industries, the conditional probability of this event is thus

$$(24) \quad \frac{m_2 \times (n - m_1 - 2)}{n(n - 1)}.$$

Again, because dt can be made arbitrarily small, the terms in $(dt)^2$ and higher order can be neglected. Hence, cases like that of a city of size 3 losing its two second-nature industries over dt can be neglected so that equation (24) accounts for all possible increases in m_1 .

The number of cities of size 1, m_1 , declines by one unit when a second-nature industry, not

located in a city of size 2, is successfully improved by research in a city of size 1. Since there are $n - m - m_2$ such industries that can be captured by m_1 industries in cities of size 1, this event occurs with the conditional probability

$$(25) \quad \frac{(n - m - m_2) \times m_1}{n(n - 1)}.$$

In the steady state defined by (23), the probability of having one more city of size 1 must equal that of having one less city of the same size. Equality of (24) and (25) implies

$$(26) \quad m_2 = \frac{n - m}{n - 2} m_1.$$

In Appendix A, the reasoning is generalized to cities of size $i \geq 2$. For any $i \geq 1$, equation (23) then becomes

$$(27) \quad i(n - i - 1)m_{i+1} = [(2i - 1)n - im - 2i(i - 1)]m_i - (i - 1)(n - m - i + 2)m_{i-1}.$$

Note that (26) is a special case of (27) for which $i = 1$. Appendix A also shows that the sequence defined by equation (27) admits the following closed-form solution:

$$(28) \quad m_i = \frac{m(m - 1)}{n - 1} \prod_{j=1}^{j=m-2} \frac{n + 1 - m - i + j}{n - m + j}.$$

The first important property of equation (28) is that the steady state of the probability distribution is fully characterized by its numbers of cities and industries, m and n . The other parameters of the model affect only the growth rate of output: a more innovative research sector (higher β and γ), larger quality improvements (higher δ), and a higher supply of investment (lower ρ) all lead to higher growth. In turn, higher growth means a faster convergence toward steady state. However, this steady state is itself determined solely by how many cities are

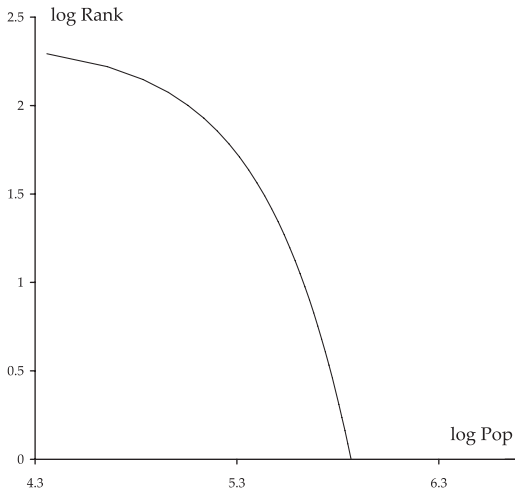


FIGURE 2. PLOT OF THE STEADY STATE OF THE PROBABILITY DISTRIBUTION

Notes: The figure plots equation (28) with $m = 232$ cities and $n = 1,500$ industries (each with 22,800 workers). These values are the same as those used below to replicate the French urban system.

competing for the existing industries. A small number of external parameters will turn out to be a very useful property to assess the model quantitatively in Section IV. Nonetheless putting more structure in the modeling of cities will lead to qualify this result in Section V.

The second important property of equation (28) is that m_i decreases monotonically with i . As shown in greater detail below, this result is driven by the fact that larger cities have more second-nature industries to lose at home and fewer to gain from outside than smaller cities. Hence, the expected growth rate of larger cities is lower and, in turn, there are fewer of them in steady state. The precise shape of the distribution in (28) is also explored more in depth below. To fix ideas at this stage, it is nonetheless useful to represent equation (28) in a log rank–log size plot. Figure 2 does this for $m = 232$ cities and $n = 1,500$ industries (each with 22,800 workers). These parameter values correspond to those used in Section IV to replicate the French city size distribution.¹⁶

¹⁶ This graph must be read with some caution. Recall that the steady state of the probability distribution ignores the indivisibility of cities. Consequently, the size corresponding to log Rank = 0 in Figure 2 is that of the fractional city of rank 1. There are other larger fractional

C. Properties of the Steady State: The Fast

According to the model, over a given time period, industries should be experiencing greater turbulence than cities, because the growth of a given city is the sum of the industries it gains minus the industries it loses. Over a given period, cities will, in general, experience both gains and losses of industries. Consequently, the sum of their employment gains and losses across industries should be larger than their overall population change. This prediction is consistent with the evidence presented in Section I.¹⁷

Second, according to the core mechanism of the model, more investment in research should lead industries to move up the quality ladder faster, and faster progress should be associated with more spatial mobility. In their study of the determinants of industry mobility in the United States, Arzaghi and Davis (2005) find some robust support in this direction. The spatial mobility of industries appears to be strongly associated with their employment share of scientists and engineers and with their expenditure share in R&D. There is also a strong positive association between the mobility of industries and their total factor productivity.

Third, the growth in industry-city employment should be negatively related to its initial level. This is because industry mobility involves a city with an established industry losing it and a city without prior employment in this industry gaining it. In his analysis of employment changes by sector across US metropolitan areas, Curtis J. Simon (2004) finds a strong negative correlation between initial sector employment and its subsequent growth.

Finally, the model also implies that larger cities are more likely both to lose and gain industries so that industry churning should in-

cities of rank 0.99, 0.98, etc. Hence, the fractional city of rank 1 is expected to be smaller than the largest indivisible city. A similar distortion, albeit smaller, also exists for cities of rank 2 and below. Unsurprisingly, in Section IV, the plot of the outcome of the simulations using the same parameter values (Figure 3, panel A) exhibits a larger largest city, a slightly larger second largest city, etc., than Figure 2.

¹⁷ This is consistent with the microfoundations of the model. R. Jason Faberman (2005), in his study of employment dynamics in cities, finds that high employment growth in cities is associated with a strong presence of young establishments that tend to experience fast employment growth.

crease with city size. Furthermore, when a cross-industry innovation takes place in a city, the likelihood that the improved industry is already located there increases with the size of that city. To take an extreme case, a city of size $n - m + 1$ can no longer grow since it already hosts all second-nature industries. For this city, the probability of losing an industry is also small, since it can occur only through cross-industry innovation originating from a first-nature industry in another city. More generally, this own-size effect implies that the probability of gaining and losing industries increases less than proportionately with city size.¹⁸

To test this important prediction, I regressed the log of the index of sectoral reallocation for each US metropolitan area between 1977 and 1997 (as defined by equation (1)) on the log of city employment in 1977. Since the index of sectoral reallocation is normalized by city population, a coefficient above -1 signifies an increase in reallocation with city size. If this increase is less than proportional to city size, this coefficient should also be below zero. The results are supportive of the prediction, since the coefficient is -0.21 with a standard error of 0.016 . When restricting the analysis to manufacturing industries, the coefficient is -0.30 with a standard error of 0.016 . Reproducing the same exercise for 217 French cities between 1985 and 1993 yields significant coefficients of -0.20 for all industries and -0.14 for manufacturing only.

