Emergent Cities:
Micro-foundations of Zipf's Law[†]

Robert Axtell and Richard Florida

The tendency for human beings, economic firms and social institutions to locate together has been the subject of intellectual inquiry for centuries. People and firms are not spread ubiquitously across the globe but have a powerful tendency to organize in well-defined geographic units, cities. First Auerbach (1913) and then Zipf (1941; 1949) observed a striking regularity in this process, noting the distribution in cities is highly skewed and follows a simple "rank-size rule" (Zipf's law). Despite considerable research on these issues, there does not as exist an explanation for Zipf's law based on behaviorally-credible micro-foundations. This article offers such a model, using the fact that business firms also have Zipf distributed sizes. The basic model is simple: people form together in firms, and firms co-locate. The agents in the model have heterogeneous abilities, exhibit bounded rationality, and interact directly with one another out of equilibrium in team production environments. Each agent is part of a firm and each firm has a spatial location. Agents periodically search for positions in other firms that would give them higher payoffs. Agents can also start-up new firms, at either their current location or a new location. Over time the movement of individual agents across firms combines with the movement of firms across locations to yield clusters of agents and firms. From these basic micro-foundations the model reproduces, under a wide range of conditions, a distribution of cities that follows Zipf's law. This model constitutes the first behaviorally-plausible, microscopic explanation of the city size distribution, on the one hand, and the mutual existence of Zipfian firm and city sizes on the other.

March 2006

[†] Axtell (raxtell@brookings.edu) is Senior Fellow at the Brookings Institution; Florida (florida@gmu.edu) is Hirst Professor of Public Policy and George Mason University Thanks are due the late Herbert Simon for stimulating our interest in this subject. Axtell gratefully acknowledges support from the National Science Foundation under grant 9820872 and the support of the John D. and Catherine T. MacArthur Foundation to the Center on Social and Economic Dynamics at Brookings. For useful comments we thank Joe Harrington and Tim Gulden in particular and seminar participants at Brookings, Carnegie Mellon, George Mason, and the Santa Fe Institute.

I Introduction

The tendency for human beings, economic firms, and social institutions to locate together and concentrate has been the subject of intellectual inquiry for centuries, captivating the curiosuity of leading thinkers such as Plato, Marx, Smith, Marshall, Weber and countless others. This heterogeneous distribution of human social and economic activity across geographic space has been a primary subject of study in geography, regional science, human ecology, and anthropology, and is of considerable renewed interest in economics, sociology, and the new field of econophysics. Despite a growing popular view that the combination of trade and technology are leading toward greater decentralization of economic activity, empirical reality indicates that human economic and social activity remains highly concentrated in geographic space. Basically, there is strong observed tendency of people and firms to locate and concentrate in tightly bounded organizational units, i.e. cities. While technological clusters of firms have captured much recent scholarly attention (Saxenian 1996), cities remain a fundamental dimension of the organization of the division of labor (Jacobs 1970). Lucas (1988) has written that tendency of people to cluster in cities generates powerful human capital externalities that account for the underlying "mechanics" of economic growth.

What is particularly striking about this spatial division of labor is not just that people and firms concentrate in well-defined cities, but that the distribution of these cities themselves follows a striking regularity. More than a half century ago, Zipf (1949) observed that the distribution follows a power law which he referred to as the "rank-size rule." Zipf's law for cities is one of the most striking regularities in the social sciences. It states that, for a given region or country, if one first ranks cities by size, a plot of log(size) versus log(rank) is well described by a line having slope exactly -1. This is equivalent to saying that city sizes are Pareto-distributed with an exponent of 1. This statistical regularity has been shown to be valid for most industrial countries today, as well as for individual countries over extended periods of time.[1] There are also well-known exceptions to the law; often the largest city in a country is 'too large' to fit well into the overall size distribution (cf. Ades and Glaeser 1997), and some countries have size distributions that deviate systematically from the Zipf distribution.

There are two basic approaches to understanding the highly heterogonous or skewed distribution of cities.[2] The first class of models derives from conventional economic theory. These models reflect Marshall's (1920) insights into the nature of urban agglomeration (see also Jacobs (1970), Lucas (1988), Henderson (1974; 1985; 1988)), and Christaller's (1966 [1933]) central place theory. Various of these ideas have been formalized in the 'new economic geography' (Krugman 1991; 1996; Fujita, Krugman et al. 1999). Yet as these researchers admit, this body of work systematically fails to yield the Zipf distribution (Fujita, Krugman et al. 1999: chapter 12)[3]. Indeed, some theoretical

[1] For a comprehensive empirical examination of the sizes of U.S. cities over time see Dobkins and Ioannides (1998). Interestingly, recent work using the 2000 U.S. Census, looking at both towns and cities find that the lognormal distribution fits the entire distribution (Eeckhout 2004).

[2] Our discussion of these two streams in the literature draws from Gulden and Hammond (2004); see also Batten (2000).

