Chapter 34
Beyond Zipf: An Agent-Based Understanding of City Size Distributions

Timothy R. Gulden and Ross A. Hammond

Abstract George Kingsley Zipf observed in 1949 that the size distribution of cities within nations tends to follow a particular kind of power-law. This distribution is often described as the "rank size rule" or simply as the Zipf distribution. While Zipf convincingly documented this distribution, he was less successful in explaining its emergence. During the ensuing half century, various theories of city formation and development have emerged that have contributed real insights into the geography and economics of cities. For the most part, however, these theories have failed to predict the Zipf distribution of sizes. Another class of theories has been put forward to explain the distribution, but these have tended to rest on unrealistic assumptions, to lack explanatory power, or, at best, to lack the ability to explain the deviations from Zipf that can be observed in many nations. In this paper, we offer a simple agent-based model of city size evolution. This model offers substantial insight into the distribution of city sizes in various countries, while complementing previous work on the economic geography of cities and offering plausible economic interpretations and logic. The model can also account for several important categories of systematic deviation from Zipf that are observed in empirical data, and offers new insights about how such deviations arise.

34.1 Introduction

George Kingsley Zipf observed in 1949 that the size distribution of cities within nations tends to follow a particular kind of power-law (Zipf 1949). This distribution is often described as the "rank size rule" or simply as the Zipf distribution. While Zipf

T.R. Gulden (✉)
Krasnow Institute for Advanced Study, George Mason University, Fairfax, VA, USA
e-mail: tgulden@gmu.edu

R.A. Hammond
Center on Social Dynamics and Policy, The Brookings Institution, Washington, DC, USA
e-mail: rhammond@brookings.edu

convincingly documented this rule in cities and many other systems (including the frequency of word usage in most languages), he was less successful in explaining its emergence. During the ensuing half century, various theories of city formation and development have emerged and have contributed real insights into the geography and economics of cities. They have, for the most part, however, failed to predict the Zipf distribution of sizes. Another class of theories has been put forward to explain the distribution, but these have tended to rest on unrealistic assumptions, to lack explanatory power, or, at best, to lack the ability to explain the deviations from Zipf that can be observed in many nations. In this paper, we offer a simple, agent-based model (ABM) of city size evolution. This model offers substantial insight into the distribution of city sizes in various countries while complementing previous work in the economic geography of cities, and offers plausible economic interpretations and logic. The model can also account for several important categories of systematic deviation from Zipf that are observed in empirical data, and offers new insights about how such deviations arise.

In essence, we find that the distribution and its variants arise naturally when people try to optimize their wellbeing by migrating between a given set of cities. We use an agent approach because the dynamics of this process depend on the citizens having imperfect information – they are more likely to move from a congested city to an uncongested one, but may also misread the situation and move to a city that is already overfull. This simulation approach also allows us to model city sizes in dynamic terms, with urban equilibrium sizes responding to population shocks with a lagged adjustment mechanism corresponding to the adaptive provision and decay of infrastructure. We represent these mechanisms in highly simplified, abstract terms beginning with a model that is so abstract that it has little to do with human behavior. We proceed, however, to develop a model that, while still abstract, does present a plausible version of these basic economic phenomena.

34.1.1   The Zipf Distribution

The Zipf distribution is neatly summarized by the expression $S_r = S_0 * r^{-1}$ where S_r is the size of city r, r is the rank of the city (i.e. for the tenth largest city, $r = 10$) and S_0 is the size of the largest city. This can be restated as the so called "rank size rule" by observing that the second largest city is half the size of the largest city, the third largest 1/3 as large, the fourth 1/4 as large, etc. One property of this distribution is that when it is plotted as an ordered histogram on log-log axes, it results in a straight line with a slope of −1 (the exponent of the power-law) as shown in Fig. 34.1.

34.1.2   Explanations for the Distribution

While the Zipf regularity has been well known for some time, it has resisted attempts at theoretical explanation. Fujita et al. (1999) directly address this fit between theory and observation in their chapter entitled "An Empirical Digression: The Sizes of Cities". They write:
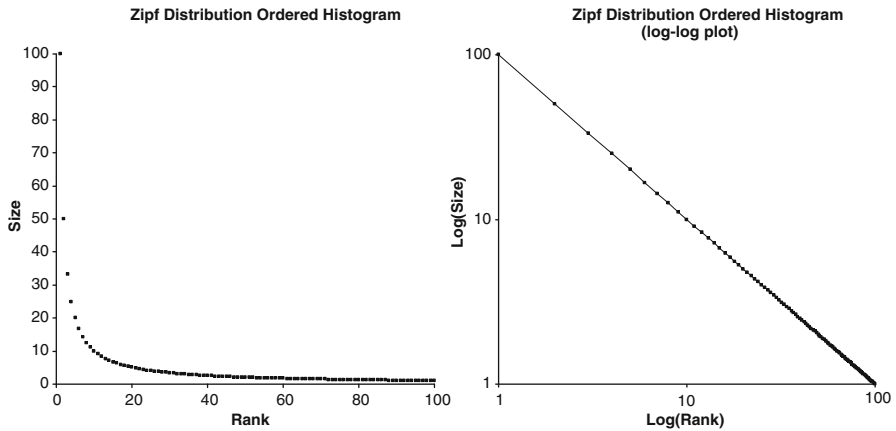
Fig. 34.1 Zipf Distribution ordered histogram on normal and log-log axes

> Attempts to match economic theory with data usually face the problem that the theory is excessively neat, that theory gives simple, sharp-edged predictions, whereas the real world throws up complicated and messy outcomes. When it comes to the size distribution of cities, however, the problem we face is that the data offer a stunningly neat picture, one that is hard to reproduce in any plausible (or even implausible) theoretical model. (p. 215).

The conclusion to this chapter begins: "At this point we have no resolution to the explanation of the striking regularity in city size distributions. We must acknowledge that it poses a real intellectual challenge to our understanding of cities…" (p. 225). Although work in this area has continued in the intervening years, there remain few behaviorally-based candidates to explain the Zipf regularity, and no consensus on how these explanations relate to one another.

Attempts to model the dynamics of city size have largely fallen into one of two categories.[1] Models in the first category extend concepts from standard economic theory to apply to city size dynamics. These include externality models, which apply the "Henry George" theorem from urban economics (Marshall 1890; Jacobs 1984; Henderson 1974; Kanemoto 1980), and models that extend Christaller's (1933) "central place" theory; see Fujita and Mori (1997). Such models are well integrated with the existing body of economic theory, and are often consistent with other economic evidence about city dynamics. Unfortunately, none of these models convincingly produce the empirical regularity of the Zipf distribution.

Models in the second category apply one or more abstract stochastic processes to represent city size dynamics. Early examples included Simon's (1957) proportional growth model and Hill and Woodrofe's (1975) application of the Bose-Einstein process. More recently, the most prominent models in this category have focused on descriptions of city growth as a "Gibrat process" (Gibrat 1931). Papers applying the Gibrat processes include Gabaix (1999a) and Reed (2002). These processes have all

---

[1] For a more comprehensive review of the literature on urban size distributions, see Gabaix and Ioannides (2003).

been shown mathematically to successfully generate a stable power-law distribution, and in many cases, to closely replicate the Zipf distribution itself. However, such models have little or no economic content. They demonstrate that the Zipf regularity can be generated using a variety of statistical mechanisms, but they do not offer a set of comparable behavioral principles or realistic economic mechanisms that are sufficient to produce Zipf. As one recent paper put it: "this collection of models is essentially statistical – they seek to *generate rather than to explain* the regularity" (Overman and Ioannides 2001). It is often unclear how the abstract mechanisms represented in many of these models can be useful metaphors for real-world social or economic processes. Indeed, in some cases, closer examination has found strong empirical evidence that mechanisms such as the Gibrat process are *not* good descriptions of real city-size dynamics; see Cuberes (2004). Abstract stochastic models have also tended to be "brittle" – they can generate the Zipf distribution, but they are "one-process-fits-all" and cannot generally account for the exceptions to or variations in Zipf that are observed in the data.

Duranton (2002) presents a model based on Grossman and Helpman's (1991) quality-ladder model of growth that produces both Zipf and certain observed variations from Zipf. This model is similar to the model presented here in that it treats urban population as a largely conserved quantity that is redistributed among interconnected cities. In this respect, these models differ sharply from other models that produce good fits, e.g. Gabaix (1999a). Most notably, this property allows these model to produce Zipf-like distributions without assuming that shocks are uncorrelated with city size (Gibrat's law). The model presented here differs from Duranton's in that it is considerably more general, while still having a strong behavioral basis. Indeed, the goal of this paper is to establish a general behavioral framework within which successful economic city size models can be built.

