ESSAYS ON LOCATION AND TRADE

by

Linda Harris Dobkins

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

APPROVED:

Yannis M. Ioannides, Chair

Kristin Butcher

Anna Hardman

Charles Michalopoulos

Daniel Nuxoll

June 1995
Blacksburg, Virginia

ESSAYS ON LOCATION AND TRADE

by

Linda Harris Dobkins

Yannis M. Ioannides, Chair

Economics

(ABSTRACT)

We observe that the location of non-primary production activities is concentrated among cities and regions. The pattern of domestic and international trade and the growth of a nation are intimately related to the country's regional and urban structure. The theoretical and empirical essays in this dissertation analyze the consequences of these observations, seeking to determine the influence of cities on national growth and vice versa.

Chapter 1 models locales which produce goods for trade outside the locale's boundaries. I use a model that assumes monopolistic competition in both service and traded goods sectors to study the impact of both a localization externality and a nation-based externality. The localization externality is related to innovation that occurs because of agglomeration in a locale. The nation-based externality reflects the idea that some nations promote competitive industries more successfully than do others. The model highlights the "success breeds success" (or "failure breeds failure") pattern associated with economies in various stages of economic development.

Many current models in urban economics assume that cities are highly specialized in the production of goods and services, and that specialization leads to growth of output and employment; Chapter 1 yields such a result. Chapter 2 is an econometric study examining those assumptions. Using a set of 63 of the largest and potentially

most specialized cities in the United States, I study data for some 40 industries and sectors in the years 1947, 1956, 1970, 1980, and 1990; I have divided the data into two periods, 1947 to 1970 and 1970 to 1990, for analyzing dynamic growth patterns. I find that static specialization is less evident in U.S. cities over time. I also find that what specialization does exist in a city is not necessarily associated with growth in an industry in a city. For some industries, increased specialization is associated with negative growth. Overall, growth in an industry in a city is more often associated with a growing labor force in that city, which I interpret as an indication of the importance of diversity.

A change in the distribution of city sizes in an economy is a potential indicator of change in that economy's structure. I consider several methods of measuring change in the city size distribution in the United States in this century, and conclude that the distribution is becoming more concentrated over that time period. I use both parametric and non-parametric distributional approaches to study the dynamics of the evolution of the U.S. city size distribution. I find that the balance between manufacturing and service sector employment, agricultural land values and the prime interest rate have statistically significant effects on the changing distribution.

TABLE OF CONTENTS

LIST OF TABLES

FIGURE

Chapter 3: Evolution of the U.S. City Size Distribution

ACKNOWLEDGMENTS

My most sincere thanks go to my committee chair, Yannis Ioannides. Amid the myriad demands of his own work, he has been more than generous in sharing his expertise and enthusiasm.

It has been a genuine pleasure to work with Chuck Michalopoulos, whose patience in explaining matters econometric has been unfailing.

My research, as well as my graduate career, has benefitted from the tutorage and friendship of Catherine Eckel, Nancy Lutz, Anna Hardman and Kristin Butcher. Special thanks go to Dan Nuxoll for stimulating discussions on the mysteries of epistemology and baseball.

One of the joys of graduate work has been the friendship and assistance of my fellow graduate students, especially my classmates Niloy Bose and Wen Mao. Joan Boyd, Barbara Barker and Irene Dowdy have provided gracious staff support over the past five years, as has Teresa Grubb at Radford University. We all get by with a little help from our friends.

My family has provided support, assistance, and most of all, a sense of humor, in regard to this enterprise, and I thank David, Brooke, and Jeff.

CHAPTER 1

LOCATION, INNOVATION AND TRADE: MICHAEL PORTER REVISITED[1]

1 Introduction

Michael Porter (1990) tells the story of the Italian ski boot industry, centered in the town of Montebelluna. The town has a history of producing leather hiking boots, specifically in the small, competitive, family-oriented businesses that the Italian milieu seems to favor. Today, Italian firms based in Montebelluna produce a number of differentiated products: plastic ski boots in a variety of designs, apres-ski boots, and plastic molding equipment. Boot component suppliers as well as the boot makers have specialized, innovated, competed, and gained international success.[2] We can recognize many of the themes of recent international trade theory in the Montebelluna story. Externalities, economic growth, the role of history, and spatial considerations are reflected in Porter's case studies. What we want is a model that ties together those themes and Porter's empirical observations.

Although we often speak, in international trade theory, of a country "producing" a good and then "trading", it is clear that nations do neither. It is *locales*, defined as cities or towns, valleys or corridors, that produce and trade. This paper presents a model built on the notion of locales as the source of the production of goods that

[1]I would like to thank Yannis Ioannides for his patient help. Nancy Lutz, Catherine Eckel, and Charles Michalopoulos made helpful suggestions, as did two anonymous referees.

[2]Porter tells this story in support of his theory linking international trade success to a nation's successful industries in *The Competitive Advantage of Nations* (1990). Many other such "success" stories are documented by case studies conducted in the 1980s.

makes a nation's economy improve. Improvement occurs through increased domestic production and associated real income benefits. These increases are the result of innovation at the locale level.

The identifying characteristic of these locales is the agglomeration of an industry cluster, a group of firms whose products are historically and/or technologically related, and that benefit from locating near one another. These pecuniary externalities are referred to in the urban economics literature as localization externalities and we will use that phrase also. With the example of Montebelluna in mind, we will model a service sector that contributes a differentiated input to a traded goods sector. The traded goods sector produces differentiated outputs that may be either final goods or intermediate ones to its consumers; the primary characteristic is that these goods are traded outside the locale.

At the same time, there is another externality at work. Each locale in a nation potentially benefits or suffers from its national identity; that is, every locale is affected by the underlying advantages and disadvantages that are part and parcel of the country to which it happens to belong. In some cases, these conditions are tangible, such as legal arrangements and trade legislation; in other cases, the conditions are cultural—and much harder to define. Porter says he found this fact repeatedly in his research, even though some conventional wisdom says that national base matters less today because of the globalization of economic activity. Porter points to "generalized factor advantages"—education is a prime example—in addition to more specific advantages and disadvantages. In the Montebelluna example, there are specific aspects of the Italian economy that have favored this industry cluster, and undoubtedly there are other aspects that are less favorable. The inclusion of a national externality in this model reflects the idea that some nations foster the growth of successfully competitive industries more surely than do others.

It is worth noting more precisely what is meant by a successfully competitive industry. Useful examples include ski boots in Montebelluna, Italy, computer goods in Silicon Valley and aircraft in Seattle in the U.S., among many others. These are indeed goods that are traded internationally and contribute to the popular notion of a country's international success. But these goods are also produced for consumption among a nation's many locales. Increased production of these goods in the locales promotes welfare. Therefore what improves a nation's economy through domestic production also helps make it more "competitive" internationally. A competitive industry is one that trades, presumably substantially, outside its locale. Put another way, an empirical look at these industry clusters would use world production share—not just export shares—as a defining characteristic.

The key to this process of becoming successfully competitive, in our model, is the externalities that exist at the locale level. These externalities arise precisely because of the concentration in one geographic area of people interested in the same types of production. Porter credits the rivalry between firms in both the service and traded goods sectors, the communication among firms, the sharing of factor advantages, and perhaps even the consumer demand conditions in a particular area for these externalities. Alfred Marshall, of course, said that a particular industry may be "in the air" in a particular locale.[3]

Marshall also suggested that the "near neighborhood" effect would cause sub-

[3]Marshall explains the situation in Book IV, Chapter 10, p. 225, *Principles of Economics*, 8th edition: "When an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from near neighborhood to one another. The mysteries of the trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously. Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas. And presently subsidiary trades grow up in the neighborhood, supplying it with implements and materials, organizing its traffic, and in many ways conducing to the economy of its material."

sidiary businesses to grow up in the locale. One way, therefore, to model the localization externality is through the variety of products produced by a service sector. The more goods produced in this sector the greater the localization externality that leads to competitiveness. In the urban literature (see in particular Francisco Rivera-Batiz, 1988a and 1988b, and Abdel-Rahman and Fujita, 1990), this is modeled by making the service sector monopolistically competitive; each firm in that sector supplies a differentiated product to the traded goods sector.[4] We extend the notion of a localization externality by specifying how and where innovation occurs, and how its impact is realized.

This model departs from the urban literature in making the traded goods sector monopolistically competitive also. This enables us to express the diversity of output that arises from each industry cluster in each locale. Furthermore, the production function of the traded goods sector is affected by what we call the nation-based externality, modeled as a shift factor. Instead of trying to make this factor a function of the diverse elements that constitute the nation-based externality, we look at "results". If output is increasing in the traded goods sector, then the shift factor reflects this success.

As we will see, there are two elements that make this shift factor important: first, the amount of output produced will increase the externality; and secondly, the factor is configured to capture the national milieu. By national milieu, we refer to the variety of elements suggested above: education in various forms; attitudes toward risk-taking, toward international trade, toward competition; institutions and their ability to adapt; and perhaps, government policy. If, for example, a nation has, in

[4]Rivera-Batiz (1988a) models, and justifies, p.127, a monopolistically competitive service sector. This treatment looks at agglomeration externalities on both the consumption and production sides. Rivera-Batiz (1988b) emphasizes the production side. Abdel-Rahman and Fujita's similar model (1990) asserts that a variety of intermediate goods increases productivity.

the past, enjoyed significant output, the shift factor may be large. Another nation may benefit from the shift factor because its national milieu has changed for the better, even though the initial output of the traded-good firms is small. Perhaps in another country, the national milieu has become less conducive to the growth of firms, in which case the actual output of competitive firms will not have as much impact as if the environment were more positive. We will consider these possibilities and others in light of the nation-based externality, below.

We make all of this specific by looking at the production functions in the locales and the demand for traded goods in Section 2. In Section 3, we consider the implications of innovation and look at two resulting effects. The final section, 4, looks at development categories and international trade cases. These cases are designed to clarify the role of the nation-based externality; a separate exposition of international trade based on the locale model presented here is the topic of another paper.

2 The Model for Locales

2.1 Production in Locales

Each locale in a nation potentially has three sectors of production. First of all, all locales produce nontraded goods, which we will bundle together and think of as a single good, denoted by G_{lt}, where the subscript l identifies the locale and t the time. This production satisfies the demand by local residents. In the urban literature, it is often thought of as housing, or services which are not easily traded (like the proverbial haircut); in looking at Standard Industrial Classification codes, we would identify many goods in the retail trade categories as non-traded.

The locales in which we are interested go beyond this local production. They produce a traded good, by which we mean traded internationally as well as nationally outside the borders of the locale. Locales do this with the help of a *service-sector*, which specializes in producing varieties of intermediate goods. These goods are sold as inputs to the firms that make up the industry cluster, or *traded goods* sector. We denote the output of a firm in the traded goods sector as x_{jlt}, where the subscripts denote the firm producing variety j in locale l at time t. The jth firm is one of m_{lt} firms in locale l; there are n such locales in a given country. We will often refer to that sector as the X-sector.

The output of the service sector, or Q-sector, by firm i, will be denoted by q_{it}. (The locale with which Q-sector firms are associated is apparent from the context.) The ith firm is one of δ_{lt} such firms in the service sector of locale l.[5]

We will use the non-traded good produced in time t, G_{lt}, as the numeraire. We assume that production of these non-traded goods occurs with constant returns to scale; it is a function only of the labor used in that sector, L_{Glt}. We write

$$G_{lt} = L_{Glt}. \tag{1}$$

2.1.1 Production in the X-sector

The most interesting sectors are the ones involved in production of traded goods. We start with the actual X-sector, where the production function for the jth firm at time t is assumed to be a product of three factors. The first factor, A_t, represents our nation-based externality, modelled as a shift factor. We will examine the form of this

[5] We are not explicitly modeling the fact that each locale supports only one industry cluster (and its attendant service sector). Following the urban economics literature we could model this based on the tension between the centripetal force of localization externalities and the centrifugal force of consumption disamenities. However, that model has been perfected by a number of authors, (see for example, Henderson (1988)), and would detract from the basic idea we are pursuing. Therefore, we are assuming the presence of only one industry cluster per locale. See also Krugman (1992) and Arthur (1990) for interesting perspectives on industry clusters.

factor more closely below; we will treat it as a parameter while we are considering localization externalities. L_{xjt} is the labor employed by the jth firm in the X (traded goods)-sector at time t. The q_{jit} is the amount of input q from firm i (in the Q or service sector) that firm j in the X-sector uses at time t. The jth firm's production function includes up to δ_{lt} such inputs, with δ_{lt} being the number of firms in the service sector in locale l at time t.

We assume that:

$$x_{jlt} = A_t L_{xjt}^{\alpha} (\sum_{i=1}^{\delta_{lt}} q_{jit}^{\rho})^{\frac{\beta}{\rho}}. \tag{2}$$

Because the time element is not useful to us until we discuss the shift factor in Section 2.3, we drop the time subscripts until then.

At any given time, the quantity employed of some input may be zero. We also assume that labor and the service sector inputs enter the production function in a Cobb-Douglas manner, so that $\alpha + \beta = 1$. [See Rivera-Batiz (1988a and 1988b), and Abdel-Rahman and Fujita, (1990), for a similar treatment.] These service sector inputs enter through a constant elasticity of substitution expression, which is equal to ρ. That allows us to say something about the amount of diversity desired among users of the intermediate goods. This is one way to express the localization externality: the greater the diversity of these products, the greater the positive spillover in the industry cluster. The diversity is measured by ρ, which we assume to be between zero and one. The smaller ρ is, the more diversity we see.

These service sector goods enter symmetrically into the production function of firm j (in the X-sector), an assumption that gives this particular set up its mathematical tractability. The X-sector firms are price takers for the input from Q because of the large number of firms in both sectors.

The fixed cost that each firm in the sector faces, denoted by f_{xl}, is of interest to us because it carries part of the "historical" explanation for industry location. As expected, these fixed costs represent a barrier to entry; however, we interpret the fixed costs to represent the history of the locale. That is, the fixed costs for the industry that has located in l stand as a barrier to firms producing the very same goods in a different locale, and thus embody an historical element; the fixed costs seem to make it cheaper for a firm interested in producing x_{jl} to be located in locale l than to be located elsewhere because of unique factor and demand conditions. This will encourage a greater number of such firms to locate in l. Thus f_{xl} is an exogenous variable in the sense that it is historically determined for locale l.[6] It will be helpful to obtain an expression for the variable cost function, c_x, for the typical firm in the X-sector, in order to determine the amount of q_{ji} that each jth firm will demand. The cost function, as a function of w, the wage paid to labor in each of the sectors (at equilibrium), p_{qi}, the price of the service sector good produced by firm i, and the amount of x_j that is produced, is defined as:

$$c_x(w, p_{qi}, x_{jl}) = min_{q_{ji}} : wL_{xj} + \sum_{i=1}^{\delta_l} p_{qi}q_{ji} \, , \tag{3}$$

subject to the production function for x_{jl}, given in equation (2). Solving for this cost function is a two-step process, considering first the CES part of the function and secondly the overall Cobb-Douglas expression. Minimizing accordingly yields

$$c_x = A^{-1}x_{jl}Bw^{\alpha}[\sum_{i=1}^{\delta_l} p_{qi}^{\frac{\rho}{\rho-1}}]^{\frac{\beta(\rho-1)}{\rho}} \tag{4}$$

$$where\ B = [(\frac{\alpha}{\beta})^{\beta} + (\frac{\beta}{\alpha})^{\alpha}].$$

[6]By definition, a fixed cost should not vary with the amount of labor or capital used in the production of the traded good. Therefore, we have to specify what inputs are truly fixed in the short run. The spirit of this historical interpretation of fixed costs would lead us to consider land-type resources. Certainly there are locales that enjoy a natural advantage, which may be continuing (such as a port), or strictly historical (as in depleted mineral resources). However, we would like to also include less tangible factors. When Marshall (1890) spoke of an industry being "in the air", he was speaking of a short-run fixed cost, in our sense of the term. He also was speaking of people. We propose that a community inclination, manifested in either interest or skills, is fixed in a locale in the short run. This is not the same as the amount of labor used, nor is it amenable to change in the short run, (although it may do so in the long run).

These are the total variable costs that a firm in the X-sector faces. [7]

The derived demand for q_{ji}, as perceived by the ith firm, is given by:

$$q_{ji} = A^{-1}\,(\frac{\beta}{\alpha})^{\alpha} p_{qi}^{\frac{1}{\rho-1}}\,(\sum_{i=1}^{\delta_l} p_{qi}^{\frac{\rho}{\rho-1}})^{\frac{-(\beta+\alpha\rho)}{\rho}}\, x_{jl}\, w^{\alpha}\,. \tag{5}$$

It follows that the elasticity of demand facing the ith producer in the Q-sector is

$$\epsilon_q = \mid (1-\rho)^{-1} \mid, \tag{6}$$

which indicates that the elasticity of the demand facing any one producer depends on the desire for diversity from X-sector firms. As noted before, ρ expresses the constant elasticity of substitution: if it is close to one, there is little desire for diversity and if it is closer to zero, there is more desire for diversity.

The derived demand in equation (5) is a function of x_{jl}, the level of output of the jth firm. That level can be determined because of the monopolistically competitive nature of the industry. Free entry and exit of firms in this industry cluster indicates that equilibrium profits of the jth firm, π_{xj}, will be driven to zero. Using p_{xj} to indicate the price charged by the jth firm, we have:

$$\begin{aligned}\pi_{xj} &= x_{jl}\, p_{xj} - f_{xl} - \\ &\quad x_{jl}\, A^{-1} B\, w^{\alpha} (\textstyle\sum_{i=1}^{\delta_l} p_{qi}^{\frac{\rho}{\rho-1}})^{\frac{\beta(\rho-1)}{\rho}} = 0.\end{aligned} \tag{7}$$

Equation (7) can be solved for x_{jl} and the resulting expression is used in equation (5) to yield the derived demand for q_{ji} as a function of the Q firm's own price, indices of the prices charged by all other Q-sector firms, wages, fixed and variable costs in the X-sector, the price charged by firm j, and parameters. We consider first production in the Q-sector.

[7]This and other mathematical derivations have been eliminated due to space considerations. All derivations are available from the author on request.

2.1.2 Production in the Q-sector

The Q-sector (service sector) is also monopolistically competitive. We assume that the sector uses only labor, L_{qi}, as input, and that the labor input involves both a fixed and variable component, such that:

$$L_{qi} = f_{ql} + c_{ql}\, q_i\,. \tag{8}$$

In reality, we should think of L_{qi} as involving both the labor and the natural resources "native" to a region, although we use labor as an approximation of these factors. Fixed costs in this sector reflect underlying, perhaps historic, conditions in the factor market. These fixed costs make it cheaper for a firm interested in producing x_l to be located in locale l than to be located elsewhere because of the unique factor conditions.[8] The variable component, c_q, will become important below as we attempt to specify the source and effects of innovation.

Profit in the ith firm in the Q-sector is

$$\pi_{qi} = q_i\, p_{qi} - w(f_{ql} + c_{ql}\, q_i)\,. \tag{9}$$

We insert the expression for q_{ji} (equation (5)), and maximize the resulting expression with respect to p_{qi} in order to find the price that a service firm will charge. In doing so, we assume that there are a large number of Q-sector firms. Therefore, they take the price index, $\sum_{i=1}^{\delta_l} p_{qi}$, for all firms in the sector as given.

Solving yields:

$$p_{qi} = \frac{w\, c_{ql}}{\rho}\,. \tag{10}$$

It is clear that the price is specific to the Q-sector (in locale l), but the price charged by each ith firm is the same. If the price charged by each firm is the same,

[8] See footnote (6) on page 8.

as are the fixed costs and the marginal coefficient of labor, then each firm produces the same amount, and q_{ji} will simply become q_j. In fact, because the inputs enter symmetrically, we may simply denote each firm's output by q. Each X-sector firm now will have the following production function:

$$x_{jl} = A L_{xj}^{\alpha} \delta_l^{\frac{\beta}{\rho}} q^{\beta}. \tag{11}$$

In order to say more about this sector, we must have a demand function for the X-sector's traded goods.

2.2 Demand for Traded Goods

Modeling demand for a good that is traded world-wide calls for heroic assumptions. The demand assumptions used here are simple and, indeed, heroic. In order to make the process clearer, we assume in this section and below, that there are only two locales in the country we consider.

First of all, the traded goods sectors in locales across the world produce goods that would be demanded by both consumers and businesses. However, we begin the derivation of demand by assuming that demand originates with identical consumers, located in any one of the locales; we begin by looking at a consumer in locale 1, but the process is identical regardless of which locale we choose. This consumer demands three goods: one of these is the non-traded good produced in locale 1, which is G_1. The consumer also has a demand for the line of traded goods produced in her own locale, and for the line of goods produced in locale 2. We express this as:

$$U_1 = G_1^{1-\sigma-\lambda} (\sum_{j=1}^{m_1} x_{j1}^{r_1})^{\frac{\lambda}{r_1}} (\sum_{j=1}^{m_2} x_{j2}^{r_2})^{\frac{\sigma}{r_2}} \qquad o < r_1, r_2 < 1, \tag{12}$$

where λ is the share of income going to goods produced in locale 1 and σ is the share going to goods produced in locale 2.