D. Properties of the Steady State: The Slow

Following industry churning, cities in the model experience small positive or negative shocks, driving them slowly up or down the urban hierarchy. This prediction is obviously consistent with the *slow* mobility of cities within their distribution documented in Section I.

Beyond this, the model also predicts that over a period of time, there should be a distribution of growth rates across cities. Some cities will receive many new industries and grow a lot. Others, on the contrary, will lose many industries and experience a strong decline. By the law

¹⁸ Going against this is the fact that first-nature industries cannot be lost and are relatively more important in small cities. However, the own-size effect can be shown to dominate in the upper tail.

of large numbers, however, most cities should receive both positive and negative shocks and enjoy a rate of growth close to the mean. To verify this prediction (as well as all the predictions that follow and do not involve the sectoral composition of cities), it is preferable to use the most comprehensive sample of cities. I thus use data for 232 French urban areas (the same sample as in Section I after reintegrating the 15 cities with data problems at the sectoral level) and 922 US metropolitan and micropolitan areas (rather than the 272 metropolitan areas used above).¹⁹ In both countries, a nonparametric estimation of the probability distribution function for city growth rates yields a single-peaked distribution skewed to the right.

It is also possible to compute the expected growth rate of cities as a function of their size. Conditional on a cross-industry innovation taking place, the probability of a city with i industries gaining an industry is equal to the probability of the innovation taking place there, i/n , multiplied by the probability of the improved industry being a second-nature industry originally located in another city, $(n - m - i + 1)/(n - 1)$. By the same token, the probability of losing an industry is equal to $[(i - 1)/(n - 1)] \times [(n - i)/n]$. After simplification, the conditional expected growth of a city is

$$(29) \quad g(i) = \frac{n - im}{(n - 1)ni}.$$

This quantity decreases with i . Hence, the model predicts that city growth should be negatively correlated with size.²⁰

The relationship between the growth and size of cities has been often investigated. Mean reversion, that is, a negative effect of initial size on subsequent growth, is a pervasive finding in the

¹⁹ As defined by the US Census Bureau, micropolitan areas must have an urban cluster with a population between 10,000 and 50,000. Micropolitan areas do not overlap with metropolitan areas and are smaller than counties, rather than entire counties or aggregates thereof. With consistent definitions for metropolitan areas, micropolitan areas can thus be viewed as an addition of smaller cities to the original sample.

²⁰ Note, however, that when n is large, the range for $g(\cdot)$ may be small. As shown below in the simulations, for a realistic amount of churning across cities, the growth rate of most (if not all) cities may not appear to be significantly different from zero over a decade.

literature (see Gabaix and Ioannides 2004 for a discussion). Black and Henderson (2003) find strong mean reversion for US metropolitan areas in their decade-by-decade analysis covering the twentieth century. They confirm earlier findings by Glaeser, José A. Scheinkman, and Andrei Shleifer (1995), who find that mean reversion also holds when controlling for other factors associated with urban growth. For corroboration, the same exercise can be conducted for the 232 largest French metropolitan areas between 1990 and 1999. When regressing average annual growth on log initial population, the coefficient is highly significant at -0.11 .

The model also provides a prediction regarding the variance of the growth of cities. Because in large cities cross-industry innovations are more likely to improve industries that are already located there, the variance of city growth is expected to decline with size. Formally, conditional on a cross-industry innovation taking place, the variance of the growth of a city of size i is

$$(30) \quad \sigma^2(i) = \frac{i}{n} \frac{n-m-i+1}{n-1} \left(\frac{i+1}{i} - 1 \right)^2 + \frac{i-1}{n-1} \frac{n-i}{n} \left(\frac{i-1}{i} - 1 \right)^2 = \frac{i(2n-m+2-2i)-n}{n(n-1)i}.$$

Simple algebra then shows that this variance is a decreasing and convex function of city size, i .

The relationship between the variance of the growth rate of city population and the initial size is far less frequently investigated than the growth-size relationship. Henry G. Overman and Ioannides (2001) find that second-tier cities experience more mobility than first-tier cities in the United States, while Eaton and Eckstein (1997) find a similar pattern for French cities. Among the 922 US metropolitan and micropolitan areas for the 1990s, only one (Las Vegas) of the ten fastest growing cities and none of the ten fastest declining cities had a 1990 population above 200,000. A similar pattern emerges for the 232 French urban areas. Only one city in the top ten (Nice) and one in the bottom ten (St. Etienne) had a 1990 population above 100,000. Thus, smaller cities seem to experience larger

relative changes, as predicted by the model. To be more systematic, I followed Overman and Ioannides (2001) and regressed the “variance” of the growth of each city for the 1990s (i.e., the square of its growth minus the cross-city mean) on the log of its initial size and its square. For both French and US cities, the two coefficients have the expected sign and are highly significant, indicating a decreasing, convex relationship between the variance of city growth and initial size.

E. Properties of the Steady State: The Still

The existence of a steady state with small-scale fluctuations around it reproduces the stability of the distribution of city sizes. The steady state characterized by equation (28) is also skewed. To investigate this skewness more in depth, m_i and m_{i+1} can be written using equation (28). Simple algebra then yields

$$(31) \quad \frac{m_{i+1}}{m_i} = \frac{n+1-m-i}{n-1-i}.$$

It is convenient to note that for any *continuous* distribution $m(i)$, the *local* Zipf’s coefficient, $\zeta(i)$, estimated in a log Population–log Rank plot is such that $(\partial m(i)/\partial i)(i/m(i)) = -1 - \zeta(i)$.²¹ The discrete equivalent of this expression for the *sequence* m_i is

$$(32) \quad \zeta(i) \approx -1 - \frac{m_{i+1} - m_i}{m_i/i} = \frac{(m-3)i - n + 1}{n-1-i}.$$

For small cities, the Zipf’s coefficient is low, which corresponds to a quasi-flat Zipf’s curve (recall that $n \gg m$ implies $m_2 \approx m_1$ through equation (26)). As larger cities are considered, the Zipf’s coefficient increases ($\partial \zeta/\partial i > 0$). Put differently, the steady-state Zipf’s curve is concave with a Zipf’s coefficient below one in the lower tail and above one in the upper tail.

The intuition for this result is the following. If cities were to gain or lose industries with a

²¹ This corresponds to a local approximation by a power law. For instance, when Zipf’s law is verified, the probability distribution function follows $m(i) = ai^{-2}$, where a is a positive constant (and the counter cumulative is classically ai^{-1}). Then, it can be readily verified that $\zeta(i) = \zeta = 1$.

probability proportional to their size, this would imply Zipf's law (Gabaix 1999; Duranton 2006). As shown above, the model here deviates from this well-known benchmark because the chances of losing an industry increase more than proportionately with city size while the chances of gaining an industry increase less than proportionately. These two features lead to a steady state, which is less skewed than Zipf's law in the upper tail. This prediction of a concave Zipf's curve is not counterfactual. A concave Zipf's curve is indeed observed for the 922 US metropolitan and micropolitan areas and for the 232 French urban areas (see Figure 3, panels A and B, below, for graphical representations). This feature is also observed in many other countries (Kenneth Rosen and Mitchell Resnick 1980; Kwok Tong Soo 2005) and it holds at the world level (Henderson and Hyung Gun Wang 2005).