34.1.3 Deviations from Zipf

While the Zipf distribution offers a remarkably good fit for many nations, the fit is imperfect in many cases. In this paper, we will examine three countries that are particularly interesting with regard to their adherence to and deviations from Zipf. These three countries are: the United States, Russia, and France. All three countries provide excellent data on urban agglomerations. The United States represents a relatively good (though significantly imperfect) fit for Zipf, while France and Russia deviate in different ways that may offer lessons applicable to broad classes of countries.

Before attempting to analyze the extent to which cities in different countries do or do not deviate from Zipf, we need to address the definition of a city. In this paper, we are interested in the city as a social and economic phenomenon, rather than as a legal entity. Our unit of analysis is thus not the population within the official city limits, but rather the population of the urban agglomeration of which the legally incorporated city is often only a part.

Consistently defining an urban agglomeration is challenging (Le Gleau et al. 1996), but in the cases we have chosen, it is possible to derive reasonably satisfying

definitions of urban agglomerations. The statistical agencies of both the United States and France have addressed this problem directly by developing various functional definitions of urban agglomerations, while Soviet central planning produced Russian cities that are clearly separated, compact and well defined. We will discuss the specifics of each of these cases in turn.

34.1.3.1   USA

The cities of the United States have generally been regarded as being very nearly Zipf distributed. Because of the sprawling nature of many US cities, and the high daily mobility of the US population, the definition of an urban agglomeration for the US has proven particularly difficult. Over the past several decades, the set of Metropolitan Statistical Areas (MSAs) developed by the US Office of Management and Budget (OMB) and the Census Bureau were the standard measure of urban agglomerations. The MSA definitions had significant limitations, however, and were often influenced by local electoral politics.

In 2003 the OMB released a new series of consistently and objectively defined agglomeration data that it terms Core Based Statistical Areas (CBSAs) or "Metropolitan and Micropolitan" areas (Federal Register 2000). This definition attempts to capture spatial and economic integration with a rigor that had not previously been attempted. The result is a consistently defined set of 922 cities. These cities follow the Zipf distribution fairly closely over a tremendous range: from greater New York City with 18.3 million people down to about the 800th city (Jennings, Louisiana) with a population of about 30,000. Although several of the largest cities are significantly smaller than Zipf would predict, the distribution generally fits a power-law exponent very close to −1.

For convenience in the analysis that follows, we will restrict our data to a subset of the 250 cities with populations over 150,000 (Fig. 34.2). This reduced set of cities looks very much like the full set, displaying a power-law exponent of 1.005.

Comparing the US city size distribution to a pure Zipf distribution for a comparable number of cities and citizens, we find the Zipf assumption misplaces about 15% of the population overall, with an error of 9.7% at the median city.[2]

---

[2] We can produce an objective measure of how well a "constructed" Zipf distribution fits the observed data by dividing the number of people which the Zipf rule misplaces relative to the data (the cumulative error) by the total population of the cities. The cumulative error is calculated as the sum of the absolute values of the errors for each city divided by two (because each citizen which is in the wrong place is also missing from the right place). We will refer to this measure as the total error.

While the overall error is well reflected by this measure, it does not give a sense of how the error is distributed. A sense of this distribution is given by the error at the median city. This is to say that we measure the error for each individual city $((abs(Data_i - Model_i)/2)/Data_i)$ and report the median of these values. This indicates whether the error is concentrated in a few large cities which fit poorly or is distributed throughout the range of the cities. We will refer to this measure as the median error.
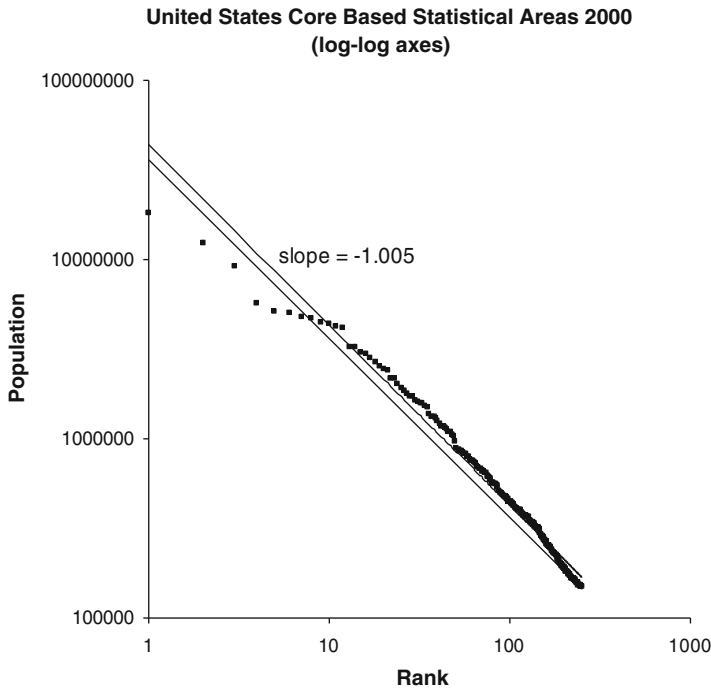
**United States Core Based Statistical Areas 2000
(log-log axes)**

Fig. 34.2 United States core based statistical areas, 2000

34.1.3.2 France

The French National Institute of Statistics and Economic Studies (INSEE) produces a variety of excellent data on French cities using various definitions. These include the municipality (*commune*); the urban pole (*pôle urbain* or *unité urbain*); and the urban area (*Aire urbain*).

One of the most salient features of the French city size distribution is the dominance of the Paris metropolitan area. Of the three ways of defining a city offered by the French statistical agency (INSEE), the "urban pole" definition is the most appropriate for our analysis, but under-represents the size of the Paris metro area. We will use a modified definition of "urban pole" which, following Le Gleau et al. (1996), we will call an "urban center". This revised definition better captures the dominance of Paris in the French urban system.

The urban center data conforms fairly closely to Zipf, displaying an overall power-law exponent of −0.98 (see Fig. 34.3). The primary deviations from Zipf are that Paris is about two and half times the size that the rest of the distribution would predict while the second agglomeration, Marseille-Aix-en-Provence, is about two thirds the size that the distribution would predict. The combination of these two factors makes Paris about seven times as large as France's second city – whereas in a
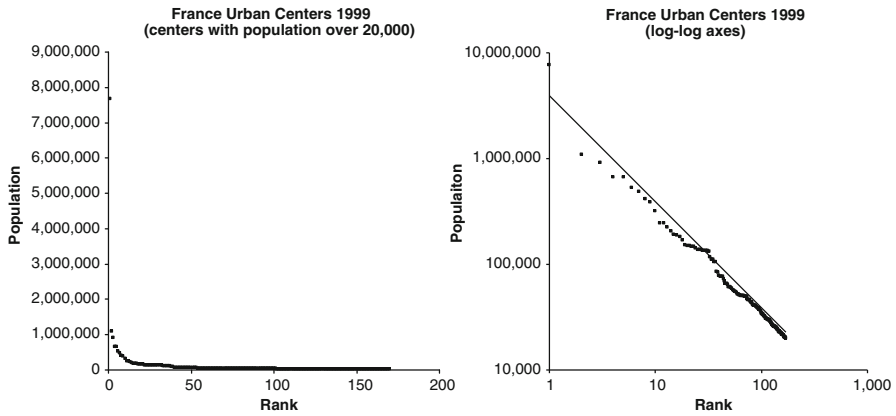
Fig. 34.3 France urban centers, 1999

strict Zipf distribution it would be twice as large. Overall, the Zipf distribution displaces 17% of the French population, but this is largely due to the very poor fit of Paris – the error at the median city is only 7%.

Because France is much less populous than the United States, its urban structure is also much smaller. Whereas the United States has about 900 cities with populations greater than 20,000, France (following the 1999 urban center definition) has only 170 cities above this size.

34.1.3.3 Russia

Unlike the United States and France, which both adhere closely to the Zipf regularity for all but their largest cities, the Russian city size distribution displays a distinct curvature on log-log axes over the entire range of its urban structure (see Fig. 34.4).

