



Munich Personal RePEc Archive

Explaining the size distribution of cities: X-treme economies

Berliant, Marcus and Watanabe, Hiroki
Washington University in St. Louis

24. October 2007

Online at <http://mpa.ub.uni-muenchen.de/5428/>
MPRA Paper No. 5428, posted 07. November 2007 / 04:43

Explaining the Size Distribution of Cities: X-treme Economies*

Marcus Berliant[†] and Hiroki Watanabe[‡]

October 2007

Abstract

The methodology used by theories to explain the size distribution of cities is contrived in that it takes an empirical fact and works backward to first obtain a reduced form of a model, then pushes this reduced form back to assumptions on primitives. The induced assumptions on consumer behavior, particularly about their ability to insure against the city-level productivity shocks in the model, are untenable. With either self insurance or insurance markets, and either an arbitrarily small cost of moving or the assumption that consumers do not perfectly observe the shocks to firms' technologies, the agents will never move. Equilibrium implies a uniform distribution of agents. Even without these frictions, our analysis yields another equilibrium with insurance that gives exactly the same utility level to consumers as the equilibrium studied in the literature, but where consumers never move. Thus, insurance is a substitute for movement. Even aggregate shocks are insufficient to generate consumer movement, since consumers can borrow and save. We propose an alternative class of models, involving extreme risk against which consumers will not insure. Instead, they will move. JEL number: R12 Keywords: Zipf's Law, Gibrat's Law, Size Distribution of Cities, Extreme Value Theory

*The first author thanks the Kyoto Institute of Economic Research at Kyoto University for their kind hospitality during the writing of the first draft of this paper. The second author thanks the Center for Research in Economics and Strategy (CRES) at the Olin School of Business of Washington University in St. Louis for financial support.

[†]Department of Economics, Washington University, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899 USA. Phone: (314) 935-8486, Fax: (314) 935-4156, e-mail: berliant@artsci.wustl.edu

[‡]Department of Economics, Washington University, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899 USA. e-mail: watanabe@wustl.edu

1 Introduction

1.1 Motivation

A small industry has developed that seeks to provide a theory to explain a singular but robust stylized fact in urban growth: the size distribution of cities. Zipf's law or the rank-size rule, as applied to the size distribution of cities, states that for any country, the rank of a city according to population (for example, New York is ranked number one in the US) multiplied by its population is constant. Thus, Los Angeles has half the population of New York, whereas Chicago has one third the population of New York. This stylized fact or factoid holds across many countries and time periods, but it is only one fact. In general, it is connected to Gibrat's law, stating that stochastic proportional growth tends to a lognormal distribution. The most compelling empirical work in this area shows that the size distribution of cities is lognormal (Eeckhout, 2004) when the data is not cut off at an arbitrary rank or population.

Explanation of this stylized fact or factoid by a theory has long been an objective of urban economists; it is quite robust, but also very difficult to theorize about. Three recent articles, Eeckhout (2004), Duranton (2007), and Rossi-Hansberg and Wright (2007), have made valiant attempts to tackle this issue head on.

The models seeking to explain this factoid are developed working backwards from empirics. That is, they take the empirical fact and attempt first to explain it using a reduced form, generally stochastic in nature, and then push this reduced form back to assumptions on primitives. As discussed below, these assumptions on primitives generally do not look natural, in the sense that if one were formulating a model of cities from scratch, it would not be obvious that one would want to begin with these assumptions. In fact, the literature of urban economics prior to the introduction of these models did not. Moreover, since these models are constructed for a single purpose, namely to explain an empirical fact, they seem incapable of explaining other empirical regularities, though they seem to be judged exclusively on the basis of how well they explain the one empirical factoid. Finally, it is worthwhile to note that there is likely an infinite number of models capable of explaining the size distribution of cities, and some of these might not even be stochastic; see, for example, Fujita and Mori (1997).

We shall focus primarily on the behavior of consumers, in particular the

degree to which they can hedge against risk. Our findings are as follows.

First, the type of risk featured in the literature is city-level risk in the form of a shock each period to the city-wide production function. The shock is known to all before they make their decisions, in particular consumer decisions about location and consumption bundle. There is no aggregate risk, since the number of cities are large and shocks are i.i.d., and this is often stated explicitly in the papers. In this context, *insurance is a perfect substitute for consumer movement*. We consider either self insurance, where a consumer insures over time by borrowing or saving to smooth consumption, or insurance markets, where a consumer insures using the fact that shocks are independent over space at a given time. With either type of insurance (or a combination), we find an equilibrium that yields the same period by period utility for each consumer as the one presented in the literature, where consumers move and generate Zipf's law or Gibrat's law. Our equilibrium features no consumer movement. Moreover, with even arbitrarily small moving costs or arbitrarily small uncertainty about shocks on the part of consumers, only our equilibrium survives. The existing literature finds that initial conditions don't matter, in that the size distribution of cities eventually tends toward lognormal. For our equilibrium, initial conditions matter in that consumers never move. The empirical truth likely lies somewhere in between.

Second, even aggregate risk is insufficient to generate consumer movement. For example, if there were a single aggregate shock common to all cities at each time, consumers could still insure by smoothing consumption through saving and borrowing over time and never moving.

Third, our proposed alternative model has aggregate risk of a specific kind. In the context of perfect competition, each city receives shocks to its productivity at each time. Only the city with the best technology in an industry produces at that time, driving out others. Our equilibrium has consumers moving to the cities producing with the best technology for some industry at that time. Insurance against shocks is too costly, as it is almost the total wage in a productive city. Our framework leads not to the central limit theorem or Gibrat's law, but rather to extreme value theory (the analog of the central limit theorem for maximal values instead of averages) and the Fisher-Tippett (1928) theorem. The implied functional form for size distribution of cities is different from the predictions in the literature.