Each representative consumer maximizes utility subject to a budget constraint, where p_{g1} is the price of the combination bundle in locale 1 and y is the income of the representative consumer:

$$G_1 p_{g1} + \sum_{j=1}^{m_1} x_{j1} p_{j1} + \sum_{j=1}^{m_2} x_{j2} p_{j2} \leq y. \tag{13}$$

We maximize this expression to find the demand for one firm's output; if we are interested in a firm in locale 2, for example, we maximize with respect to $x_{j,2}$, the output of a single firm in the X- sector of locale 2. We get:

$$x_{j2} = \frac{\sigma y}{(\sum_{j=1}^{m2} p_{xj2}^{\frac{r_2}{r_2-1}}) p_{xj2}^{\frac{1}{1-r_2}}}. \tag{14}$$

This give us the individual firm's price elasticity of demand:

$$\epsilon_{x2} = \mid (1-r_2)^{-1} \mid, \tag{15}$$

indicating that the elasticity is a function of the desired diversity among consumers of the traded good produced in locale 2.

Profit maximization requires that the jth producer set marginal revenue equal to marginal cost. We have an expression for total variable cost in equation (4) above; if we divide by x_{jl}, we

have an expression for average variable cost (equals marginal cost),

$$\bar{c}_{x2} = A^{-1}\, B\, w^{\alpha}\, \delta_2^{\frac{\beta(\rho-1)}{\rho}}\, p_q^{\beta}, \tag{16}$$

so that

$$p_{xj2}\,(1 - \frac{1}{\epsilon_{x2}}) = \bar{c}_{x2} \tag{17}$$

Solving for p_{xjt} we find that

$$p_{xj2} = \frac{\bar{c}_{x2}}{r_2} \quad or \quad p_{xj2} = \frac{B w^{\alpha} p_q^{\beta}}{r_2 A \delta_2^{\frac{\beta(1-\rho)}{\rho}}}. \tag{18}$$

Because all firms in the sector have the same variable cost,

$$p_{xj2} = p_{x2} \,. \tag{19}$$

By the symmetrical argument, based on a good produced in locale 1, we would find that:

$$p_{xj1} = p_{x1} \,. \tag{20}$$

If the price charged by each firm in locale 1's X-sector is the same, we would expect the quantity demanded of each firm to be the same. The symmetrical argument holds for a firm located in locale 2. Therefore, the demand for a locale 2 firm's output may be expressed as:

$$x_2 = \frac{\sigma y}{p_2 m_2} \,. \tag{21}$$

and the demand for a firm in locale 1 would be:

$$x_1 = \frac{\lambda y}{p_1 m_1} \,. \tag{22}$$

This says that demand for the product of a firm in the X-sector of locale 2 varies directly with the market share enjoyed by the locale 2 industry, and with the income of a representative consumer; and inversely with its own price, and with the number of firms in the locale 1 industry. The latter implies (in Chamberlinian fashion) that increased demand for one locale 1 firm's product may reduce the demand for another's, a point to which we will return later.

The expressions above give us the demand for traded goods in from the two locales. We want the expressions for goods supplied, and we use the free entry zero- profit condition to see that each firm in the X-sector produces

$$x_l = \frac{f_{x,l}\, r_l}{\bar{c}_{xl}(1 - r_l)}, \qquad l = 1, 2. \tag{23}$$

That is to say that an X-firm's output is inversely related to both its average variable costs and the desired diversity among the industry cluster's products. (Recall that a decrease in r represents an increased desire for diversity, as was the case with ρ.) Setting the expressions in equations (21) and (22) equal to the expression in equation (23) gives us the number of firms, m_l in each locale:

$$m_1 = \frac{\lambda y(1 - r_1)}{f_{x1}}, \tag{24}$$

and

$$m_2 = \frac{\sigma y(1 - r_2)}{f_{x2}}. \tag{25}$$

These equations say that the number of firms in a locale's industry cluster is directly proportional to the locale/industry's market share, and to the preference for diversity for the goods produced in the industry; and inversely proportional to the fixed costs in the locale.

Given these expressions, we can find the entire output of a traded good in a locale, by multiplying the output of a single firm times the number of firms. We let the capital letters denote output for the locale:

$$X_1 = \frac{\lambda y r_1}{\bar{c}_{x1}} \text{ and } X_2 = \frac{\sigma y r_2}{\bar{c}_{x2}}. \tag{26}$$

These expressions for output include, in the average variable cost expression, δ_l, the expression for the number of firms in the Q-sector in each locale, which we can now determine. The free entry, zero profit condition in the Q-sector yields the quantity that will be supplied in the locale (where the S indicates supply):

$$Q_{Sl} = \delta_l \frac{f_{ql}}{c_{ql}} \frac{\rho}{1 - \rho}. \tag{27}$$

We define the quantity demanded of the entire Q-sector in a locale by multiplying the number of firms in the sector times the demand per firm from equation (5). Instead of x_{jl} in equation (5), we use X_2 from equation (26) above. Therefore,

$$Q_{D2} = p_q^{-1} \beta \sigma r_2 y, \tag{28}$$

and the expression for Q_{D1} is symmetric. Setting the expressions for supply and demand equal, we can find the number of firms in the Q-sector in each locale:

$$\delta_2 = \frac{\beta\sigma r_2(1-\rho)}{w f_{q2}}, \tag{29}$$

and

$$\delta_1 = \frac{\beta\lambda r_1(1-\rho)}{w f_{q1}}. \tag{30}$$

We see that the number of firms in the service sector is inversely related to the desire for diversity *for* X-sector goods, but directly related to the desire for diversified inputs *from* the X-sector. The parameter indicating the share of Q-inputs in the manufacturing process, β, has the expected relation. The higher the fixed costs, f_{ql}, an arbitrary constant according to our historical explanation, the lower the number of firms in the sector.

The total output in the X-sector in locale 2 now becomes:

$$X_2 = \frac{\sigma r_2 A \rho^{\beta}}{B w c_{q2}^{\beta}} y [\frac{\beta\sigma r_2(1-\rho)}{w f_{q2}}]^{\frac{\beta(1-\rho)}{\rho}}, \tag{31}$$

and

$$X_1 = \frac{\lambda r_1 A \rho^{\beta}}{B w c_{q1}^{\beta}} y [\frac{\beta\lambda r_1(1-\rho)}{w f_{q1}}]^{\frac{\beta(1-\rho)}{\rho}}. \tag{32}$$

It is interesting to note that a decrease in fixed costs would actually *lower* output by a single firm. At the same time, it would increase the *number* of firms. In equations (31) and (32), we do not see f_{xl}, because total output is the product of output per firm times the number of firms. This indicates that a decrease (or increase) in fixed costs is a "wash", industry-wide. We use this expression in Section 3, but before we move on, we should look at the role of the nation-based externality.

2.3 The nation-based externality

The nation-based externality has been designated by the shift factor A_t thus far. To amplify the remarks at the beginning of the paper, we note that we are interested in those aspects of a national milieu that promote growth in industries in various ways. Marshall noted this externality, asserting that:

> In sketching the history of free industry and enterprise we have already incidentally traced the outline of the causes which have localized the industrial leadership of the world now in this country and now in that. We have seen how physical nature acts on man's energies; ... we have also seen how the use he makes of these advantages depends on his ideals of life, and how inextricably therefore the religious, political and economic threads of the world's history are interwoven; while together they have been bent this way or that by great political events and the influence of the strong personalities of individuals. (*op. cit.*, Book IV, Chapter 10, pp. 224-25.)

Porter (1990) describes several such situations in modern economies. He notes that some countries do a better job than others in providing generalized advantages, such as quality public education. This is a *generalized* advantage because it applies fairly equally to all industries. Some nations, he adds, also provide specific advantages such as university programs and institutes that provide training and research in a particular industry specialty. We are concerned with nation-wide, generalized factor advantages in this formulation.

Because of the diverse aspects of this externality, we measure the nation's advantage by the growth of output in the traded goods sectors of either (or both) of our locales. In other words, because we cannot possibly specify all the factors that promote growth in various national environments, we consider the result of those environments, a technique suggested in the learning by doing models of Arrow (1962), Uzawa (1965) and Romer (1986). We define $\bar{X}_t$ to be the sum of the production of the traded-good industries in all (both) locales at time t, and therefore,

$$A_t = \gamma \int_0^t \bar{X}_\tau \, d\tau \,, \tag{33}$$

where γ is a positive-valued parameter designed to reflect those factors that promote or hinder growth as discussed above. A at time t reflects past outputs across locales and across time from $\tau = 0, \ldots, t$ As we have said above, there are a variety of such factors: education, attitudes, government policy, etc. One nation may enjoy a significantly large and positive γ at a point in time because of one subset of factors; another country may enjoy an advantage because of a different subset of factors. A lower value of γ implies that a country's policies are actually reducing the output growth of its industries, and this will help explain our "failure breeds failure" scenario below.

Clearly, $\bar{X}_t$ is proportional to A_t, and we denote the factor of proportionality as k. For our purposes below, we let

$$k_t = (c_{q1t}^{-\beta} + c_{q2t}^{-\beta})\,(y_t w_t^{\frac{\beta(\rho-1)}{\rho}-1})I \,, \tag{34}$$

where I is an index of time-invariant parameters and fixed costs in the Q-sectors. We note that

$$\dot{A} = \gamma \, k_t \, A_t \,, \tag{35}$$

and therefore

$$A_t = A_0 \, e^{\gamma k t} \,, \tag{36}$$

which is to say that the growth rate of the nation-based shift factor, $\dot{A}/A$,is γk.

We see that A_t, the nation-based externality shift factor can change for either of two reasons: unanticipated shocks cause changes in either k or γ. We discuss the comparative statics of this formulation below.

3 Innovation and Comparative Statics

In order to consider the benefits of both localization and nation-based externalities to a locale's economy, we need to think carefully about the role of innovation in this model, and about the interrelationships of the variables we have developed.

Key equation are (31) and (32), which give us total production in the traded good sector of a locale. We use this equation to see the impact of A_t, the shift factor, and c_{ql}, the variable cost factor from the Q-sector.

3.1 Specifying innovation

Rivera-Batiz, and Abdel-Rahman and Fujita, *op.cit.*, use a monopolistically-competitive service sector to convey the idea that a larger variety of services increases the *productivity* of the traded-goods sector. The Dixit-Stiglitz production that they (and we) use embodies this assumption. We can see this most clearly in our model in equation (23), the equation which gives us the output of a single X-sector firm. We rewrite it here in expanded form to highlight the role of δ_l, the number of firms in the Q-sector in locale l:

$$x_l = \frac{f_{xl} r_l}{(1 - r_l)} \frac{A\delta^{\frac{\beta(1-\rho)}{\rho}} \rho^{\beta}}{Bwc_{ql}^{\beta}}, \tag{37}$$

The presence of more Q-sector firms, which means that there are more inputs available for the X-sector, increases the output of each X-sector firm (and of the X-sector as a whole). However, the question remains: why do a greater variety of inputs increase productivity? We follow the suggestion of Porter—which is related to the earlier suggestion of Marshall, as quoted in the footnote in Section 1—that the key is innovation. New products (new inputs) are generated in the Q- and X-sectors because the concentration of people in related businesses leads them to innovate. While we recognize that there are potentially other forces at work, and other benefits of the localization externality to be had, we assume that the localization externality in this model depends on innovation. It is helpful in this case to distinguish between *process* innovation and *product* innovation. Process innovation is the advantage that comes from *producing* a good in an innovative way. Product innovation is the advantage that results when firms make a better *end-product.* We will use process innovation to illustrate our comparative static effects.

Put most simply, process innovation reduces production costs. Perhaps a new technology is developed that reduces costs. Perhaps someone simply sees a better way of applying existing technology so as to reduce costs. In the real world, we might see the results of process innovation in increased output per worker. Whatever the particular nature of the process, we have several options in expressing it in this model.

We could assume that an increase in process innovation decreases either the variable cost or fixed cost components in the Q-sector, or the fixed cost components in the X-sector.[9] However, we have chosen in this model to let f_{ql} and f_{xl} carry the "history" of the locale. Therefore, we will assume that process innovation lowers the variable cost component in producing Q-sector goods.

If we say that an increase in innovation reduces c_{qlt}, we can show precisely the increase in production of x_{jl} by the jth firm in the X-sector. Equation (10) tells us that a reduction in c_{ql} lowers the price charged by a representative Q-sector firm. Equation (16) shows us that the reduction in this price lowers average variable cost for the X-firm, (because the price of q is part of the variable cost component). We can see this effect summarized in another way in equations (31) and (32), in which a decrease in c_{ql} clearly increases output throughout the X-sector in both locales.

[9]For another treatment of fixed and variable cost components in a monopolistic competition model, see Ioannides, 1994. There the fixed and variable cost components in the single monopolistic competition sector are functions of manufacturing capital.

Therefore, by specifying the role of innovation in the localization externality, we can see clearly the predicted increase in productivity in the X-sector.

But there is another impact on output because of innovation: A_t increases through the impact on the auxiliary variable, k, of decreasing variable costs, increasing X-output. Note the role of A_t in equation (36). If either c_{q1} or c_{q2} decreases, k increases. This increases A_t, and further increases output. Of course, for this positive feedback to have its optimal effect, γ must be significantly large. It is in this sense that a "good" national environment acts positively on the potential for increases in output. We compare situations of "good" and "bad" national environments, below.

3.2 Resulting Effects

We designate the initial effect of process innovation in terms of efficiency; we designate the second result of process innovation in terms of its national externality effect, and discuss each in turn.

3.2.1 The efficiency effect

By the efficiency effect, we mean the kind of increase in productivity that initially comes about because of process innovation. As we noted above, a decrease in c_{ql} increases equilibrium output by a representative X-firm, by the entire X-sector, by a representative Q-firm, and by the Q-sector. More output is achieved without using more of our fundamental resource, labor. To see this, consider once again at equation (8), which tells us the amount of labor need by each Q-firm, and equation (28), which gives us the demand for the Q-sector's product. We see in equation (28) that the variable cost, c_{ql}, cancels out, leaving us with no change in labor required in the Q-sector. Employment does not increase with the increased productivity of intermediate inputs in our Cobb-Douglas formulation of the X- sector, and the number of firms in the X-sector remains the same.[10] Only output increases with process innovation, telling us that this one locale produces more efficiency with its share of the world's resources. However, another benefit follows from the increase in efficiency.

The increased output is reflected in a decline in p_x, the price of the X-sector good. If c_{ql} falls due to process innovation, p_q falls (equation (10)). That lowers the cost of producing traded goods in the X-sector and their price (equation (18)). Hence, real incomes increase.

3.2.2 The national externality effect

Process innovation potentially leads to what we may call the national externality effect. Because innovation leads to an increase in k, A_t increases—*provided* that the national environment supports a significantly large γ. See equation (36). Because the increase in A_t boosts output in each locale, prices of all traded goods fall. Again, real incomes increase.

We emphasize that increased innovation in one locale, made possible by the geographic concentration of an industry, benefits the entire economy—the innovating locale and the *other* locale in our two-locale case. This happens in this model because an increase in the output in any locale is understood to be an indication that the national milieu favors innovation. Whatever is good in the national milieu is therefore available to each and every locale; we do not know the cause of this favorable milieu in this model, but we see its impact in this national externality effect.

To summarize our two effects, process innovation initially increases output, not

[10] We can easily derive the total labor supply needed in the locale in all three sectors; that expression is $L_{lt} = \frac{\sigma_{lt} r}{w_t} + G_{lt}$. We present this to emphasize that labor force does not increase with a decline in c_{qlt}.

by increasing employment, but through a more efficient use of resources. Process innovation first improves output directly; process innovation also works through the national externality shift factor to further increase output in the innovating locale and in all other locales. The price of an X-sector good falls, increasing real income to the extent that the representative consumer buys traded goods.

In this model, an increase in one nation's standard of living is not part of a zero-sum game; opportunities exist for specialization, localization externalities and growth of income in a number of countries at the same time. *Which* countries gain and which gain the most may depend on γ-type factors as well as historical antecedents embodied in f_{ql} and f_{xl}. We will show this specifically in Section 4.2 as we consider the possibilities for trade between locales and between countries.

4 National Production and Trade

The purpose of this section is not so much to provide a model of international trade as it is to clarify, by example, the workings of the locale model presented above. In so doing, we can illustrate Porter's categories of economic development. These categories can usefully be related to our nation-based externality. Furthermore, the categorization helps us discuss the patterns in international trade that may be examined through the locale model. We present Porter's four categories of production in the first subsection and then consider a brief application to international trade.

4.1 Types of Production

Porter (*op.cit.*, Chapter 10) asserts that nations may be (roughly) grouped into four categories: factor-driven, investment-driven, innovation-driven, and wealth-driven production.

Factor-driven production depends on favorable natural resources or a large semi-skilled labor pool. Investment-driven production implies that locales in a country adopt, and improve upon, foreign technology. Innovation-driven production is the ideal for Porter. Locales produce more output as innovation proceeds. Furthermore, the number of locales producing traded goods may increase. In our model, this represents the γ factor at work, although we do not specify a mechanism for developing new X-producing locales.

In short, this is the production category that our model was designed to represent: a "success breeds success" mechanism based on production and agglomeration at the locale level and aided by national factors that enhance productivity. If γ is significantly large, then we have set the stage for "success breeds success". And of course, we expect to see such a pattern if innovation is occurring at the locale level. *How much* success we see depends on the size of γ. In a very successful economy, such as the Japanese, we would expect to have a large γ factor, reflecting the environment that "works" for the Japanese. [11]

Innovation-driven economies account for many of the most successful international trading nations, as well as many of the most productive economies. Porter asserts that innovation involves thinking in terms of global markets. We will consider international trade within the framework of our model below.

The last of Porter's categories is not one to which nations should aspire. The wealth-driven stage is one in which innovation wanes and the economy is driven—less well—by wealth accumulated in a previous stage. In our terms, the key characteristic

[11] We emphasize once again that certain basic factor enhancements are probably important to any economy, education being a good example. On the other hand, other aspects of the national environment may well be appropriate only to that culture; what works for the Japanese economy might not work for a Western economy.

is that γ is falling. The γ factor may decline due to less productive institutions, poorer education and training, less effective risk-taking, and other cultural issues. In a sense, γ falling represents a sort of "hardening of the innovation arteries". This is the stage in which our model allows "failure to breed failure" because output is directly related to A_t. Perhaps market shares are being captured by locales engaged in innovative production in another country. Porter says this category describes Great Britain (and warns that Switzerland, Germany and the U.S. may be in danger of entering this category).

Not every country goes through each category, according to Porter: some never leave the factor-driven category and others combine factor-based advantages with innovative locales. However, we may take these categories as useful for the discussion of international trade which follows.

4.2 International Trade Scenarios

Porter asserts—and a casual survey of international trade numbers confirms—that most of the world's trade is carried on by countries in the innovative stage, which is to say, by the industrialized nations.

Using insights from our model of production, we can suggest possible trade scenarios.

Let us consider first trade by two countries, each having a number of locales that produce a cluster of goods available for trade. One country has n traded-good industries, the other n^*. The n's, of course, indicate not only the number of industries, but also a set of products in the industry cluster. So, we will let the n's also represent sets of products in the following discussion, and whenever no confusion arises.

The degree to which these two countries can profitably trade depends on the degree of overlap between the sets n and $n^\star$. If n does not coincide with $n^\star$, then the industries (locales) of the two countries can "mesh". The traditional benefits of specialization and trade accrue.

However, if $n \subset n^\star$, trade conflicts must result. The advantage in these conflicts goes to the nation which most actively supports its industries—which is to say, the country with the larger shift factor, as measured by γ.

Suppose that locale 1 in Country 1 produces the same cluster of products as locale 2 in Country 2. Innovation occurs in Country 1's locale 1, increasing output and lowering prices directly from that locale. Let us assume further that the country's economy, through a very positive national environment, encourages a substantial increase in production through the shift factor; γ is a large positive number.

As a result of innovation in locale 1, Country 1, output also increases (and prices fall) in other locales in that country. Real incomes increase. To the extent that consumers in Country 2 buy goods from Country 1 (in the non-competing industries), real incomes in Country 2 rise also. The negative impact is on the producers in locale 2 of Country 2, who are now at a price disadvantage relative to similar producers in Country 1.

Here, the national environment may come into play: Porter identifies a willingness to think globally, a mindset which may be part of the national milieu, or which may be part of the innovation mentality in particular locales. In any case, innovation is the response to competition in these economies.

Free trade is clearly welfare-enhancing for consumers as well as firms if $n \neq n^\star$. In a truly innovative economy, free trade is also welfare-enhancing if $n \subset n^\star$: consumers benefit from the competition, and firms benefit if competition forces them to innovate.[12]

If we enlarge this scenario to include trade among all the globe's innovation-driven economies, we have a continually shifting picture of many firms in many locales, across countries, innovating and competing both domestically and internationally. That different countries pursue success in different patterns reflects differences in γ; there are many roads to success in this model. Failure is what happens when a country's firms, by and large, fall out of the innovative-production mode.

What about that part of the world's trade that happens between newly-industrialized and industrialized nations? In our terms, trade exists between countries whose locales are in the investment-driven stage and those whose locales are primarily in the innovation-driven stage. It may well be that the developing country is taking over a developed nation's industry in which innovation is no longer occurring. This clearly sets up a conflict between locales within one country and the country which is seeking the investment-type production. This shifting of industrial strength goes back to the earliest discussions of comparative advantage, of course; but casting the problem in terms of locales helps us to understand conflicts such as political opposition in the United States to the North American Free Trade Agreement.