The substantially different Russian urban structure is not surprising given the radically different physical, social, and economic environment in which it developed. Much of Russia's urbanization took place during the Soviet period when internal migration was intensely managed by the central government, which pursued objectives such as the extraction of natural resources, the occupation of territory that might be claimed by China, and the movement of industrial production away from the potential front with Western Europe. Policies of forced and incentivized migration, costly investments in infrastructure, and intensive subsidies to far-flung cities in inhospitable locations increased both the number and the size of cities in remote parts of the Soviet Union (Hill and Gaddy 2003). A basic reality of this system, which we will make use of in our model, was that moving down the urban hierarchy was generally easier than moving up it. A person living in Moscow might be assigned a job in a minor industrial center in Siberia, but a person living in that Siberian city would be unlikely to be assigned to Moscow. This system insured that the smaller (and often colder and
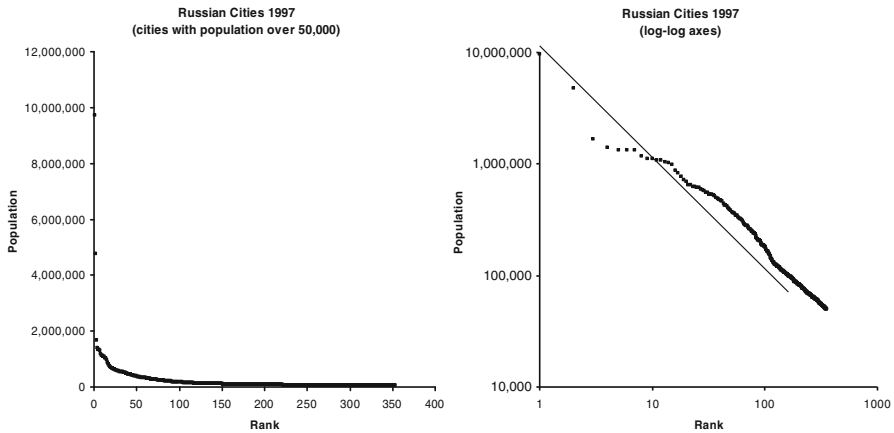
Fig. 34.4 Distribution of Russian city sizes

generally less hospitable) industrial cities of Siberia remained populated in spite of Russian citizens' inclinations to move elsewhere (Hill and Gaddy 2003; Iyer 2003).

Russian urban agglomerations are easier to define than their US and French counterparts because of the way that Soviet planners designed the Russian urban structure (Hill and Gaddy 2003). The desire to spread population over the vast territory of the Russian empire created large distances between cities, while the planned nature of these cities reduced or eliminated urban sprawl in most cases. Because Russian cities tend to be distinct and compact, Russian city population numbers and urban agglomeration numbers tend to coincide, requiring the aggregation of suburbs with central cities only for Moscow and St. Petersburg. The data generated by the Russian census are therefore appropriate for our purpose without adjustment beyond the agglomeration of these suburbs.

The overall best fit power-law for this data has an exponent of −0.92 – a number close enough to unity that some authors have failed to remark on it. Our measure of total error indicates that the fit between the Russian distribution and the Zipf distribution is similar to that for the US and France, misplacing 16% of the population (as compared to 15% and 17% respectively), but this apparent similarity is misleading. This shows up in a median error figure of 17% (as compared to about 10% for the US and 7% for France). While distributions for the US and France are generally Zipf like, with departures in the largest cities, the Russian distribution is distinctly curved as shown in Fig. 34.5.

We can demonstrate this curvature by dividing the Russian city distribution into two parts and examining the exponents of the best-fit power-law that describes each part, measuring the power law exponent for cities larger than 500,000 separately from those between 500,000 and 100,000. These sets of cities display two distinct exponents. The upper part of the curve has a slope of −0.68 while the lower part has a slope of −1.19. These slopes are significantly different with p<<0.001. Similar tests on data from the US and France yield slopes that are not significantly different.
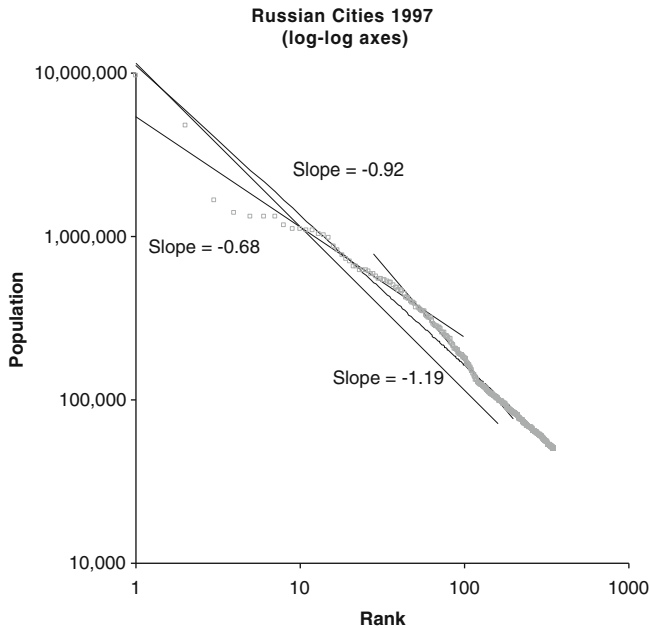
Fig. 34.5 Curvature of the Russian city size distribution

By inspecting the graph (Fig. 34.5) we can see that our cut off of 500,000 between the two groups is an arbitrary one, and that the distribution of cities larger than 100,000 is better described by a curve that is concave toward the origin. In this sense, the Russian distribution departs from the Zipf distribution for all of the 161 cities in this range.

34.2   A Simple, Abstract Model: Jars and Beans

34.2.1   Model Description

In the sections that follow, we will attempt to explain both the tendency of urban systems to approximate Zipf, and the reasons why various countries depart from it, by constructing a model that is as simple as possible while capturing the essential features of the systems in question.

We begin with an abstract model that produces remarkably good agreement with real city size distributions. This model is designed to explore the way in which power-law distributions can emerge from systems involving stochastic exchange. Because the abstract model does not itself contain sufficient detail to capture plausible urban dynamics, we describe it in terms of "jars" (rather than cities) exchanging

"beans" (rather than citizens). In the next section, we will extend the model in such a way that it demonstrates a plausible relationship to social and economic realities.

The rules of the abstract model are simple. The model begins with some number of jars, each of which contains some number of beans. The jars interact in random pairings. In each interaction, the jars exchange some number of beans ("the bet") equal to half of the beans in the smaller jar. In the base case, both jars have an equal probability of winning the bet. Once the winner is determined, the beans are exchanged and a new random pairing of two different jars is made. There is a floor size of 1 bean. If a jar of size 1 loses a bet, nothing happens and it remains at size 1. If it wins a bet, it wins a whole bean (rather than half a bean).

One important feature of this model is that it assumes that urban population is conserved, unlike other models (e.g. Gabaix 1999a; Fujita et al. 1999) that assume people freely enter and leave the urban system. Our assumption of conservation fits with empirical evidence that once people have migrated to a city and have traded their rural skills for urban ones, they tend to remain in the urban system – migrating from one city to another in search of opportunities, but seldom returning to live in the hinterlands. In our simple abstract model, this is reflected in a strict conservation law: beans are neither created nor destroyed, they simply move from jar to jar.

The model also differs from other stochastic models (typified by Gabaix (1999a)) by not needing to assume that the growth rates of cities are independent. These previous models generally depend on a Gibrat process for their results, in which cities grow (or shrink) by random amounts that are uncorrelated but share a common mean. In our model, growth rates *are* correlated (one city's gain is another city's loss). We believe that this is a more plausible assumption for modeling city size dynamics. We also assume that growth rates depend on city size. When a small city faces a larger city, it faces a gain or loss of half its size, whereas the larger city faces a gain or loss that comprises a smaller fraction of its population. Small cities, therefore, face greater size volatility than large ones, a fact that also coincides with real world observation (Gabaix 1999b).

34.2.2   Results from the Abstract Model

Although the model is very simple, it can produce statistically robust Zipf distributions as well as some interesting variations on the distribution. If the model is run with the appropriate number of beans[3] for the given number of jars, it will approach the Zipf distribution regardless of the initial distribution of the beans between jars. Initializing the model with "too many" beans – more beans than would be required

---

[3] From the definition of the distribution, it follows that a certain number of jars requires a certain number of beans to fill the distribution. When the floor size (the size of the smallest jar) is one bean, the largest jar should contain a number of beans equal to the number of jars. The sizes of all the jars between the largest and the smallest are then given by the rank/size rule, rounding to the nearest whole bean. For example, for 100 jars, 516 beans are required to fill the distribution.

to fill a Zipf distribution for the given number of jars – produces instability in the top of the distribution with large fluctuations in the sizes of the largest jars, with the excess beans tending to float among the top few jars. Radical overfilling of the distribution tends to produce "jamming" at the top, where the largest jar ends up with the majority of the excess beans. Initializing the model instead with "too few" beans produces a curvature of the distribution, maintaining the power-law exponent in the lower tail and progressively lowering it in the upper tail. As we will see below, these two related results have important parallels to the real world deviations from Zipf's law observed in the cases of France and Russia.