The paper is organized as follows. First, in the balance of this section, we shall discuss the literature that attempts to refine the stylized fact and

explain it. Then in section 2 we shall raise specific objections, involving insurance against city-level risk, to these models. In section 3 we introduce Eeckhout's model and modify it to make the objections formal. In section 4, we propose an alternative type of model to explain the size distribution of cities, and implement it empirically. Finally, in section 5, we shall discuss our conclusions and directions for future work.

1.2 The Older Literature

The innovative work of Gabaix (1999a, 1999b) is the source from which the modern literature on the size distribution of cities flows. This work uses an overlapping generations structure where consumers live for two periods. It is assumed that moving costs are so high that consumers can only choose their location (city) when they are young. This location decision is made after shocks to production and amenities are realized for that period, and known to all. The consumer/workers cannot move again when old. The wages or income for the old in a city are never even specified, and it is simply assumed that the young make their decisions in a myopic manner. Moreover, the availability of insurance or capital markets is never discussed, so it is unknown whether the young can hedge against uncertainty about their wage when they are old in the city they choose.

If the old people are immobile, why is this important? It is important because when the young make their decisions, they can anticipate what happens when they are old, and might change their mind about their location decision when young. In other words, they won't behave myopically. Without myopia, insurance becomes important.

1.3 Recent Literature

Chief among recent work are Rossi-Hansberg and Wright (2007), Duranton (2006), Eeckhout (2004) and Duranton (2007). We focus on the latter two.

Eeckhout's model has consumers who are infinitely lived with foresight and who can move each period. There are technological shocks to production in each city in each time period. It is movement of consumer/worker population in response to these shocks that generates Gibrat's law. On p. 1445, the following statement is made: "Moreover, because there is no aggregate uncertainty over different locations, and because capital markets are perfect, the location decision in each period depends only on the current period utility.

The problem is therefore a static problem of maximizing current utility for a given population distribution, and the population distribution must be such that in all cities, the population $S_{i,t}$ equates utilities across cities.”

Here we wish to make an important distinction between transfers of consumption across time, namely perfect capital markets, and across states, namely complete and perfect futures markets.

The actual consumer optimization problem does not involve state-dependent assets nor does it allow state-contingent transfers of income. If it were to allow this, as in a standard model of complete futures or insurance markets, then agents would never move. They would simply buy assets at the start of time that would pay them under a bad state in their city at a particular time, and such that they would pay under a good realization in their city. In other words, they would insure against the state of nature in their city.¹

The basic model of Duranton (2007) has consumers maximizing an intertemporal utility function subject to an intertemporal budget constraint, without facing uncertainty. However, once the detailed urban features are added (in Section V and Duranton, 2006), the model looks similar to Eeckhout’s at least in terms of the urban features. One simply needs some dependence of local prices (land rents or wages) on the state of nature. Then utility equalization implies that people will move depending on the state realization, but this movement disappears if one allows insurance.

There isn’t enough detail about the urban market in Duranton (2006, 2007) to make specific statements about how insurance would work, but the consumers in a city face uncertainty about employment due to the uncertainty about innovations in various industries, so similar insurance arguments should work if the details of the model are filled in.

The bottom line here is not that complete and perfect futures markets are needed to upset the purpose of these rather fragile and contrived models. Rather, it is that any insurance at all will do the trick. Then the question becomes whether moving or buying insurance is cheaper for the consumers. If insurance is incomplete, it can still hinder mobility. Typically in these models and the real world, if moving costs are positive, it makes sense for consumers to stay put.

¹It is important to recognize that in this model there are two factors determining a worker/consumer’s productivity, namely the city-specific shock, and the externality in production induced by total population in the city. Even if capital markets are perfect, the production externality is not internalized (even with a land market), so the equilibrium allocation is not necessarily first best.

2 The Issue

2.1 How Insurance Reduces Population Movement

So how might this insurance occur in practice? Let's assume either that consumers cannot perfectly observe the technology shocks to cities, or moving has a small cost, or both.

- Self insurance. Since consumers can transfer consumption across time, and they know that shocks are i.i.d., then they can borrow or use their savings in bad times and save (or pay off their loans) in good, staying in the same city. In the literature, the intertemporal uncertainty faced by consumers does not show up in their objective function, whereas the possibility of self insurance does not show up in the budget constraint. The earlier quote from Eeckhout seems to imply that this is allowed, but the formal statement of the consumer budget constraint makes it clear that this is not allowed. This type of insurance exploits the fact that for any given city, the shocks are i.i.d. over time. Empirically, the place to look for self insurance is in the savings response to local employment shocks.
- Insurance markets. In all of these models, at each time the state of nature (the random shock to each production function for each city) is known to all and verifiable before consumers make their decisions about consumption bundles and location. So this is a perfect setting for a viable insurance market. An insurance firm can step in or the continuum of consumers can simply pool resources in each period, smoothing their consumption without changing location so it is independent of the state in their city.² This type of insurance exploits the fact that at any given time, the shocks are i.i.d. across cities. Empirically, one place to look for insurance is a cross-country comparison of how varying benefits of unemployment insurance affect mobility in response to local employment shocks.
- Futures markets. Consumers formulate plans to sell labor and buy consumption commodity and housing contingent on every possible state

²Although landlords (and the destination of land or housing rent) are not made explicit in these models, they might wish to participate in the insurance market as well, since their incomes fluctuate with the state of nature in their city. Of course, risk aversion on the part of landlords simply requires that their utility as a function of rental income be concave.

in every time period. There is no empirical complement. We mention this for completeness.

For our criticism to apply, insurance via either of the vehicles mentioned above need not be perfect. It only need be enough so that it throws off the mobility result, which requires that the response to shocks is *only* in consumer movement, rents and congestion.