[3] These authors write "…at this point nobody has come up with a plausible story about the process that generates the rank-size rule…"

work by economists even fails to acknowledge that data of such extreme regularity exist (for instance, Page 1999). [4]

The second class of models for the distribution of cities derives from modern stochastic processes. Simon's proportional growth model (Simon 1955), together with successors (e.g., Berry and Garrison 1958; Steindl 1965; Gell-Mann 1992; Marsili and Zhang 1998; Gabaix 1999) represent aggregate explanations: they demonstrate that certain growth characteristics of cities are *sufficient* to explain the Zipf law, but provide no behavioral or microscopic justification for the postulated growth characteristics. Most of these models treat city formation and growth as following a "Gibrat process" [after Gibrat (1931)]. Gabaix (1999) reviews much of this literature and unifies it with a particular stochastic process due to Kesten (1973). These contributions generate a stable power-law distribution mathematically, but do not offer a compelling micro explanation to account for the results. As one review of the field put it: "this collection of models is essentially statistical—they seek to generate rather than explain the regularity" (Overman and Ioannides 2001).

Another domain of human affairs in which a similar power law exists involves business firm sizes (Ijiri and Simon 1977; Axtell 2001). Indeed, essentially the same Zipf distribution that governs city sizes describes firm sizes! Heretofore, Zipfian city and firm sizes have been treated as independent phenomena. It is as if by pure coincidence that these two extremely skew structures coincide. Common sense tells us there is *some* relationship between firm sizes and city sizes: there are few large firms in small cities. It is also well known that in pre-industrial times, city sizes were less skew than now. For instance, in feudal times the boundaries of castle walls played an important role in limiting population agglomeration. For these reasons, we think it reasonable to hypothesize that, far from being independent phenomena, there are a deep and important connections between firm and city size distributions. In this paper we describe a model with explicit microeconomic foundations that yields both Zipf-like firm and city size distributions.

The model is built around two types of autonomous actors: people and firms. The core mechanism behind the model is that people form firms, and firms in turn select locations in which to locate. Cities are locations that have high populations of firms and, of necessity, people. The number of people is fixed in the model, but the number and sizes of firms and cities is endogenous. Furthermore, while we have described firms as 'choosing' their location, it is also the case that, conversely, successful cities attract people who in turn create new firms.

---

[4] For these authors this is not an issue requiring, say, better parameterization of existing models, but rather is conceptual; they write:

> The stochastic models that have been proposed all rely fundamentally on the assumption of constant returns to city size, so that a city's expected growth rate is independent of its size. Yet all existing economic models of cities involve returns that are anything but constant. Rather, they involve a tug-of-war between increasing returns and decreasing returns, which for any given type of city…determines a characteristic size. Perhaps there is some way that we do not currently understand to reconcile the tension between centripetal and centrifugal forces that we believe determines city sizes at the micro level, and the as-if-constant-returns dynamics that seem to apply at the macro-level. We hope future research will resolve this puzzle.

In the model, agents form organizational units or firms. Agents are heterogeneous, operate under bounded rationality, and interact directly with one another out of equilibrium. The agents form teams and engage in joint production, which following Lucas's notion of human capital externalities, increase each other's productivity. In doing so, these productive teams of agents become firms. Agents have the ability to start new firms or switch between exiting firms. Firms display increasing returns to human capital and dynamic processes of formation and evolution.

The second step in the model is that cities *emerge* from the interactions of agents and firms. When many such firms have the same location, in our model, we call the resulting agglomeration a 'city.' Cities, in the model, have no agency, but are able to attract people from other cities to work within their population of firms. Cities are also able to house new firms, as when an agent decides to start up a new firm and stay in its present location. A finite set of locations is assumed, and the initial actors locations are random. People can join firms, adopt firms locations, or create new firms either in their current or a new random location. Thus, cities emerge through the interactions of purposive agents through the institution of the firm.

The results of the model are straightforward. From these basic micro-foundations and under a wide range of conditions, the model co-generates Zipf distributions of firms and cities. These two famously skewed size distributions, among the most regular statistical features of modern societies, have always been treated as individual objects in isolation, are shown here to be fundamentally related.

The rest of the paper is organized as follows. Section II outlines the analytical structure of the model. Section III describes the typical realizations of the model. Section IV summarizes the main findings and draws conclusions.

II Model Structure

The point of departure for this work is a multi-agent model of endogenous firm formation (Axtell 1999; 2002). This model is a quantitative explanation of several empirical features of U.S. firms, including: (1) the extreme right skewness of the firm size distribution, (2) the non-normal distribution of firm growth rates, (3) the decrease in growth rate variance with firm size, (4) the wage premium associated with working in larger firms, and (5) approximately constant returns to scale at the aggregate level. In summary terms, it is a model in which heterogeneous individuals engage in team production, each periodically adjusting its input to production in order to improve utility non-cooperatively. Permitting agents to migrate between teams when it is utility enhancing to do so yields a constantly evolving set of coalitions— microeconomic disequilibrium—yet stationary statistics at the aggregate level. It is known that this model has no stable fixed point equilibria for a wide class of effort level adjustment dynamics.