The model is reasonably robust to changes in key parameters. For example, while it is important that the "bet" be related to the size of the smaller jar in a given interaction, the exact proportion used generally affects only the speed with which the system approaches equilibrium, not the nature of that equilibrium.[4]

We can make a first analogy from this abstract model to urban dynamics by thinking of the jars as cities and the beans as groups of citizens. Each bean represents the number of citizens in the smallest city (size = 1) in the sample. Actual population data can therefore by translated for use in the jars and beans model by dividing the total population of the urban system by the size of the smallest city in the system. This translation means that the units of exchange in the model are the size of the smallest city. This coarse assumption leads to discontinuities in the lower tail of our graphs, but it produces some interesting initial results and we will subsequently refine them.

Population figures for the United States, inserted into this simple model, produce a distribution that bears a noticeable resemblance to real data. In the year 2000, according to the Census Bureau data discussed above, the US had 250 cities with a population larger than 150,000 and these cities were home to a total of 220,227,293 people. We translate this for use in the jars and beans model by dividing the total population by the size of the smallest city (150,000), giving 1,468 beans in total. We can then get a first approximation of the US urban distribution by initializing the model with 250 jars and 1,468 beans (initially distributed randomly). Running the model with these parameters gives a fit that is quite suggestive.

Figure 34.6 shows the discretized version of the US data compared to output from 100 runs of the simple model using 250 jars and 1,468 beans. The heavier, central line on the graph indicates the median size for the city of each rank across all model runs; the lighter lines represent a 90% confidence interval around this median. The US data do not fit precisely within this envelope, but it is not far off. The gray circles in the figure represent one of the hundred sample runs that is particularly suggestive. We will return for a more careful analysis with a more complex model in the next section of the paper.

---

[4] Extremely small bet sizes can begin to cause the lower tail of the distribution to collapse. This does not occur with a bet sizes close to 50% of the smaller jar (the setting used throughout this section of the paper). We will discuss the sensitivity of the model to bet sizes in more detail in the next section of the paper.
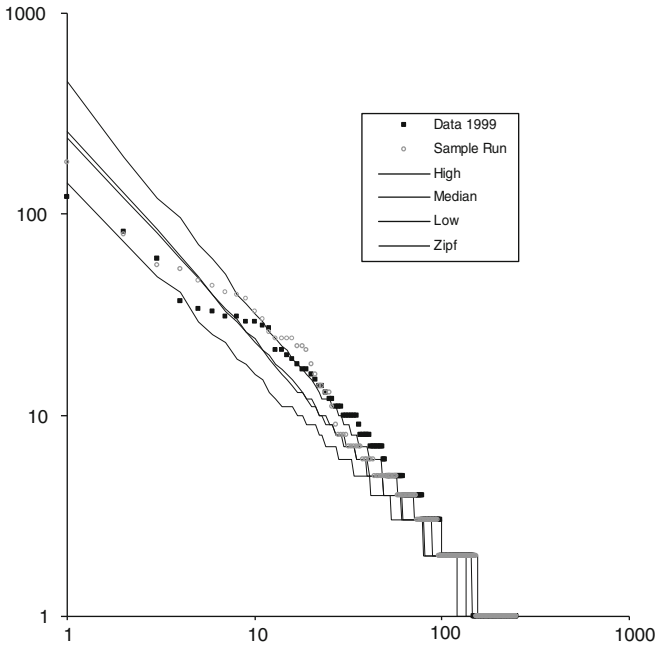
Fig. 34.6 Simple model output compared with discretized US data

Conducting the same exercise for France produces similarly provocative (although again not entirely realistic) results. Using our definition of an urban center, France has 170 cities with populations larger than 20,000 that collectively contain 22,386,598 people. We thus initialize the model with 170 jars and 1,119 beans (again distributed randomly).

Again, we see that the real data generally fit within the range of model results (Fig. 34.7). We can see from the representative sample run (grey dots) that in a case where the first two cities are of the proper size, the fit of the rest of the distribution is also very close. Although the simple model does not fully predict the primacy of Paris in the French urban system, the median model run does reflect an increase in slope in the top three or four positions. This is consistent with the notion that a small urban system with a relatively large population will tend to see disproportionately large cities at the top of its range.

Finally, we can obtain intriguing results for Russia by applying the model with a slight variation. In 1997 Russia had 161 cities with populations over 100,000 that collectively contained 70,282,100 people. This yields 703 beans in 161 jars.

Initializing the model with these values gives us a distribution that is concave toward the origin on log-log axes, but which has a somewhat different shape than we see in the data from Russia. If, however, we approximate Soviet era restrictions on internal migration by introducing a slight bias into the process (simulating the asymmetry in difficulty between moving up and moving down the urban
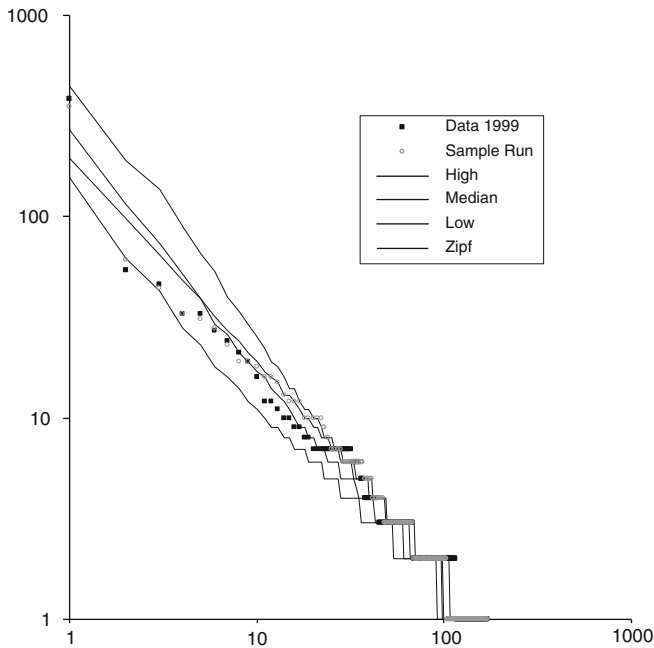
Fig. 34.7 Simple model output compared with discretized France data

hierarchy by giving the smaller city in each pairwise interaction a small advantage) the shape of the distribution comes to match the Russian case much more closely, as shown in Fig. 34.8.

34.2.3   *Limitations of the Abstract Model*

While the abstract model offers a simple mechanism that is capable of generating distributions resembling real city size distributions, it suffers from several serious limitations in interpretation. Although this model incorporates more realistic assumptions (such as correlated growth rates) than other stochastic models have employed, the dynamics of the model still bear little resemblance to those of real cities: cities do not engage in "tournaments" of flipping coins for half of their citizens. In addition, the floor assumption of the abstract model provides a subsidy to the smallest jars – in each interaction they stand to either remain unchanged or to double their number of beans. This mechanism tends to move beans from the upper parts of the distribution into the lower tail in a way that has no clear analog in the dynamics of urban migration.

Also, the dynamics of the simple model involve a high churn rate, with cities changing rapidly changing their rank within the distribution and the largest individual cities varying tremendously in size over time. In the time scale that is required
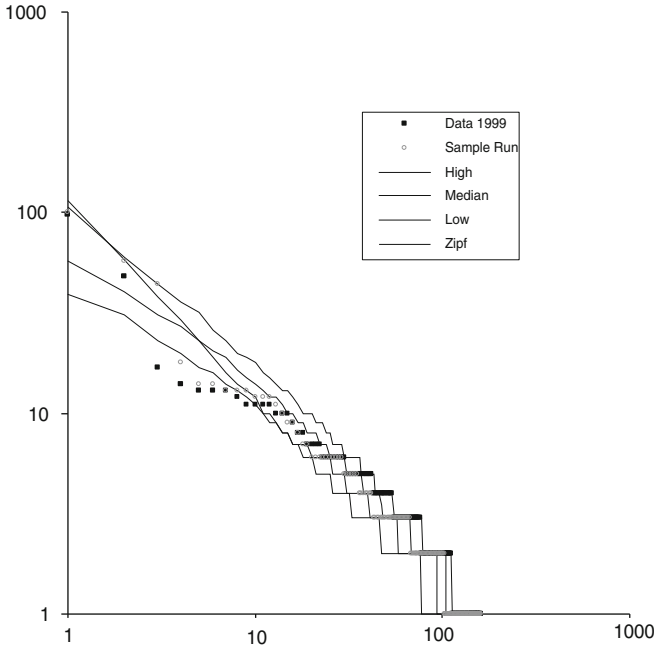
Fig. 34.8 Simple model output compared with discretized Russia data

to achieve the power-law distribution, Chicago might change places with Peoria
several times. This unrealistic dynamic highlights the fact that the abstract model
has no place in it for differences in site suitability. Some sites (natural ports, for
example) are inherently better than others for large cities, and any plausible model
of urban dynamics should be able to reflect this fact.