Given that for Gibrat's law to hold, the shocks to each city in each period must be "small" (see Eeckhout, 2004, p. 1447), it seems reasonable to think that insurance would yield higher consumer utility than movement, if moving costs are at all significant or if consumers cannot observe shocks to firms perfectly, and thus face even a small amount of uncertainty in their optimization problems.

2.2 Possible Objections

The usual cause of a breakdown of insurance markets is adverse selection, represented for example by cream-skimming on the part of insurance companies. In the models discussed here, the state is assumed realized and observable to all before decisions are made in a given time period. So there is no issue of adverse selection. But one can easily imagine variations of these models that incorporate some form of information asymmetry. It would not be natural for, say, only consumers to know the shock to the local economy, since the technology shock really affects firms. If only firms knew the realization of the shock before making their decisions, then consumers could draw inferences from firm behavior, or the consumers could self insure or insure. It is not clear what hidden information or hidden action on the part of consumers would cause an insurance market breakdown in this context. Amenities are usually observable.

Another objection that could be raised is the commitment required on the part of consumers. In fact, commitment to a plan or contract is a requirement of models that feature self insurance, insurance or futures markets generally. For example, a consumer might experience regret over the purchase of a long-term health insurance contract after the state of the world that tells them that they are healthy is realized. Or the insurance company might experience regret if the consumer turns out to be unhealthy. In the models of the size distribution of cities, for example, one could begin the random process of technological change and at any point in time, allow insurance to begin. Then the population distribution will not change from that point on.

Finally, one might easily object to even small moving costs or even a small amount of noise in consumer observations of shocks. Then what we present is another equilibrium, that yields exactly the same period by period utility as the equilibrium studied in this literature. This alternative equilibrium features a uniform distribution of consumers, and does not generate Zipf's law.

3 A Model from the Literature Modified to Include Insurance

3.1 Notation

We use the model of Eeckhout (2004) as the basis for the analysis because it is explicit about consumer behavior, in the form of an optimization problem, as well as endogenous urban variables, namely local wages and land rents. We conjecture that the other models in the literature can be modified in a similar fashion.

The original model is specified as follows. For complete detail, see Eeckhout (2004, pp. 1445-1446).

Time is discrete and indexed by t . The set of cities is indexed by $i, j \in I$. Production is constant returns to scale. The measure of population in city i at time t is $S_{i,t}$, and let $A_{i,t}$ be the technological productivity parameter of city i at time t . This parameter follows the law of motion:

$$A_{i,t} = A_{i,t-1}(1 + \sigma_{i,t})$$

where $\sigma_{i,t}$ is the exogenous technological shock to city i at time t . It is assumed that $\sigma_{i,t}$ is i.i.d. with mean 0, symmetric, and satisfies $1 + \sigma_{i,t} > 0$. The positive local externality (spillover) function is given by $a_+(S_{i,t}) > 0$, where $a'_+(S_{i,t}) > 0$. The marginal product of a worker in city i at time t is given by

$$y_{i,t} = A_{i,t}a_+(S_{i,t})$$

Consumers are infinitely lived and identical. In city i at time t , consumption good is $c_{i,t}$, housing or land consumption is $h_{i,t}$ whereas leisure is $1 - l_{i,t}$ for labor supply $l_{i,t} \in [0, 1]$. Utility for a consumer in city i at time t is Cobb-Douglas:

$$u(c_{i,t}, h_{i,t}, l_{i,t}) = c_{i,t}^\alpha h_{i,t}^\beta (1 - l_{i,t})^{1-\alpha-\beta}$$

with $\alpha, \beta, \alpha + \beta \in (0, 1)$. For prices, let the consumption good be numéraire, the price of housing or land in city i at time t be $p_{i,t}$, and let the wage in

city i at time t be $w_{i,t}$. The local negative externality function is given by $a_-(S_{i,t}) \in [0, 1]$, where $a'_-(S_{i,t}) < 0$. The optimization problem of a consumer in city i at time t is:

$$\begin{aligned} \max_{\{c_{i,t}, h_{i,t}, l_{i,t}\}} & c_{i,t}^\alpha h_{i,t}^\beta (1 - l_{i,t})^{1-\alpha-\beta} \\ \text{subject to} & \\ & c_{i,t} + p_{i,t} h_{i,t} \leq w_{i,t} L_{i,t} \end{aligned}$$

where $w_{i,t} = A_{i,t} a_+(S_{i,t})$ and $L_{i,t} = a_-(S_{i,t}) l_{i,t}$. Total land or housing in a city is H . Using the first order conditions from this optimization problem and market clearance, equilibrium (denoted by asterisks) in city i at time t as a function of population $S_{i,t}$ can be found:

$$\begin{aligned} p_{i,t}^* &= \frac{\beta A_{i,t} a_+(S_{i,t}) a_-(S_{i,t}) S_{i,t}}{H} \\ w_{i,t}^* &= A_{i,t} a_+(S_{i,t}) \\ c_{i,t}^* &= \alpha A_{i,t} a_+(S_{i,t}) a_-(S_{i,t}) \\ h_{i,t}^* &= \frac{H}{S_{i,t}} \\ l_{i,t}^* &= \alpha + \beta \end{aligned}$$

The last equation in particular, indicating that labor supply is independent of population, is an artifact of the Cobb-Douglas specification.

Substituting back into the utility function, indirect equilibrium utility as a function of population $u^*(S_{i,t})$ can be written as

$$u^*(S_{i,t}) = [\alpha A_{i,t} \cdot a_+(S_{i,t}) a_-(S_{i,t})]^\alpha S_{i,t}^{-\beta} H^\beta [1 - \alpha - \beta]^{1-\alpha-\beta}$$

Under free mobility of consumers, indirect utility is equated across cities in each time period, determining their populations as a function of their productivity and their realized history of shocks, summarized by $A_{i,t}$. Instantaneous utility is constant over both time and location in equilibrium. Again using Eeckhout's notation, call this instantaneous utility level U . For later use, notice that this implies that

$$K \equiv \alpha A_{i,t} \cdot a_+(S_{i,t}) a_-(S_{i,t}) S_{i,t}^{-\beta/\alpha} = c_{i,t}^* \cdot S_{i,t}^{-\beta/\alpha} \quad (1)$$

is the same across cities and time.