The way this model has been modified to study and city formation processes is very simple. We endow agents with locations; each firm has the location of its founder. We then require each agent to assume the location of any firm it joins. Then, when an agent starts-up a new firm, that firm has, with probability $p = 1 - \varepsilon$, the same location as the agent's previous firm. However, with probability ε the new firm's location is selected at random. For $\varepsilon \ll 1$, i.e., $p \sim 1$, there results a very skewed distribution of location sizes: most have small numbers of agent occupants, while progressively fewer locations have progressively larger numbers of occupants. It is demonstrated that the stationary configuration of this model is approximately a Zipf-type distribution.

Endogenous Firms

Consider a heterogeneous group of agents engaged in team production. Each agent has human capital level $\theta_i \in [0, 1]$. There are increasing returns and agents choose how much to contribute to production, $e_i \in [0, 1]$, based on the behavior of other agents in the group. Agents provide relatively more input to the extent that they gain utility from doing so. That is, each agent adjusts its e_i non-cooperatively. Postulate that there is in place in the group a compensation scheme that is nondecreasing in e_i, i.e., more effort is not penalized. For such a group it is possible to prove the following:

Proposition 1: There exists a unique Nash equilibrium of production contributions by the agents

Proposition 2 (Holmstrom 1982): Agents under-supply effort (with respect to a cooperative solution that maximized total payoffs) at Nash equilibrium

Proposition 3: There exists a maximum stable group size, beyond which groups are dynamically unstable

The first of these just means that in any team it is possible for the agents to behave as we have specified they should. The second proposition asserts that there are other levels of agent contributions to production that Pareto dominate the Nash levels. Unfortunately, as first pointed out by Hölmstrom, these Pareto superior levels are not individually rational. The final proposition means that if groups beyond a certain size have their Nash equilibrium effort levels perturbed slightly, then the resulting individual adjustments will generally be an unstable process, never settling down to the Nash levels.

The final ingredient of this model is that agents are permitted to leave their current group if they can either find a different group where they will have greater utility, or if starting up a new firm yields higher utility than their present firm.

Overall, this specification leads to a non-equilibrium model at the agent level of endogenous team formation. Consider a heterogeneous population of agents arranged into a particular constellation of groups. If some agents can get greater utility by leaving their groups they do so. When agents join groups they may induce modifications of the behaviors of the existing group members, possibly causing subsequent migrations to other firms by some of the existing group members. Because of proposition 3, such microeconomic dynamics will always be present in the model. That is, there is constant flux at the agent level, although in large populations stationarity will emerge at the aggregate level.

In particular, aggregate stationary configurations of firms are quite closely related to empirical data on firms. These data are as follows:

- o the size distribution of the largest firms is approximately a power law (e.g. Ijiri and Simon (1977), and this is reproduced by the model;
- o firm growth rates are Laplace-distributed (Stanley, Amaral et al. 1996);
- o the variance in growth rates decreases with firm size according to a power law (Stanley, Amaral et al. 1996);

o wages are an increasing function of firm size, a well-known puzzle in labor economics (Brown and Medoff 1989);

o constant returns to scale obtain at the aggregate level (Basu and Fernald 1997).

The model of firms described above is capable of reproducing all of these data, as shown in figure 1.