34.3   A Richer Model: Cities and Citizens

34.3.1   Model Overview

To address these deficiencies, we will now introduce a richer model that comes
closer to representing real urban dynamics. This model preserves and improves
upon many of the desirable qualities of the abstract model while remedying some of
its shortcomings. The richer model relies on the notion that a city has a short-term
equilibrium size that balances economies of agglomeration (reasons to move into
the city) with diseconomies of congestion (reasons to move out). A city can be
thought of as being oversized if it moves above this equilibrium value and under-
sized if it moves below it. This short-term equilibrium is subject to shocks that result
from the bounded rationality of citizens. The equilibrium reacts to these shocks over

the longer term according to a lagged adjustment mechanism. Finally, the model introduces a conception of "core size" – a size below which it is not economically feasible for a given city to shrink.

34.3.2 Bounded Rationality

The concept of bounded rationality underlies the exchange mechanism in the abstract model and provides us with guidance in refining it in terms of both the size of and bias in exchanges. We can see the central role of imperfect information in the model by assuming (temporarily) that all cities are at their equilibrium sizes. In this case, with each city at its optimal size, perfectly informed and rational agents would have no incentive to move from one city to another because any move would leave their home city underfilled and their new city overfilled – making the mover worse off. Under the assumption of perfect information, any distribution of city sizes in which all cities are in local equilibrium would be stable indefinitely.

The citizens in our model, however, have imperfect information and bounded rationality. Some citizens, therefore, will move from city to city *even at* an "equilibrium" distribution of sizes. People are more likely to move from a more crowded city to a less crowded city, but the reverse is also possible. The size of the exchange between cities, therefore, is a parameter of the model. It represents the degree to which the rationality of the citizens is bounded – the percentage of the citizenry that will move between two equally attractive cities because of imperfect information (which we are modeling only in the abstract). With full information and no bounds on rationality, the exchange between two cities at their equilibrium sizes would always be zero. In the extended model we present below, the expected value of the exchange is zero, but the actual exchange amount varies symmetrically around zero. In this sense, the exchange mechanism is "unbiased".[5] This principle of unbiased exchange differs from the abstract (jars and beans) model discussed above. In that model, the floor mechanism provides a significant subsidy to small jars. With 100 Zipf distributed jars, an exchange size of 50% of the smaller jar, and a floor of one, about 1/3 of the jars face positive expected returns – and the rest face negative expected returns.

As with the abstract model, the primary effect of changing the size of the bet in this richer model is simply to change the speed with which the system moves. However, when the bet is small enough, very few small cities face positive expected returns. Over the long run, this leads the lower tail of the distribution generated by the model to sag (i.e. to bend toward the origin) and produces long oscillations in the extent of this sagging. These features are not observed in real data. A closer

---

[5] When the exchange amount is decreased to 1% of the smaller jar (as it is in the runs of the model that follows), only the single smallest jar can be expected to be within 1% of its floor, and the bias that it introduces into the system is vanishingly small.

match to empirical reality can be achieved by introducing a small amount of growth into the system. When all cities grow by a tiny amount each round, the lower tail restabilizes near a slope of −1. The amount of growth does not need to be carefully tuned to achieve this result. The growth rate needs only to be large enough to keep the tail from sagging, and small enough that the system can "digest" the new citizens. Within that range, the growth rate can vary by an order of magnitude without significant impact on model output.

This assumption of growth is consistent with the real world, where all of the world's major urban systems are still growing. This is most apparent in developing countries, which are experiencing both population growth and urbanization. It is also true, however, of OECD countries such as France, where urbanization continues even as population has stabilized (Julien 2001a).

34.3.3   Lagged Adjustment

Cities' equilibrium sizes adapt to the shocks imposed by the bounded rationality of their citizens through a lagged adjustment mechanism. If a city grows above its equilibrium size, it will become congested in the short run. If it remains congested for long enough, however, the city will adapt. Firms will move in to hire idle workers. New housing, roads and facilities will be built. Once these things happen, the city can comfortably accommodate more people than it did before – its equilibrium size has increased. Similarly, if citizens move out and stay out for long enough, firms will leave and infrastructure will deteriorate, leaving the city able to comfortably accommodate fewer people than it once could.

Adding an adjustment lag does not change the dynamics of the model, but does impact the rate at which individual cities change size over time and therefore the rate at which the distribution changes. Because the parameters of this mechanism only influence the speed with which the model changes (and we are not attempting to calibrate the model to real time), we will not dwell on the lagged adjustment mechanism here. Any mechanism that retains the unbiased quality of the exchange system from the simple model, and that does not introduce excessive noise into the model will produce similar results.

34.3.4   Inherent Suitability

A further requirement for the richer model is to account for the influence of geographic suitability and the persistence of great cities. We accomplish this by positing "core size", determined according to more conventional economic logic, which is one component of observed size.

We begin with the assumption that only some fraction of the population of a city is tied to the city's specific geographic location. Chicago, for instance, is in a unique location to serve as a port for a huge section of the American Midwest. Many of the

jobs in Chicago need to be located exactly where they are geographically – at the base of Lake Michigan. Many other jobs in Chicago, however, do not have to be in that location. But they do have to be somewhere. We thus divide the population of a city into a "core" population, which is dependent on the city's geographic location (and is subject to more or less standard microeconomic rules for its size), and a floating population, which is subject to the mechanisms of the model.

A recurring problem for theorists of city sizes has been that models containing appropriate economic content (e.g. Fujita et al. 1999) tend to predict distributions that look quite different from those that are actually observed. The model presented here solves this problem and dovetails nicely with such models by freeing them from the need to predict a Zipf like distribution. A model like that of Fujita et al. (1999) is probably well suited for predicting the core sizes of cites. These core sizes should be much more readily subject to "rational" analysis. The core sizes, however, are not the only component of the observed distribution. The sizes we observe are based on the sum of the core size and the size of the floating population that can potentially live elsewhere. We will equate the "core" size of a city to its "floor" (i.e. minimum size) in the model.

Remarkably, the presence of some cities with higher floors (larger core sizes) does not change the basic dynamics of the model. It still produces Zipf distributions and the aforementioned characteristic departures from Zipf. However, the cities with higher floors tend to stay in the upper part of the distribution, thus reflecting much more realistically the persistence of major cities that we observe in the real world.

An analogy to a cake with icing is a useful way to visualize the relationship between the core and observed distributions of city sizes. The core distribution is the cake, while the floating population is the icing. All that we observe in city size data is the height of the top of the icing. While the cake of the core size distribution might be rather lumpy and vary depending on economic and geographic structure, the icing of the floating population flows smoothly over the cake and finds its level. In the case of cities, the attractor is not flat – as it is in the case of a physical cake – but rather follows the shape of the Zipf distribution and its related departures as outlined above.

Because this study is concerned with the overall shape of the various city size distributions, it is sufficient to note that adding heterogeneous core sizes does not change the distributions that emerge from the model. The simulations that follow will use uniform core sizes unless otherwise stated, with the core size being equal to the size of the smallest city in the system. The results would not be changed if a more complex or dynamic core distribution were used.

34.3.5 Results from the Richer Model

34.3.5.1 USA

This richer model produces a fit for United States core-based statistical area data that is significantly *better* than the Zipf approximation. The only significant
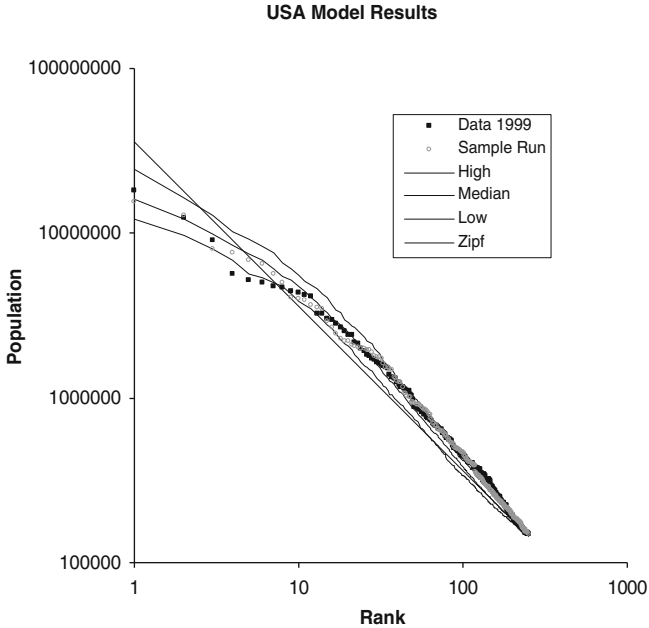
USA Model Results

Fig. 34.9 USA model results

parameters in this model are the number of cities with populations over 150,000 (250 of them), the number of people in these cities (220,227,293 in total), the rate by which each city grows at the end of every round, and the fraction of the smaller city which will serve as the exchange amount in interactions. The first two (cities and citizens) are given by the data. The results are insensitive to the growth rate over a broad range of values (roughly and order of magnitude). The size of exchanges alters the degree of variance between runs, but does not have a noticeable impact on their median outcome. The model has no other free parameters.