3.2 Insurance

Let the discount factor be denoted by $\rho \in (0, 1]$. In correspondence with the assumption of complete capital markets, it is assumed that all consumers can borrow or lend at rate $\frac{1}{\rho} - 1$. The consumer optimization problem (at time 0) becomes:

$$\begin{aligned} & \max_{\{c_{i,t}, h_{i,t}, l_{i,t}\}} \sum_{t=1}^{\infty} \rho^t \cdot c_{i,t}^{\alpha} h_{i,t}^{\beta} (1 - l_{i,t})^{1-\alpha-\beta} \\ & \text{subject to} \\ & \sum_{t=1}^{\infty} \rho^t \cdot (c_{i,t} + p_{i,t} h_{i,t}) \leq \sum_{t=1}^{\infty} \rho^t \cdot w_{i,t} L_{i,t} \end{aligned}$$

As stated by Eeckhout, the problem reduces to the one period optimization problem if there is no insurance or futures markets. Formally, there should be an expectation in the objective function and a requirement that the budget constraint hold for every state of nature. However, this is omitted in the literature since the problem is reduced to a static optimization problem where the state of nature is observed before consumers make their choices.

There are several important points to be made at this juncture. First, it is useful to imagine the consumers stepping back at $t = 0$ and making decisions about their cities of residence and their consumption bundles for the entire time stream of their infinite lives. Second, and more important, *it does not matter which interpretation of the model one employs*. Specifically, resources can be transferred across states of the world (at any given time) in one or more of several ways. In the end, what a consumer is choosing is their residence and consumption bundle for every time and for every possible state of the world, optimizing utility subject to the budget constraint. The state of the world at time t affects the optimization problem through the prices, $p_{i,t}$ and $w_{i,t}$, and income (through $a_-(S_{i,t})$ and $L_{i,t}$) only. These variables depend on $A_{i,t}$ both directly and indirectly, the latter because $S_{i,t}$ depends on $A_{i,t}$ in equilibrium. The state of the world at time t does not enter into the consumer optimization problem otherwise. For example, it does not enter into the utility function. We could index these prices and incomes by the state of the world, but that would only serve to complicate notation.

As already mentioned, what will matter are only the lifetime choices of residence and consumption bundles, contingent on the state of the world in each period. The method used to actually implement them, via *transfers*

across states in a time period as opposed to across time periods, does not matter; there are many possibilities. With complete futures markets, at time $t = 0$ the consumers can sell their labor in every future time period and state, buying consumption good and housing in every future time period and state. With insurance markets, at $t = 0$ the consumers can buy actuarially fair insurance against price and income changes. With self insurance, they can commit to a plan of borrowing and saving under all possible scenarios, namely realizations of states in each time period.

When any of these is possible, independently or in combination, we propose the following equilibrium solution. For notational purposes, let \bar{S} be the mean population of cities, that is $\bar{S} = \frac{\sum_{i \in I} S_{i,t}}{|I|}$, where $|I|$ is the cardinality of the set I .³ Let A_0 denote the common initial technology level for all the identical cities before the process begins. With insurance, self insurance, or a futures market (or some combination of all 3), we propose the following solution:

$$\begin{aligned}\bar{p}_{i,t} &= \frac{\beta A_0 a_+(\bar{S}) a_-(\bar{S}) \bar{S}}{H} \\ \bar{w}_{i,t} &= A_0 a_+(\bar{S}) \\ \bar{c}_{i,t} &= \alpha A_0 a_+(\bar{S}) a_-(\bar{S}) \\ \bar{h}_{i,t} &= \frac{H}{\bar{S}} \\ \bar{l}_{i,t} &= \alpha + \beta\end{aligned}$$

In other words, this is the allocation generated by a constant, over both time and state, allocation with a uniform distribution of consumers. Of course, this distribution can be considered lognormal, but only in a trivial sense. It is more interesting to note that this is something like another manifestation of Starrett's spatial impossibility theorem, though here markets are incomplete due to the presence of unpriced local externalities, both positive and negative.

If one were to allow the random shocks to affect technology but keep consumers fixed at the proposed locations, the allocation would be

$$\begin{aligned}\hat{p}_{i,t} &= \frac{\beta A_{i,t} a_+(\bar{S}) a_-(\bar{S}) \bar{S}}{H} \\ \hat{w}_{i,t} &= A_{i,t} a_+(\bar{S}) \\ \hat{c}_{i,t} &= \alpha A_{i,t} a_+(\bar{S}) a_-(\bar{S}) \\ \hat{h}_{i,t} &= \frac{H}{\bar{S}} \\ \hat{l}_{i,t} &= \alpha + \beta\end{aligned}$$

³There are technical issues concerning the cardinality of I , but we shall ignore them here.

whereas the instantaneous utility level of consumers is

$$u(\widehat{c}_{i,t}, \widehat{h}_{i,t}, \widehat{l}_{i,t}) = [\alpha A_{i,t} a_+(\bar{S}) a_-(\bar{S})]^\alpha \left[\frac{H}{\bar{S}}\right]^\beta [1 - \alpha - \beta]^{1-\alpha-\beta}$$

Notice that this varies with $A_{i,t}^\alpha$. With insurance or futures markets, we obtain the proposed allocation, generating instantaneous utility

$$\begin{aligned} u(\bar{c}_{i,t}, \bar{h}_{i,t}, \bar{l}_{i,t}) &= [\alpha A_0 a_+(\bar{S}) a_-(\bar{S})]^\alpha \left[\frac{H}{\bar{S}}\right]^\beta [1 - \alpha - \beta]^{1-\alpha-\beta} \\ &= u^*(S_{i',t'}) \end{aligned}$$

for all i, t, i', t' .