This model, and modifications to it, suggests a number of international trade scenarios, some of which are pursued elsewhere. It suffices to say here that if we want to consider the sources of production within an economy engaged in international trade, we need a model that makes locales a primary feature.

[12]The details of this trade scenario are the subject of another paper.

5 Conclusions

We have developed a model that integrates both localization externalities and a nation-based externality to see how nations may increase their output and standard of living. One advantage of the two, integrated, monopolistically competitive sectors is that the model provides the opportunity to see most clearly where innovation occurs, and where and how it gives rise to the localization externality. Increases in efficiency and output follow from the existence of externalities at the locale level. We have suggested that these two types of effects, the efficiency effect and the national externality effect are each related to process innovation.

The combination of two monopolistically competitive sectors with fixed costs gives us the opportunity to emphasize the role of history in attracting firms to a given locale. This locational feature is an ongoing topic of study in the urban economics and economic history literature. We have let our fixed costs in the service sector carry the idea of superior resources, the factor advantage that is part of a location's draw. We let the fixed cost in the traded goods sector express the idea that firms interested in the production of a good in a particular industry cluster will find it cheaper to produce in a location that favors that production—for any number of reasons.

The model of monopolistic competition in two sectors also sheds light on the relationship between locales within an economy. We show that differences between nations lie in the different ways in which the national environment affects locales.

One contribution of this paper is to put parts of Michael Porter's theory and research into a mathematical model that highlights certain of his key ideas. The role of innovation, geographical concentration, historical factors, categories of development and the national contribution to competitiveness are given specific expression in the model. The construction of the model is especially suitable to industrialized economies, the most successful of which are in an innovative stage, according to Porter. It is clearly welfare-enhancing for as many countries as possible to be in this stage, and our model suggests one rather vague avenue to achieve that goal: through nationally-appropriate externality factors. Our model also suggests how some economies may prosper by moving into the investment-driven stage, although this process probably requires even greater precision in finetuning the nation-based externality factor, from a governmental perspective. (That is because policies that favor such investment may have to be industry-specific as opposed to generalized.)

There are a number of issues in international trade with which this model does not directly deal at present, although its construction may suggest a format for discussing those matters. The model does not provide a mechanism for goods that "move" from the service-sector to the traded-goods-sector, like the plastic molding equipment that is sold internationally from Montebelluna. However, we can certainly view the traded goods sector in a different light if we realize the possibility of such movement.

This model does not allow for foreign direct investment or location of manufacturing outside the home-base nation. But it does present the opportunity to note that foreign-based manufacturing, while clouding our notion of "locale", does not negate it. A company that locates in a foreign country chooses its "extended" locale for reasons of factor abundance or market advantage, either of which is consistent with our basic formulation. Furthermore, a company plant in a foreign country may well be operating within the confines of its national personality just as it did back home. Legal arrangements change, but the motivations and characteristics that govern the management of a foreign location may not vary with the move.

There are undoubtedly other issues of international concern not raised here, but the model represents an attempt to see international trade in "locale terms", by putting the emphasis on where production occurs and why.

REFERENCES

Abdel-Rahman, Hesham, and Masahisa Fujita (1990) "Product Variety, Marshallian Externalities, and City Sizes", *Journal of Regional Science* 30:165-183.

Arrow, Kenneth J. (1962) "The Economic Implications of Learning by Doing", *Review of Economic Studies* 29:155-73.

Arthur, W. Brian (1990) " 'Silicon Valley' Locational Clusters: When Do Increasing Returns Imply Monopoly?", *Mathematical Social Sciences* 19:235-251.

Dixit, A.K., and J. Stiglitz (1977) "Monopolistic Competition and Optimum Product Diversity" *American Economic Review* 67:297-308.

Elizondo, Raul Livas, and Paul Krugman (1992) "Trade Policy and the Third World Metropolis", NBER Working Paper No. 4238, December 1992.

Henderson, J. Vernon (1988) *Urban Development: Theory, Fact and Illusion*, (New York: Oxford University Press).

Krugman, Paul (1992) "A Dynamic Spatial Model", NBER Working Paper No. 4219, November 1992.

Ioannides, Yannis M. (1994) "Product Differentiation and Growth in a System of Cities", *Regional Science and Urban Economics* 24:461-484.

Marshall, Alfred (1920) *Principles of Economics*, Eighth Edition, (Philadelphia: Porcupine Press, 1982).

Porter, Michael E. (1990) *The Competitive Advantage of Nations*, (New York: The Free Press).

Rivera-Batiz, Francisco L. (1988a) "Increasing Returns, Monopolistic Competition, and Agglomeration Economies in Consumption and Production", *Regional Science and Urban Economics* 18:125-153.

________ (1988b) "Modeling Urban Agglomeration: Producer Services, Linkage Externalities and Specialization Economies", in *Modeling and Simulation*, M.H. Mickel and V. Vogt, eds. (Research Triangle Park, North Carolina: Instrument Society of America, 1988): 113-119.

Romer, Paul (1986) "Increasing Returns and Long- run Growth", *Journal of Political Economy* 94:1002-1038.

Uzawa, Hirofumi (1965) "Optimum Technical Change in an Aggregative Model of Economic Growth", *International Economic Review* 6:18-31.

CHAPTER 2

SPECIALIZATION AND DIVERSITY IN CITIES[1]

1 Introduction

The issue of specialization in the production of goods and services by specific industries in specific cities appeals to a wide spectrum of interests. In a narrow sense, the degree of specialization is an important empirical fact which should guide our urban model-building. In a more general sense, the degree of specialization informs our perceptions of national growth, international trade competitiveness, and urban policy decisions.

There has been an interest in the matter of specialization (and therefore, diversity) of late in both theoretical and empirical work in urban economics. Specialization of production in cities is the outcome of the phenomenon of increasing returns thought to underly many important production processes. Models of increasing returns imply that the externalities involved in production a class of goods would naturally lead contiguous geographic areas to specialize in production of that product, however broadly the term "product" may be interpreted empirically. On the other hand, there exists some theoretical work which suggests that large cities grow because of diversity in production, not because of specialization.

At the same time, empirical research has been proceeding, and has not always

[1] I would like to thank Charles Michalopoulos and Yannis Ioannides for their help and suggestions. All errors in formulations and executions are, of course, my own.

confirmed theoretical results. One goal of this paper is to look at the varying empirical work that has attempted to trace the importance of specialization, as opposed to diversification, in U.S. cities. I will discuss three studies, noting their similarities and differences, and then present results of similar tests of my own. The tests I run are differentiated from other work in that they attempt to assess changes in patterns of urban specialization over a broader time period, including the possibility of a structural break at 1970. I refer extensively to papers by Edward Glaeser, Hedi Kallal, José Scheinkman and Andrei Shleifer (1992), J. Vernon Henderson, Ari Kuncoro and Matt Turner (1992), and Breandán Ó hUallacháin (1989).

It is impossible to evaluate these papers, or the present paper, in a theoretical vacuum. In Section 2, I present a theoretical context for issues of specialization and diversification in cities. Section 3 discusses the similarities and differences of each of the three empirical papers mentioned above. Sections 4 and 5 present my own empirical investigations of these issues, Section 4 explaining method and Section 5 explaining results.

2 Theoretical background

Increasing returns to production, which are experienced when production is geographically concentrated, can help explain the formation of cities. Producers gather to specialize in production of a particular good in a particular city to take advantage of positive spillovers, which may involve knowledge spillovers, more general labor pool and resource considerations, innovation-sharing, etc. See for example Francisco Rivera-Batiz (1988b), Abdel-Rahman and Fujita [(1990), (1993)], Henderson [(1974), (1983), (1988)], Henderson and Ioannides (1981), Ioannides (1994) and Dobkins (1994). (In some models, these agglomeration economies are also experienced by consumers. See

F. Rivera-Batiz (1988a) for a combination of production and consumption economies of scale.) Many of these papers trace their theoretical roots to the increasing returns models of Arrow (1961), Romer (1986), Lucas (1988) and ultimately to Alfred Marshall (1890); and their method to Dixit and Stiglitz (1977).

In Dobkins (1994), I show that geographic concentration supports innovation, which leads to increased output. I assume there that each locale is specialized in the production of one cluster of goods (one industry), while the production of non-traded goods proceeds with the usual mix of retail and service activity. How does a system of cities then form? The answer to that has been well-defined in a series of papers by J. Vernon Henderson. In Henderson (1988), he provides a system in which each city is specialized in an industry, since there are no positive benefits from locating more than one industry in a city. In fact, there is a benefit in avoiding congestion, so that it is optimal for each city to specialize in one industry to take advantage of increasing returns in production and to avoid the possibility of undue congestion. City sizes differ because different industries have different economies of scale.

Dobkins (1994) shows increasing returns to production. The combination of geographic concentration and innovation does not support increased employment, however, because the innovation leads only to *process* innovation. Workers simply innovate, finding better ways to produce a given output with the same labor force.[2] The labor supply in each of the locales would respond positively to an increased market share for that locale's traded good (industry), but the model does not pursue that possibility, largely because changing preferences is a tricky matter. However, it is certainly possible in a real world scenario that innovation could lead to *product* inno-

[2]The emphasis on innovation as the source of increasing returns reflects the theory and empirical findings of Michael Porter (1990).

vation, and that therefore market share could increase. An increased demand for the industry's product could increase employment in the locale, as well as output. I raise this issue because empirical studies often look at employment rather than output, as a matter of practicality; I will return to this matter below.

Standard urban economic theory recognizes that specialization is one reason for the formation and growth of cities. However, the theory distinguishes between localization externalities and urbanization externalities. Localization externalities occur because of the kind of industry-specific factors discussed above. Urbanization externalities are more general, and allow for the spillover to be between industries. The increasing returns literature cited above emphasizes localization externalities. Henderson (1988) offers empirical evidence that localization externalities predominate. The presence of urbanization externalities would diminish the role of specialization in production activities among cities; empirical studies often attempt to measure this distinction.

Why would cities exhibit urbanization externalities? This idea is attributed to Jane Jacobs in Glaeser *et al.* (1992), and Jacobs' works (1969 in particular) are certainly a forceful, though not quantitative, assertion of that point. Jacobs (1969) describes how innovation is likely to occur *between* industries, as differing techniques and needs stimulate new ideas.

On another level, the diversity notion can be supported by a theory that distinguishes between places engaged in local trade and long distance trade, and between manufacturing and nodal cities. As presented by John Marshall (1989) this rather intricate theory involves both central place and mercantilist considerations, along with a taxonomy of city functions. Marshall's thesis is that certain North American cities function as the major "places", the most central nodes, in a central place theory picture. Growth in these cities is not dependent on the convergence of manufacturing functions, but on the convergence of trading functions, denoted by employment in services, wholesaling and retailing activities, with manufacturing employment being relatively low and diverse. In functional terms, most cities are either "nodal", "mixed", or "manufacturing". Nodal cities are identified by a high percentage of employment in trade and service activities. Manufacturing cities are at the other end of the scale with high employment in manufacturing activities. Mixed cities are, of course, somewhere in the middle on the continuum from nodal to manufacturing. By Marshall's definitions, the number of nodal and mixed cities far outweighs the number of manufacturing centers. In our terms, we might say that while specialization is usually associated with manufacturing, diversity is usually associated with the characteristics of nodal and mixed cities. "Diversity" therefore emerges as a much more complicated factor than specialization, or perhaps the appropriate conclusion is that an urban system is more complicated than our theoretical models, especially those that only emphasize the importance of specialization in explaining a system of cities. (See Appendix B for a more detailed explanation of John Marshall's work.)

Because I want to consider empirical investigations of these theories, I propose a testable relationship, in the form of an equation, which would include many of the theoretical aspects of the issue. I can then relate existing empirical work to this relationship and explain my current empirical results in light of the relationship. The usual method to test for the presence of localization and urbanization externalities is to look at employment in certain industries, defined at a relatively aggregate (two-digit code) or disaggregate (three-and four-digit code) level. Output in an industry *by city* is difficult information to obtain. Therefore, the left hand side of the relationship will be the change in employment in an industry i in a city c in a time interval between t and $t+1$, which I will denote $E_{i,c,t,t+1}$.

What does determine employment? The theories above suggest that we should consider localization and urbanization economies of scale, usually expressed by means of indices of specialization and diversity of production, respectively. We include such indices as explanatory variables, denoted by $S_{i,c,t}$ and $D_{c,t}$: $S_{i,c,t}$ measures how specialized a city c is in industry i at time t; $D_{c,t}$ measures the amount of diversity of production in a city c at time t, and is not, of course, related to a specific industry.

Another possible explanation of employment growth in an industry in a given city is the changing demand for that industry's product in the national economy. We denote this by $M_{i,t}$, where the M is mnemonic for market share; this is an industry-wide characteristic, so the c subscript is not appropriate.[3]

The nature of increasing returns have been attributed to several factors in recent theory: such as education, or human capital accumulation. Often the externality is expressed in terms of R&D expenditure, or perhaps there is an historical element involving natural resource availability that explains initial advantage. Regional considerations are obviously at play in the U.S., and that fact leads us to the complicated matter of real wages as factor in location of production, (because wages are often thought to vary among regions). I sum up these various explanations in a catchall variable, $R_{i,c,t}$, where the t may be particularly important in picking up historical background. It will be the case that the catchall variable I use includes past employment in an industry as a historical explanation.

In its most general form, then, the relationship looks like this,

$$E_{i,c,t,t+1} = \alpha\, S_{i,c,t} + \beta\, D_{c,t} + \gamma\, M_{i,t} + \delta\, R_{i,c,t}. \quad (1)$$

[3]Of course, if every industry occupies only one city, industry and city are synonymous terms. While many theoretical models assume or justify this position, it takes very little in the way of real world analysis to see that this is not the case. In fact, on the theoretical front, Arthur (1990) shows that industries will locate in a single place only if the increasing returns to scale are unbounded.

The interpretation of each of these variables varies, of course, with the authors I have chosen to study, and that is the topic of the next section.

3 Empirical implementation

Several recent papers in the literature attempt to look at similar issues, but the techniques involved are subtly different. Glaeser *et al.* (1992) consider a full array of industries across 170 cities, the array covering two-digit Standard Industrial Classification (SIC) industries in all categories including manufacturing, retail, and service. They are looking for specialization versus diversity, among other things, within a time span from 1956 to 1987. A unit of analysis in this study is the "city/industry", that is, an industry located in a specific city. Apparel in New York City in 1956 is one observation; business services in New York City in 1956 is another observation. The authors limit their study to the largest six industries (by employment) in each city.

Glaeser *et al.* attempt to test three different theories, which they attribute, respectively, to Marshall/Arrow/Romer, Jacobs, and Porter. They construe the increasing returns theories of Alfred Marshall, Arrow and Romer to mean that localization externalities will predominate and that industries in specific cities will tend toward monopoly in order to exploit increasing returns. They construe Jacobs' work to say that urbanization externalities will predominate and that competition will be the tendency. They take Porter's work to imply that a combination of localization externalities and competition is most productive.

Glaeser *et al.* test for the presence of localization as well as urbanization externalities in a dynamic sense. That is, they assert that cities are specialized at a point in time. They then look to see if this specialization in a given time period (1956) is a predictor of growth in that city/industry in the future (1987). Therefore, their version of my $E_{i,c,t,t+1}$ in Equation (2) is the log of employment in the city/industry in 1987 divided by employment in the city/industry in 1956. Specialization is measured as the percentage of the city/industry's labor force, divided by the city's labor force, adjusted for the percentage employed in that industry nationally. Diversity is measured by employment in the city's other five top industries in 1956 as a percentage of total city employment. Market share is measured by the log of U.S. employment in 1987 divided by U.S. employment in 1956 in the industry outside the city. As other control factors, they consider the actual employment in the city/industry in 1956, the real wage in 1956, and a regional dummy variable. Because they are interested in the competition versus monopoly issue, they also include an additional index: competition is measured by the number of firms in the industry in a given city compared to the average number of firms in the industry nationally.

They find that specialization is not as important as diversity. (They also find that competition, as opposed to monopoly, is significant for growth.) Therefore, they find full support for Jacobs' theory, no support for the Marshall/Arrow/Romer theories, and mixed support for Porter. Although they do not note it specifically, lack of support for "MAR", as they denote it, also implies lack of support for the whole host of models mentioned above that depend on industry-specific localization externalities, including the theories of Henderson.

In contrast, Henderson *et al.* (1992) show that specialization is more important than diversity. Specifically, they test for the existence of dynamic localization externalities in five specific industries: machinery, electrical machinery, primary metals, computers, and electronic components in 1987 compared to 1970, across 224 cities. The first three are considered traditional manufacturing industries, while the other two are newer "high tech" industries. This study uses many more variables than Glaeser *et al.*, (*op. cit.*) all of which fit into one of the categories in my Equation

(2) above; the more critical difference with Glaeser *et al.* is the formulation of the dependent variable.

Henderson *et al.* (1992) use the log of 1987 employment in the industry in a given city as their dependent variable. Among the right hand side variables, they then include the log of 1970 employment. This contrasts with Glaeser *et al.* using the log of the ratio of employment in 1987 to employment in 1956 as the dependent variable, and including employment in the city/industry in 1956 as a right hand side variable. Henderson *et al.* suggest that this difference in technique accounts for some of the differing results, a point to which we return below.

Henderson *et al.* use a simple specialization index that is the ratio of 1970 city/industry employment to total employment in the city. Their diversity index comes in two parts: the log of all other manufacturing employment in the city, and a Hirschman-Herfindahl index on the lack of diversity in the local manufacturing mix. Also related are their inclusion of the log of population (presumably in 1987) and the ratio of durable manufacturing to all manufacturing.

As for their "other" variables, my $R_{i,c,t}$, Henderson *et al.* include a most interesting set. Along with regional dummies, they use a spatial measure, the distance to the nearest national business center. They also include measures of education and the log of the manufacturing wage.

Henderson *et al.* find in general that concentration in these industries in 1970 stimulated growth in employment levels between 1970 and 1987, a finding supporting the notion of localization externalities. But their more complex analysis requires more explanation. On the diversity issue, an increase in the level of other manufacturing activity increases employment growth in the three traditional industries, but diversity as measured by the Hirschman-Herfindahl index does not matter for growth of primary metals, machinery and electrical machinery. They suggest that this finding may reflect a product demand dimension for the scale variable. As for the other economic and geographic variables, they find that, "Apart from own industry employment levels and concentration in 1970, with one exception, no measures of levels or diversity of other past general or specific economic activity persist in significant influence in 1987, after 1987 characteristics are accounted for," in the three traditional industries. (pg. 27)

Results for the two newer industries are similar, with the 1970 industrial mix being very significant. They note that, "Past concentrations (but not levels) in electronic components and related activities have a strong impact on locational probabilities." (pg. 33) Diversity had an interesting impact for these industries, where their results suggest that, "increased diversity is important in attracting an industry but more concentration is important in retaining it." (pg. 37)

All in all, this paper provides strong support for localization externalities, particularly in the traditional industries, where the index of 1970 specialization was the only number from the "past" to have a significant influence on current activity. On the other hand, this formulation allows for current conditions to have a major impact on current employment.

Henderson *et al.* compare their dependent variable specifications to that of Glaeser *et al.*. Using the method from Glaeser *et al.* on their own data, they find that that "would lead us to conclude that prior concentration and current other manufacturing are harmful to current employment", (pg. 22), and they clearly reject this interpretation. This is surely one of the more interesting issues in comparing the two studies, which on the surface give us conflicting results.

The third empirical study, that of Ó hUallacháin (1989), is of particular interest because the industries studied are all in the service sector. He looks at 27 such industries in 264 U.S. metropolitan areas and compares 1977 and 1984. Ó hUallacháin argues that there are a variety of factors affecting the growth of these service industries, but that all of the factors can be identified with either urbanization or localization externalities. Those factors include a market penetration effect (reflecting urbanization externalities), and a return to the early phases of product cycles (reflecting localization externalities). In addition to employment growth, the results I report here, he is also interested in explaining the changing number of establishments for these industries. The industries involved include some in the transportation and public utilities categories such as trucking and warehousing and communication; in the FIRE (finance, insurance and real estate) category including banking and insurance; in services including hotels, advertising, and business services including R&D labs, management and public relations, and professional services; and hospitals, and health and allied services. Ó hUallacháin uses as a dependent variable the natural log of the absolute growth in employment in each city/industry from 1977 to 1984.[4] To measure localization externalities, he uses the natural log of employment in the city/industry in 1977; to measure urbanization externalities, he uses the natural log of the size of the local labor force in 1977.

The results on employment growth are straightforward: almost all of the service industries exhibit localization externalities. The only exceptions are communications except radio and T.V., advertising, hospitals and health and allied services. (Urbanization externalities more often explained the growth of establishments.)

These results are interesting in their simplicity. The key in accepting the formulation lies in judging the validity of expressing all factors, including $M_{i,t}$ and $R_{c,i,t}$ of Equation (2), as either localization or urbanization effects.

[4]He notes that using absolute growth instead of percentage growth means that small initial observations don't exaggerate the growth rate.