Finally, with all industries being symmetric and having only one location, an obvious prediction of the model is that larger cities should host more industries. Indeed, small cities tend to be highly specialized in their manufacturing employment, while large cities are more diversified (Black and Henderson 2003).

In conclusion, the model is successful at replicating the three stylized facts mentioned in the introduction. Furthermore, there is also some empirical support for its other predictions.

IV. Simulations and Quantitative Predictions

With the model able to match many qualitative features of existing city size distributions, the important issue at this stage is to know how good an approximation it provides in quantitative terms. Because the steady state of the probability distribution explored above is just a limit case, a quantitative evaluation may be carried out only by means of Monte Carlo simulations. The size distributions of cities for France in 1999 and for the United States in 2000 are used as benchmarks.

To perform these simulations, recall that only two parameters must be exogenously set, m and n . To limit the number of degrees of freedom, the values for these two parameters should be obtained from the empirical distributions. Regarding the number of cities, m , the data for both countries are quasi-exhaustive, so that $m_{\text{France}} = 232$ and $m_{\text{US}} = 922$. Obtaining a value for the number of industries, n , is more

difficult. The procedure to derive n from the data is detailed in Appendix B. It relies on the idea that the system is close to its steady state and thus the smallest cities have only one industry. This procedure yields $n_{\text{France}} = 1,500$ and $n_{\text{US}} = 12,820$ (or a population per industry of 22,800 in France and 20,400 in the United States).

For both sets of parameters, 1,000 sequences of simulations are run. Each sequence takes place as follows. Initially the n industries are distributed across the m cities proportionately to their existing population.²² In the US case, this allocation implies 898 industries for "New York," 606 for "Los Angeles," etc. Then, note that the only events of interest are the cross-industry innovations, because own-industry innovations leave the distribution of city sizes unchanged. It is therefore convenient to redefine the time scale and discretize it so that consecutive cross-industry innovations occur at consecutive dates. This "normalization" is without loss of generality, since the previous section showed that the rate of innovation did not affect the steady state. At the first date, the first "innovating" industry is picked at random, with all industries equally likely to innovate. Its corresponding "innovated" industry is also drawn from the pool of (mobile) second-nature industries. The innovated industry then relocates to the city of the innovating industry. The process is repeated for subsequent dates. By simulating a large number of innovations (e.g., one million), one can check that the steady state is reached, since any number of extra draws (e.g., 10,000) has only a minimal effect on the *average* Zipf's coefficient for the 1,000 distributions (less than 0.01).

Each sequence of simulations thus yields a fictitious population for each city. After ranking cities by decreasing log size within each sequence, the mean log size across simulations is computed for each rank. This quantity can then be plotted and compared to the empirical Zipf's curve. Panels A and B in Figure 3 represent the results for France and the United States. In both cases, the fit is good, albeit seemingly better for the United States than for France. These two panels also plot the 5- and 95-percent confidence

²² Each city receives, first, the integer part of the ratio of its population by the size of an industry (22,800 for France in 1999 and 20,400 for the United States in 2000). The remaining industries are allocated according to the largest remainder.

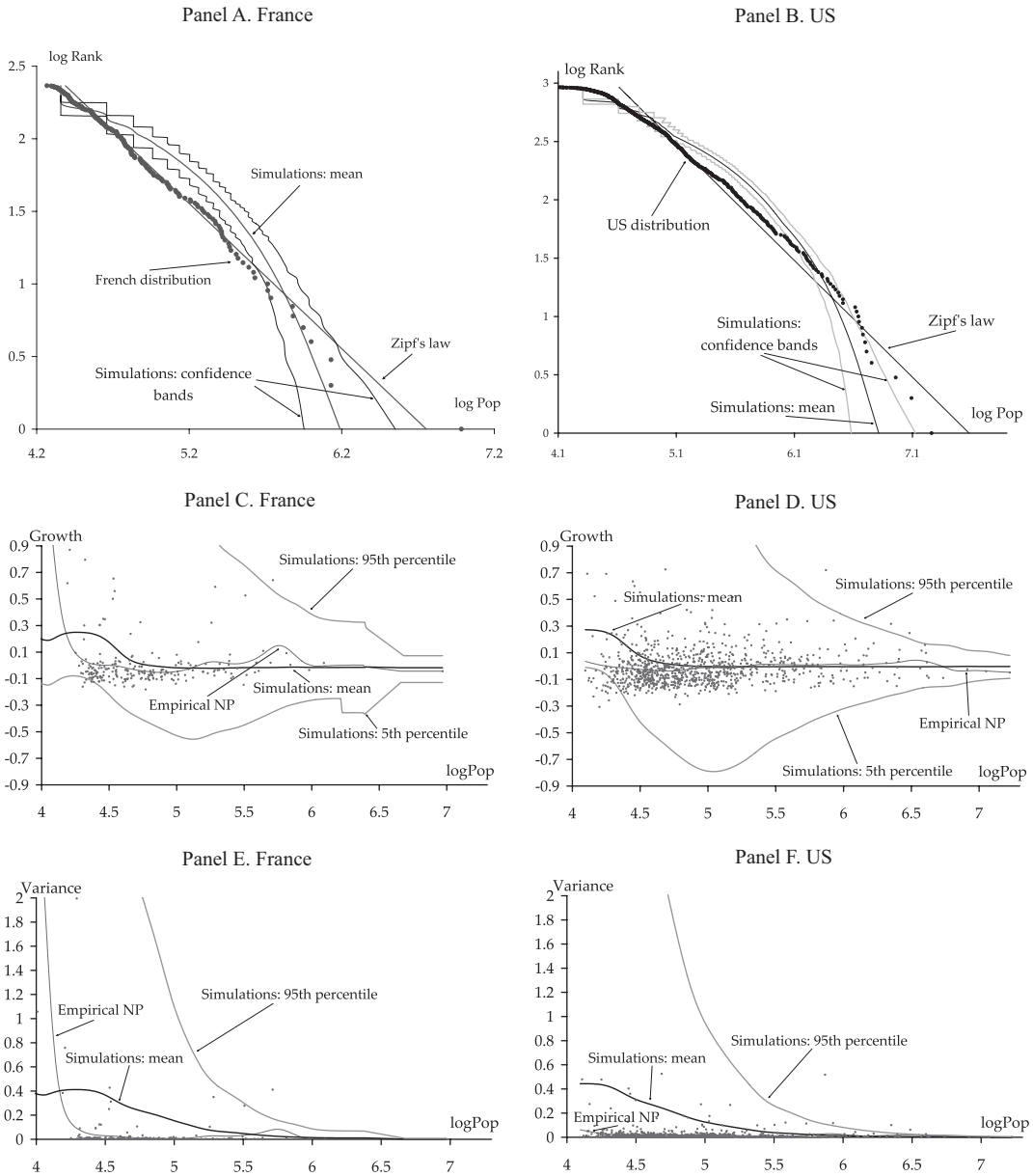


FIGURE 3. BASELINE SIMULATIONS

Notes: Panels A and B plot the actual distributions of city sizes, Zipf's law, the mean rank-size distributions from the simulations, and their 5- and 95-percent confidence bands for France and the United States, respectively. Panels C and D plot the actual observations for the normalized growth rates of French and US cities during the 1990s against their 1990 populations, the nonparametric estimations of these growth patterns (labelled NP), the mean growth rates for each size from the simulations, and their 5- and 95-percent confidence bands. Panels E and F plot the actual observations for the normalized squared-growth of cities during the 1990s against their 1990 populations, the nonparametric estimations of these (labelled NP), the mean-squared growth for each size from the simulations, and their 95-percent confidence band. The 5-percent confidence band is 0 throughout for both countries.