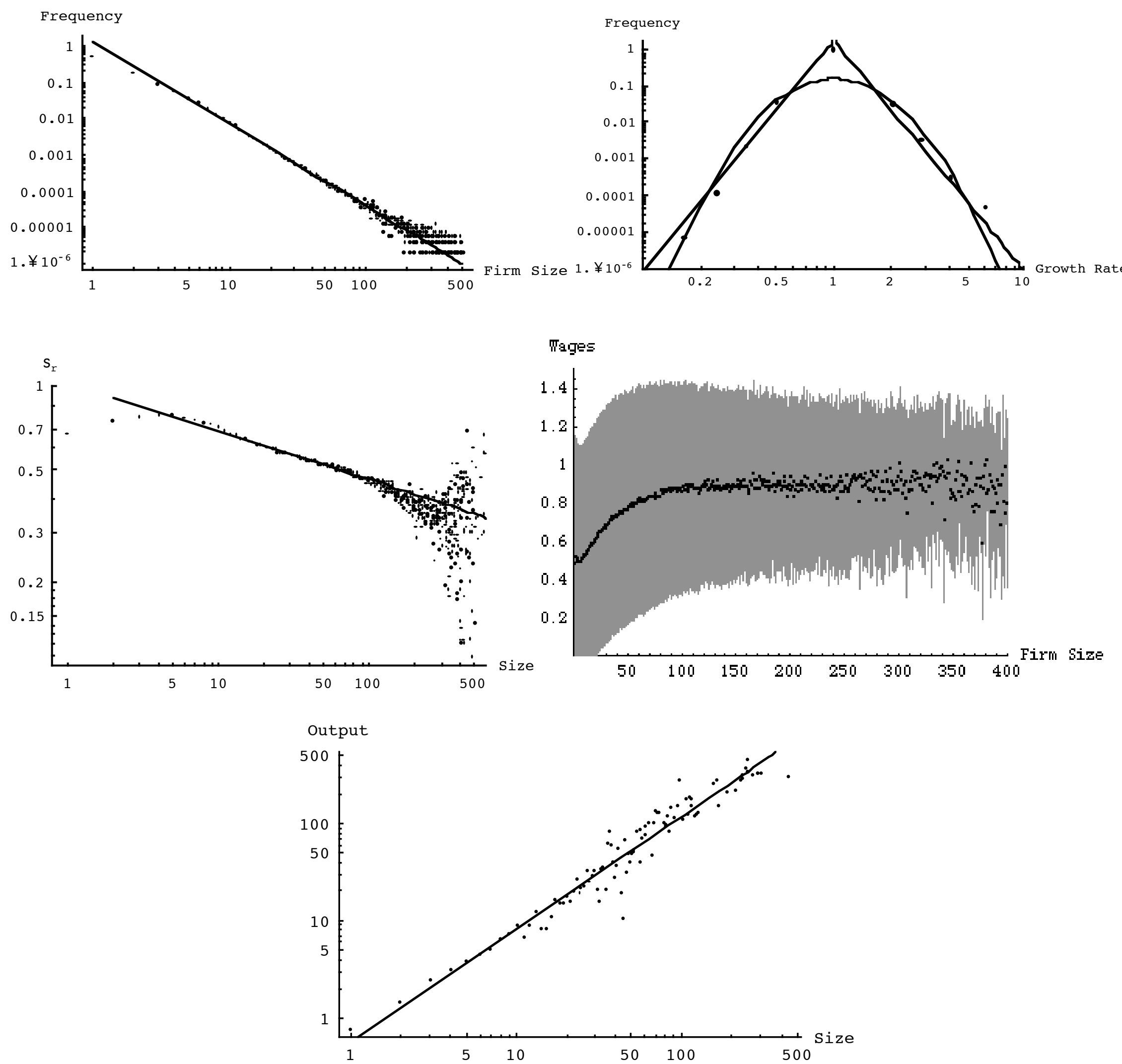

Figure 1: For firms, (1) Pareto size distribution with exponent near unity (Zipf distribution) and (2) growth rate distributions better fit by Laplace than normal distribution, (3) dependence of growth rate variance and (4) wages on firm size (error bars are ±1 standard deviation), and (5) returns to scale

It has been shown (Axtell 1999; 2002) that the *qualitative* structure of these results are robust to a wide range of alternative specifications of preferences, bounded rationality, social networks, compensation schemes, and firm management practices. Varying the parameters of the model changes only the *quantitative* aspects of the model. This model will now be extended in order to make it relevant to city formation.

Firms and Locations

In order to model city formation processes, it is necessary to have some sense of location, and perhaps distance. That is, different cities occupy different locations, so a model must necessarily reflect this. So, to begin, let us define a finite set of locations, $L = \{a, b, c, \ldots, z\}$. We require that each firm have such a location and that all agents in the firm adopt the location of the firm.[5] When an agent changes firms it alters its location to that of the new firm. Finally, when an agent decides to start-up a new firm it does so in its current location with probability $1 - \varepsilon$, but with probability $\varepsilon \ll 1$ it selects a location at random.

Note that the 'flavor' of this model is quite similar to Simon's city formation model (Simon 1955). But where the probabilities in Simon's model governed the disposition of 'lumps' of individuals, here it is a parameter in agent decision-making, mediated through the institution of the firm. Furthermore, instead of people migrating in a way that is proportional to city sizes, here we given only behavioral specifications and agent migrations are therefore endogenous. We now turn to realizations of the model.

III Results

In this section the overall performance of an implementation of this model is described in detail. First, a particular parameterization will be introduced, closely related to an intensively studied specification of the endogenous firms model (Axtell 1999; 2002). Then, in order to build up the reader's intuition about the overall functioning of the model, a picture of its dynamics will be developed. We find that a stationary size distribution emerges and does indeed possess the Zipf property.

Model Configuration

We begin our investigations of the performance of this model with a population of 1,000,000 agents, heterogeneous by tastes for income and leisure. There are 100 locations, and at time zero each agent is working alone in its own firm, at a location randomly assigned from the uniform distribution. There are quadratic increasing returns to effort and output is divided equally among all members of a firm. Each agent has a fixed social network of two other agents, assigned randomly—the overall social network is a random graph. Agents are activated at random and engage in purposive behavior, seeking utility improving changes to their effort levels and group membership. In the event that an agent finds it welfare improving to move to a different firm it adopts the location of the new firm. When an agent can receive the most welfare from going off on its own and starting up a new firm it does so at its current location with probability $p = 0.99$, while 1% of the time it selects a new location at random. These specifications are summarized in the following table.

[5] In this model no firms have multiple establishments, so it is not possible for firms to operate in more than one location.