4 Method

This paper aims at a closer look at some of the results of the works discussed above, although it certainly does not attempt to replicate that work. In particular, I look at which cities are "specialized". I use various indices to measure specialization, and I look at various time periods to assess the dynamic impact of specialization. I restrict the analysis to certain two and three digit industries in the manufacturing categories, to the broad categories of services, retail, and wholesale, and to financial services and mining operations.

In contrast to Glaeser *et al.*, I use a smaller sample of cities, either 61 (in 1947) or 63; and more years, 1947, 1956, 1970, 1980, and 1990. My source of data for specialization indices is the same: the Census Bureau's *County Business Patterns.* Therefore, I have the same problems of censored data; the Census Bureau does not reveal exact employment numbers if there are only a few firms in the industry in a city. I describe my choice of cities, years, and data construction in Appendix A.

All of the papers discussed above are testing for the presence of dynamic specialization and assert the existence of static specialization. They assume that specialization is a fact in a given city in a given time period. Henderson *et al.* are most specific: they tell us that in 1987 "most" 2-digit industries registered employment of 250 workers or more in only 60-90 % of metro areas, and that "most" 3-digit industries registered employment in under 40 % of metro areas. They also point out that "many metro areas exhibited a very high degree of specialization with a significant fraction (8-15 %) of the local labor force in just one industry."

The definition of the words "many" and "most" here are unclear. In fact, it is not stated why 8-15 % is significant, although I would agree that any reasonable multiplier acting on jobs totalling 15 % of the work force would be significant for a city's employment picture. The point is that specialization, in a static sense, is not necessarily proven to be important, or even factual. Because I use a sample biased toward larger cities, we do not, in fact, see this static specialization as strongly. (Larger cities are generally considered to be less specialized than smaller cities and towns.) This raises issues of which cities (if any) are driving national growth: the larger, not-so-specialized cities, or the smaller, more specialized ones. Answering that question is beyond the scope of this paper, but it is possible to shed some light on the subsidiary question of specialization.

I investigate the role of specialization using three different indices, each one of which yields slightly different information. (Two of these indices are used below to examine static specialization; another one is used in the regressions discussed below.) The simplest specialization index is the percentage of a city's labor force involved in a particular industry. (This is calculated as the ratio of the number of employees in a particular industry to the number of employees in all industries, which is to say, the city's labor force.) I refer to this as the *city* specialization index, and use it to gauge static city specialization. It is the index used by Henderson *et al.*

A similar index is used by Glaeser *et al.* with a modification. They consider how specialized a city is in an industry relative to what one would expect if employment in that industry were randomly scattered across the country. They point out that this corrects for cases in which an industry in a city is large only because the city is large. (Although they do not note it, this adjustment also should help distinguish between specialization and central place/nodal considerations. If large cities are not highly specialized, it may be because they serve a nodal function.) Technically, this index is the ratio of the city index above to the fraction of the industry's employment over total national employment. I call this the *adjusted city* index.

A third index, which I will call the *industry* index, is simply the ratio of the number of employees in an industry in a city to the number of employees in that industry nationally. Clearly, this index says less about the city than about the industry. I find this index of interest if for no other reason than the stylized fact it yields: industries in the U.S. are not geographically very concentrated, as we will see below. Again, this result is particularly interesting in light of theoretical work by Brian Arthur (1990); and it certainly sheds a different light on the notion of specialization.

I want to know, first of all, if the cities in my sample are specialized in the production of a good or goods, based on any or all of the indices described above. This is simply a matter of interpreting the evidence provided by the indices: the data are presented in Tables 1A and 1B and are discussed in the next section.

Secondly, I want to know what causes growth in an industry in a city; whether specialization is paramount or if other factors are associated with growth. I run simple regressions to see if specialization in period t led to employment growth in period $t+1$ in the industry in a given city, or whether growth is associated with growth of the city's labor force overall. In terms of Equation (2), I regress growth in employment against two factors, specialization and diversity. The growth measure I use is the simple percentage change in growth, using t period employment as the base. The specialization index is the adjusted city index described above, based on employment in year t. The "diversity" index is less a true measure of diversity than a method of measuring the possibility that employment in certain industries grow with the city size. It is calculated simply as the percentage change in labor force from period t to period $t+1$. To consider this a measure of diversity, I have to assume that changing labor force is an indicator of underlying factors as various as metropolitan dominance, central place theory, and mercantile theory. This is a direction suggested by J. Marshall (1989) and Ó hUallacháin.

Finally, I use the employment in industry i in city c at time t as a control, denoted by $e_{c,i,t}$. Therefore, I estimate

$$E_{c,i,t,t+1} = S_{c,i,t} + D_{c,t,t+1} + e_{c,i,t} + \epsilon_t. \quad (2)$$

first for $t = 1947$ and $t + 1 = 1970$; and again for $t = 1970$ and $t + 1 = 1990$. The error term is denoted by ϵ_t. There are a few three-digit-code industries that become available in 1956, so that a handful of the first set of regressions are actually for the period 1956 to 1970; one industry, the aerospace portion of the transportation equipment industry, is included for the period from 1980 to 1990 in the second set of regressions.

The break at 1970 is appropriate for several reasons, the first being that it is simply near the midpoint of the available post World War II data. The second reason is more subtle. Borchert (1983) sees structural changes occurring about 1970 that may have changed urban patterns. Borchert is often quoted in the geography literature for his naming of several epochs in U.S. growth patterns; those epochs begin in 1830 with the "Iron Horse" epoch and ended with the "Auto-Air-Cheap Oil-Amenity" epoch that stretched from 1920 to 1970. In Borchert (1983) he asserts that three factors (at least) may mark a new epoch: expensive oil, immigration and the emergence into the work force of the baby-boom generation. That epoch would begin in 1970. Furthermore, the U.S. city size distribution is known to have reacted to these or other factors, as reflected in the 1970 census, a phenomenon discussed in Dobkins and Ioannides (1995).

The results of the regressions are reported in Tables 2A and 2B and discussed below.

5 Results

Tables 1B and 1A list the industry specialization indices above 5 % for 1956 and 1990, and the city specialization indices above 8 %, a figure taken from the Henderson *et al.* analysis noted above. The point here is that the cities I considered were not highly specialized, nor were the industries highly concentrated. I look at the industry concentrations first.

The industry index list in Table 1B includes New York City, Los Angeles and Chicago numerous times. However, we note that in 1956 those three cities had 11.6%, 4.6%, and 5.1% of the labor force, respectively. In 1990, they had 7.3%, 6.1%, and 3.4%, respectively. Concentrations near those figures may simply reflect the production of non-traded goods; the numbers would have to be higher than the labor force percentages to show industry concentration. This is particularly true for the major classifications of manufacturing, service, etc. Therefore, I do not report those categories. However, it may well be significant that New York City had 12.7 % of the apparel industry employment in 1990, and less than 5% of the textile industry employment. Percentages greater than 5% in any of the two or three digit codes are therefore reported for any applicable city, and the significance of those percentages can be assessed case by case.

What I find is that even the largest cities do not have high concentrations of industry employment, particularly at the two-digit level. A notable exception is Richmond, which had 45.4 % of the tobacco industry employment in 1990. At the three digit level, the numbers are slightly larger. New York City had 13.5% of the drug manufacturing employment in 1990; Chicago had 14% of the iron and steel mills in 1990 (compared to 5.8% for Pittsburgh); Detroit had 18.3 % of the automotive trade (which includes both motor vehicles and parts); Norfolk had 14% of the shipbuilding employment;

and Seattle and Los Angeles each claimed about 21% of the aircraft industry. The only other really large numbers are the two obvious ones: for 1990, 62.2% of the movie industry was located in Los Angeles, and Rochester, New York, had 48.6% of the photographic equipment industry. Los Angeles had 31% of the aerospace industry, but the other "high tech" fields are not nearly as concentrated as many would expect. San Francisco had 13.6% of the electronic components employment, followed by Los Angeles (8.1%) and Boston (5.1%). (Other industry codes that might be considered high tech are the ones that identify computers as office equipment, the communications equipment code, and the medical instruments code.)

Were industries more concentrated in 1956? Well, New York City had more of the apparel industry (32%) and the drug industry (25%). Boston had more of the footwear industry (17.7%). Pittsburgh boasted almost 14% of the steel and iron mills, as did Chicago, and Detroit had over 38% of the auto industry in 1956. So there is some evidence that industries were more concentrated in 1956 (tobacco is an exception), but even then the degree of concentration for most industries is smaller than might be expected.

I turn to the static data on city specialization in Table 1A, and ask what percentage of a city's labor force works in a specific industry. Clearly, the largest cities are unlikely to show up in this city specialization index listing. Although people may associate certain industries with New York City, Los Angeles or Chicago, the large labor force numbers in those cities overwhelm any one industry's work force. I report those specializations that account for more than 8% of a city's labor force as well as those which have interest for comparison purposes.

Using the city specialization index, we see even more contrast between 1956 and 1990. Partly, this is due to the fact that I have included many manufacturing categories; it may be that cities have dealt with declining employment in those industries by diversifying into other fields. (Labor force does not decline in any of these cities in these time periods.) The cities in this study do have heavy concentrations in the FIRE (finance, insurance, and real estate) category, showing a significant increase from 1956 to 1990. J. Marshall (1989) notes that "significant financial control" is a mark of metropolitan dominance (pg.293); perhaps it is not surprising that in 1990, 30 of these 63 large cities had more than 8% of their labor force involved in financial and related sectors. All cities in this study also show a large increase in the service sector. In 1956, the average for these cities was 11% of the labor force in the service sector compared to a U.S. average of 9.8%. In 1990, the average for these cities was 32.9% in service sectors compared to 30.8% for the whole country. The listing shows the percentage of the entire U.S. labor force involved in each industry for both 1956 and 1990, and almost all manufacturing categories show a decrease from 1956 to 1990.

The most extreme city example is Greenville, South Carolina, where employment in textiles drops from 42.8% of its labor force in 1956 to 7.9% in 1990. In less dramatic ways, many other cities seem to have diversified. Most numbers in 1990 are small: the largest are 10% for photographic equipment in Rochester, New York, 18% for aircraft in Wichita, and the predictable 28.2% for hotels in Las Vegas. However, in 1956, many cities exceeded these numbers. Note, in particular, that cities in textile, primary metals, machinery and transportation industries tended to have very large percentages of their labor force concentrated.

I also see clearly the contrast between the industry specialization index and the city specialization index. Detroit had 18.3% of the motor vehicle employment in 1990, which accounts for 7.3% of the city's labor force that year. I believe that it is very important to note that both numbers underestimate the impact of other activities that take place in Detroit because of the motor vehicle employment there. On the other hand, those other industries may well be reflected in the diversity I apparently see between 1956 and 1990.

Before discussing the results of the regression, I note some of the descriptive characteristics of the data, characteristics that shed some light on the regression results. First of all, labor force increases in all of these cities over the relevant time periods, even though it may be the case that some cities experienced declining population. The mean of manufacturing employment, in these cities, increases at a decreasing rate from 1947 through 1956, 1970, and 1980. In 1990, it falls back, coming close to the 1956 level. The other sectors, wholesale, retail, FIRE, and services all increase employment substantially from 1947 to 1990. The mean for service employment goes from 36,532 in 1947 to 298,675 in 1990.

Among the manufacturing industries, there were increases from 1947 to 1990 in publishing (according to J. Marshall, (1989, pg. 69) the most ubiquitous of all manufacturing industries), chemicals, rubber products, fabricated metals, electrical machinery, transportation equipment, and instruments. These were, of course, more than offset by the losses in the food industry, textiles, apparel, and primary metals, to name a few large, declining industries. Within the transportation industry, motor vehicles and parts employment was down, but there were large increases in aircraft employment. Within the chemical industry code, increases in drug manufacturing employment increased. Within the electrical equipment industry, employment in communication equipment was down from 1956 to 1990, while employment in the electronic components group was up from 1970 to 1990. As expected, the employment in computers as office equipment, a division under machinery, was up. Within the instruments grouping, employment in medical instruments was up from 1970 to 1990, and employment in photographic equipment was up from 1956 to 1990.

With these static and descriptive results in mind, the dynamic situation is particularly interesting. I find, across almost all industries and sectors and all time periods, that specialization in an industry is negatively correlated with growth in that (own) industry. In the handful of cases in which the coefficient on the specialization index is positive, the coefficient and the t-ratio are quite close to zero. The number of industries and sectors in which specialization is statistically significant in explaining own industry growth, even in a negative sense, falls from the 1947-1970 period to the 1970-1990 period. On the other hand, growth in an industry is either positively associated with growth in a city's total labor force, or not significantly related at all.

Regression results are reported in Table 2A for the 1947 (1956) to 1970, and in Table 2B for the 1970 to 1990 periods. I report coefficients and t-ratios for the specialization index ($S_{c,i,t}$) and labor force growth ($D_{c,t,t+1}$) variables. The variable controlling for initial industry size is never significant, and is not reported. The entries at the top of the tables report those industries for which the specialization index is statistically significant; those entries at the bottom of the tables report industries for which the labor force growth variable is statistically significant. To put these entries in perspective, I would point out that, counting sectors as well as two and three digit codes, there are 42 and 46 regressions for the two time periods, respectively, and half of those regressions yielded statistically significant coefficients..

I emphasize that results are obtained for both industries *and* sectors, because sectors including manufacturing, services, FIRE, retail and wholesale, are prominent in these tables, as are ubiquitous activities such as insurance, publishing and fabricated metals. In general, but not completely, I find that the explanatory power of the models, as measured by the R^2 and F values reported in the tables, is higher for the broader categories.

The role of the high tech industries is notable by its absence. There is not much evidence of the three-digit industries such as computers, electronic components, medical instruments, aircraft and aerospace, in these tables, implying that while specialization in these industries may not increase own industry growth, there is no negative relationship, such as exists with the more traditional industries.

Given that neither specialization nor labor force growth explains own industry growth (or decline) for most of the industry codes in these regressions, I would suggest that two possible factors are at work. First, of all, because these are manufacturing industries, the role of specific service sector industries is ignored. However, the Ó hUallacháin paper fills in this gap nicely, finding that specialization is significantly related to own industry growth in that sector. (How sensitive these analyses are to our slightly different specifications is unclear.)

Secondly, I am leaving out market share and regions as specific variables. Henderson *et al.* find some regions significant in explaining employment in some industries; Glaeser *et al.* find both regions and the measure of market share significant in explaining own industry growth. Availability and inclusion of those variables might, of course, alter coefficients on the variables included in these present regressions.

6 Conclusion

The results of these regressions as well as an informed study of the descriptive statistics of the data, highlighted against the backdrop of the earlier research described above, paint a picture, perhaps tentative, of changing urban dynamics. In the larger cities in the United States growth seems to be associated with their role as central places rather than with specializations that result from economies of scale. To the extent that specialization is important in this group of cities, the effect is likely to be felt mainly in the service sector.

This does not imply that specialization is not important in explaining the growth or decline of other groups of cities. I would note here that there are certainly smaller cities and towns that are highly specialized, and which would show very high city specialization indices simply because they are small. It would require a wider review of the data on cities of various sizes to verify whether those smaller cities have grown because of that specialization. However, a quick review of the same type of data used above, from counties in Southwest Virginia shows a clear pattern: the counties that exhibit the greatest specialization have suffered the most in terms of growth. Of course, these counties have specialized in coal, apparel and furniture, none of which are high tech, growth industries. Nevertheless, one suspects that their stories are repeated in many small towns and counties across the country, while the high tech, growth industries are located in the very kinds of cities I study in this paper.

The breadth of this matter of urban growth makes a comprehensive study difficult: one would like to know the relationship of smaller cities to larger, both in terms of specialization and of diversity, perhaps within a hierarchy of cities. One would like to incorporate changing market structures, shifting demand and evolving supply technologies, into the mix. International trade considerations should certainly make a difference in those market structures. And the intriguing question of whether city growth drives national growth, or vice versa, remains unsettled.

BIBLIOGRAPHY

Abdel-Rahman, Hesham and Masahisa Fujita (1990), "Product Variety, Marshallian Externalities, and City Sizes," *Journal of Regional Science* 30, 165-183.

Abdel-Rahman, Hesham and Masahisa Fujita (1993), "Specialization and Diversification in a System of Cities," *Journal of Urban Economics*, 33,189-222.

Arrow, Kenneth J. (1962) "The Economic Implications of Learning by Doing," *Review of Economic Studies* 29, 155-73.

Arthur, W. Brian (1990), " 'Silicon Valley' Locational Clusters: When Do Increasing Returns Imply Monopoly?," *Mathematical Social Sciences* 19, 235-251.

Borchert, John R. (1983), "Instability in American Metropolitan Growth," *The Geographical Review* 73, 127-149.

Christallier, W. (1933), in translation, (1966), *Central Places in Southern Germany*, Translated by C.W. Baskin. (Englewood Cliffs, NJ: Prentice-Hall.)

Dobkins, Linda Harris (1994), "Location, Innovation and Trade: Michael Porter Revisited," Ph.D. dissertation, Virginia Polytechnic Institute and State University.

Dobkins, Linda Harris and Yannis M. Ioannides (1995), "Evolution of the U.S. City Size Distribution," mimeo, Virginia Polytechnic Institute and State University.

Dixit, A. K., and J. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67, 297-308.

Glaeser, Edward L., Hedi D. Kallal, José Scheinkman, and Andrei Shleifer (1992), "Growth in Cities," *Journal of Political Economy*, 100, 1126-1152.

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

Henderson, J. Vernon (1983), "Industrial Bases and City Size," *American Economic Review*, 73, 164-168.

Henderson, J. Vernon (1988), *Urban Development: Theory, Fact and Illusion*, (Oxford: Oxford University Press.)

Henderson, J. Vernon, Ari Kuncoro and Matt Turner (1992), "Industrial Development in Cities," mimeo, Brown University, March.

Henderson, J. Vernon and Yannis M. Ioannides (1981), "Aspects of Growth in a System of Cities," *Journal of Urban Economics*, 10, 117-139.

Ioannides, Yannis M. (1994) "Product Differentiation and Economic Growth in a System of Cities, " *Regional Science and Urban Economics*, 24, 461-484.

Jacobs, Jane (1969) *The Economy of Cities*, (New York: Vintage)

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

Marshall, Alfred (1890) *Principles of Economics*, Eighth Edition (Philadelphia: Porcupine Press, 1982)

Marshall, John U. (1989) *The Structure of Urban Systems*, (Toronto: University of Toronto Press.)

Porter, Michael E. (1990) *The Competitive Advantage of Nations*, (New York City: The Free Press)

Ó hUallacháin, Breándan (1989), "Agglomeration of Services in American Metropolitan Areas," *Growth and Change* 20(3), 254-70.

O'Sullivan, Arthur (1993) *Urban Economics*, Homewood, IL: Irwin Press.

Rivera-Batiz, Francisco L. (1988a), "Increasing Returns, Monopolistic Competition, and Agglomeration Economies in Consumption and Production," *Regional Science and Urban Economics* 18, 125-153.

Rivera-Batiz, Francisco L. (1988b), "Modeling Urban Agglomeration: Producer Services, Linkage Externalities and Specialization Economies," in *Modeling and Simulation*, M.H. Mickel and V. Vogt, eds. (Research Triangle Park, North Carolina: Instrument Society of America) 113-119.

Romer, Paul (1986), "Increasing Returns and Long-run Growth," *Journal of Political Economy* 94, 1002-1038.

Vance, J.E., Jr. (1970), *The Merchant's World: The Geography of Wholesaling* (Englewood Cliffs, NJ: Prentice-Hall.)

TABLE 1

STATIC SPECIALIZATION RESULTS

Table 1A

City Specializations

The city specializations listed are generally above 8%, a figure used because of the Henderson, *et al.* analysis, as described in the text. Other cities are listed for comparison purposes. These are percentages of a city's labor force employed in the given category. With each industry or sector category is listed the percentage of the U.S. labor force involved in that category's employment, first in 1956 and then in 1990. All data are from the Census Bureau's *County Business Patterns 1956* and *County Business Patterns 1990.*

CATEGORY (Employment %)	CITY SPECIALIZATIONS 1956	CITY SPECIALIZATIONS 1990
FIRE(1) (5.6%; 7.4%)	New York City, 9.5%; Boston, 8.2% Dallas, 8.3%; Hartford, 11.1% Jacksonville, 11.3%; Raleigh, 9.5% Richmond, 9.1%; Sacramento, 13% San Antonio, 8.2%; San Francisco, 8.6%	New York City, 12%; Denver, 14% Hartford, 14.5%; Richmond, 11% Jacksonville, 11.7%; Omaha, 11.4% Columbus, 10.7%; San Antonio, 10%
Hotels (1.2%; 1.6%)	Las Vegas, 17.5%; Miami, 7.7%	Las Vegas, 28.2%; Orlando, 7% Honolulu, 6.7%
Mining (2%; .77%)	Houston, 9.5%; Salt Lake City, 8.6%	Anchorage, 6.3%; Houston, 4.5%
Crude petroleum mining (.8%; .31%)	Houston, 5%; Tulsa, 11.4% Oklahoma City, 7.7%	Anchorage, 6.2%
Food products (3.5%; 1.6%)	Kalamazoo, 9%; Omaha, 16.6%	Fresno, 4.4%
Tobacco (.2%; .04%)	Raleigh, 8.2%; Richmond, 7.3%	Richmond, 5.2%
Textiles (2.6%; .7%)	Charlotte, 29.5%; Greensboro, 28.5% Greenville, 42.8%; Huntsville, 9.2% Providence, 15.5%; Raleigh, 9.2%	Greenville, 7.9%; Greensboro, 7.3%
Apparel (3.1%; 1.1%)	Greenville, 16.7% New York City, 8.6%	None
Paper products (1.4%; .7%)	Kalamazoo, 12.9%; Sacramento, 9.6%	Kalamazoo, 3.1%

(1) FIRE stands for finance, insurance and real estate.