Therefore, the equilibrium time path of utility for every consumer is the same, and constant, under insurance and under the equilibrium that generates movement and eventually becomes lognormal. At the very least, a discussion of why the latter equilibrium is selected should be offered in the literature.

With any moving cost, the insurance or futures market equilibrium (the one denoted with bars) clearly dominates the path with asterisks, the one put forth in the literature. Given a choice between moving along the equilibrium path or insuring at $t = 0$, each consumer will individually choose to insure.

A second, and perhaps more reasonable possibility, is that consumers observe A_{it} imperfectly when they make their location decisions each period. In that case as well, the consumers will insure rather than move, since they are risk averse.

4 Models of the Size Distribution of Cities

4.1 A Stochastic Non-Sarcastic Model

This model is loosely based on Duranton (2007), but in the context of perfect competition⁴ instead of monopolistic competition. In contrast with the other models in the literature, there is *economy-wide* risk in addition to *city-level* risk. But this in itself is not sufficient to generate consumer movement. For example, if all cities faced the same shock at each time, namely A_{jt} is independent of j , consumers could still insure against this risk by smoothing their consumption through borrowing and saving. Thus, we employ a more extreme form of aggregate risk.

⁴Since there is no market failure built into this model, equilibrium allocations will be Pareto optimal.

Time is discrete and all consumers are infinitely lived. Assume that there are many cities (indexed by $j = 1, \dots, m$) and many industries, each producing one consumption commodity (indexed by $i = 1, \dots, n$). All commodities are freely mobile. The production function for commodity i in city j at time t is given by

$$y_{ijt} = A_{ijt} \cdot l_{ijt}$$

where y_{ijt} is the output of commodity i in city j at time t , and l_{ijt} is labor input. The random variable $A_{ijt} \in \mathbb{R}_{++}$ will be discussed in detail shortly. Suppose that each consumer supplies 1 unit of labor inelastically and that the total number of consumers as well as total labor supply is given by N .⁵

In each time period t , each city j receives a random draw for its productivity in producing commodity i , namely A_{ijt} . Since we will be using the Fisher-Tippett limit theorem from extreme value theory rather than the central limit theorem, there is no requirement that these random variables be independent. It is assumed that with probability 1, the random draws for 2 industries at time t for city i are not both maximal among all cities for these given industries. In equilibrium, only the cities with the highest draw of the random variable for some industry will have employees and population. (Alternatively, we could simply classify cities exogenously by industry, and assume that a city in an industry receives only a draw for that industry.)

The wage rate for the (freely mobile) population of consumers is given by $w(t)$. In equilibrium, it will be the same across industries.

As is standard in this literature, the utility function of a consumer at time t is given by

$$u(t) = \sum_{i=1}^n \frac{1}{n} c_i(t)^\gamma$$

where $c_i(t)$ is the consumption of commodity i by a consumer at time t and $\gamma \in (0, 1)$. Let $p_i(t)$ be the price of commodity i at time t . Then a consumer's budget constraint is

$$\sum_{i=1}^n p_i(t) \cdot c_i(t) = w(t)$$

Let $\lambda(t)$ be the Lagrange multiplier associated with the budget constraint in the consumer optimization problem. Standard calculations yield demand for

⁵The assumption of Starrett's spatial impossibility theorem that is violated by this model is the assumption of location-independent production sets.

commodity i at time t for a single consumer $d_i(t)$:

$$d_i(t) = \left(\frac{\gamma}{-\lambda(t)n \cdot p_i(t)} \right)^{\frac{1}{1-\gamma}}$$

Aggregate demand is given by

$$N \cdot d_i(t) = N \left(\frac{\gamma}{-\lambda(t)n \cdot p_i(t)} \right)^{\frac{1}{1-\gamma}}$$

Profit optimization yields, for each t :

$$\begin{aligned} \text{For } i = 1, \dots, n, \text{ for } j^* \text{ with } A_{ij^*t} &= \max_{1 \leq j \leq m, 0 \leq t' \leq t} A_{ij't'} \\ p_i(t) \cdot A_{ij^*t} &= w(t) \end{aligned}$$

Here we are assuming total recall, in that the best technology from the past is remembered, so new technologies are not used unless they are better than all the old ones. Also, only the best technology in industry i survives, where the best is across all cities and previous time periods. This assumption is made for convenience. We discuss it more below.

Hence

$$\begin{aligned} \text{For } i = 1, \dots, n, \text{ for } j^* \text{ with } A_{ij^*t} &= \max_{1 \leq j \leq m, 0 \leq t' \leq t} A_{ij't'} \\ p_i(t) &= \frac{w(t)}{A_{ij^*t}} \end{aligned} \quad (2)$$

In other words, even though wage is constant across occupied cities, output price varies inversely with the production shock. Consumption commodity market clearance requires, for each t :

$$\begin{aligned} \text{For } i = 1, \dots, n, \text{ for } j^* \text{ with } A_{ij^*t} &= \max_{1 \leq j \leq m, 0 \leq t' \leq t} A_{ij't'} \\ l_{ij^*t} \cdot A_{ij^*t} &= N \cdot d_i(t) = N \left(\frac{\gamma}{-\lambda(t)n \cdot p_i(t)} \right)^{\frac{1}{1-\gamma}} \end{aligned} \quad (3)$$

This is the key equation for our analysis.⁶

Setting the constant $\kappa(t)$ to be

$$\kappa(t) = N \left(\frac{\gamma}{-\lambda(t) \cdot n \cdot w(t)} \right)^{\frac{1}{1-\gamma}}$$