We begin the simulation of the United States city size distribution with 250 cities and a reduced population of 50 million citizens (about 1/5 of the actual population) distributed evenly between the cities. The initial population size is not significant so long as it is small enough to allow the model to approach equilibrium before the full population is reached. We run the simulation forward with each city growing by a small amount (1/20,000th) at the end of each round, stopping when the population reaches the year 2000 total urban population of 220,227,293 (Fig. 34.9).

For the sake of simplicity, we begin these simulations with a uniform distribution and with a fixed number of cities, although in the real world the urban system is always in the neighborhood of the Zipf distribution (with the number of cities increasing along with their populations). Such a growth pattern is supported by history (Zipf 1949; Pumain 2004) and emerges from certain theoretical formulations (Simon 1957; Gabaix 1999a; Axtell and Florida 2001). When the initial state is close to Zipf, the growth rate becomes much less critical. It needs to be great enough
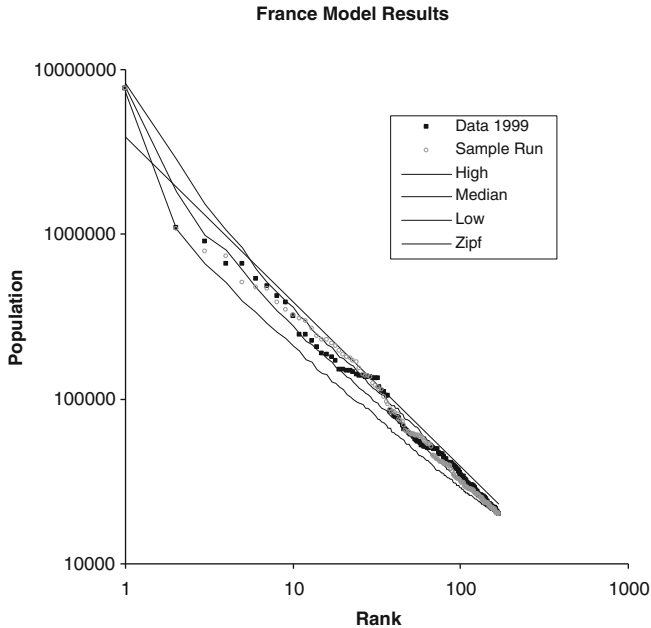
France Model Results

Fig. 34.10 France model results

to prevent the collapse of the lower tail, but more rapid growth is not a problem because the system does not need to produce major structural changes.

34.3.5.2 France

As discussed above, France is generally characterized by a Zipf distribution with Paris considerably larger than the distribution would predict. Although the abstract model was capable of producing results that were consistent with French data, this occurred in only a small fraction of model runs. The richer model performs considerably better in this respect, although it requires a somewhat more complex assumption about growth[6] (see Appendix). The distributions are shown in Fig. 34.10.

---

[6] The previously discussed issue with collapse of the lower tail in the absence of growth is particularly problematic in this case because of the very large size of Paris. Starting from a uniform distribution of city sizes, any growth rate large enough to prevent the collapse of the lower causes the population to reach its target size before the model has had time to grow Paris to its full size. We therefore begin the model with approximately 90% of the total population of France and run it forward until Paris has reached 90% of its actual population. We then introduce growth at the same rate of 1/20,000th per iteration used for the US simulation and run it until the model population is equal to the French population.
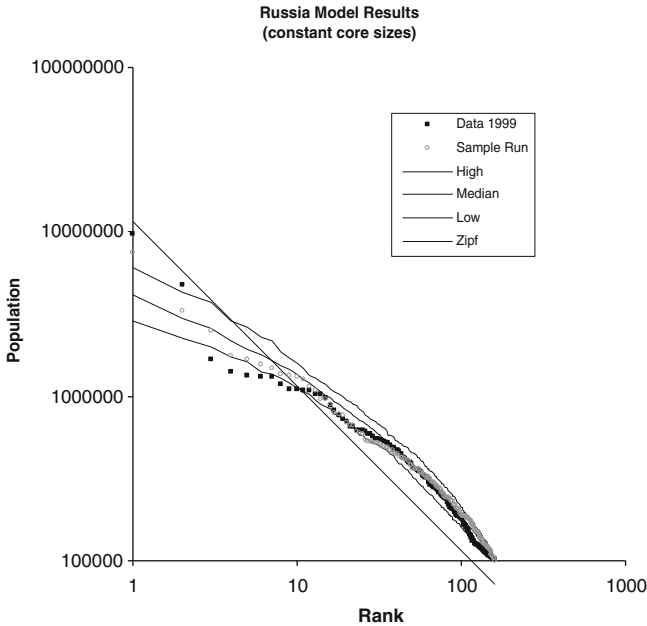
Fig. 34.11 Russia model results with constant core sizes

Whereas Zipf produced a total error of 17% and a median error of 7%, the model produces a total error of 5.3% and a median error of 3.3%.

34.3.5.3 Russia

To simulate Russia, we initialize the model with 161 cities, a population of 70,282,100 in these cities and a floor of 100,000 (the size of the 161st city). As with the simple model, we introduce a bias into the migration probability to simulate the effects of internal movement restrictions. The degree of this bias is a free parameter of the model, which we calibrate to 0.0025 in favor of the smaller city in each pairwise interaction.[7] The slight bias toward smaller cities eliminates the tendency of the lower tail of the distribution to collapse and makes the model behavior invariant in the presence or absence of population growth (Fig. 34.11).

---

[7] Given that the model does not attempt to represent the urban system in actual space and time, it is not possible to calculate this movement bias using actual data. Because it is the only free parameter in the model, however, we can calibrate it by comparing model results to the observed data. We obtain a good fit by assuming a bias of 0.25% in favor of the smaller city in each pair-wise interaction. That is to say that, in each interaction, the probability of the larger city receiving the migration is 49.75% while the probability of the smaller city receiving the migration is 50.25%.
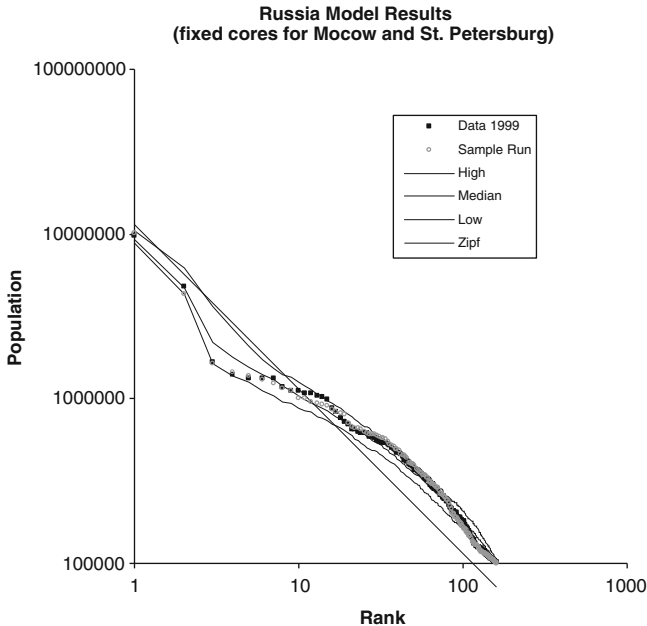
Fig. 34.12 Russia model results with larger cores for Moscow and St. Petersburg

When run with these parameters, the model captures the basic shape of the Russian city size distribution, but misses the primacy of Moscow and St. Petersburg. These cities have each played unique roles in Russia's economic and political history, serving as capitals of highly centralized political systems under both the Czars and the Soviet system. St. Petersburg is also unique in serving as European Russia's only ice-free port. The continuing pressure of internal immigration on these cities – even in the face of falling population in Russia generally (Iyer 2003), indicates that these cities remain at or below their equilibrium size in the collective mind of the Russian people. We incorporate the unique economic and geographic appeal of these two cities by assigning them core sizes that are 90% of their observed sizes, while leaving the cores of the remaining cities uniform at 100,000 people.