Source: US Census Bureau (1990 and 2000 Censuses), INSEE (1990 and 1999 Censuses), and author's simulations.

bands. These bands give an indication of the variation generated by the model. Their interpretation in terms of statistical test should be downplayed, since by construction no simulated city has a population below 20,000 (because of fixed industry size), while this is the case for actual French and US cities at the bottom of the distribution. Hence, the smallest simulated cities (with only one industry) always lie outside the confidence bands.

To evaluate more precisely how well the model fits the data, the empirical variances for log city sizes, var_{France} for 1999 and var_{US} for 2000, are first computed. Then, the mean squared differences between the actual distribution and the mean simulated distribution, msd_{France} and msd_{US} for the same dates, are computed in the following fashion:

$$(33) \quad msd = \frac{1}{m} \sum_{j=1}^m [\text{Actual log Size}(j) - \text{Mean log Simulated Size}(j)]^2,$$

where j is the rank of cities. Finally, the R^2 is calculated according to $R^2 = 1 - msd/var$. Like for the R^2 used in regression analysis, it takes a value of 100 percent when the mean simulated values reproduce exactly the actual values and a value of 0 when the mean simulated values are all equal. However, unlike “regression R^2 ,” which result from an optimization, the R^2 used here can be negative.²³

As an alternative to the model (though it ignores industry churning), Zipf’s law can be evaluated in the same way. Under Zipf’s law, the expected distribution of city sizes on a log size–log rank plot is a straight line of slope -1 . It should therefore be that the population in a “Zipf urban system” is the same as that of the empirical urban system it seeks to replicate. These two features are enough to characterize a counterfactual Zipf urban system. Thus, the fit of Zipf’s law can be evaluated without simulations. The Zipf counterfactuals for France and the United States are also plotted in Figure 3, panels A and B.

For France, the R^2 of the simulations is a satisfactory 85 percent. Zipf’s law nonetheless does better since its R^2 is 97 percent. This last result is not very surprising since the OLS Zipf coefficient for France is -0.97 with a regression R^2 of 98 percent. Interestingly, the simulated Zipf’s curves are concave like the French Zipf’s curve, though more so. For the United States, the performance of the simulations is better than in the French case, with an R^2 of 91 percent. This is also better than Zipf’s law ($R^2 = 89$ percent). The simulated Zipf’s curves also replicate the concavity of the US Zipf’s curve but exaggerate it again.

In a nutshell, in a log size–log rank plot, the Zipf’s curve generated by Zipf’s law are “too straight” while those generated by the model are too concave. Since the US Zipf’s curve is more concave than the French, the model does better in this former case and worse in the latter. Note finally that because industries are symmetric, the simulated Zipf’s curves are stepwise continuous in their lower tail. Having some heterogeneity in industry sizes would, of course, get rid of this counterfactual property, since the smallest cities would no longer bunch at the employment size of the representative industry. Introducing some heterogeneity in industry employment could also introduce the possibility of more dispersion in the upper tail when a number of large industries locate in the same city. This realistic extension would, however, require a very detailed estimation of industry size heterogeneity and is best left for future work.

Simulations are also useful to examine how well the model can replicate other moments of existing city size distributions, namely the growth and variance of French and US cities during the 1990s. The simulations here are constructed as follow. For each country, industries are initially distributed across cities proportionately to their size in 1990. Then a series of cross-industry innovations is drawn and the simulations proceed as before. A key question, of course, is how many shocks to draw. I assume that a third of industries are hit by cross-industry innovations (i.e., 500 industry shocks for France and 4,270 for the United States). This is a conservative figure, since for the United States between 1987 and 1997, the average of the index of reallocation across sectors and cities, $Churn_c$ (calculated over a ten-year interval instead of annual intervals as in equation (1)),

²³ Negative R^2 occur when the variance of the difference between the empirical distribution and the simulations is greater than the variance in the empirical distribution. This can happen if the largest simulated city becomes much larger than its empirical counterpart.

net of population growth is around 40 percent. These simulations are repeated 1,000 times. For each simulated city, the initial and final sizes are compared to compute the growth rate and its square. Then, the mean growth of cities for each initial size, as well as the 5- and 95-percent confidence bands, can be nonparametrically estimated. This information is plotted in Figure 3 (panel C for France and panel D for the United States). These two figures also plot the (normalized) growth of all cities and a nonparametric estimate of the empirical growth-size relationship for France and the United States. The exercise is then repeated for the variance. It is plotted in Figure 3, panels E and F.

The important conclusion here is that a rate of sectoral reallocation across cities of a third per decade yields (more than) enough variation to be consistent with the observed patterns of growth. In Figure 3 (panels C to F), only very few cities are outside the confidence bands. For France, only five cities in 232 had a rate of growth outside the bands, while only two had too large a variance. For the simulated US size-growth relationship, the number of cities outside the bands is larger but still well below 5 percent (26 in 933), while only four cities were outside the variance bands.²⁴

V. Adding Urban Features to the Benchmark

The main drawback of the model in Section II is its simplistic modeling of cities with neither costs nor benefits to size. This fits uneasily with the large literature documenting that cities face a trade-off between various forms of agglomeration economies and crowding costs.²⁵

²⁴ Using the same R^2 approach as previously, the performance of the model can be compared to that of Gibrat's law (i.e., all cities face a random growth process with the same mean and variance), which is the most popular growth process underlying Zipf's law. For both growth and variance, the model has a better performance than Gibrat's law for French cities and a slightly worse performance for US cities, reversing the results above regarding the steady-state distribution. Not too much emphasis should be put on this comparison, since the differences stem mostly from the lower tail. The high growth and variance in the French lower tail is better captured by the model, whereas the lower growth and variance in the US lower tail is better captured by Gibrat's law.

²⁵ See Duranton and Puga (2004) for a recent survey of the theoretical literature, and Rosenthal and Strange (2004) for a review of its empirical counterpart. A single location for each industry could, however, be interpreted as agglomeration economies within industries (a.k.a. localization economies).

Adding agglomeration economies in final production would make labor more efficient in larger cities, while crowding costs would make it more costly. The profit of monopolists may increase or decrease depending on which force dominates. Monopoly profits then determine the incentives for research firms to invest, which in turn affect the growth of cities. Agglomeration economies in research would have a more direct effect and make research labor more innovative in larger cities, while crowding costs would make it more costly.