⁶Labor market clearance requires, for each t :

$$\sum_{i=1}^n \sum_{\{j^* \text{ with } A_{ij^*t} = \max_{1 \leq j \leq m, 0 \leq t' \leq t} A_{ij't'}\}} l_{ij^*t} = N$$

and using (2) and (3), we obtain

$$\text{For } i = 1, \dots, n, \text{ for } j^* \text{ with } A_{ij^*t} = \max_{1 \leq j \leq m, 0 \leq t' \leq t} A_{ijt'} \\ l_{ij^*t} \cdot (A_{ij^*t})^{\frac{\gamma}{\gamma-1}} = \kappa(t)$$

Hence

$$\text{For } i = 1, \dots, n, \text{ for } j^* \text{ with } A_{ij^*t} = \max_{1 \leq j \leq m, 0 \leq t' \leq t} A_{ijt'} \quad (4) \\ l_{ij^*t} = \kappa(t) \cdot (A_{ij^*t})^{\frac{\gamma}{1-\gamma}}$$

Since $\gamma < 1$, labor usage l_{ij^*t} and the shock A_{ij^*t} are positively correlated. Notice that cities that do not have an industry with the largest shock in that industry at time t are empty.⁷

The original work on the asymptotic distribution of maxima drawn from a distribution is due to Fisher and Tippett (1928). Modern, more general treatments are given in Coles (2001) and Embrechts *et al* (1997). We shall return to a discussion of extreme value theory momentarily, but first we will draw the implications for our analysis.

The bottom line from this literature is that A_{ij^*t} has an asymptotic distribution of the following form, known as the generalized extreme value distribution:

$$F_{GEV}(x) = \begin{cases} \exp\{-[1 + \xi b(x - u)]^{-\frac{1}{\xi}}\} & \text{when } \xi \neq 0 \\ \exp\{-\exp(-b(x - u))\} & \text{when } \xi = 0 \end{cases}$$

Notice that there are 3 free parameters to be estimated here, namely b , u , and ξ . Also notice that to use rank as the left hand side variable in the regression, one simply computes $1 - F_{GEV}(x)$. But from a pragmatic point of view, it is easier to simply use $\ln(F_{GEV}(x))$ as the left hand side variable.

If there are no upper or lower bounds on the distribution, then $\xi = 0$ and the distribution is Gumbel. If there is an upper bound on the distribution, then $\xi < 0$ and the distribution is reverse Weibull. If there is a lower bound on the distribution, for example 0 in our case, then $\xi > 0$ and the distribution is Fréchet.

Substituting (4),

$$\ln(F(l)) = \begin{cases} -[1 + \xi b((\frac{l}{\kappa(t)})^{\frac{1-\gamma}{\gamma}} - u)]^{-\frac{1}{\xi}} & \text{when } \xi \neq 0 \\ -\exp(-b((\frac{l}{\kappa(t)})^{\frac{1-\gamma}{\gamma}} - u)) & \text{when } \xi = 0 \end{cases} \quad (5)$$

⁷Existence of an equilibrium is not an issue here, since the equilibrium prices and quantities can be solved analytically. For example, at $t = 1$, setting $p_1(1) = 1$, then $w(1) = A_{1j^*1}$, $p_i(1) = A_{1j^*1}/A_{ij^*1}$, $\lambda(1) = -\frac{\gamma}{nA_{1j^*1}}(\sum_{i=1}^n A_{ij^*1}^{\frac{\gamma}{1-\gamma}})^{1-\gamma}$, $l_{ij^*1} = N(\frac{\gamma}{-\lambda(1)nA_{1j^*1}})^{\frac{1}{1-\gamma}} A_{ij^*1}^{\frac{\gamma}{1-\gamma}}$, and so forth.

Notice that if we use cross section data, then t and hence $\kappa(t)$ is constant. Thus, in addition to the 3 parameters for the distribution of A_{ij^*t} (namely b , u and ξ), for the distribution of l_{ij^*t} there are two additional parameters, namely κ and γ .

In conclusion, we note that consumers will not want to insure against this risk. If only a small percentage of cities produce at any time, then insurance would cost only slightly less than the wage, so they might as well move and receive the wage in each period. The consumers still might want to insure against aggregate wage volatility (namely movement in $w(t)$ over time) by saving and borrowing to smooth consumption, but their spatial distribution is unchanged.

Returning now to our assumptions and extreme value theory, the original theory of Fisher and Tippett presumed that, fixing i , the random variables, A_{ijt} in our case, were i.i.d. across j and t . Of course, in our context this makes little sense. In general, the city with the best technology for some good i at a particular time t is more likely to innovate and produce a better technology for the next period than an arbitrary city. Moreover, it is possible that cities nearby are more likely to innovate than an arbitrary city. Fortunately, much progress has been made in extreme value theory since 1928. The modern versions of the Fisher-Tippett theorem, as given by Coles (2001, Theorem 5.1) and Embrechts *et al* (1997, Theorem 4.4.1) allow some dependence. Specifically, what is required is that the sequence of random variables be stationary and that a form of asymptotic independence (as blocks of random variables become farther apart in time) hold.⁸ It is also important to note that the model and results can be extended to the case where more than one city in an industry produces. This could happen, for example, if there is transportation cost for consumption goods between cities, so a city with a high realization of productivity for a commodity, but not the highest, might serve a local market. It turns out that extreme value theory applies not only to the maximum of a sequence of random variables, but also to the upper order statistics. A detailed discussion of the results can be found in Embrechts *et al* (1997, Section 4.2).