The cities in the this study averaged 8% for the FIRE category in 1990; 30 of the 63 cities were above 8%. I report only those above 10%.

Continued on next page.

Chemicals (2%; .92%)	None	Beaumont, 8.7%
Industrial organic chemicals (.7%; .53%)	Beaumont, 25.2%	None
Petroleum and coal refining (.7%; .12%)	Beaumont, 24%	Beaumont, 4.7%
Rubber products (.6%; .94%)	Cleveland (Akron), 5.5%	None
Primary metal (3.3%; .77%)	Allentown, 16.6%; Youngstown, 32% Pittsburgh, 17.7%	Youngstown, 4.5%
Iron and steel mills (1.6%; .3%)	Youngstown, 24.8%; Allentown, 15.9% Pittsburgh, 14.5%; Buffalo, 8%	None
Fabricated metal (2.7%; 1.6%)	Grand Rapids, 8%	None
Machinery (4.2%; 2.1%)	Grand Rapids, 15.2%; Milwaukee, 14.8% Syracuse, 12.3%; Louisville, 11.5%	None
Electrical machinery (2.6%; 1.7%)	Syracuse, 10%	None
Communication equipment (1.2%; .27%)	Syracuse, 7.7%; Indianapolis, 5.9% Boston, 4.7%	None
Transportation equipment (4.4%; 1.9%)	Wichita, 32.5%; San Diego, 26.7% Detroit, 25.5%; Sacramento, 19% Hartford, 15.6%; Seattle, 14.8%; Los Angeles, 12.3% ; Dallas, 11.7% Indianapolis, 10.8%; Cincinnati, 8.6% Cleveland, 8.6%; Columbus, 8.4%	Wichita, 18%; Norfolk, 8.8% Youngstown, 8.6%
Motor vehicles and and equipment (2%; .76%)	Detroit, 24.5%	Youngstown, 8.5%; Detroit, 7.3%
Aircraft and parts (1.9%; .67%)	Wichita, 32%; San Diego, 26% Sacramento, 18.8%; Hartford, 15.6% Seattle, 12.9%; Los Angeles, 10.4% Tulsa, 10.2%; Dallas, 10% Fresno, 10%	Wichita, 18%
Photographic equipment (.15%; .09%)	Rochester NY, 14.4%	Rochester, NY, 10%
Motion pictures (.15%; .2%)	Los Angeles, 1.6%	Los Angeles, 2.1%
Computers (as office equipment) (NA;.29%)	NA	Huntsville, 7%; Raleigh, 5.4%
Electronic components (NA; .59%)	NA	Austin, 6%
Aerospace (NA, .22%)	NA	Huntsville, 3.7%; Denver, 3.5%

Table 1B

Industry Specializations

This table gives the percentage of an industry's nationwide employment concentrated in one city. Cities with more than 5% of an industry's employment are listed. The sources are given above.

CATEGORY (SIC CODES)	INDUSTRY SPECIALIZATIONS 1956	INDUSTRY SPECIALIZATIONS 1990
FIRE	New York City, 19% Chicago, 6%; Los Angeles, 5%	New York City, 11.9% Los Angeles, 6.5%
Hotels (70)	New York City, 10% Chicago, 6%	Las Vegas, 6.2%
Mining	None	Houston, 8.3%
Crude petroleum mining (13)	Houston, 5.6% Los Angeles, ***	Houston, 7.2%
Food products (20)	New York City, 8.1%	None
Tobacco (21)	Greensboro, 11%; Richmond, 10% Richmond, 10%; Syracuse, 5.8% Philadelphia, 5.7%	Richmond, 45.4% Greensboro, 13.7% Louisville, 9.3%
Textiles (22)	New York City, 6.7%	Greensboro, 5.4%
Apparel (23)	New York City, 32.3%	New York City, 12.7% Los Angeles, 10.7%
Furniture and fixtures (25)	New York City, 8.3% Chicago, 5%	Los Angeles, 9.4%
Paper products (26)	New York City, 9.2%	New York City, 5.4%
Publishing (27)	New York City, 19.2%	New York City, 10.8% Los Angeles, 5.7%
Chemicals (28)	New York City, 9.3%	New York City, 8.9%
Industrial inorganic chemicals (281)	Cleveland, 6.1%	None
Industrial organic chemicals (282)	New York City, 9% Beaumont, 7.1%; Buffalo, 5%	Richmond, 5.8%
Drugs (283)	New York City, 25% Indianapolis, 6.6% Chicago, 6.4%; Philadelphia, 5.6%	New York City, 13.5% Los Angeles, 6.2% Philadelphia, 5.6%
Petroleum and coal refining (29)	Beaumont, 9%; Houston, 8% Chicago, 7.5% New York City, 5.7%	Houston, 7.6%; Los Angeles, 7.5%
Rubber products (30)	Cleveland (Akron), 11% New York City, 8.8%	Los Angeles, 7%
Leather products (31)	New York City, 14.2%	New York City, 6.7%

Continued on next page.

Footwear (leather) (314)	Boston, 17.7%	None
Stone, clay, and glass products (32)	None	Los Angeles, 5.5%
Primary metals (33)	None	Chicago, 6.9%
Iron and steel mills (331)	Pittsburgh, 13.7%; Chicago, 13.6% Youngstown, 6.5%	Chicago, 14%; Detroit, 6.2% Pittsburgh, 5.8%
Iron and steel foundries (332)	Cleveland, 5.2%	Chicago, 9%
Nonferrous mills (335)	New York City, 17%; Los Angeles, 5%	New York City, 6.1%
Nonferrous foundries (336)	Philadelphia, 19.9%; Cleveland, 12.2% Chicago, 9.3%; New York City, 6.7% Los Angeles, 5.8%	Los Angeles, 8%; Chicago, 5.3%
Fabricated metals (34)	None	Los Angeles, 7.1%
Machinery (35)	New York City, 6.2%; Detroit, 6%	None
Electrical machinery (36)	New York City, 13.3%; Chicago, 12%	San Francisco, 7.2% Los Angeles, 6.5%
Communication equipment (366)	New York City, 16.7% Chicago, 16.6%; Boston, 8.7%	Boston, 8.6%; Chicago, 8.4% San Francisco, 7.3%; Dallas, 6% Los Angeles, 5.4%
Transportation equipment (37)	New York City, 5.9%	Los Angeles, 12.2%; Detroit, 7.5%
Motor vehicles and equipment (371)	Detroit, 38.4%; Cleveland, 6.4%	Detroit, 18.3%
Aircraft and parts (372)	Los Angeles, 24.9% New York City, 7.9%; Dallas, 5.6% Seattle, 5%; San Diego, 5% Hartford, 5%	Seattle, 21.5%; Los Angeles, 20.5% Dallas, 8.6%; Wichita, 6.1% Hartford, 5.6%
Shipbuilding (373)	New York City, 8.2% Baltimore/Washington D.C., 8.1%	Norfolk, 14%; New Orleans, 5.3%
Instruments (38)	New York City, 19.5%	Los Angeles, 11% New York City, 7.2% Boston, 6%; Rochester, NY, 5.1%
Photographic equipment (386)	Rochester, NY, 50.1% New York City, 14%	Rochester, NY, 50.1% New York City, 14%
Motion pictures (781)	Los Angeles, 50.3% New York City, 24%	Los Angeles, 62.2%
Medical equipment (384)	NA	Los Angeles, 9% New York City, 6%; Boston, 5.8%
Computers (as office equipment) (357)	NA	Raleigh, 7.1%; Los Angeles, 6.3% Boston, 6%; Minneapolis, 5.1%
Electronic components (367)	NA	San Francisco, 13.6% Los Angeles, 8.1%; Boston, 5.1%
Aerospace (376)	NA	Los Angeles, 31.3% San Francisco, 17.1%; Denver, 8.7% St. Louis, 6.5%; Seattle, 6.1%

TABLE 2

REGRESSION RESULTS

Results are reported for the regressions described in the text and below for two time periods, 1947 to 1970, in Table 2A, and 1970 to 1990, in Table 2B. Some industries appear in the *County Business Patterns* in 1956, and therefore these few industries are reported for the period from 1956 to 1970. The only results recorded in the two tables below are those in which the coefficient for the city specialization index were significant at a .05 level. There were a number of industries, as explained in the text, for which the labor force growth was positive and significant, and the specialization coefficient was not significant. These industries and sectors are listed at the bottom of each table, below the double line.

Coefficients in Table 2A are obtained by regressing growth in industry or sector employment from 1947 (1956) to 1970 on the specialization index for 1947 (1956) and growth of labor force from 1947 (1956) to 1970. Coefficients in Table 2B are obtained by regressing growth in industry or sector employment from 1970 to 1990 on the specialization index for 1970 and growth of labor force from 1970 to 1990. t-ratios are given in parenthesis. Adjusted R^2 and F values are given in the last column. (The F values are goodness of fit tests formed by the ratio of the sum of squared deviations due to the model to the sum of squared deviations due to the residuals, adjusted for degrees of freedom.) Asterisks indicate that coefficients are significant at the .05 level based on t-ratios; or that F values would occur by chance with a probability below .05. Tables follow, on the next two pages.

Table 2A (1947 - 1970)

Industry or sector (Number of cities with 1947/1956 employment)	Coefficient of specialization index (t-ratio)	Coefficient of labor force growth (t-ratio)	Adjusted R^2 F-value
Services (61)	-4.589* (-4.917)	2.849* (6.416)	0.4044 14.577*
FIRE (61)	-2.335* (-5.897)	2.197* (18.519)	0.8601 123.915*
Retail (61)	-.6653* (-3.128)	0.7624* (14.569)	0.8228 93.872*
Wholesale (61)	-0.8505 (-5.511)	0.8348 (17.765)	0.8519 116.024*
Insurance (59)	-1.357* (-3.662)	1.606* (4.123)	0.3539 11.590*
Foods (60)	-0.3068* (-4.818)	0.4600* (9.490)	0.6784 42.491
Publishing (59)	-6.879* (-2.687)	0.4453 (0.458)	0.075 2.586
Chemicals (59)	-.7520* (-2.153)	0.1457 (0.529)	0.0897 2.905*
Stone, glass, ceramic (58)	-1.818* (-2.251)	0.3288 (0.755)	0.1094 3.335*
Fabricated metals (58)	-1.435* (-2.084)	2.207* (3.129)	0.3249 10.145*
Transportation equipment (47)	-4.831* (-1.976)	13.287* (4.944)	0.3467 9.136*
Leather shoes (20)	-0.5811* (-3.194)	2.322* (3.707)	0.3819 4.914
Manufacturing (61)	-0.6076 (-1.674)	0.8798* (7.680)	0.6828 44.060*
Hotels (60)	-0.2653 (-1.502)	2.450* (8.490)	0.5910 29.423*
Apparel (58)	-1.515 (-0.448)	11.058* (4.536)	0.2584 7.622*
Tobacco (32)	-0.035 (-1.265)	1.582* (4.869)	0.4032 7.981*
Furniture (58)	-0.3394 (-1.074)	0.8888* (2.021)	0.0906 2.892*
Paper products (49)	-0.2783 (-1.376)	1.865 (5.002)	0.4368 13.411*
Machinery (59)	-1.394 (-1.096)	6.034* (4.141)	0.3026 9.388*
Electrical machinery (47)	-4.750 (-0.775)	16.743* (2.884)	0.2223 5.384*
Instruments (54)	-0.0849 (-0.162)	3.650* (2.790)	0.1039 3.049*
Nonferrous milling (40)	0.0188 (0.026)	2.791* (5.845)	0.4629 12.203*
Drugs (31)	-0.0576 (-0.703)	2.404* (2.722)	0.1750 3.122*

Table 2B (1970 - 1990)

Industry or sector (Number of cities with employment in 1970)	Coefficient of specialization index (t-ratio)	Coefficient of labor force growth (t-ratio)	Adjusted R^2 F-value
Services (63)	-7.713* (-6.027)	3.057* (5.008)	0.3917 14.309*
Retail (63)	-6.003* (-7.222)	1.461* (5.121)	0.4804 20.107*
Wholesale (63)	-0.8168* (-6.476)	0.8270* (15.616)	0.8206 95.547*
Lumber products (59)	0.5098 (-2.097)*	0.8920 (3.805)*	0.2439 7.238*
Publishing (63)	-0.6138* (-2.535)	0.5452* (4.655)	0.3641 12.831*
Stone, glass, ceramic (62)	-.6668* (-2.245)	0.2889 (1.378)	0.1339 4.143*
Electrical equipment (59)	-1.985* (-2.425)	3.339* (3.817)	0.2699 8.147*
Communications equipment (42)	-1.118* (-2.736)	3.469* (5.817)	0.4779 13.510*
Manufacturing (63)	-.0072 (-0.023)	0.8684* (5.412)	0.4638 18.875*
FIRE (63)	-0.1572 (-1.253)	0.8437* (16.130)	0.8247 98.222
Hotels (63)	-0.1051 (-1.543)	1.748* (7.827)	0.5195 23.343*
Insurance (63)	-0.0590 (-0.827)	0.6340* (7.384)	0.4808 20.137*
Mining (57)	-0.2747 (-1.066)	1.344* (3.222)	0.1417 4.081
Foods (63)	-0.0767 (-0.918)	0.2490* (4.348)	.03052 10.077*
Apparel (59)	-0.1657 (-1.621)	0.3998 (3.083)	0.1666 43866*
Paper products (58)	-0.1971 (-1.451)	0.5926* (2.659)	0.1819 5.224
Chemicals (61)	-0.0514 (-0.257)	1.522* (5.148)	0.3267 10.703*
Leather products (33)	-0.3614 (-1.116)	1.156* (2.136)	0.0780 1.902
Fabricated metals (62)	0.2097 (0.391)	2.623* (4.634)	0.2924 9.402*
Machinery (61)	-0.6215 (-1.884)	1.374* (3.844)	0.3184 10.343*
Transportation equipment (59)	-0.1622 (-0.789)	1.182* (3.303)	0.1782 5.193*
Instruments (52)	-0.1510 (-0.196)	21.465* (6.447)	0.4453 14.646
Medical instruments (28)	-2.633 (-0.904)	22.191* (4.449)	0.4725 90.63

CHAPTER 3

EVOLUTION OF THE U.S. CITY SIZE DISTRIBUTION[1]

1 Introduction

The empirical study of trends in city size distributions has engaged economists at least since the early years of this century, in part because the system of cities in an economy is a fundamental feature of its structure. The last fifteen years have seen a variety of theoretical models aimed at explaining the sizes and types of cities that arise in a market economy.

We propose here a new empirical approach to test several existing theories. Our approach attempts to integrate macroeconomic and economy-wide spatial considerations. We find that the U.S. city size distribution seems to have been moving toward more concentration, with the counterurbanization of the 1970s being something of a digression. Our empirical results suggest that agricultural land values, the prime interest rate, and the balance between manufacturing and services have a significant effect on the city size distribution. We also attempt to explain actual populations given the prevailing distribution, and find that foreign immigration, the distance between cities, and growth in real Gross National Product are significant explanatory

[1]The research reported here is coauthored with Yannis M. Ioannides. The authors would like to thank the following people for their helpful comments: Kristin Butcher, Anna M. Hardman, Vernon Henderson, Charles Michalopoulos, and Dan Nuxoll.

factors.

Some of the theoretical work of the last few years involve the evolution of a single city; others involve a system of cities. Those in the latter category yield predictions about the distribution of city sizes within the system. We consider one line of inquiry that would allow for changing city size distributions based on national output patterns and demand conditions, and a second line that allows for parallel growth (no change in the distribution) by emphasizing human capital accumulation.

Many prominent contributions in the city size distribution literature are found in the work of J. Vernon Henderson, who obtains a determination of city size by integrating international trade theory with urban economics. As summarized in Henderson [(1984); (1987); (1988)], this fundamental approach proposes a model of optimum city size, which balances the benefits of larger city size with its costs in terms of disamenities. Several of the implications of these theories have been tested, especially by Henderson himself. To the best of our knowledge, Henderson's theories and others we discuss below have never been tested directly, that is, in terms of systems of cities. We think such an endeavor is timely in view of the enormous attention which city sizes have attracted in the economic geography, regional economics and urban economics literature.

Empirical interest has been revived recently by Eaton and Eckstein (1994), who use data from France and Japan, and find their urban systems to be characterized by parallel growth. That is, in those economies city size distributions remain unchanged over many years. We discuss their results as well as the theoretical basis for our empirical work from Henderson, and Abdel-Rahman and Fujita (1993) in Section 3 below. Section 2 reviews significant literature from early in the century to the present.

We examine the evolution of city size distributions in the U.S. in this century by using Census Bureau data for the decennial years from 1900 to 1990. The definition of a city in this work is critical and we describe the conventions we use in generating our data in Section 4. We discuss our method in Section 5, and present our results in Section 6.

2 An historical perspective on city size distribution

Studies of city size distribution follow two tracks, and we consider the relationship of these studies to our present work. Some economists (as well as geographers and other social scientists) have found it useful to invoke the *rank size rule*, an alleged statistical regularity that is an outgrowth of the application of the Pareto distribution to city size data. Stated in its various forms, the rank size rule has been debated, calculated, and dismissed several times over since its first mention (that we can find) in Auerbach (1913). At the root of economists' ambivalence on the matter is the simplicity of the rule, which asserts that the population of a city, multiplied by the rank of that city in its system, equals a constant. This implies that the size of the second largest city would be one half the size of the first; the size of the third largest city would be one-third the size of the first, and so on.

There are several writers, including Auerbach (1913), who see the rank size rule as part of an overall scheme of things, a view that is characteristic of that period. Lotka (1925) noted the rule as he sought similarity in the "laws" of various disciplines, including economics and physics. Zipf (1949) made the rule as example of his overriding "Principle of Least Effort" in the conduct of human behavior. The latest in this history of fascination with the rank size rule is Paul Krugman's (1994) suggestion that the rank size rule may be evidence of complexity theory at work. (Krugman links the "self-organization properties" of complex systems to the statistical regularity implied by the rank-size rule.)

In 1936, H.W. Singer suggested that city sizes follow a Pareto distribution, by which he meant that his law of populations (*courbe des populations*) was a parallel to Pareto's (1906) law of income distribution (*courbe des revenues*). If city sizes follow a Pareto distribution, the parameter α, expresses the degree of concentration of population among cities in a system. (See section 6.1 for a detailed discussion of the Pareto distribution.) Proponents of the rank size rule assert that this coefficient is equal to 1. (Strictly speaking, $\alpha = 1$ is inconsistent with the Pareto law. In that case, the rank size rule may hold on its own.)

Whether the critical coefficient, α, is actually 1; whether the α coefficient equal to 1 is the true test of the rank size rule; and whether a Pareto distribution is at all an appropriate law of city size distributions, are all questions addressed in the literature. We refer interested readers to the works of Alperovich (1982, 1984, 1988, 1989); Cameron (1990), and Rosen and Resnick (1980). The latter paper makes international comparisons of the Pareto coefficient and tries to identify variables that explain why the rule might "work." In that sense, Rosen and Resnick's work lies in the tradition of Beckmann (1958), who shows that the rank size rule is consistent with central place theory. Empirical results using the rank size rule are certainly mixed in regard to the questions above, although it is the general conclusion of Dobkins (1994) that the strict rank size rule result—that an estimate of α is always equal to 1—is not robust to varying configurations of the data.[2]

[2]In Dobkins (1994), the θ coefficients of a Pareto distribution for city sizes are estimated with

A second and somewhat related body of work looks at changes in the α coefficient, and that work is relevant to the research at hand. As noted above, Singer (1936) used estimates of the Pareto coefficient to compare city size distributions across various time periods and countries. In Singer's interpretation, smaller Pareto coefficients indicated a greater proportion of larger places (cities) in the distribution. Thus, Singer was able to say that France in 1921 was relatively less "metropolitanized" than Germany in 1925 and the U.S. in 1920.

Singer's seminal work was followed by a comparative study by Allen (1954). Allen used data up to 1951 and compared estimates of the Pareto coefficients over time and among countries to conclude that metropolitanization had increased over time and then stabilized during the 20th century for most countries. According to Allen, the degree of concentration in U.S. city size distribution was fairly constant since 1790. Parr (1985) continued this line of research, using the methods of Singer and Allen, and found a U-shaped pattern. He hypothesized that Allen was correct for data available by 1951, but that increasing economic development would push systems toward less concentration. Parr uses "city proper" data: that is, data on entities defined in terms of legal boundaries of cities and not urban areas. Therefore, Parr suggests suburbanization as a possible explanation of the changes he sees.[3] We continue this line of inquiry by performing a rank size regression with our data set. We are look,

city proper (legal boundary) data organized by rank and organized in frequency tables. Different numbers of cities and different divisions of frequencies are used, and in each case, different α values are obtained, most of which are not unity. SMSA data for 1960 to the present were added to the exercise, and did not improve performance. The estimates of α obtained from any of these regressions are very sensitive to the definition and organization of the data.