To solve the model with agglomeration economies, one would need to derive a solution to the size-innovation relationship accounting for these features. Because of worker mobility, general equilibrium effects imply that the expected growth of any city depends (nonlinearly) on the entire distribution of city sizes so that no closed-form solution can be obtained.²⁶

Instead, the objective of this section is more modest. I consider a reduced form, whereby the probability of an innovation taking place in a city with i industries in steady state is proportional to $i \times \psi(i)$, where $\psi(i)$ is the innovativeness of the city relative to its size. This reduced form can encapsulate any trade-off between agglomeration economies in research and crowding costs. Crowding costs alone imply $\psi(i)' < 0$, whereas agglomeration economies in research alone imply $\psi(i)' > 0$.²⁷ Since this func-

tion is not a function of city size, it captures agglomeration economies within industries (a.k.a. localization economies).

²⁶ Because of the complex stochastic forward-looking nature of the dynamics, even simulations are well beyond the scope of this paper. To see this, note that the underlying dynamics is non-Markov in the sense that the growth of a city between t and $t + \Delta t$ affects its expected growth between $t + \Delta t$ and $t + 2\Delta t$, etc. In turn, this affects *ex ante* investment.

²⁷ This exercise is "partial" since agglomeration economies in production are ignored (or assumed to cancel out directly with crowding costs). This omission is not as important as it seems. First, with agglomeration effects, the level of employment in industries will depend on the size of the city. With each industry receiving a constant share of expenditure and modest agglomeration economies, this effect is unlikely to be large. Second, agglomeration economies in production affect the profit of monopolists and thus the incentive to invest in research. This second effect can be ignored since it is redundant with the direct effect of agglomeration economies in research. Finally, note that ignoring agglomeration effects in production makes it possible to focus the simulations on the key function $\psi(i)$, without introducing too many degrees of freedom.

TABLE 2—SUMMARY STATISTICS FOR THE SIMULATIONS

	France			United States		
	ζ_{100}	ζ_{232}	R^2 (percent)	ζ_{100}	ζ_{232}	R^2 (percent)
$\varepsilon = 0$	1.28	0.83	85	1.49	0.71	91
$\varepsilon = 0.005$	1.25	0.84	88	1.40	0.73	93
$\varepsilon = 0.01$	1.21	0.84	89	1.22	0.88	80
$\varepsilon = 0.02$	1.15	0.85	92	1.11	1.09	44
$\varepsilon = 0.025$	1.09	0.95	97	1.07	1.16	30
$\varepsilon = 0.03$	1.07	1.02	92	1.04	1.23	16
$\varepsilon = 0.04$	1.04	1.11	80	1.00	1.35	-5
$\varepsilon = 0.05$	1.04	1.17	67	0.96	1.44	-25
Zipf's law	1.00	1.00	97	1.00	1.00	89
Empirical values	1.02	0.97	—	1.07	0.82	—

Notes: Empirical values are for 1999 (France) and 2000 (United States). The ζ coefficients reported in the different columns are estimated as in equation (5) for the number of cities in subindex. The R^2 are evaluated as described above.

tion, $\psi(i)$, is not derived explicitly from an underlying model, some external information should be used for calibration purposes.

Unfortunately different proxies for urban innovativeness (such as patenting per capita, rates of firm creation, or various measure of city education) give fairly different estimates for $\psi(i)$.²⁸ It is thus better to take values for $\psi(i)$ that are within an empirically plausible range and look at how changes in $\psi(i)$ affect the steady-state distribution, instead of using a particular data source and a “carefully selected” estimate for the size-innovation relationship. In their review of the empirical literature, Rosenthal and Strange (2004) argue that most estimates for agglomeration economies imply an elasticity of productivity to size in the 3- to 8-percent range. These estimates concern *gross* agglomeration economies. After taking crowding costs into account, a 0- to 5-percent range is probably realistic for *net* agglomeration economies.

²⁸ Patenting (i.e., average number of utility patents per capita between 1990 and 1999 for US metropolitan areas from the US Patent and Trademark Office) is arguably the best proxy, but the estimates are sensitive to the chosen functional form. This is in part because there is a peak of innovativeness for population between half a million and a million (the five most innovative US cities are San Jose, California; Boulder, Colorado; Rochester, New York; and Saginaw and Ann Harbor, Michigan). Looking at firm creation leads to the same broad pattern (a per capita increase with city size that tails off for larger cities), but the elasticities are much smaller (around 1 to 2 percent as opposed to 10 percent or more). The proportion of university graduates in the city work force also implies high elasticities, whereas for the average years of education the elasticities are much smaller.

Using the same simulation approach as previously, the effect of agglomeration economies in research net of crowding costs can be assessed quantitatively. The only difference with the baseline simulations is that the probability of a city generating an innovation is no longer proportional to its size, but rather is proportional to its size times $\psi(i) = i^\varepsilon$, with ε being allowed to vary over the 0- to 5-percent range. The parameter ε captures net dynamic economies of scale: the probability of innovating within a given industry increases by $1 + \varepsilon$ percent when the size of its city increases by 1 percent.

An increase in net agglomeration economies in research, ε , makes larger cities more likely to innovate relative to smaller cities. Hence, the expected growth of larger cities relative to smaller cities should increase with ε . As a result, the steady-state distribution should become more skewed in the upper tail. Smaller cities will also find it more difficult to grow so that there should be more of them in steady state.

This is exactly what we observe in the simulations. A summary of the results is given in Table 2. As ε increases, the upper tail becomes more skewed, as shown by the decline of the Zipf's coefficients for the upper tails (ζ_{100} for France and ζ_{150} for the United States). As ε increases, there are also more smaller cities. This implies a higher Zipf's coefficient for the lower tail. Given that there are many more cities in the lower tail than in the upper tail, the second effect dominates and a rise in ε leads to higher Zipf's coefficients for the entire

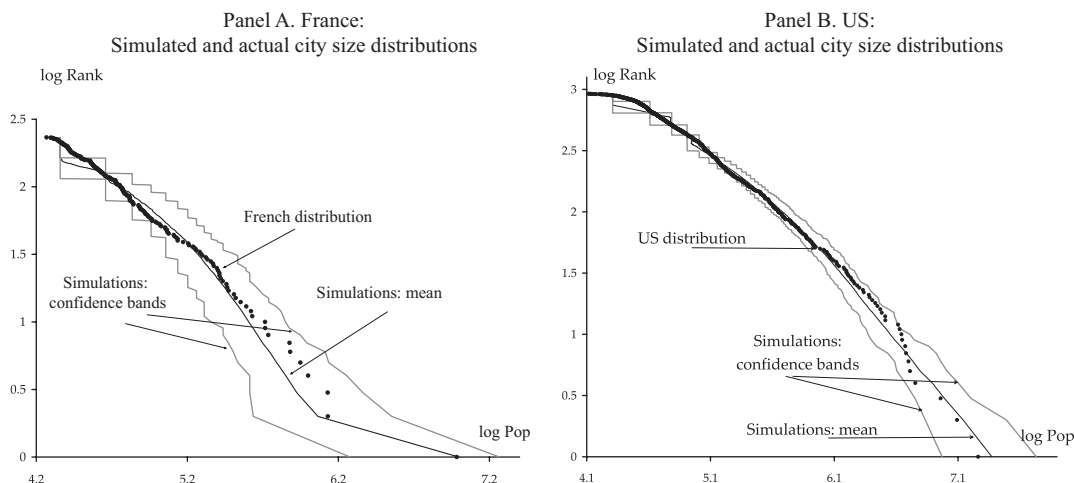


FIGURE 4. SIMULATIONS WITH SCALE EFFECTS IN INNOVATION

Notes: The figures plot the actual distributions of city sizes, the mean rank-size distributions from the augmented simulations, and their 5- and 95-percent confidence bands for France and the United States, respectively. Panel A was produced with 232 cities, 1,500 industries each with 22,800 workers, and $\varepsilon = 0.025$. Panel B was produced with 932 cities, 12,820 industries each with 20,400 workers, and $\varepsilon = 0.007 + 0.05t^{0.2}$.