Model Attribute	Value
agents, A	1,000,000
constant returns coefficient, a	1
increasing returns coefficient, b	1
increasing returns exponent, a	2
distribution of preferences, q	U(0, 1)
compensation system	equal shares
number of neighbors, n	2
agent activation	random
initial team condition	all agents in singleton firms
number of locations	100
initial location	uniformly distributed
new firm adopts random location	p = 0.01

Table 1: 'Base case' configuration of the agent-based computational model

In Axtell (1999; 2002), the effects of varying all but the last three of these specifications are exhaustively studied. It turns out that highly skewed firm size distributions obtain for all variations as long as (1) agents are heterogeneous and (2) purposive, and (3) significant increasing returns are present (exponent greater than about 1.5).[6]

Model Dynamics

When the model begins running, agents immediately form firms, small ones at first, then larger ones progressively. After some time, this initial transient period gives way to a stationary distribution of firm sizes, and the locational distribution begins to stabilize. While all locations are initially occupied by approximately 10,000 agents, over time some regions become relatively barren as agents and firms leave them for jobs in other locations. Once in these new locations, agents who wish to start up new firms usually do so in their existing location, and so there results an agglomeration, especially of new, rapidly growing firms, in a relatively small number of specific locations. Over longer periods of time, these agglomerations of agents and firms grow and decline, but stationary distributions of various kinds emerge.

Model Size Distributions

Most notably, after the transient has dissipated, this model yields a Zipf distribution of city sizes, as shown in figure 2, a typical result. Note that some deviations from the pure rank-size rule obtain, but this is true of real data as well; see Gabaix (1999) for an analysis of the 'typical' character of such deviations.

[6] We have instantiated populations of size 10 million, which at $O(100)$ bytes/agent requires just over a gigabyte of memory. Such models take several days to reach stationarity from uniformly random initial conditions. The million size populations run merely overnight on a typical workstation. We have plans to make realizations from a realistically-sized population (300 million) using a cluster facilities.

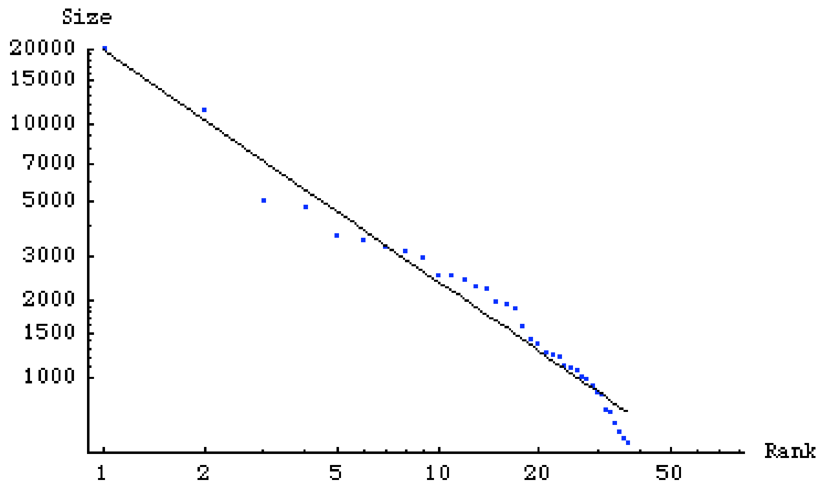

Figure 2: Data from a particular realization of the model (in blue), along with superposition of a Zipf distribution (black line)

No parameter 'tuning' of the model is necessary to achieve this result. Each realization always produces the Zipf relation. As long as the model specification is reasonable, in the sense of creating skewed firm size distributions, and the relocation parameter, ε, is not too large, skewed city sizes of this kind always obtain.

A spatial distribution of city sizes is generated by the model. A typical one is shown in figure 3.

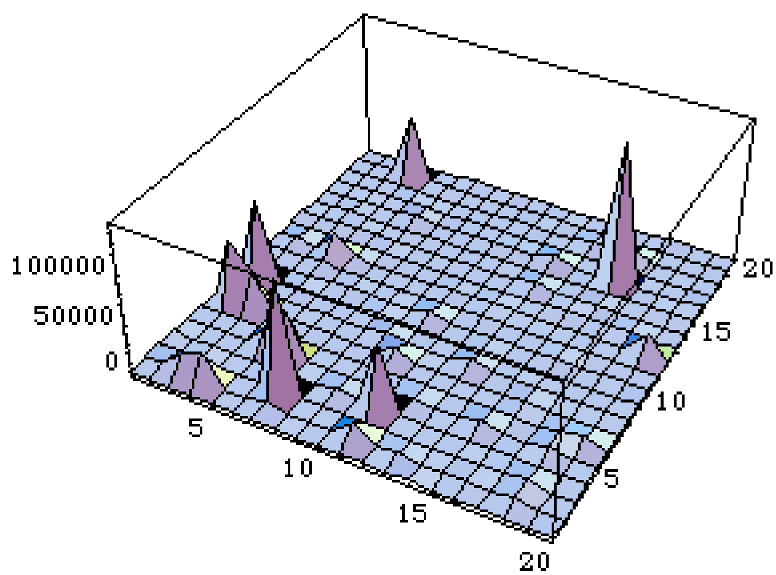

Figure 3: Spatial distribution of cities having a Zipf distribution, where height refers to absolute population

Note that there are a few large cities and more small ones, with the largest located at approximately (17, 17). Most locations have few people and even fewer firms.