[3]For an excellent survey of city size distribution literature through the mid-eighties, see Suarez-Villa (1988).

however, at Standard Metropolitan Statistical Areas—as Singer would have said, "conurbations"—and see a different pattern than Parr.

The question of city size distribution also underlies Eaton and Eckstein (1994), who use a different method than that of Singer, Allen and Parr, but ask the same question: how has city size distribution changed over a long historical period? Using data nonparametrically from France for 1876 to 1990, and from Japan for 1925 to 1985, they find evidence of "parallel" growth; that is to say, city size distributions in those countries have remained almost the same. We discuss the Eaton and Eckstein method in more detail below, as we attempt to apply a more complex analysis to what we consider the more complex situation of city size distributions in the U.S.

Eaton and Eckstein selected those two countries because of extremely stable geographical boundaries and because of the consistent availability of data. In contrast, the United States has grown by continuous expansion into a well defined hinterland. Numerous new regions and cities have been brought into the U.S. urban system, older regions have grown and declined, and the spatial distribution of economic activity has undergone some remarkable changes. The economic forces at work may be the same as in other economies. However, to the extent that "history matters", the U.S. urban system has developed with initial conditions quite different from those of other countries. It is for this reason, too, that a fresh approach to the U.S. case is of particular interest.

3 Theoretical aspects

The dynamics of city size distributions when cities of different sizes and types coexist is still not very well understood. Several seminal studies by Henderson [Henderson (1983; 1987; 1988)] rest on the notion that cities differ because of the demand for their products either as final goods or intermediate goods. However growth in a Henderson-type system of cities would consist of the economy's producing an increasing number of cities, with the number of each city type growing at the rate of growth of national population [Henderson and Ioannides (1981)]. The Achilles' heel of such an approach is the fact that national space is ignored. This is also true of Ioannides (1994), where the Dixit-Stiglitz-Ethier-Krugman of monopolistic competition model is employed to motivate the existence of many city types.[4] Eaton and Eckstein (1994) also work with the assumption that the price of non-urban land use remains constant over time, or at least is exogenous as far as urban growth is concerned, and thus exclude national space considerations as well. Fujita (1993) introduces a more complicated system of preferences than Ioannides (1994) but does not avoid this problem.[5]

The dynamics of city growth in Eaton and Eckstein (1994) depend critically on knowledge flows across a given number of cities. Each individual's learning productivity depends on a linear combination of the average levels of human capital. A

---

[4]Even though that model is symmetric, it is straightforward to see that a model of asymmetric preferences would produce different city types and not very different dynamics.

[5]Specifically, individuals value a (homogeneous) agricultural good (A-good) and m groups of manufactured goods (M-goods) and each group consists of a continuum of differentiated goods of size $n_i (i = 1, \ldots, m)$; n_i is to be determined endogenously, and the substitutability across different goods in each group may vary across groups:

$$U = \alpha_0 \ell n z_0 + \sum_{i=1}^{m} \alpha_i \ell n \left(\int_0^{n_i} z_i(\omega)^{\rho_i} d\omega \right)^{\frac{1}{\rho_i}}, \omega \in [0, n_i], 0 < \rho_i < 1, \ \alpha_0 + \sum_{i=1}^{m} \alpha_i = 1.$$

This basic approach is pursued further in a series of papers [Fujita and Krugman (1993); Fujita, Krugman and Mori (1994); Fujita and Mori (1994)] and leads to a theory of the evolution of urban systems. A particularly notable result obtained by this approach is that the transition to a new urban system when a new city becomes necessary involves a bifurcation of the underlying dynamic system. [Fujita and Mori (1994)] For a simpler approach to this problem, see Part II in Henderson and Ioannides (1981).

second source of the dependence of each city's growth on all others' is the assumption of a nationwide capital market. The model does not yield equilibrium city sizes for different cities. Equilibrium city sizes depend critically on the condition that at the steady state residents of different cities have no incentives to migrate. At best, human capital is general; at worst, it is perfectly city-specific. Eaton and Eckstein show that the general case where human capital is partly city-specific imply lower and upper bounds on city size distributions. These bounds share some common determinants, including, in particular, the ratio of human capital at the corresponding cities at the steady state, which is of course endogenous.[6] Models of the Henderson *genre*, on the other hand, imply a theory of city size distribution that directly reflects preferences. For similar sets of reasons, such a theory implies that all determinants are highly interdependent.

Another reason why an approach as proposed here, that relies less on specific parametric assumptions, is potentially more powerful than previous studies has to do with the fact that theory implies, in general, different equilibrium city sizes for different city types. City types change over time and, as urbanization evolves along with suburbanization, we see that city identification becomes a serious problem over time as formation of new cities causes neighboring urban areas to fuse into urban agglomerations. Henderson (1983) points out that ". . . because different types of cities have different sizes, there is a direct link between national output patterns and the size composition of cities" (p.165). Henderson goes on to explain that the types of

[6]Human capital accumulation and learning have also been invoked to explain growth of nations [Quah (1994)].

goods produced in cities helps determine the size of those cities; if the type of goods currently in vogue changes, then we would expect urban concentrations to change. Thus city-specific factors combine with aggregate ones to determine the distribution of city sizes.

Much of Henderson's work involves a system of cities in which each city specializes in the production of a single good. Abdel-Rahman and Fujita (1993) show that diversified and specialized cities can co-exist in a system, with diversified cities being larger. The specialized cities vary in size depending on the fixed costs involved in producing their specialized good. Urban hierarchies are the result of variation in industrial structures among cities and demand for the products produced.

Clearly, several distribution patterns could emerge from the Abdel-Rahman and Fujita model. If industrial structures change to favor smaller, more specialized cities, we would see less concentration, *ceteris paribus*. If industrial structures remain the same over time, we might see a parallel growth pattern, as in Eaton and Eckstein. And if industrial structures change to favor larger cities, or demand conditions change to favor the larger, more diversified cities, we might well see increasing concentration.[7]

4 Data

The very nature of change in cities suggests that consistency in data conventions will have to be handled with care. Cities have changed their boundaries, in some cases

[7]Uniqueness and stability of equilibrium city sizes rest on rather stringent assumptions. As Bernard and Durlauf (1992) and Durlauf and Johnson (1993) have emphasized in the context of growth of nations, multiplicity of equilibria of the national growth process can cause havoc upon interpretation of cross-sectional convergence in the style of Barro and Sala-i-Martin (1990; 1991).

multiplying their populations in the process. Even more challenging for our attempts at consistency is the change in transportation technology that has pushed the effective boundaries of cities out to the suburbs, and, in some cases, into surrounding cities.

The U.S. Bureau of the Census has, in its attempts to gauge population changes in light of these difficulties, created a variety of definitions for cities over the years. The Bureau moved to the Standard Metropolitan Statistical Area (SMSA) concept in 1950, and to the Metropolitan Statistical Area-Primary Metropolitan Statistical Area-Consolidated Statistical Area format between the 1980 and 1990 censuses (in 1983). These latter definitions are denoted MSA, PMSA and CMSA, respectively.[8]

The Bureau, prior to 1950, kept data only on "city proper" sizes, reflecting legal city boundaries. This type of data is still available, but ignores the very real fact of suburban integration. Cities actually are larger than their legal boundaries, in a real sense, and metropolitan area data reflect this fact. Bogue (1953) used the 1950 SMSA definitions to *reconstruct* what populations would have been in those areas in each of the decennial years from 1900 to 1940. These numbers may be the best possible bridge from the city proper data to metropolitan area data.

Most of the metropolitan area data identify city units by counties. (In New England metropolitan area definitions may involve parts of counties.) The most pressing issue involves these changing definitions within the SMSA/MSA structure. With each succeeding wave of new census information, the Bureau may adjust the grouping of counties, thus defining cities differently—or even changing the names of

[8]Census definitions may be found in Bogue (1953), and theee 1960, 1970, 1980, and 1990 *Census of Population* volumes.

groupings.

James Fitzsimmons, of the Population Distribution Branch of the Census in Suitland, Maryland, suggested to us, in a private conversation, that there are three appropriate approaches to assembling "consistent" data. First, it would be appropriate for some uses of the data to have populations for past years drawn up under a consistent set of rules, presumably the 1990 standards of the Bureau of the Census. [9]

A second way of generating consistent data would be to use the areas as defined at the time of the appropriate census. That is, we use the 1960 definitions for 1960; 1970 definitions for 1970 data, etc. This requires returning to original data sources for those years. It is quite easy to find data for 1970, 1980, and 1990 reported in the 1990 format, as is often done in the various volumes of the *U.S. Statistical Abstract.* This is part of the confusion that we see when numbers that are purportedly for the same year seem to differ, and researchers seem to be using different data sets!

A third way to generate the data would be to pick a *geographical* area that defines a city, and to use it consistently. For example, we might use the counties that define a city in 1990 (or in 1991 after the 1990 census numbers were analyzed), and then assign those counties to the cities for each year from 1960 forward. This is essentially what Bogue did in 1953. The trick here is to make this method fit Bogue's work, since the only source we have for the 1900-1940 data is Bogue. In other words, we do not want a pronounced jump in the data between 1950 and preceding years on one hand, and 1960 and succeeding years on the other hand. Because of the latter

[9] This project has been proposed in the Bureau, but has not been attempted; it apparently has a rather low priority.

consideration, we use the second method.

The biggest problem that arises in using each year's data in contemporary definition is in the span between the most recent censuses, from 1980 to 1990. The Bureau has, since 1983, considered MSAs and CMSAs to be the primary metropolitan area definitions. The Bureau first determines the outer boundaries of every metropolitan area. If that metropolitan area is larger than one million and contains identifiable PMSAs, it is labeled a CMSA. Otherwise it is an MSA.[10] The problem for us is jumping from SMSAs in 1980 to CMSAs in 1990. The 18 large metropolitan areas which are CMSAs would seem to take an enormous jump in size. Therefore, we reassembled the 1990 data to fit the 1980 definitions (by county). We therefore have SMSA data from 1900 to 1990. The 1900 to 1940 data are constructed using the 1950 definitions of SMSAs according to Bogue (*op. cit.*). The 1950 to 1980 data are consistent with the definitions in those years. The 1990 data are a reconstruction using the 1980 definitions. We believe that this method, while not perfect, is the most consistent way to construct urban area sizes. Appendix A includes a listing of the cities used in the study, the years in which they enter the system of cities structure, and other city-specific data we use in regressions reported below.

A second issue is how many metropolitan areas to use. If we adhere to the Bureau definition of metropolitan areas as having populations of more than 50,000 people, then we see a change in the number of cities each year. In 1900, there are 112 urban areas that qualify; by 1990, there have become 334. We believe that the number of

[10]That is why large metropolitan areas such as Atlanta and Minneapolis are MSAs; there are no identifiable PMSAs among their constituent counties.

new, "entering" cities, as defined in each decade, is a key feature of the U.S. urban system. Our task and our method therefore contrasts with Eaton and Eckstein, who premiss their study on an assertion that the number of cities in Japan and France remain the same over the time periods involved. That assertion is probably appropriate given the size, ages, and political geography of the economies involved.

5 Methods of studying city size distributions

Let X_{it} denote the size (in terms of different but not equivalent measures, such as population, or employment, or labor force) of city i and time t, and assume that data are available for cities $i = 1, \ldots, I$, and time periods $t = 1, \ldots T$. Let: $X_{ut}, t = 1, \ldots, T$, the total urban population, $X_{ut} = \sum_{i=1}^{I_t} X_{it}$, and $u_t = \frac{X_{ut}}{I_t}$ mean city size at time t. A theory of city size distribution, such as say Henderson's, *op. cit.*, would predict the number and sizes of different cities in the economy at any particular point in time. For completion, such a theory should predict the allocation of population between the urban and rural sectors of the economy.[11] The literature of the size distribution of cities has addressed itself to the frequency distribution of city sizes at any point in time. In general, there will be a number of types of cities, $j = 1, \ldots, n_t$; let m_j be the number of cities of type j at time t. Therefore, $I_t = \sum_{j=1}^{n_t} m_j$. In a Henderson-type model of a system of cities, behavioral parameters, β, technology characteristics, γ, and economy-wide prices, P_t, are determinants of city sizes by type,

[11] No existing theory is complete in this sense. Most theories require some fundamental friction. E.g., Krugman has typically made frictions an important part of his dynamic models. In the most extreme cases, say Krugman (1993), he assumes that farmers are perfectly immobile and (manufacturing) workers are mobile. This and related issues are pursued elsewhere in Ioannides (1994).

N_j:

$$N_j = N_j(\beta, \gamma; P_t). \tag{1}$$

For example, the particular set of assumptions made by Henderson for a static economy imply that, at any point in time:

$$\frac{m_j}{m_1} = \frac{a_j}{a_1}\frac{\alpha_j}{\alpha_1}\left(\frac{\frac{\alpha_j}{\phi_j}}{\frac{\alpha_1}{\phi_1}}\right)^{-1-\frac{1}{\psi}}, \tag{2}$$

where a_j is related to the budget share for goods produced by a city of type j; α_j is the labor share parameter in the production function of traded goods produced by a city of type j, ϕ_j is the external localization economy parameter for city j, and ψ is a function of parameters common to all city types. [Henderson (1988), p.45] This model may be augmented to allow for growth [Henderson and Ioannides (1981); Ioannides (1994)], which would introduce time dependency for the variables entering the right hand side of (2). The absolute number of cities depends upon total urban population, which is, of course, endogenous as the economy evolves over time.

The empirical approach we adopt below may be contrasted with that of Henderson's study of the determinants of urban deconcentration[12] in a cross section of countries [Henderson (1988), Chapter 10, "Empirical Determinants of Urban Concentration."]. We study, instead, the evolution of urban concentration within the United States since 1900. Henderson finds an important role for the ratio of manu-

[12]This is measured by means of an index

$$UD \equiv [\sum_{i=1}^{n}(\frac{N_i}{N})^2]^{-1},$$

where N_i denotes population of urban area i, $i = 1, \ldots, n$, N total urbanized population, and n the number of urban areas. [*ibid.*, p.194].

facturing to services and our results agree with his. As he put it, this reflects, "the spatial dispersion of footloose urban production that services resource-oriented production" [*ibid.*, p. 198]. Henderson finds that holding urbanized population fixed, urban deconcentration rises as the extent of agricultural activity rises. He interprets that increase in agricultural activity as being associated with the spreading out of the footloose segment of the urbanized populations in order to serve the agricultural population with urban services. Since we do not control for urbanized population, (but do control for GNP growth), we interpret our finding as implying that in the United States an increase in the opportunity cost of urban land use causes greater urban concentration. (See Section 6 below for results.)

5.1 Non-parametric approaches

City sizes are the outcome of economic processes associated with interactions of thoroughly open economic entities. Data on city sizes exhibit extensive variation simultaneously in both cross-section, across i for given t, and time-series, across t for given i, dimensions. In most econometric time-series analyses, one studies the dynamics of a vector of random variables, whose dimension is fairly small and fixed. Time-series analysis aims at understanding of the dynamic behavior of such a vector and patterns of interactions among its components. Time-series techniques utilize time averaging and other curve-fitting techniques, but do not involve averaging across components of a vector. Cross-section and panel-data analyses involve investigation of the behavior of the average (or representative) member of the each cross-section and deviation of each individual observation from the average across all cross section units. Typical cross-section or panel data techniques do not allow us to make inference about the patterns in the intertemporal evolution of the entire cross-section distribution. Making such inferences requires that one models directly the full dynamics of the entire distribution of city sizes. In contrast, typical panel data analyses involve efficient and consistent estimation of models where the error consists of components reflecting individual effects (random or fixed), time effects and purely random factors. The evolution of urbanization and suburbanization may affect individual cities so drastically as to render conventional methods of accounting for attrition totally inappropriate. As smaller urban units fuse[13] to create larger ones and given the small number of time series observations, non-parametric or semi-parametric distributional approaches such as the one proposed here would be the only appropriate ones. [14]

For city size distributions, in particular, accurate data are available from the U.S. Census, which takes place every 10 years. Therefore even for the United States, the availability of data are severely restricted in the time dimension. Similar is the case for the cross-section dimension, where data are available for well-defined (and hopefully consistent over time) metropolitan areas.

From population data one may construct, following Quah (1993a;b) and Eaton and Eckstein (1994), a fairly low-dimensional vector indicating the frequency of cities in each of a number of arbitrarily defined cells. All these studies use 6 cells, defined relative to the mean.[15] Let f_t denote the frequency (density) distribution of X_{it} across

[13]*C.f.* similar phenomena as addressed by the literature in economies with interacting agents.

[14]Quah (1993a; 1994) has argued forcefully that interdependence across nations through international trade and technology transfers is reflected on statistical dependence across the error structure and among the sort of variables several researchers have used (perhaps following the lead of Barro and Sala-i-Martin (1992)) as explanatory variables in regressions. Such dependence may not be adequately handled by standard methods, such instrumental estimation, GMM, etc.

[15]Quah (1993a;b) defines the end points within each distribution as the mean in the respective period times (0, .25, .50, 1, 2, ∞); Eaton and Eckstein (1994) define them as the mean times (0, .30, .50, .75, 1, 2, 20).

i at time t. Eaton and Eckstein (1994) assume that f_t evolves according to

$$f_{t+1} = M \cdot f_t, \tag{3}$$

where M is a matrix of parameters. Equ. (3) is very much like a standard first-order autoregression, except that it applies to entire distribution functions (rather than scalars or vectors of numbers) and it contains, for the time being, no explicit disturbance or innovation. Iterating forward yields:

$$f_{t+s} = (M \cdot M \cdot \ldots \cdot M) \cdot f_t = M^S \cdot f_t. \tag{4}$$

When $s \to \infty$, one may characterize the likely long-run distribution of city sizes. Divergent, convergent or parallel growth may be ascertained by the properties of $\lim_{\to\infty} : f_{t+s}$, which in turn depend upon M and f_t. If a limit distribution f_∞ exists, then it is given by the non-zero solution associated with the homogeneous matrix equation $[M - I]f_\infty = [0]$, where $[0]$ denotes a column vector of zeroes.

According to the Perron-Frobenius theorem, f_∞ is obtained in terms of the unique eigenvector of M corresponding to its single unitary eigenvalue. Parallel growth is understood to occur if f_∞ tends to a limit with non-zero probability over the entire support. Convergent growth would occur if f_{infty} tends to a mass point, and divergent growth if f_∞ tends to a polarized or segmented distribution.

Equation (3) may be generalized to allow for a generalized disturbance, u_t:

$$f_{t+1} = M^*(f_t, U_{t+1}), \tag{5}$$

where M^* is an operator that maps (f_t, U_{t+1}) to a probability measure. If F_t were restricted to be measures defined over a discrete set, then M in (3) is a Markov transition matrix. More generally,

$$\forall A \in \mathbf{R}: \; F_t(A) = \int_A M(x, A) F_{t-1}(dx), where M : R \times \mathbf{R} \to [0,1]$$

maps the Cartesian product of the real line R with its Borel sets $\mathbf{R}$ to the unit interval, then M would be a mixed-discrete continuous analogue of a transition probability matrix. There is a fairly well-developed literature on the invariant distributions of such generalized Markov chains [See Futia (1982), Quah (1994a)]. Quah (1994a) considers data for the growth of countries and conditions his non-parametric estimation on a number of "exogenous" variables. The transition dynamics are obtained for OLS residuals of pooled cross-section and time-series observations. As he explains, no individual effects are allowed for in those regressions. Permitting such individual effects would remove, and thus leave unexplained, the very object of analysis, the relative growth rates of cities. Eaton and Eckstein (1994) do not allow for any conditioning. They just compute the long-run average transition probabilities. Unlike Quah, they use the data only descriptively. It is, of course, hard to believe that economic forces shaping city size data have so conspired as to leave city size distributions unaffected. They estimate M by computing the average $M_{i,i+1}$ for all periods in the sample.[16]

Below we implement a simplified version of Quah's approach, with data for sizes of cities in the United States.

[16]They do not make clear how they deal with the non-urban population. It would be most appropriate to include in the bottom category.

6 Results

6.1 The Rank Size Rule Revisited

We start our study of the determinants of city size distribution by first returning to the methods of Singer (1936) and Lotka(1925). Seen more broadly, those methods involve working with a Pareto law for city size X_{it}.

Specifically, a random variable x is said to be Pareto distributed if it has a density function given by $f(x;x_0,\alpha) = (\frac{\alpha}{x_0})(\frac{x_0}{x})^{\alpha+1}$, with support $[x_0, +\infty)$ and $\alpha > 0$. The mean exists and is given by $E(x) = \frac{\alpha}{\alpha-1}x_0$, if $\alpha > 1$; the variance exists and is given by $Var(x) = \frac{\alpha}{(\alpha-1)^2(\alpha-2)}x_0^2$, if $\alpha > 2$. [Spanos (1989), pp. 339-340.] The corresponding cumulative distribution function is given by $F(x;x_0,\alpha) = 1 - (\frac{x}{x_0})^{-\alpha}$. Therefore, from the countercumulative probability function, $y = (\frac{x}{x_0})^{-\alpha}$, by taking logarithms, we have $\ell ny = \alpha \ell n x_0 - \alpha \ell n x$.