Source: US Census Bureau (2000 Census), INSEE (1999 Census), and author's simulations.

distributions (ζ_{232} for France and ζ_{922} for the United States). Further increases in ε lead to some convexity in the upper tail and further reductions of the concavity in the lower tail.

Since the baseline case ($\varepsilon = 0$) generated overly concave Zipf's curves, a positive ε may thus improve on the performance of the baseline. For France, the best fit is obtained for $\varepsilon = 2.5$ percent. With this value of ε , the performance of the simulations is better than that of Zipf's law. With $R^2 = 97.2$ percent, there is little room for improvement because the lumpiness of industries makes it impossible to get an R^2 above 98.5 percent. Panel A of Figure 4 shows that the mean simulation is indeed very close to its empirical counterpart. It also shows that all French cities with a population above 30,000 are within the confidence bands of the simulations. For the United States, the R^2 is maximized for $\varepsilon = 0.5$ percent, offering only a modest improvement on the baseline case. The reason is that, with further increases of ε , the largest city in the simulations becomes too large in comparison to New York. Because of the large number of second-nature industries in our simulated US urban system, even a small degree of unmitigated agglomeration economies can lead to a very large dominant city. This problem

is avoided by allowing (realistically) for agglomeration effects to tail off when cities get very large. For $\varepsilon = 0.007 + 0.05t^{0.2}$, the R^2 for the entire US distribution is 96 percent, and above 99 percent for the 300 largest cities. This functional form corresponds to net economies of agglomeration of about 6 percent for the smallest cities and about 2 percent for a city the size of New York. Panel B of Figure 4 shows that these simulations are indeed very close to the US urban system.

In conclusion, these richer simulations show that the model can accommodate important features of cities that were hitherto ignored. Interestingly, for empirically plausible values of agglomeration effects, the simulations can get very close to existing urban systems.²⁹

²⁹ One should not attempt, however, to read too much into these simulations. In particular, agglomeration effects cannot be precisely identified by applying a maximum likelihood approach to the simulations. First, the confidence bands get fairly wide for the largest cities. Second, higher net agglomeration economies in the upper tail can be compensated by even higher agglomeration economies in the lower tail to yield very similar distributions of cities sizes.

VI. Concluding Comments

To summarize, the contributions of this paper are the following. First, it provides a simple mechanism (innovation-driven shocks at the level of industries and cities) to explain the growth and decline of cities. This model is able to replicate three major stylized facts about urban evolutions: the fast, the slow, and the still. Through simulations, it can also replicate effectively various moments of the French and US city size distributions. Second, this paper shows that the fit of the model is improved when

standard urban features like agglomeration economies and crowding costs are considered. The last contribution of the paper is more subtle. Replicating existing patterns of city size distributions may not be as difficult as previously thought. This implies that the real test to distinguish between different mechanisms like the one highlighted here, that of Gabaix (1999), and any potential alternative is not whether they can replicate observed patterns. Instead the key issue regards the empirical importance of these mechanisms as determinants of the growth of cities.

APPENDIX A: STEADY STATE OF THE PROBABILITY DISTRIBUTION

The number m_i of cities of size $i \geq 2$ grows by two units when a second-nature industry located in a city of size $i + 1$ is successfully improved by a research firm located in a city of size $i - 1$. Since there are $n(n - 1)$ possible cross-industry innovations, im_{i+1} second-nature industries in cities of size $i + 1$, and $(i - 1)m_{i-1}$ potential sources of innovations in cities of size $i - 1$, this event occurs with probability

$$(A1) \quad \frac{[im_{i+1}] \times [(i - 1)m_{i-1}]}{n(n - 1)},$$

conditionally on a cross-industry innovation taking place. Then, m_i grows by one unit when an industry relocates from a city of size $i + 1$ to a (different) city of any size, except i and $i - 1$. Since there are im_{i+1} industries that can be captured this way by $n - (i + 1) - im_i - (i - 1)m_{i-1}$ industries, this event takes place with the conditional probability

$$(A2) \quad \frac{[im_{i+1}] \times [n - i - 1 - im_i - (i - 1)m_{i-1}]}{n(n - 1)}.$$

Finally, m_i also increases by one unit when research in a city of size $i - 1$ successfully improves a second-nature industry located in a different city of any size but i and $i + 1$. Since there are $(i - 1)m_{i-1}$ industries in cities of size $i - 1$ that can capture one of $n - m - (i - 2) - (i - 1)m_i - im_{i+1}$ industries, this event takes place with the conditional probability

$$(A3) \quad \frac{[n - m - (i - 2) - (i - 1)m_i - im_{i+1}] \times [(i - 1)m_{i-1}]}{n(n - 1)}.$$

The number m_i of cities of size $i \geq 2$ decreases by two units when a second-nature industry located in a city of size i is successfully improved by a research firm located in another city of size i . Since there are im_i industries that can capture one of $(i - 1)(m_i - 1)$ industries, the conditional probability of this event is

$$(A4) \quad \frac{[(i - 1)(m_i - 1)] \times [im_i]}{n(n - 1)}.$$

Then, m_i declines by one unit when an industry relocates from a city of size i to a city of any size except i and $i - 1$. With $(i - 1)m_i$ industries that can be captured this way by $n - im_i - (i - 1)m_{i-1}$ industries, this event takes place with the conditional probability

$$(A5) \quad \frac{[(i-1)m_i] \times [n - im_i - (i-1)m_{i-1}]}{n(n-1)}.$$

Finally, m_i also decreases by one unit when research in a city of size i successfully improves a second-nature industry located in a city of any size but i and $i+1$. Since there are im_i industries in cities of size i that can capture one of $n - m - (i-1)m_i - im_{i+1}$ industries, this event takes place with the conditional probability

$$(A6) \quad \frac{[n - m - (i-1)m_i - im_{i+1}] \times [im_i]}{n(n-1)}.$$

Using (A1)–(A6), the steady-state condition (23) for $i \geq 2$ yields equation (27) in the text. The general solution to the sequence defined by (27), $m_{i+1} = \{[(2i-1)n - im - 2i(i-1)]m_i - (i-1)(n - m - i + 2)m_{i-1}\} / [i(n - i - 1)]$, is

$$(A7) \quad m_i = \frac{\Gamma(n-2)}{\Gamma(n-m)} \frac{\Gamma(n-m-i+1)}{\Gamma(n-i-1)} m_1.$$

A well-known property of Γ -functions is $\Gamma(x) = (x-1)\Gamma(x-1)$. Using this property in equation (A7) yields, after simplification,

$$(A8) \quad m_i = m_1 \prod_{j=1}^{i-2} \frac{n+1-m-i+j}{n-m+j}.$$

(For $i=2$, it is easy to verify that equation (A8) boils down to equation (26). For $i \geq 2$, it can also be verified that m_{i+1} , m_i , and m_{i-1} , as derived from (A8), satisfy equation (27).)