A few more remarks are in order. First, the role of having different industries i , as in the other models in the literature, is to generate a full distribution of limiting populations rather than just one realization of the asymptotic dis-

⁸An easy way to fit our structure into the theory is to fix an industry i and imagine that at each time t , there are m subperiods. A city j draws its random variable A_{ijt} in subperiod j of time t .

tribution of city populations. Second, in contrast with other models in the literature, the cities without the best technology for some industry at a given time have zero population, so they don't show up in the data because they are rural. Third, the derivations above work fine if there is not complete recall of previous technologies. For instance, if there were no recall, then the realizations of random variables for all cities at one time are independent of those at another time, so we have a form of block independence that is commonly used in extreme value theory.

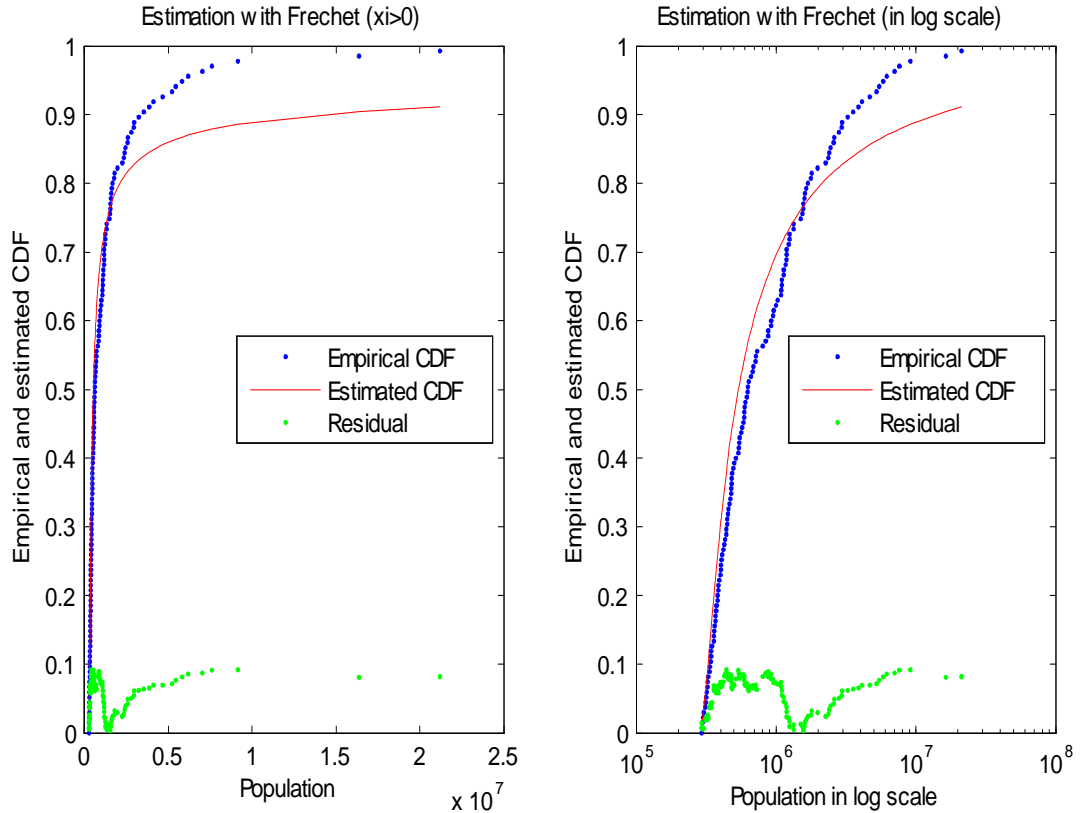
4.2 Empirical Implementation

Notice that we are not overly concerned with identification of the 5 parameters in equation (5). As in the balance of the literature on the size distribution of cities, these are just the parameters of a reduced form equilibrium distribution implied by a theory. In essence, the parameters are identified by the functional form itself.

We use the Census 2000 data set for US metropolitan areas with the 135 highest populations.

As noted in the sources we cite for extreme value theory, the most common method of estimating extreme value distributions is to use maximum likelihood. The maximum likelihood estimator (MLE) does not yield the smallest Kolmogorov-Smirnov (KS) statistic in our data set. The KS statistic measures the maximum distance between a sample distribution and its estimate. Most cities are of medium or smaller size. Since the MLE weights each data point equally, it produces a better fit in the range of smaller cities but not in the range of larger cities. As a result, larger cities such as New York City or Los Angeles are left with large residuals, leading to a large KS statistic.

We have fitted (5) by minimizing the KS statistic. The estimated values are described below.



Generalized Extreme Value Estimate of the Size Distribution of Cities

The estimated parameters are $b = 6$, $u = 3$, $\kappa = 22$, $\gamma = .9$, and $\xi = 1$. The KS statistic is .091.

4.3 A Non-Stochastic Sarcastic Model

It is well known that the lognormal distribution applies not only to the size distribution of cities, but to many phenomena in the physical sciences as well; see the survey by Limpert *et al* (2001) with applications as broad as the length of sentences in the writings of George Bernard Shaw to geology and plant physiology. Our particular interest is in the literature showing that rainfall (see Meneghini *et al*, 2001), soil moisture (see Janowicz *et al*, 2003), and crop abundance (see Halloy, 1999) are distributed according to lognormal distributions. In fact, Halloy (1999) displays graphs (see for example his figure 4) for crop abundance in New Zealand from 1842 to 1990 that look very much like the graphs used to explain the size distribution of cities. He also explains dynamics as in Duranton (2007).

After discussing explanations similar to ones advanced in the literature we have discussed, he concludes (p. 54):

In our case, external determinants are also apparent at the level of functional groups: first the dominance of potato and wheat relating to the need to feed the new population; then a prolonged period of dominance of feed crops, to feed the working horse; finally the rise and dominance of longer lived timber crops.

These are not necessarily random events, but can be driven by technological advances and the spread of technology. For our purposes, we take the distribution of crop abundance to be exogenous and lognormal.