Using the data set described above, we estimate the *alpha* coefficient for each cross-section of cities in the ten decennial years in our sample. The estimation is based on the equation:

$$\ell n\, y_{it} \;=\; A_t \;-\; \alpha_t\, \ell n\, x_{it} \;+ \epsilon_{it}\,, i = 1, \ldots, I;\; t = 1, \ldots, T; \tag{6}$$

where y_{it} is the proportion of cities of population greater than or equal to x_{it} at time t, and ϵ_{it} is a random variable that is identically normally distributed across i for every t. That is, y_{it} is defined as the empirical countercumulative distribution of x_{it}. A_t and α_t are possibly time-varying parameters. This is, of course, a more general version of the equation used in the rank size rule literature, and implies, as a special case, α_t equal to unity. A_t should be the logarithm of the size of the largest city, if the strict rank size rule is to hold. [See Alperovich (1984).]

The results, reported in Table 1, show a decreasing estimate of α_t over time. That implies an increasing concentration over time, which is to say that more cities are getting larger relative to the increasing mean. The constant values reflect the changing proportion of "rank" to the number of cities over time. The *alpha* coefficients offer indications of subtle change in the U.S. urban structure.

The jump from 1900 to 1910 may reflect heavy foreign immigration during the first decade of the century. The stagnation of the 1930s is suggested as is the resurgence of economic activity in the 1950s. One of the most obvious movements is the much touted counterurbanization of the 1970s.[17] A more subtle reflection may be the move from manufacturing to service industries in U.S. cities, a trend that begins in the 1950s and accelerates in the 1960s.

These small but significant changes in the *alpha* coefficients prompt further investigation. We chose to move to a nonparametric method, suggested in part by the empirical methods of Eaton and Eckstein *op. cit*, and Quah *op. cit.*

6.2 Dynamics of city size

By coding the position of each city relative to the others within the distribution, we are able to see that specific cities are moving either up or down in the distribution

[17]Our analysis of *alpha* coefficients throws an interesting light on counterurbanization, in that our numbers indicate that it is—so far—a one-decade digression from the trend. That is not the view of some geographers, writing in the 1970s. We refer in particular to the words of Brian J.L. Berry (1976) that, "A turning point has been reached in the American urban experience. Counterurbanization has replaced urbanization as the dominant force shaping the nation's settlement pattern." (p.17) "The trend has been one leading unremittingly toward the reversal of the processes of population concentration unleashed by technologies of the Industrial Revolution, a reversal finally achieved after 1970." (pg.24) As Parr (1985) wrote in response to Allen (1954), "developments . . . have rendered previous conclusions. . . not necessarily wrong but incomplete." (pp.200- 201) Cadwallader (1991) noted that the trend might be reversing itself as early as 1984.

over time. We constructed transition matrices, presented in Appendix B, in which each cell gives the proportion of cities which start in a given quantile (row) in a particular year (representing 1900 in the first matrix) which move to a particular quantile (column) in the next year (representing 1910 in the next census). Entries in the diagonal indicate that cities are staying in the same category as in the previous time period. Our categories are based on .30, .50, .75, 1.00, 2.00, and 20.00 times the mean.[18] We have thus redefined X_{it} in terms of a coarser information set involving each year's mean, u_t:

$$\begin{aligned} Q_{it} &= 1, \quad \text{if} \quad X_{it} \in (0, .3\,u_t]; \\ Q_{it} &= 2, \quad \text{if} \quad X_{it} \in (.3\,u_t\,, .5\,u_t]; \\ Q_{it} &= 3, \quad \text{if} \quad X_{it} \in (.5\,u_t\,, .75\,u_t]; \\ Q_{it} &= 4, \quad \text{if} \quad X_{it} \in (.75\,u_t\,, 1\,u_t]; \\ Q_{it} &= 5, \quad \text{if} \quad X_{it} \in (1\,u_t\,, 2\,u_t]; \\ Q_{it} &= 6, \quad \text{if} \quad X_{it} \in (2\,u_t\,, \infty]\,. \end{aligned} \tag{7}$$

The first matrix in Appendix B is for the first decade of the century, and shows that some cities moved to a higher category while none moved "down". On the face of it, this transition pattern shows us, in another form, what the rank size regression results reported above yielded: an increase in concentration in the U.S. city size distribution. However, one other factor is at work: the mean is also changing, increasing as a matter of fact, from 260,851 at 1900 to 287,367 at 1910, so that these matrices will also reflect

[18]We experimented with categories based on the median; we report these results based on these categories because they conform to those used by Eaton and Eckstein.

a shift in the distribution. We averaged these decennial movements to get an average transition matrix for comparison to the Eaton and Eckstein results. It is presented as Table 3, below, and shows, as we might expect in the U.S. data, somewhat more movement off the diagonal (compared to the French and Japanese data). Most of that movement is toward greater concentration in the time period from 1900 to 1990.

Our computation of the invariant distribution, which is associated with the average transition matrix (c.f. matrix M in (3)) for 1900 to 1990, is reported in Table 3. It suggests inescapably increasing concentration at the upper end of the distribution of city sizes. This is in great contrast to the computed invariant distributions reported by Eaton and Eckstein, (*op. cit.*), for France and Japan. So, it is not just an upward trend in the mean city size, but an overall, and sharp, tendency of the city size distribution that we see.

6.3 Dynamic effects

However, these transition matrices have limitations. They do not pick up the full effect of "entering" cities and they do not offer us any more insight into why such changes might occur. We are currently working along the lines of Equation (3) to implement an estimation methodology pioneered by D. Quah. Short of having a full exploration based on such a semi-parametric approach, we think that a number of fairly ad hoc approximations may throw additional light on this phenomenon. One of these approximations considers the dynamics of movements of cities across quantiles by modelling the propensity of such movements as a function of observable characteristics of individual cities and of the state of the economy. The second approximation attempts to explain the dynamics of city sizes, as measured by city population.

We recall the discrete categories of size based on the mean which we defined above in order to compute transition matrices. We define a new "coarser" variable that tells us whether a given city moves to a "higher" or "lower" cell, or stays in the same cell. That is, we define a "change-in-quantile" variable, Z_{it}, defined below, based on the quantiles in equation (7) above:

$$
\begin{aligned}
Z_{it} &= 1 \quad \text{if } Q_{i,t} = j \text{ and } Q_{i,t-1} = j - 1 \,, \\
Z_{it} &= 0 \quad \text{if } Q_{i,t} = Q_{i,t-1} = j \,, \\
Z_{it} &= -1 \quad \text{if } Q_{it} = j - 1 \text{ and } Q_{i,t-1} = j \,;
\end{aligned} \tag{8}
$$

where $j = 1, \ldots, 6$. Cities that enter the distribution in a given year (usually because they pass the 50,000 population threshold definition) are denoted as rising in category. The change-in-quantile variable is used as a dependent variable in regressions which explore the determinants of changes in the city size distribution.

Descriptive statistics for the change-in-quantile variable are given in Table 6, along with statistics for the regressors that we use below. The mean of the change variable is 0.26439, indicating that more cities are moving up in the distribution than are falling in the distribution, which of course, bears out our rank size regression results.

We study the dynamics of the evolution of city size distribution by means of a simplified version of equation (1). That is, we regress this change variable against three types of regressors: those macroeconomic variables that vary by time period, (the decennial years from 1910 to 1990), are denoted by the vector N_t; those that vary by city but are invariant over time are denoted by C_i. The quantiles defined above express a city's position within the distribution and thus vary over both cities and time. The quantile in which a city is positioned is measured at the beginning of the decade so that $t - 1 = 1900$ to 1980. Our regression therefore takes the form:

$$
Z_{it} = a + \beta N_t + \gamma C_i + \delta Q_{i,t-1} + \epsilon_{it}^{Z}, \; i = 1, \ldots, 334; \; t = 1910 \text{ to } 1990 \,, \tag{9}
$$

where Z_{it} is the change variable for city i at time t; the random variable, ϵ_{it}^{Z} is independently and identically normally distributed for all i and t.

The theoretical relationships explored in Section 3 suggest candidate regressors to explain these distribution changes. Henderson's results that city size distributions depend on demand and output conditions and that different sizes of cities produce different types of goods led us to consider the balance between service and manufacturing employment nationwide. We suspect, as Henderson (1983) did, that the move from manufacturing to service dominance might be significant.[19] Thus we regress our change variable against the changing balance between manufacturing and service industries' employment. Because the number expressing this balance declines as manufacturing employment decreases, we would expect to see a negative coefficient.

We also include as a regressor the growth in real GNP, as an indicator of output conditions. One could hypothesize several relationships here, and we were curious to see what results would obtain. It is possible that GNP is reflected in the other variables we have included.

Many theoretical urban models include the value of agricultural land at the city's edge as a variable in accounting for the pattern and rate of urban spread (e.g., O'Sullivan (1993)). Surely the value of rural land is a factor in city growth. The value of agricultural land depends, of course, on the demand for, and the technological supply conditions of, its products. That value is related, however, to urban spread, particularly in those regions in which new cities are springing up and old ones

[19] Dobkins (1995) shows a decline in city specialization and particularly in manufacturing specialization; in fact, manufacturing specialization may have a negative impact on growth in some industries.

spreading out. Moreover, the value of agricultural land introduces a suggestion of national space considerations, as modeled by Fujita (1993). Consequently, we use the changing average value of farm land and buildings as a regressor. We would expect a positive number here, implying that as agricultural land gets more expensive, cities move up in the distribution, given that we have more cities. Fewer cities would also be a possibility, with no change in the distribution. However, given that cities are entering, we would expect the distribution to tip toward larger cities instead of toward more and smaller cities.

Eaton and Eckstein, *op.cit.*, assume a nationwide capital market that links cities. We agree that the cost of capital is potentially significant and we include the national real prime rate as a regressor.[20] Theory is not totally clear on the predicted impact of the national interest rate. The model in Ioannides (1994) suggests that equilibrium city size might be negatively related to the interest rate for high values of the latter; in the case of low interest rates, the relationship could be either positive or negative. For higher rates, equilibrium city size would decrease due to the dependence of city size on manufacturing capital. Eaton and Eckstein's emphasis on city specific human capital is more difficult data to obtain for 334 cities, and so we do not attempt, at present, to account for it.

Urban history is not complete without noting the ebb and flow of foreign immigration. Because the majority of immigrants move to cities, and many to the largest cities, we include a measure for immigration. [Chudacoff and Smith (1994)]. We would expect large immigration flows to lead to larger cities in the concentration, in

[20]Eaton and Eckstein assume for convenience that the national rate does not change.

part because of the immigration flows to our very largest cities[21]

Each of these variables applies to all cities equally and vary only across time periods. We also note the possibility of regional effects, coding each city in the sample by the Census Bureau's nine regions: New England, Middle Atlantic, South Atlantic, East North Central, East South Central, West North Central, West South Central, Mountain and Pacific. We also note the distance (measured as driving mileage) from the nearest node for cities in the sample. A city that serves as a primary node was given a zero for distance; cities in the next rung were judged by their distance to the nearest primary node, etc. Actual driving distances were used, with no attempt to judge "economic" distance, which might be more in keeping with the intent of central place theory.[22] A description of the distance variable is provided in Appendix A, along with the list of cities, regions, and years of "entry." There are undoubtedly other variables that might impact on city size distribution. Collecting these data is often constrained either by the number of cities involved or by the time range.

Table 2 presents descriptive statistics for real Gross National Product, the real prime rate, immigration, the percentage of total employment in manufacturing and services, respectively, and the average real value of agricultural land and buildings, to give a sense of the change over the century. The values actually used in one of our regressions are amended as appropriate to indicate change over the decade. A map indicating the boundaries of the nine regions and the number of cities located

[21] New York City early in this century and Miami and Los Angeles in the 1980s are examples.

[22] Cameron (1990) attempted to measure the effect of distance on population change for 121 cities over 100,000 population in 1984. The variable was significant, but she, too, was unhappy with its definition.

in each is Figure 1. Table 6 reports the descriptive statistics for all variables used in our regressions. We see that the mean population in these cities *over time* is close to a half million people. We also see that the swing in the balance of manufacturing to service is a fairly recent thing, as reflected in a mean higher than one. (See Table 2 for the numbers on which this balance variable is based.)

6.3.1 Dynamics of U.S. city size distribution

Because our dependent variable is descriptive, falling into one of three categories of growth, stasis or decline, we performed homogeneous ordered probit regressions. (We are interested in more general stochastic structures, such as exploiting the panel aspect of the data. However, for the time being, we do not have appropriate software to carry out an ordered probit panel regression.) We also performed ordered logit, and linear probability regressions with the panel aspects of the data. We obtained very similar results. While the panel structure was important, the performance of the remaining explanatory variables was very similar to what we report here. Because we want to look at change from decade to decade, the data set actually includes 1,876 observations from 1900-1910 to 1980-1990.

The $Q_{i,t-1}$ regressors include dummy variables for the quantiles. That is to say, we control for *which* quantile a city starts out in at the beginning of each decade. The omitted dummy here is the largest quantile. Also included among the C_i regressors are eight regional dummies; the omitted dummy is for the West North Central region. Results are reported in Table 4 and discussed below.[23] The estimates of the intercept

[23]It is difficult to judge the explanatory power of these models; certainly the traditional R-squared

and of the threshold, which define the ordered structure is very significant.

We find that the change in the real value of agricultural land is significantly and positively associated with cities moving up in the distribution. This accords with theory: the higher the value of rural land, the less urban spread and the more urban concentration one would expect.

The relative importance of manufacturing employment to service sector employment is measured by a variable that expresses the balance (ratio) between manufacturing and service employment; that variable is significant and has the expected negative sign. This variable is set up to reflect the changing ratio of manufacturing to service and therefore declines as service employment comes to dominate. The economic interpretation is that service industries have economies of scale that allow for—and require—greater city size concentration. Specialization in manufacturing may have become, increasingly, the province of smaller cities, possibly due to product-cycle considerations. Therefore, as service industries come to dominate (26% of employment in 1990 compared to 17% in manufacturing), we see the trend to larger cities emerge. Whether or not the service sectors will change in nature as their own product-cycles take over is a question for future data to resolve.

The third significant, time-varying, factor is the real prime rate, our proxy for the cost of capital. It is negatively related to growing concentration. If we are to interpret this result in light of Ioannides (1994), we would say that the rates over the century were high enough to be associated with increasing equilibrium city size.

is not appropriate. Looking at the predictions of the model in terms of the three dependent values is not considered appropriate, either, since there is no well-defined criteria for such predictions. See the discussion in Kennedy (1992), Chapter 15.

Immigration is not highly significant, although it does have the expected positive sign, and the t-statistic is 1.495. Growth in real GNP has a negative coefficient, but is not significant in explaining the change in concentration, suggesting that the other variables themselves are explaining the components of GNP growth. Compared to the highest quantile (largest size category), all quantiles except the smallest show significance. This confirms our emphasis on a non-parametric approach to the study of the dynamics of the U.S. city size distribution.

The regions show significance in a pattern which regional economists and geographers have come to expect. Recall that our change variable reflects cities moving into the urban category as well as existing cities moving up. The New England and Middle Atlantic regions are negatively related to the concentration process, but not significantly so. The East North Central "rust belt" is positive but not significantly so. But the southern and western regions show highly positive and significant impact. These regions include the South Atlantic, East South Central, West South Central, Mountain and Pacific. (The referent here is the West North Central.)

6.3.2 Dynamics of Urban Population

We now look also at determinants of actual city populations. Understanding the possible explanations for changing city sizes may shed some additional light on our question of changing distributions. Cameron (1990) discusses in detail the appropriate distribution for city sizes, and concludes that a lognormal distribution is preferred. We follow Cameron and regress the logarithm of population against our variables, as follows:

$$\ell n\, P_{it} = a + \beta N_t + \gamma C_i + \epsilon_{it}^{P}, i = 1, \ldots, 334;\ t = 1910 \text{ to } 1990, \qquad (10)$$

where P_{it} is the population of city i in time t, and the random variable ϵ_{it}^{P} is independently, identically normally distributed for all i and t. All other variables are as defined above. We find that several of our regressors are significant. The results are reported here and in Table 5. Like Cameron, we find that distance between cities is significant in explaining city size; unlike her, we find that distance has a *positive* effect, meaning that the greater the distance from the nearest "node", the larger the city. In our case, a one mile increase in distance is associated with a .05 percent increase in size. We note that Cameron defined distance to be the mileage from the center of one city to nearest city in her sample, which included only cities of population 100,000 and larger. Our sample includes smaller cities and our distance variable is based on driving mileage to the nearest "node". Cameron's results seemed to contradict central place theory or notions of hierarchy. To the extent that central place theory is the correct theory, one would expect more isolated cities to be larger. Those more isolated cities are serving nodal functions as well as the long-distance trade and manufacturing functions served by less isolated cities. That is what our results do suggest.[24]

As for the regional dummy variables, we see highly significant, positive impact of a city being located in New England, the Middle Atlantic, South Atlantic, East North Central, and Pacific regions (in reference to the West North Central region).

The balance between manufacturing and service has a negative coefficient in relation to population. This variable is now expressed as the number employed in

[24]See Marshall (1989) for a discussion of the relevance of central place theory to local as opposed to long distance trade.

manufacturing over the number employed in service industries, so that a one-unit change is rather large. (See Table 3.) These regression results tell us that a one-unit decrease in the balance variable leads to a 14 percent increase in population. The result here accords with our understanding of this variable in the previous regression.

The results regarding immigration are less clear cut. In the previous regression, the t-ratio was positive but not quite significant. Here, it is highly significant and negative. Our results tell us that a one percentage point increase in immigration yields a 6.4 percent decrease in city size. This is puzzling in view of the general notion that immigrants move to the largest cities. However, Filer (1992) points out that immigration inflow to metropolitan areas is negatively correlated with native outflow, in the 1970s. So, it may be the case that immigrants do move to the larger cities, but that this does not increase population.

The real prime rate and agricultural land values are no longer significant in this regression, but the growth in real GNP is. The results tell us that a one percentage point growth in GNP (measured over a decade) is associated with a .72 percent increase in city population. This result, of course, does not indicate causation; it is still an open and intriguing question whether economic growth in cities causes national growth, or vice versa, to which we plan to return elsewhere.

6.3.3 Dynamics of City Size Distribution for Cities Extant in 1900

Because entering cities exercise a potentially strong influence on the regression results, we also ran regressions that use data from only the 112 cities that met the criteria for inclusion in the sample in 1900. Based on equation (11) below, we regressed the quantile in which each of the 112 cities was located in each decennial year (1900 to 1990) against the same set of explanatory regressors used above minus distance.

For these purposes, the "quantile" was binary: cities were either above or below the median. Therefore, our Q_{pt} in this equation indicates that a city is in either the lower half or upper half of the distribution; $p = 1$ if the city is in the upper half and $p = 0$ if the distribution is in the lower half:

$$Q_{pt} = a + \beta N_t + \gamma C_i + \epsilon_{it}^{Q}, i = 1, \ldots, 112; t = 1900 \text{ to } 1990; p = 1, 2, \quad (11)$$

where the random variable, ϵ_{it}^{Q} is independently and identically normally distributed for all i and t. The other variables are as defined above. (The mean for the quantile variable is 0.6741.)

There are several ways to proceed with such regressions: we used both a linear, panel regression so that we could consider fixed effects, and a probit regression. Considering all these methods, we found one consistent result: among our time-varying variables, only the balance between manufacturing and service employment is significant, and always negatively so. This agrees with results for the full set of cities; however, with the smaller set, the other variables are not significant. In the restricted set regressions, all the regions show (varying) significance. Overall, we interpret this to mean that among the set of 112 cities that had population over 50,000 in 1900—cities that by and large are older and heavily represented in the top half of the distribution over time—the changing balance between manufacturing and services has been particularly important in explaining the distribution.

7 Conclusions

We are concerned with city size distributions in the United States in this century and the determinants of those distributions. The rank size-Pareto distribution indicates small but significant movement toward increasing concentration, based on the declining exponent of the Pareto tail and increasing means.

In a nonparametric way, we see similar results in the transition of size distributions across each of the last ten decades for the U.S., in contrast with results from Eaton and Eckstein for France and Japan. The average transition matrix expresses changing distributions, but not the increase in the number of cities that is an important feature of the American urban structure.

We see the impact on the distribution of new cities entering by regressing the change in each city's position within the distribution against a variety of time-varying and city specific factors. We find that increasing agricultural land values, decreases in the real prime rate, the changing balance from manufacturing to service industries, as well as location in western and southern regions make it more likely that a city will move upward within the distribution.

We also regress the logarithm of population against our variables in an attempt to explain city size. We find that the growth in real Gross National Product, the shifting balance from manufacturing to service sector employment, increased distance from each city to the nearest nodal city, and declining immigration are significant in explaining increased city size.

Using the 112 cities of the 1900 Census, we regress the quantile (upper or lower) against our macroeconomic variables and against regions. We find that only the shift from manufacturing to service sector employment is significant in explaining the distribution of cities between these two quantiles.