To derive m_1 , it must be noted first that summing industries across all possible sizes for cities from 1 (the smallest) to $n-m+1$ (the largest) implies $\sum_{i=1}^{n-m+1} m_i = m$. Using equation (A8) then yields

$$(A9) \quad m = \sum_{i=1}^{n-m+1} m_i = m_1 \sum_{i=1}^{n-m+1} \left(\prod_{j=1}^{i-2} \frac{n+1-m-i+j}{n-m+j} \right).$$

To simplify this expression, it is convenient to define the series $s_0 = 0$ and $s_i = m_i + s_{i-1}$ for all $i \geq 1$. The general solution to this series is

$$(A10) \quad s_i = \frac{n-1}{m-1} m_1 - \frac{n-i-1}{m-1} \frac{\Gamma(n-i-1)}{\Gamma(n-1)} \frac{\Gamma(n-m+1)}{\Gamma(n-m+1-i)} m_1.$$

For the largest possible city size, $i = n - m + 1$, we have $\Gamma(n - m + 1 - i) = \Gamma(0) = +\infty$, so that the last term in equation (A10) is zero. Consequently, this yields

$$(A11) \quad s_{n-m+1} = \frac{n-1}{m-1} m_1,$$

since, by definition, $s_{n-m+1} = m$, m_1 is given by

$$(A12) \quad m_1 = \frac{m(m-1)}{n-1}.$$

Finally, inserting equation (A12) into (A8) yields equation (28) in the text.

APPENDIX B: SIMULATIONS: DERIVING AN EMPIRICAL VALUE FOR n

Note first that equation (A12) can be rewritten as

$$(B1) \quad \frac{1}{n} = \frac{1}{\frac{m(m-1)}{m_1} + 1}.$$

Then, assume that the smallest m_1 cities in the empirical distribution have only a single industry. Denote $\overline{Size}(m_1)$ the average size of the m_1 smallest cities in the distribution. In this case, the model implies that $\overline{Size}(m_1)$, which is also the size of an industry, is such that $\overline{Size}(m_1) \times n = Pop$, where Pop denotes the overall urban population. Inserting this equality into (B1) implies that m_1 can finally be computed as the (empirical) fixed-point of

$$(B2) \quad \overline{Size}(m_1) = \frac{Pop}{\frac{m(m-1)}{m_1} + 1}.$$

In France, the population of the 232 largest cities is 34.2 m. The 36 smallest cities have an average population of circa 22,800. These numbers approximately verify equation (B2) (subject to the integer constraint regarding the number of cities). We thus obtain $n_{\text{France}} = 1,500$. For the United States, the overall urban population of the 922 largest areas is 261.5 m. The average population of the 66 smallest areas is around 20,400. Again these numbers approximately verify equation (B2). They imply $n_{\text{US}} = 12,820$.

REFERENCES

- Arzaghi, Mohammad, and James C. Davis.** 2005. "Spatial Mobility and Geographic Concentration." Unpublished.
- Barrios, Salvador, Luisito Bertinelli, Eric Strobl, and Antonio-Carlos Teixeira.** 2005. "The Dynamics of Agglomeration: Evidence from Ireland and Portugal." *Journal of Urban Economics*, 57(1): 170–88.
- Beardsell, Mark, and J. Vernon Henderson.** 1999. "Spatial Evolution of the Computer Industry in the USA." *European Economic Review*, 43(2): 431–56.
- Berliant, Marcus, and Ping Wang.** 2004. "Dynamic Urban Models: Agglomeration and Growth." In *Urban Dynamics and Growth: Advances in Urban Economics*, ed. Roberta Capello and Peter Nijkamp, 533–81. Amsterdam: Elsevier.
- Black, Duncan, and J. Vernon Henderson.** 1999. "Urban Evolution of Population and Industry in the United States." *American Economic Review (Papers and Proceedings)*, 89(2): 321–27.
- Black, Duncan, and J. Vernon Henderson.** 2003. "Urban Evolution in the USA." *Journal of Economic Geography*, 3(4): 343–72.
- Blanchard, Olivier Jean, and Lawrence F. Katz.** 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, 1: 1–61.
- Carlino, Gerald A., Robert H. DeFina, and Keith Sill.** 2001. "Sectoral Shocks and Metropolitan Employment Growth." *Journal of Urban Economics*, 50(3): 396–417.
- Combes, Pierre-Philippe, Thierry Magnac, and Jean-Marc Robin.** 2004. "The Dynamics of Local Employment in France." *Journal of Urban Economics*, 56(2): 217–43.
- Córdoba, Juan-Carlos.** 2003. "On the Distribution of City Sizes." Unpublished.
- Coulson, N. Edward.** 1999. "Sectoral Sources of Metropolitan Growth." *Regional Science and Urban Economics*, 29(6): 723–43.
- Davis, Steven J., and John C. Haltiwanger.** 1998. "Measuring Gross Worker and Job Flows." In *Labor Statistics Measurement Issues*, ed. John C. Haltiwanger, Marilyn E. Manser, and Robert H. Topel, 77–119. Chicago: University of Chicago Press.
- Dobkins, Linda Harris, and Yannis M. Ioannides.** 2000. "Dynamic Evolution of the Size Distribution of U.S. Cities." In *Economics of Cities: Theoretical Perspectives*, ed. Jean-Marie Huriot and Jacques-François Thisse,