Much as in the model of endogenous firms (Axtell 1999; 2002), we can query the model output and build up empirical distributions of various other kinds. In particular, the model yields (a) growth rate distributions quite similar to those reported in Ioannides and Overman (2001) for U.S. cities, (b) variance in growth rates that decline with city size, much like those reported by Glaeser *et al*. [1998], (c) a city size-wage effect—average wages are higher in larger cities—and (d) constant returns to city size, i.e., output grows linearly with city size. Due to space constraints here we will present these results elsewhere.

Alternative Model Specifications

In Axtell (1999; 2002) the model of firm formation that underlies the present model of city formation is systematically explored with a variety of alternative specifications. It is found that the *qualitative* characteristics of the model output—i.e., right-skewed size distribution, Laplace-distributed growth rate distribution, variance in growth rates that decline with size, wages increasing with firm size, and constant returns to scale at the aggregate level—are invariant over wide ranges of model parameters; rather, varying the model parameters alters the model output *quantitatively*, e.g., changing the slope of the firm size distribution. These variations include altering the extent of increasing returns, modifying the size and nature of agent job search networks, increasing the extent to which rationality is bounded, changing the system of compensation, and introducing the possibility that agents are loyal to their firms.

Here we study the effect on the city size distribution of the several parameters introduced in the city formation model, i.e., the number of locations, the probability of choosing a new location randomly when starting up a new firm, and the rate of population growth, the latter being a new variation.

Effect of the number of agents

The main effect of changing the number of agents is to make the statistical properties of the model more robust. Below about *O*(1000) agents/location, there can be significant deviations from the Zipf law.

Effect of ε, the probability of choosing a random location

The effect of increasing p is to increase the rate of growth of the peripheral areas, and to generally increase the migration of firms up and down the size distribution. However, if ε is too large, then there is so much 'bleeding' of the productive firms out of the large cities that there is insufficient density of human capital to support a Zipf-, yielding 'too many' small cities. Results for $\varepsilon = 0.5\%$ are shown in figure 4.

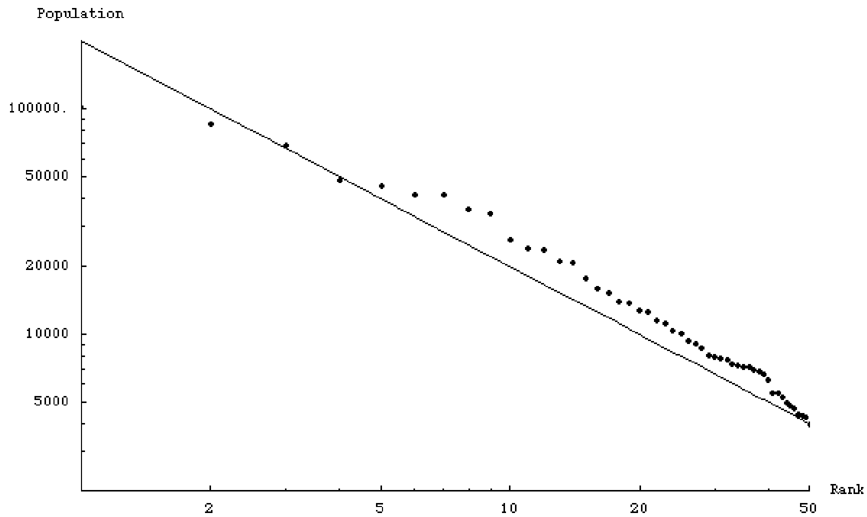

Figure 4: Data from a particular realization of the model (dots) for ε = 0.5%, along with superposition of a Zipf distribution (black line)

This parameter can be increased substantially while preserving the character of the Zipf distribution.

Effect of the number of locations

Changing the number of locations is closely related to changing the number of agents; as long as the number of agents/location is sufficiently large for robust statistics, there is no significant effect of adding locations. Results for 400 locations are shown in figure 5, with only the largest 100 cities plotted.

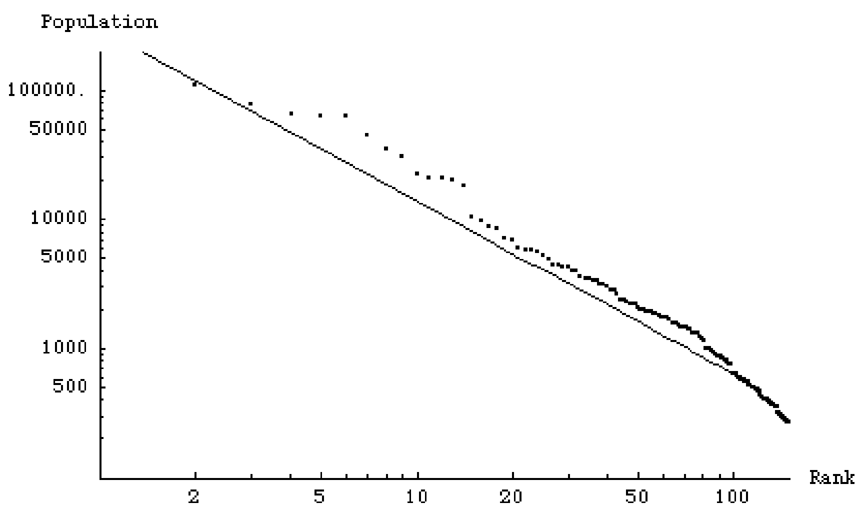

Figure 5: Data from a particular realization of the model (dots) for 400 locations, along with superposition of a Zipf distribution (black line)