Presume that, for reasons of increasing returns to scale or because soil composition or weather is spatially autocorrelated, the same crop is planted at close sites. Further, if we presume a fixed proportions production technology in agriculture, we shall obtain that employment is perfectly correlated with the land devoted to a particular crop, thus generating a lognormal population distribution of agricultural towns. Once industry develops, the insurance ideas of the previous section imply that nobody moves.

5 Conclusions

If one takes the last model in the previous section seriously, then one needs help badly. What we did, of course, was to take the empirical factoid that cities follow a lognormal distribution and work backwards. In the end, the assumptions employed about individual behavior are set up to explain the factoid. If we had more factoids to explain, then we would need some more free parameters.

So what's the point? Well, actually, there are several related points.

- First, when a model, markedly different from those found previously in the literature, is constructed to explain a specific empirical phenomenon, the microeconomic, structural assumptions about individual behavior and markets must make sense. Here, there is a rather obvious problem that self-insurance and insurance markets are assumed not to be functional. Models in the literature feature city-level risk, and it is generally possible to insure against such risk through many vehicles, barring asymmetric information. The latter does not arise naturally in

these models, since consumers are assumed to know the state of nature before making their location and consumption decisions.

- With time in the model, it is even possible to insure against aggregate risk through borrowing and saving.
- However, it is much more difficult to insure against extreme aggregate risk, so we propose such a model. Our model begins with microfoundations and implies a different functional form for the size distribution of cities than has been used in the literature.
- When one tests a model of the size distribution of cities, the real test is *not* whether it can explain the phenomenon it was designed to explain, but rather whether it is also consistent with regularities that it was not explicitly designed to explain.

Regarding the last item, it is convenient to use an analogy with the agglomeration literature. Cities happen. But that does not provide evidence that any particular model of agglomeration and city formation is correct; there is now a large variety of such models. Further testable hypotheses and evidence, such as predictions about trade, land rent, and wages, are necessary to tease out the contributions of various forces and models. The literature on the size distributions of cities could learn from this example.

In summary, we began with a criticism of the literature based on the fact that a primitive assumption in previous work, that consumers cannot insure, either by borrowing and saving or by pooling resources, against the random productivity variable for each city that is observable to all. Taking Eeckhout's (2004) model as an example, we showed that if insurance is allowed, there is another equilibrium of the model with a uniform distribution of consumers where there is never any migration. Instead, consumers insure against the risk, and the utility stream they obtain in this manner is the same as that in the equilibrium used in the literature. If there is any moving cost or residual uncertainty, the equilibrium used in the literature disappears. Finally, we propose an alternative model based on primitive assumptions, not designed to match any particular empirical factoid. Insurance is allowed, but consumers will never use it, as it is very costly. Instead, they move. The new model is based on extreme value theory and yields a functional form for the size distribution of cities different from the other models, and this prediction is empirically competitive with the ones in the literature.

Future work includes testing further predictions of the model, for example the wage distribution, and applying the model in new (but appropriate) contexts, such as finance.

References

- [1] Coles, S., 2001. *An Introduction to Statistical Modelling of Extreme Values*. Springer-Verlag.
- [2] Duranton, Gilles, 2006. “Some Foundations for Zipf’s Law: Product Proliferation and Local Spillovers.” *Regional Science and Urban Economics* 36(4), 542-563.
- [3] Duranton, Gilles, 2007. “Urban Evolutions: The Fast, The Slow, and the Still.” *American Economic Review* 97(1), 197-221.
- [4] Eeckhout, J., 2004. “Gibrat’s Law for (All) Cities.” *American Economic Review* 94(5), 1429-1451.
- [5] Embrechts, P., Kluppelberg, C., and T. Mikosch, 1997. *Modelling Extremal Events*. Springer-Verlag.
- [6] Fisher, R.A., and L.H.C. Tippett, 1928. “Limiting Forms of the Frequency by Distribution of the Largest or Smallest Members of a Sample.” *Proceedings of the Cambridge Philosophical Society*, 180-190.
- [7] Fujita, M., and T. Mori, 1997. “Structural Stability and Evolution of Urban Systems.” *Regional Science and Urban Economics* 27, 399–442.
- [8] Gabaix, Xavier, 1999a. “Zipf’s Law for Cities: An Explanation.” *Quarterly Journal of Economics* 114(3), 739-767.
- [9] Gabaix, Xavier, 1999b. “Zipf’s Law and the Growth of Cities.” *American Economic Review Papers and Proceedings* 89(2), 129-132.
- [10] Halloy, Stephan R.P., 1999. “The Dynamic Contribution of New Crops to the Agricultural Economy: Is it Predictable?” *Perspectives on New Crops and New Uses*, J. Janick (ed.), ASHS Press, Alexandria, VA.
- [11] Janowicz, J.R.; Gray, D.M. and Pomeroy, J.W., 2003. “Spatial Variability of Fall Soil Moisture and Spring Snow Water Equivalent Within

a Mountainous Sub-Arctic Watershed.” *60th Eastern Snow Conference*, Sherbrooke, Québec, Canada.

- [12] Limpert, Eckhard; Staehl, Werner A. and Abbt, Markus, 2001. “Log-normal Distributions across the Sciences: Keys and Clues.” *Bioscience* 51(5), 341-352.
- [13] Meneghini, R.; Jones, J.A.; Iguchi, T.; Okamoto, K.; and Kwiatkowski, J., 2001. “Statistical Methods of Estimating Average Rainfall over Large Space-Timescales Using Data from the TRMM Precipitation Radar.” *Journal of Applied Meteorology* 40(3), 568-585.
- [14] Rossi-Hansberg, E. and M.L.J. Wright, 2007. “Urban Structure and Growth.” *The Review of Economic Studies* 74(2), 597-624.