- 217–60. Cambridge: Cambridge University Press.
- Dumais, Guy, Glenn Ellison, and Edward L. Glaeser.** 2002. “Geographic Concentration as a Dynamic Process.” *Review of Economics and Statistics*, 84(2): 193–204.
- Duranton, Gilles.** 2006. “Some Foundations for Zipf’s Law: Product Proliferation and Local Spillovers.” *Regional Science and Urban Economics*, 36(4): 542–63.
- Duranton, Gilles, and Diego Puga.** 2001. “Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products.” *American Economic Review*, 91(5): 1454–77.
- Duranton, Gilles, and Diego Puga.** 2004. “Micro-Foundations of Urban Agglomeration Economies.” In *Handbook of Regional and Urban Economics Volume 4: Cities and Geography*, ed. J. Vernon Henderson and Jacques-François Thisse, 2063–2117. Amsterdam: Elsevier.
- Eaton, Jonathan, and Zvi Eckstein.** 1997. “Cities and Growth: Theory and Evidence from France and Japan.” *Regional Science and Urban Economics*, 27(4-5): 443–74.
- Eckhout, Jan.** 2004. “Gibrat’s Law for (All) Cities.” *American Economic Review*, 94(5): 1429–51.
- Faberman, R. Jason.** 2005. “What’s in a City? Understanding the Micro-Level Employer Dynamics Underlying Urban Growth.” United States Bureau of Labor Statistics Working Paper 386.
- Fujita, Masahisa, and Ryoichi Ishii.** 1998. “Global Location Behavior and Organizational Dynamics of Japanese Electronics Firms and their Impact on Regional Economies.” In *The Dynamic Firm: The Role of Technology, Strategy, Organization, and Regions*, ed. Alfred D. Chandler, Jr., Peter Hagström, and Örjan Sölvell, 343–83. Oxford: Oxford University Press.
- Gabaix, Xavier.** 1999. “Zipf’s Law for Cities: An Explanation.” *Quarterly Journal of Economics*, 114(3): 739–67.
- Gabaix, Xavier, and Yannis M. Ioannides.** 2004. “The Evolution of City Size Distributions.” In *Handbook of Regional and Urban Economics, Volume 4: Cities and Geography*, ed. J. Vernon Henderson and Jacques-François Thisse, 2341–78. Amsterdam: Elsevier.
- Glaeser, Edward L., José A. Scheinkman, and Andrei Shleifer.** 1995. “Economic Growth in a Cross-Section of Cities.” *Journal of Monetary Economics*, 36(1): 117–43.
- Grossman, Gene M., and Elhanan Helpman.** 1991. “Quality Ladders in the Theory of Growth.” *Review of Economic Studies*, 58(1): 43–61.
- Henderson, J. Vernon.** 2006. “Urbanization and Growth.” In *Handbook of Economic Growth*, Vol. 1B, ed. Philippe Aghion and Steven N. Durlauf, 1543–91. Amsterdam: North-Holland.
- Henderson, J. Vernon, and Hyung Gun Wang.** 2005. “Urbanization and City Growth: The Role of Institutions.” Unpublished.
- Jacobs, Jane.** 1970. *The Economy of Cities*. New York: Random House.
- Overman, Henry G., and Yannis M. Ioannides.** 2001. “Cross-Sectional Evolution of the U.S. City Size Distribution.” *Journal of Urban Economics*, 49(3): 543–66.
- Romer, Paul M.** 1990. “Endogenous Technological Change.” *Journal of Political Economy*, 98(5): S71–102.
- Rosen, Kenneth T., and Mitchel Resnick.** 1980. “The Size Distribution of Cities: An Examination of the Pareto Law and Primacy.” *Journal of Urban Economics*, 8(2): 165–86.
- Rosenthal, Stuart S., and William C. Strange.** 2004. “Evidence on the Nature and Sources of Agglomeration Economies.” In *Handbook of Regional and Urban Economics Volume 4: Cities and Geography*, ed. J. Vernon Henderson and Jacques-François Thisse, 2119–71. Amsterdam: Elsevier.
- Rossi-Hansberg, Esteban, and Mark L. J. Wright.** Forthcoming. “Urban Structure and Growth.” *Review of Economic Studies*.
- Scherer, Frederic M.** 1984. “Using Linked Patent and R&D Data to Measure Interindustry Technology Flows.” In *R&D, Patents, and Productivity*, ed. Zvi Griliches, 417–61. Chicago: University of Chicago Press.
- Simon, Curtis J.** 2004. “Industrial Reallocation across US Cities, 1977–1997.” *Journal of Urban Economics*, 56(1): 119–43.
- Simon, Herbert.** 1955. “On a Class of Skew Distribution Functions.” *Biometrika*, 42(2): 425–40.
- Soo, Kwok Tong.** 2005. “Zipf’s Law for Cities: A Cross-Country Investigation.” *Regional Science and Urban Economics*, 35(3): 239–63.
- Wallace, Nancy.** 2004. “Agglomeration Economies and the High-Tech Computer Cluster.” Unpublished.

This article has been cited by:

1. Corinne Autant-Bernard, James P. LeSage. 2010. QUANTIFYING KNOWLEDGE SPILLOVERS USING SPATIAL ECONOMETRIC MODELS. *Journal of Regional Science* no-no. [[CrossRef](#)]
2. Rafael González-Val. 2010. THE EVOLUTION OF U.S. CITY SIZE DISTRIBUTION FROM A LONG-TERM PERSPECTIVE (1900-2000)*. *Journal of Regional Science* **50**:5, 952-972. [[CrossRef](#)]
3. Mark D. Partridge. 2010. The duelling models: NEG vs amenity migration in explaining US engines of growth. *Papers in Regional Science* **89**:3, 513-536. [[CrossRef](#)]
4. I. S. Gill, C.-C. Goh. 2010. Scale Economies and Cities. *The World Bank Research Observer* **25**:2, 235-262. [[CrossRef](#)]
5. M. ROSE OLFERT, MARK D. PARTRIDGE. 2010. Best Practices in Twenty-First-Century Rural Development and Policy. *Growth and Change* **41**:2, 147-164. [[CrossRef](#)]
6. Jesús Clemente, Rafael González-Val, Irene Olloqui. 2010. Zipf's and Gibrat's laws for migrations. *The Annals of Regional Science* . [[CrossRef](#)]
7. Klaus Desmet, Esteban Rossi-Hansberg. 2010. ON SPATIAL DYNAMICS. *Journal of Regional Science* **50**:1, 43-63. [[CrossRef](#)]
8. Michael Storper. 2010. AGGLOMERATION, TRADE, AND SPATIAL DEVELOPMENT: BRINGING DYNAMICS BACK IN. *Journal of Regional Science* **50**:1, 313-342. [[CrossRef](#)]
9. J. Vernon Henderson. 2010. CITIES AND DEVELOPMENT. *Journal of Regional Science* **50**:1, 515-540. [[CrossRef](#)]
10. Stephen J. Redding. 2010. THE EMPIRICS OF NEW ECONOMIC GEOGRAPHY. *Journal of Regional Science* **50**:1, 297-311. [[CrossRef](#)]
11. Xavier Gabaix. 2009. Power Laws in Economics and Finance. *Annual Review of Economics* **1**:1, 255-294. [[CrossRef](#)]
12. Mark D. Partridge, Dan S. Rickman, Kamar Ali, M. Rose Olfert. 2009. Do New Economic Geography agglomeration shadows underlie current population dynamics across the urban hierarchy?. *Papers in Regional Science* **88**:2, 445-466. [[CrossRef](#)]
13. Gilles Duranton. 2008. Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics/Revue canadienne d'économique* **41**:3, 689-736. [[CrossRef](#)]
14. M. D. Partridge, D. S. Rickman, K. Ali, M. R. Olfert. 2008. Lost in space: population growth in the American hinterlands and small cities. *Journal of Economic Geography* **8**:6, 727-757. [[CrossRef](#)]
15. M. Batty. 2008. The Size, Scale, and Shape of Cities. *Science* **319**:5864, 769-771. [[CrossRef](#)]
16. Tomoya Mori, Koji Nishikimi, Tony E. Smith. 2008. THE NUMBER-AVERAGE SIZE RULE: A NEW EMPIRICAL RELATIONSHIP BETWEEN INDUSTRIAL LOCATION AND CITY SIZE*. *Journal of Regional Science* **48**:1, 165-211. [[CrossRef](#)]