Effect of a growing population

In some sense, 'growing' Zipf from a uniform initial condition is seemingly a much harder requirement than what is called for empirically. For at no time in human history has there been a uniform population over any significant fraction of the Earth's surface. Rather, early on in the evolution of societies there always exist more or less skewed size distributions. Subsequent population from this initial condition simply amplifies this skewness to full-blown Zipf behavior.

We have built the model to accommodate population growth. Beginning with a relatively small population and letting it grow by an order of magnitude, say, while rearranging itself into firms and cities, yields Zipf-type behavior *much more quickly* than starting from a large population homogeneously distributed across locations.

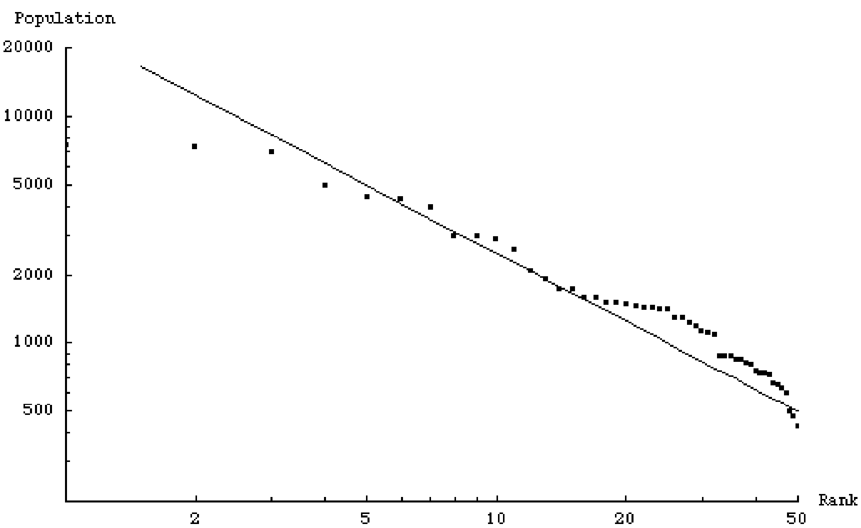

Figure 6: Data from a particular realization of the model (dots) for a growing population, along with superimposition of a Zipf distribution (black line)

It is interesting to watch the evolution of this distribution over time as population grows <maybe add QuickTime movie online>.

IV Conclusions

A great puzzle in the social sciences has been the heterogeneous distribution of people across time and space. Economists have recognized this uneven, hierarchal structure but have been unable to generate empirically-salient explanations for it, based on rationality and equilibrium. Other researchers have provided stochastic models for this heterogeneous city structure but these models have lacked compelling theoretical foundations.

This article has developed a model which shows how cities and firms form and co-evolve from a simple micro-foundation: people form firms and these firms agglomerate to make cities. From these basic micro-foundations, the model generates the Zipf distribution of cities. Agents have heterogeneous abilities, exhibit bounded rationality, and interact directly with one another out of equilibrium in team production environments. Each agent works in a firm and each firm has a location. Agents periodically search for positions in other firms that would give them higher utility. Moves between firms are migrations when they involve changes in location. Agents can also start-up new firms. Over time, the movement of individuals across firms combines with the movement of firms across locations to yield clusters of agents and firms in particular locations, i.e., cities. It is demonstrated that under a wide range of conditions these locational clusters reproduce the so-called 'Zipf law' for city size while simultaneously generating Zipf firm sizes. These two famously skewed size distributions, among the single most regular statistical features of modern societies, are shown here to be fundamentally related.

While this paper provides what appears to be the first microeconomic explanation of Zipf's law, it would surprise us if it turned out that there were *not* other models that have similar properties to the one elaborated above. The existence of other, microeconomically distinct yet macroeconomically similar mechanisms is a consequence of *universality*. While it is difficult to say just what such other mechanism will look like, it is almost certainly the case that they will feature heterogeneous agents who interact directly with one another out of equilibrium. This defines a universality class, whose specific properties await the further exploration and elaboration of empirically-relevant models of the Zipf distribution.

References

Ades, A. F. and E. L. Glaeser (1997). "Trade and Circuses: Explalining Urban Giants." Quarterly Journal of Economics **110**(1): 195-227.

Auerbach, F. (1913). "Das gesetz der bevolkerungskonzentration." Petermanns Geographische Mitteilungen **59**: 74-76.