We observed earlier that introducing heterogeneous floor sizes alters the stability of individual cities but does not change the shape of the overall distribution *unless* floors are set so high as to make a city "protrude" from the distribution. In this case, we are conjecturing that political and geographic forces have caused the core sizes of Moscow and St. Petersburg to protrude from the Russian city size distribution.

When we incorporate these larger core sizes for Moscow and St. Petersburg into the model, it produces an excellent fit for the data (Fig. 34.12). Overall, the median model run misplaces only 3.25% of the population. This is much better than Zipf, which displaces 12.5%. The error at the median city similarly drops yet further to 2.5% as compared to 16.5% for Zipf.

34.3.6   Limitations

The richer model presented above displays a good deal of success in reproducing the distribution of city sizes in the United States, France, and Russia, but does have some limitations.

While the model can predict the overall shape of the urban distribution for various countries, it does not predict the movements of particular cities within that distribution. In order to control volatility of city size in the model, we have employed the "core size" concept – but we do not model explicitly how such core sizes evolve. The model produces similar distributions over an extremely broad range of possible core size including, we believe (but do not show here), core sizes that are compatible with observed levels of volatility.

A second, related, limitation of the model is that its current formulation does not lend itself to calibration to real time scales. Real urban systems generally expand simultaneously in both population and number of cities, whereas we hold the number of cities fixed. We believe that this assumption, although unrealistic in the long term, can yield insights in the shorter term by keeping the model simple enough for ready analysis and insight.

A third limitation is that the model uses a simple but highly unrealistic interaction network. Cities in the model interact randomly, regardless of their size or location – indeed, location is not represented in the model at all. We do not explore here the sensitivity of the model to different interaction regimes.

34.4   Discussion

34.4.1   Implications for Developing Nation Megacities

One of the more interesting and policy-relevant insights generated by the model is that the primacy of Paris (and, by extension, other disproportionately large capitals) might have more to do with the number of small cities than it does with the nature of the large city. Previous efforts to explain urban primacy (e.g. Ades and Glaeser 1995) have tended to focus on the political economy of the capital as the reason that it grows disproportionately large. These theories would attribute the massive size of Paris to the highly centralized nature of the French political system and the fact that it is "the capital of everything" including politics, finance and culture, for the nation. This contrasts with the United States where the political capital (Washington) is different from the finance capital (New York and to some extent Chicago) and the cultural capital (which one might argue is split between New York and Los Angeles). Our model allows for such theories – we invoke this kind of reasoning to explain the size of Moscow and St. Petersburg in Russia – but the model suggests that this kind of explanation may not be required to explain the size of Paris. While the central role that Paris plays in French political, economic, and cultural life undoubtedly

does endow it with a substantial core size, it is not clear that this role requires the city to be as large as it actually is.

The stylized result from the model is that a country with a large population and relatively few cities will tend to produce a Zipf distributed population in all but the largest city (or few cities) with the "overflow" population collecting at the top of the distribution (here, in Paris). Our framework suggests more generally that there is a relationship between number of cities and number of people in an urban system. This relationship has important implications for urban planning in the developing world. Our analysis presents a reason to expect the emergence of megacities such as Sao Paulo in Brazil, Dhaka in Bangladesh, and Jakarta in Indonesia. These countries generally have highly centralized governments and severely constrained capital availability. These factors make it very difficult for their urban systems to expand in terms of number of cities at a rate that bears any resemblance to their rates of population growth and urbanization. Developing nations are therefore left with a small number of cities and a large urban population. While a person's first move from rural to urban life may be from the countryside to a nearby city (a tendency that would tend toward balanced urban growth), our model suggests that the next step of inter-urban migration will tend to concentrate the urban population.

Megacities create numerous policy challenges, involving growth management and the provision of adequate infrastructure for a rapidly growing population. Failure to meet these challenges can create disastrous situations in the areas of environmental protection, public health, and human development and can lead to social unrest, political instability and violence (Bugliarello 1999).

The model further suggests that efforts to encourage migration from the first tier cities to middle sized cities are not likely to succeed over the long term. A government hoping to stem the growth of a primate city would do better to focus limited resources on providing the infrastructure and economic base that would allow large towns to become full participants in the urban system – thus expanding the number of cities and thereby reducing pressure on the capital.

34.4.2 Implications for Russian Urban Structure

The model suggests that two factors have played a role in creating the odd distribution currently observed in the Russian urban structure: a large urban system relative to its population and movement restrictions that have historically biased movements toward smaller cities. Unlike the urban structures of the US and France, the Russian urban structure was not created by free mobility and free markets. Soviet central planning created, instead "a structure of production – location, capital, employment, materials, energy use, etc. without any regard for economic opportunity costs, in an environment free of economic valuation" (Ericson 1999).

The result of this non-market resource allocation was an extensive urban structure that post-Soviet leaders have continued to work hard to preserve through subsidies and other measures. For a host of ideological and security related reasons,

Soviet central planners aimed for relatively even dispersal of cities of fairly uniform size while at the same time creating a highly centralized system of power (Demko and Fuchs 1984). These factors contributed heavily to the creation of the odd urban structure that we see today.

One of the major Soviet era policies used to maintain this sprawling urban structure was a system of permits that were required for one to move from the hinterlands into an industrial center, and from a smaller industrial center to a larger one. This policy may be likened to biasing migration toward the smaller city in our model. While these policies are officially no longer in place since the fall of the Soviet Union, traces of them remain – particularly with regard to migration into Moscow and St. Petersburg. President Putin remained committed to avoiding Siberian "ghost towns" at almost any cost and many subsidies to these towns are in place long after the end of the Soviet system (Gaddy and Ickes 2002).

While the current form of our model is not useful for estimating the speed with which the size distribution might change with the relaxation of these restrictions and subsidies, we can use it to speculate about their general nature. We expect that unmanaged movement would lead to continued growth pressure on Moscow and St. Petersburg. We would further expect strong growth in a small handful of second tier industrial cities with current populations between 1 and 1.5 million (Novosibirsk is a typical example). However, we would expect this growth to extend to only three or four such cities, with the vast majority of cities with population between 100,000 and 1.5 million experiencing a prolonged period of population decline.

34.5   Conclusions

Our adaptive agent framework has allowed us to design and explore a framework for understanding city size distributions which, in spite of its extreme simplicity, is able to generate close approximations of the actual city size distributions for the US, France, and Russia. Although simple, this model is hard to examine analytically because of the high degree of interaction among its parts. Previous attempts to explain the Zipf distribution have, in general, gained analytical tractability by assuming independence of the growth rates of cities. While it is possible to generate the Zipf distribution using such assumptions (Gabaix 1999a) it is hard to imagine how the departures that we have reproduced could be derived in such a setting, and independent growth rates seem implausible for real city interactions. Our agent-based simulation methodology allows us to drop this restrictive assumption.

The use of this approach has made it possible for us to make real progress in understanding a phenomenon that has puzzled economists, geographers and others for over 50 years. Our model establishes a basis for moving beyond the assignment of mystical significance to the Zipf distribution of city sizes and allows us to see city size distributions as the result of straightforward behavioral rules. We can further

understand Zipf as only a special case of city size distributions and see deviations from Zipf not as noise or error of some sort, but as the products of differing policies and situations.

Appendix: City Definition for France

The French National Institute of Statistics and Economic Studies (INSEE) produces a variety of excellent data on French cities using various definitions. These include the municipality (*commune*); the urban pole (*pôle urbain* or *unité urbain*); and the urban area (*Aire urbain*).

Of these three ways of defining a city, the first and third are inappropriate for use in this analysis. The municipality definition is not useful because most major cities are composed of many municipalities. The municipality of Paris, for example, had a population of only about 2.1 million people in 1999. The urban pole of Paris, in contrast, was composed of 396 such municipalities and was home to over 9.6 million people (Chavouet and Fanouillet 2000). While the legal definition of a municipality reflects historical and administrative realities, it tells us little about the urban agglomerations that we are studying.

Where the city as municipality definition is too restrictive, the city as urban area definition seems to be too broad. French urban areas are defined as those areas where at least 40% of the workers commute into an urban center that employs at least 5,000 people (INSEE 2004). These areas can be very large, often many times the area of the urban pole. A major problem with this definition for our purposes is that this surrounding area mixes people who commute into the city center with people whose social and economic lives are not integrated with the city. This commuting based definition also creates the impression of rapid growth for many cities, not because the cities have changed significantly, but because French commuting patterns have been changing, with workers traveling increasing distances to work (Julien 2001b). French cities have therefore been expanding their areas of influence more rapidly than they have been growing in terms of employment, built area, or other measures of city size (Julien 2001a).

The French definition of an urban pole strikes something of a balance between these two definitions. An urban pole is defined as a collection of contiguous communes in which more than half of the population lives in an area where buildings are separated by no more than 200 m. This definition is thus a reasonably close approximation of the built up area of the city. However, because this definition includes whole communes that are only partly urbanized, it tends to over count the urban population at the edges of cities. Because the circumference of a circle increases more slowly than its area, this bias tends to inflate the size of smaller cities.

In an effort to avoid this problem, we adopt a slightly more restrictive definition of a French city, that we will call an "urban center". Our definition follows the spirit of the one described by Le Gleau et al. (1996) while adapting it to better
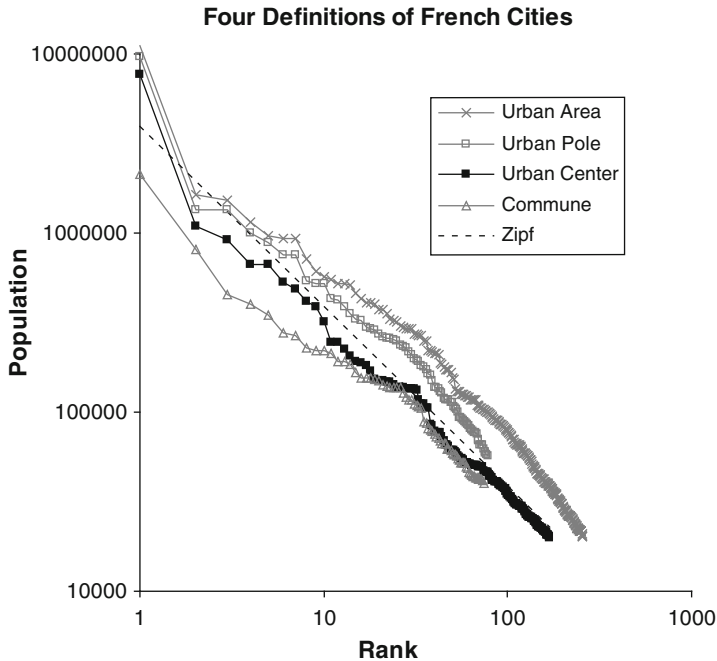
Fig. 34.13 Four definitions of French city sizes

capture the dominance of Paris in the French urban system. Le Gleau defines an urban center such that, if a single commune within an urban pole contains more than half of the pole's population, then this commune is the urban center. If the central commune contains less than half of the population of the pole, then it is agglomerated with the other communes of the pole that have at least half of the population of the largest commune. This definition has the effect of making the urban centers of France appear very nearly Zipf distributed (Fig. 34.13) – but it makes little sense as a definition of a city. Most notably, the central commune of Paris is much larger than any of the other 395 communes that make up the Parisian urban pole. This means that, by Le Gleau's definition, the urban center of Paris is represented by only this one commune, putting its size at 2.1 million people (as compared to 9.6 million in the urban pole).

We retain Le Gleau's concept of omitting the fringe areas by changing the criteria for agglomerating secondary communes, but refine it to avoid distorting large cities (particularly Paris). Under our definition, we agglomerate all of the communes in the pole that have a population greater than 20,000 people. Because communes tend to be of roughly uniform size, this is a reasonable proxy for density. We choose the number 20,000 because it is also the minimum size of a city in our dataset. Thus, any commune within an urban pole that would qualify as a city in its own right by virtue of its population of 20,000 is agglomerated into the

urban center. This definition eliminates the inflation of the urban periphery that is present in the urban pole definition while retaining the basic idea of a city as a contiguous built-up area. Our analysis uses this definition of a French urban center.

References

Ades, A. F., & Glaeser, E. L. (1995). Trade and circuses: Explaining urban giants. *Quarterly Journal of Economics, 110*(1), 195–227.

Axtell, R. L., & Florida, R. (2001). *Zipf's law of city sizes: a microeconomic explanation* (Working Paper). Washington, DC: The Brookings Institution. www.brook.edu/dynamics/papers/cities

Bugliarello, G. (1999). Megacities and the developing world. *The Bridge, 29*(4), 19–26.

Chavouet, J. M., & Fanouillet, J. C. (2000). *Forte extensions des villes entre 1990 et 1999* (INSEE Paper #707).

Christaller, W. (1933). *Die zentralen Orte in Süddeutschland*. Jena: Gustav Fischer (Charlisle W. Baskin, Trans. in part)). (1966). *Central places in Southern Germany* . Englewood Cliffs: Prentice Hall.

Cuberes, D. (2004). *The rise and decline of cities*. Unpublished manuscript. http://economics. uchicago.edu/download/cities3.pdf

Demko, G. J., & Fuchs, R. J. (1984). Urban policy and settlement system change in the USSR, 1897–1979. In D. Harris, G. J. Demko, & R. J. Fuchs (Eds.), *Geographical studies on the Soviet Union: Essays in honor of Chauncy*. Chicago: Department of Geography, University of Chicago.

Duranton, G. (2002). *City size distributions as a consequence of the growth process* (CEPR Discussion Paper 3577).

Ericson, R. (1999). The structural barrier to transition hidden in input-output tables of centrally planned economies. *Economic Systems, 23*(3), 199–244.

Federal Register. (2000). Vol. 65, No. 249, 27 Dec 2000.

Fujita, M., & Mori, T. (1997). Structural stability and evolution of urban systems. *Regional Science and Urban Economics, 24*(4–5), 399–442.

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: Cities, regions and international trade*. Cambridge, MA: The MIT Press.

Gabaix, X. (1999a). Zipf's law for cities: An explanation. *Quarterly Journal of Economics, 114*, 739–767.

Gabaix, X. (1999b). Zipf's law and the growth of cities. *American Economic Review Papers and Proceedings, 89*(2), 129–132.

Gabaix, X., & Ioannides, Y. M. (2003). The evolution of city size distributions. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of urban and regional economics* (Cities and geography, Vol. IV). Amsterdam: North-Holland.

Gaddy, C. G., & Ickes, B. W. (2002). *Russia's virtual economy*. Washington, DC: Brookings Institution Press.

Gibrat, R. (1931). *Les inégalités économiques; applications: aux inégalités des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d'une loi nouvelle, la loi de l'effet proportionnel.* Paris: Librairie du Recueil Sirey.

Grossman, G. M., & Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies, 58*, 43–61.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review, 64*, 640–656.

Hill, F., & Gaddy, C. (2003). *The Siberian curse: How communist planners left Russia out in the cold*. Washington, DC: Brookings Institution Press.

Hill, B., & Woodrofe, M. (1975). Stronger forms of Zipf's law. *Journal of the American Statistical Association, 70*(349), 212.

INSEE. (2004). http://www.recensement.insee.fr/EN/RUB_MOT/ACC_MOT.htm

Iyer, S. D. (2003). Increasing unevenness in the distribution of city sizes in post-Soviet Russia. *Eurasian Geography and Economics, 44*(5), 348–367.

Jacobs, J. (1984). *Cities and the wealth of nations: Principles of economic life*. New York: Random House.

Julien, P. (2001a). *Les Grandes Villes Francaises Etendent Leur Influence* (ISEE Paper No. 766).

Julien, P. (2001b). *Les Deplacements Domicile-Travail: de plus en plus d'actifs travaillent loin de chez eux*. ISEE Paper No. 767. (Trans. (part) More and more workers are working far from home).

Kanemoto, Y. (1980). *Theories of urban externalities*. Amsterdam: North-Holland.

Le Gléau, J. -P., Pumain, D., & Saint-Julien T. (1996). Villes d'Europe: à chaque pays sa definition. *Économie et Statistique*, no. 294-295,1996-4/5 [in English as "Towns of Europe: To each country its definition"]. http://www.tu-chemnitz.de/phil/geographie/material/WS_03_04/BK/villeseuro.pdf

Marshall, A. (1890). Principles of Economics. London: Macmillan.

Overman, H. G., & Ioannides, Y. M. (2001). The cross-sectional evolution of the US city size distribution. *Journal of Urban Economics, 49*, 543–566.

Pumain, D. (2004). *Scaling laws and urban systems.* Santa Fe Institute (Working Paper No. 04-02-002).

Reed, B. (2002). On the rank-size distribution for human settlements. *Journal of Regional Science, 41*, 1–17.

Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge: Addison-Wesley.