Axtell, R. L. (1999). The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions. Working paper. Santa Fe, N.M., Santa Fe Institute.

Axtell, R. L. (2001). "Zipf Distribution of U.S. Firm Sizes." Science **293**(5536): 1818-20.

Axtell, R. L. (2002). Non-Cooperative Dynamics of Multi-Agent Teams. Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems. C. Castelfranchi and W. L. Johnson. Bologna, Italy, ACM Press. **Part 3:** 1082-1089.

Basu, S. and J. G. Fernald (1997). "Returns to Scale in U.S. Manufacturing: Estimaes and Implications." Journal of Political Economy **105**: 249-283.

Batten, D. F. (2000). Discovering Artificial Economics: How Agents Learn and Economies Evolve. Boulder, Colorado, Westview Press.

Berry, B. J. L. and W. L. Garrison (1958). "Alternative explanations of urban rank-size relationships." Annals of the Association of American Geographers **48**: 83-91.

Brown, C. and J. Medoff (1989). "The Employer Size-Wage Effect." Journal of Political Economy **97**: 1027-1059.

Christaller, W. (1966 [1933]). Central Places in Southern Germany. London, Prentice Hall.

Dobkins, L. H. and Y. M. Ioannides (1998). Evolution of the U.S. City Size Distribution. The Economics of Cities. J.-M. Huriot and J.-F. Thisse. New York, N.Y., Cambridge University Press.

Eeckhout, J. (2004). "Gibrat's Law for (All) Cities." American Economic Review **94**(5): 1429-1451.

Fujita, M., P. Krugman and A. J. Venables (1999). The Spatial Economy: Cities, Regions, and International Trade. Cambridge, Mass., MIT Press.

Gabaix, X. (1999). "Zipf's Law for Cities: An Explanation." Quarterly Journal of Economics **114**(3): 739-767.

Gell-Mann, M. (1992). Zipf's Law and Related Mysteries. First Santa Fe Winter School on Complex Systems. Tucson, Arizona, University of Arizona/Santa Fe Institute.

Gibrat, R. (1931). Les Inegalities Economiques; Applications: Aux Ineqalities des Richesses, a la Concentartion des Entreprises, Aux Populations des Villes, Aux Statistiques des Familles, etc., d'une Loi Nouvelle, La Loi de l'Effet Proportionnel. Paris, Librarie du Recueil Sirey.

Gulden, T. and R. Hammond (2004). Beyond Zipf: An Agent Based Understanding of City Size Distributions. Working paper.

Henderson, V. J. (1974). "The Sizes and Types of Cities." American Economic Review **64**: 640-656.

Henderson, V. J. (1985). Economic Theory and the Cities. San Diego, Calif., Academic Press.

Henderson, V. J. (1988). Urban Development: Theory, Fact, and Illusion. New York, N.Y., Oxford University Press.

Holmstrom, B. (1982). "Moral Hazard in Teams." Bell Journal of Economics **13**: 324-340.

Ijiri, Y. and H. A. Simon (1977). Skew Distributions and the Sizes of Business Firms. New York, N.Y., North-Holland.

Jacobs, J. (1970). The Economy of Cities. New York, N.Y., Vintage Books.

Kesten, H. (1973). "Random Difference Equations and Renewal Theory for Products of Random Matrices." Acta Mathematica **CXXXI**: 207-248.

Krugman, P. (1991). Geography and Trade. Cambridge, Mass., MIT Press.

Krugman, P. (1996). The Self-Organizing Economy. New York, N.Y., Blackwell.

Lucas, R. E., Jr. (1988). "On the Mechanics of Economic Development." Journal of Monetary Economics **22**: 3-42.

Marshall, A. (1920). Principles of Economics. London, Macmillan.

Marsili, M. and Y.-C. Zhang (1998). "Interacting Individuals Leading to Zipf's Law." Physical Review Letters **LXXX**(12): 2741-2744.

Overman, H. G. and Y. M. Ioannides (2001). "The Cross-Sectional Evolution of the U.S. City Size Distribution." Journal of Urban Economics **49**: 543-566.

Page, S. E. (1999). "On the Emergence of Cities." Journal of Urban Economics **45**: 184-208.

Saxenian, A. (1996). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, Mass., Harvard University Press.

Simon, H. A. (1955). "On a Class of Skew Distribution Functions." Biometrika **42**: 425-440.

Stanley, M. H. R., L. A. N. Amaral, S. V. Buldyrev, S. Havlin, H. Leschhorn, P. Maass, M. A. Salinger and H. E. Stanley (1996). "Scaling Behaviour in the Growth of Companies." Nature **379**(29): 804-806.

Steindl, J. (1965). Random Processes and the Growth of Firms. New York, N.Y., Hafner Publishing Company.

Zipf, G. K. (1941). National Unity and Disunity. Bloomington, Indiana, Principia.

Zipf, G. K. (1949). Human Behavior and the Principle of Least Effort. Cambridge, Mass., Addison-Wesley Press.