Our data set is constructed to reflect changing definitions of Standard Metropolitan Statistical Areas from 1900 to 1990. While no ideal data set for cities exists, we feel fairly confident that the one we have constructed anew reflects changing numbers of cities, changing city sizes and the changing distribution. Our methods are designed to reflect the complexities of our panel of observations, which vary across both time and cities.

There are undoubtedly other variables which might be influential. We would like to have had consistent data on educational levels in cities over time in order to say more about Eaton and Eckstein's emphasis on human capital accumulation. We would like to be able to say something about commuting costs, because standard urban theory suggests that the changing technology of commuting has contributed to urban spread. We will continue to search for these and other variables that might have an impact. It will be of interest to see what insight the census of 2000 may yield. We have work in progress which adopts a more formal method of non-parametric analysis along the line of Quah's work. Nevertheless, we have prescribed a method and described results which we feel shed light on the shifting distribution of city sizes in the United States to date.

BIBLIOGRAPHY

Abdel-Rahman, Hesham and Masahisa Fujita (1993),"Specialization and Diversification in a System of Cities ", *Journal of Urban Economics*, 33,189-222.

Allen, G.R. (1954) "The 'Courbe Des Populations,' A Further Analysis, *Oxford University Institute of Statistics Bulletin*, 16, 179-189.

Alperovich, Gershon (1982) "Scale Economies and Diseconomies in the Determination of City Size Distributions," *Journal of Urban Economics*, 12, 202-213.

Alperovich, Gershon (1984) "The Size Distribution of Cities: On the Empirical Validity of the Rank-Size Rule," *Journal of Urban Economics*, 16, 232-239.

Alperovich, Gershon (1988) "A New Testing Procedure of the Rank Size Distribution," *Journal of Urban Economics*, 23, 251-259.

Alperovich, Gershon (1989) "The Distribution of City Size: A Sensitivity Analysis," *Journal of Urban Economics*, 25, 93-102.

Auerbach, F. (1913) "Das Gesetz der Bevölkerungskoncentration," *Petermanns Geographische Mitteilungen*, 59, 74-76.

Bairoch, Paul (1988), *Cities and Economic Development*, Chicago: University of Chicago Press.

Barro, Robert J., and Xavier Sala-i-Martin (1991), "Convergence across States and Regions," *Brookings Papers on Economic Activity*, 1, 107-182, Apne.

Barro, Robert J., and Xavier Sala-i-Martin (1992), "Convergence," *Journal of Political Economy*, 100(2), 223- 251.

Beckmann, M. J. (1958) "City Hierarchies and the Distribution of City Size", *Economic Development and Cultural Change*, 6, 243-248.

Bernard, Andrew B., and Steven N. Durlauf (1992), "Interpreting Tests of the Convergence Hypothesis," mimeo, MIT and Stanford University, February.

Bogue, Donald J. (1953) *Population Growth in Standard Metropolitan Areas 1900-1950*, Oxford, Ohio: Scripps Foundation in Research in Population Problems.

Cadwallader, Martin (1991) "Metropolitan Growth and Decline in the United States: an Empirical Analysis," *Growth and Change*, Summer 1991, 1-16.

Cameron, Trudy Ann (1990) "One-Stage Structural Models to Explain City Size," *Journal of Urban Economics*, 27, 294-307.

Chudacoff, Howard P. and Judith E. Smith (1994) *The Evolution of American Urban Society*, Englewood Cliffs, NJ: Prentice Hall, Chapter 4.

Dobkins, Linda Harris (1994) "A Historical Perspective on City Size Distribution Studies", mimeo, Virginia Polytechnic Institute and State University.

Dobkins, Linda Harris (1995) "Specialization in Cities", Ph.D. dissertation, Virginia Polytechnic Institute and State University, June.

Durlauf, Steven N., and Paul A. Johnson (1992)," Local versus Global Convergence across National Economies," mimeo, Stanford University and University of Oregon, January.

Eaton, Jonathan, and Zvi Eckstein (1994), "Cities and Growth: Theory and Evidence from France and Japan," NBER, Working Paper No. 4612, January.

Filer, Randall K. (1992), "Immigrant Arrivals and the Migratory Patterns of Native Workers," in *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, George J. Borjas and Richard B. Freeman, eds. Chicago: University of Chicago Press.

Fujita, Masahisa (1993), "Monopolistic Competition and Urban Systems," *European Economic Review*, 37, 308-315.

Fujita, Masahisa and Paul Krugman (1993), "A Monopolistic Competition Model of Urban Systems and Trade," Working Papers in Regional Science No. 163, University of Pennsylvania, October.

Fujita, Masahisa, Paul Krugman, and Tomoya Mori (1994), "On the Evolution of Hierarchical Urban Systems," presented at the North American Meetings of Regional Science International, Niagara Falls, November.

Fujita, Masahisa and Tomoya Mori (1994), "Structural Stability and Evolution of Urban Systems," presented at Special Conference, Tel-Aviv University, May.

Futia, Karl (1982), "Invariant Distributions and the Limiting Behavior of Markovian Economic Models," *Econometrica,* 50, 377-408.

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

Henderson, J. Vernon (1983), "Industrial Bases and City Size," *American Economic Review*, 73, 164-168.

Henderson, J. Vernon (1987), "General Equilibrium Modeling of Systems of Cities," Ch. 23, 927-956, in Edwin S. Mills, ed. *Handbook of Regional and Urban Economics*, Volume II, Amsterdam, North-Holland.

Henderson, J. Vernon (1988), *Urban Development: Theory, Fact and Illusion,* Oxford: Oxford University Press.

Henderson, J. Vernon and Yannis M. Ioannides (1981), "Aspects of Growth in a System of Cities," *Journal of Urban Economics*, 10, 117-139.

Ioannides, Yannis M. (1994) "Product Differentiation and Economic Growth in a System of Cities, " *Regional Science and Urban Economics,* 24, 461-484.

Ioannides, Yannis M. (1995) "Urbanization and Economic Growth," (in progress).

Kennedy, Peter (1992) *A Guide to Econometrics* Cambridge, MA: The MIT Press.

Knox, Paul L.(1994) *Urbanization: An Introduction to Urban Geography*, Englewood Cliffs, NJ: Prentice Hall.

Krugman, Paul (1994) "Complex Landscapes in Economic Geography," *American Economic Review, Papers and Proceedings*) 84, 412-416.

Krugman, Paul (1993) "On the Number and Location of Cities," *European Economic Reviews* 37, 293-308.

Marshall, John U. (1989) *The Structure of Urban Systems*, Toronto: University of Toronto Press.

Lotka, Alfred (1925) *Elements of Physical Biology*, Baltimore: Williams and Wilkins Company.

O'Sullivan, Arthur (1993) *Urban Economics*, Homewood, IL: Irwin Press.

Pareto, V. (1906) *Cours d'Economie Politique*, Libraire Droz, Geneva, 1964.

Parr, John B. (1985) "A Note on the Size Distribution of Cities over Time," *Journal of Urban Economics*, 18, 199-212.

Quah, Danny (1993a) "Empirical Cross-section Dynamics and Economic Growth," *European Economic Review*, 37 (2/3), 426-434.

Quah, Danny (1993b) "Galton's Fallacy and Tests of the Convergence Hypothesis," *The Scandinavian Journal of Economics*, 95, 4, 427-443.

Quah, Danny (1994a) "Convergence Empirics across Economics with (Some) Capital Mobility," LSE Economics Department mimeo, January.

Quah, Danny (1994b) "Ideas Determining Convergence Clubs," LSE Economics Department mimeo, September.

Rosen, Kenneth T. and Mitchell Resnick (1980), "The Size Distribution of Cities over Time", *Journal of Urban Economics*, 8,165-180.

Singer, H. W. (1936) "The 'Courbes des Populations: A Parallel to Pareto's Law," *Economic Journal*, 46, 254-263.

Spanos, Aris (1989) *Statistical Foundations of Econometric Modelling*, Cambridge: Cambridge University Press.

Suarez-Villa, Luis (1988) "Metropolitan Evolution, Sectoral Economic Change, and the City Size Distribution", *Urban Studies*, 25, 1-20.

United States Bureau of the Census (1960, 1970, 1980, 1990), *Census of Population.*

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

Table 1

ESTIMATES OF THE PARETO EXPONENT

This table gives the number of cities available for each year, the mean size for those cities, and the estimated constant, and the Pareto *alpha* exponent.

Year	Number of cities	Mean city size	Constant	*Alpha* value (s.e.)
1900	112	260,851	11.419	-1.04398 (0.010145)
1910	139	287,367	11.106	-1.01431 (0.009261)
1920	149	338,549	11.214	-1.01019 (0.008561)
1930	157	411,378	11.075	-0.98543 (0.010481)
1940	160	438,433	11.263	-0.99463 (0.010709)
1950	162	528,223	11.523	-0.99860 (0.011742)
1960	210	536,158	11.278	-0.97706 (0.011092)
1970	243	573,742	10.986	-0.94940 (0.011929)
1980	322	526,177	11.378	-0.98518 (0.009611)
1990	334	576,383	11.977	-0.94869 (0.010119)

Source: 1900-1950:D. Bogue (1953); 1960-1990: U.S. Census Bureau publications.

Table 2

ECONOMIC DATA
1900 – 1990[1]

Year	Real GNP in billions(2)	Real prime rate(2)	Manufacturing percent(3)	Service percent(3)	Average real value of agricultural land(2)	Immigration percent(4)
1900	$71.2	5.92%	36%	11.5%	$81.58	4.85%
1910	$107.5	3.82%	36.1%	11.1%	$135.08	9.52%
1920	$135.9	-.50%	39%	11.3%	$105.98	4.38%
1930	$184.8	6.39%	32.5%	11.5%	$98.32	3.25%
1940	$229.2	1.76%	33.9%	11.4%	$72.20	.40%
1950	$354.9	7.45%	33.7%	11.9%	$81.01	.68%
1960	$497.0	1.35%	31%	13.7%	$107.07	1.39%
1970	$747.6	5.12%	27.4%	16.3%	$139.65	1.62%
1980	$963.0	7.97%	22.4%	19.8%	$271.29	1.97%
1990	$1277.8	5.71%	17.4%	25.6%	$154.54	2.94%

1. All figures are taken from *Historical Statistics of the United States from Colonial Times to 1970*, Volumes 1 and 2, and *Statistical Abstract of the United States, 1993.*
2. Adjusted by the implicit price deflator constructed from sources above; 1958=100.
3. Manufacturing, as well as service industries, employment is given as a percentage of the total employment for each year.
4. Foreign immigration is summed over the preceding decade, and reported as a percentage of the total population of the respective decennial year.

Table 3

AVERAGE TRANSITION MATRIX
U.S. CITIES 1900 - 1990

The cells in this matrix are identified by the upper endpoints of the categories, as explained in the text; that is, .3, .5, .75, 1, 2, and 20 times the mean. Entries in the cells are the averages over nine matrices that define decade to decade changes. See Appendix B for the decade matrices. The total cities column and row give the actual distributions for 1900 (summing to 112) and 1990 (summing to 322).

1990/1900	0.3	0.5	0.75	1	2	20	Total cities
0.3	79.84	6.90	0.41	0	0	0	24
0.5	19.79	66.36	9.07	0	0	0	31
0.75	0.37	25.66	62.54	7.62	2.53	0	15
1	0	1.09	21.48	53.47	4.48	0	14
2	0	0	6.49	38.91	79.23	.85	15
20	0	0	0	0	13.76	98.76	13
Total cities	110	64	51	23	38	36	322/112

Table 4

DYNAMICS OF CITY SIZE DISTRIBUTION

Ordered Probit Model for LHS = Distribution Change
Maximum Likelihood Estimates
Log-Likelihood.............. -1241.663
The time-varying macroeconomic variables are reported in the top portion of the table, with the city specific variables in the lower portion.

VARIABLE	Coefficient	Std. Error	t-ratio
Constant	2.3515	0.2923	8.044
GNP Growth Rate	-0.25332E-02	0.3281E-02	-0.772
Real Prime Rate	-0.75554E-01	0.1473E-01	-5.132
Balance Change, manuf./service	-0.78750	0.2581	-3.052
Immigration	0.22560E-01	0.1511E-01	1.493
Rural land value change	0.12078E-01	0.7345E-03	17.077
Distance	0.55965E-04	0.3206E-03	0.175
New England	-0.16167	0.1478	-1.094
Middle Atlantic	-0.16648	0.1448	-1.149
South Atlantic	0.45355	0.1329	3.413
East North Central	0.87278E-01	0.1334	0.654
East South Central	0.36021	0.1646	2.188
West South Central	0.55574	0.1405	3.956
Mountain	0.87589	0.1849	4.737
Pacific	0.76356	0.1577	4.843
Q1	0.61035E-01	0.1373	0.445
Q2	0.29788	0.1381	2.157
Q3	0.38210	0.1444	2.646
Q4	0.55070	0.1607	3.428
Q5	0.33032	0.1547	2.135
Threshold	2.6170	0.7045E-01	37.149

Table 5

DYNAMICS OF CITY POPULATION

Ordinary least squares regression.
Dependent Variable = Logarithm of city population
Observations = 1876
The time-varying macroeconomic variables are reported in the top portion of the table, with the city specific variables in the lower portion.
Adjusted R-squared= 0.1238008E+00 F[14,1861] = 0.1992317E+00
Log-likelihood = -0.2508446E+04 Durbin-Watson stat.= 1.9159334

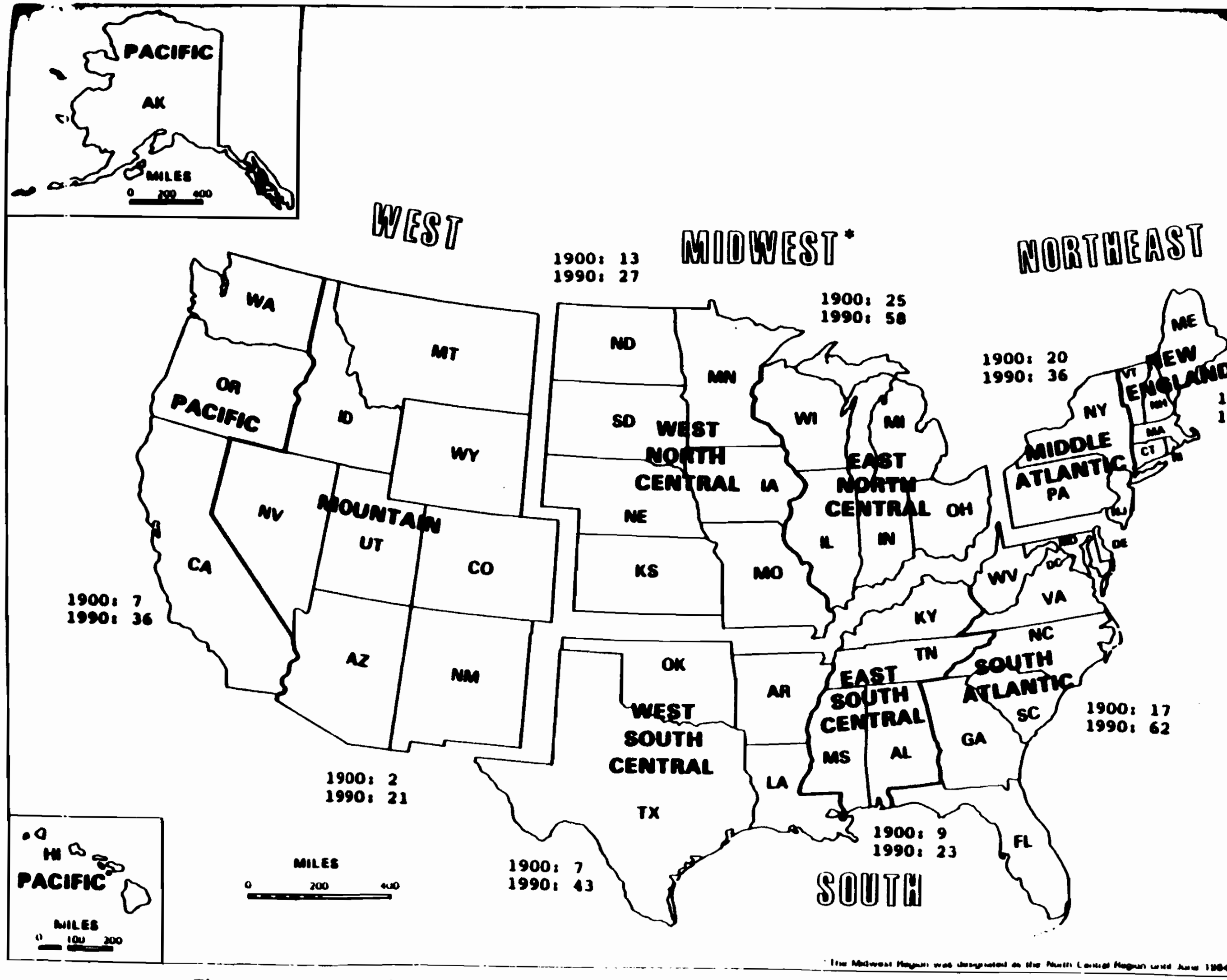
VARIABLE	Coefficient	Std. Error	t-ratio
Constant	12.420	0.1349	92.082
GNP Growth	0.71605E-02	0.2314E-02	3.095
Real Prime Rate	-0.12489E-01	0.1116E-01	-1.119
Balance, manuf./service	-0.18436	0.2943E-01	-6.264
Immigration	-0.63963E-01	0.1018E-01	-6.282
Rural land value change	0.13536E-03	0.5131E-03	0.264
Distance	0.54127E-03	0.2185E-03	2.477
New England	0.30434	0.1060	2.872
Middle Atlantic	0.84347	0.9759E-01	8.643
South Atlantic	0.21458	0.9027E-01	2.377
East North Central	0.23389	0.8858E-01	2.640
East South Central	0.14277	0.1105	1.292
West South Central	-0.61289E-01	0.9365E-01	-0.654
Mountain	-0.75230E-01	0.1213	-0.620
Pacific	0.70562	0.1031	6.844

Table 6

DESCRIPTIVE STATISTICS
1900 - 1990

The time-varying macroeconomic variables are reported in the top portion of the table, with the city specific variables in the lower portion.

VARIABLE	Mean	Stnd. Deviation	Minimum	Maximum
Real GNP (billions)	$629.70	417.20	$107.50	$1277.80
Real Prime Rate (%)	4.77	2.67	-0.50	7.97
Immigration (%)	2.64	2.22	0.40	9.52
GNP Growth(%)	37.44	10.21	24.0	54.8
Balance change, manuf./service	0.81	0.16	0.60	1.06
Rural Land Value Change	17.08	46.20	-42.0	92.0
Balance, manf./service	2.03	0.97	0.68	3.45
Rural Land Value (per acre)	$143.96	63.53	$72.20	$271.29
Change	0.2643923	0.53	-1.00	1.00
Population	492883.38	1022918.46	50731	12912000
Distance (miles)	135.5	107.98	0	600.0
New England	0.0868870	0.2817445	0	1.0
Middle Atlantic	0.1199360	0.3249733	0	1.0
South Atlantic	0.1700426	0.3757704	0	1.0
East North Central	0.2041578	0.4031923	0	1.0
East South Central	0.0655650	0.2475862	0	1.0
West North Central	0.0895522	0.2856153	0	1.0
West South Central	0.1257996	0.3317118	0	1.0
Mountain	0.0479744	0.2137691	0	1.0
Pacific	0.0900853	0.2863802	0	1.0
Quantile 1	0.3491471	0.4768276	0	1.0
Quantile 2	0.2153518	0.4111758	0	1.0
Quantile 3	0.1439232	0.3511055	0	1.0
Quantile 4	0.0724947	0.2593743	0	1.0
Quantile 5	0.1183369	0.3230927	0	1.0
Quantile 6	0.1007463	0.3010727	0	1.0



The number of cities in our sample from each region are given for 1900 and 1990.

FIGURE 1 Census Divisions of the United States

Source: U.S. Department of Commerce, Economics and Statistics Administration
Bureau of the Census

APPENDIX A

Cities, Regions and Distances

This Appendix lists the cities used in our regressions, the year in which they enter the data set, the region in which they are located, and the distance variable. A word about the latter variable is in order.

This variable is the actual driving mileage from each city in the sample to the nearest larger city in a higher tier. In order to construct the tiers, we took as our basic classification a listing of U.S. cities by function (nodal centers) from Paul Knox, *Urbanization: An Introduction to Urban Geography*, 1994. Because we are considering a central place-type classification, we combine cities in top tiers. (New York City would stand alone in most category systems because of its status as "world" city; we include it among several cities in our top tier for geographic purposes. We also "promoted" Atlanta, Denver, Houston, Miami and Seattle for purposes of geographic convenience.) Our amended listing is as follows:

Top tier cities: New York City, Chicago, Los Angeles, Houston, Miami, San Francisco, Washington D.C., Atlanta, Denver, Seattle.

Regional nodal centers: Baltimore, Boston, Cincinnati, Cleveland, Columbus OH, Dallas, Indianapolis, Kansas City, Minneapolis, New Orleans, Philadelphia, Phoenix, Portland OR, and St. Louis.

Subregional nodal centers and industrial centers: Birmingham, Charlotte, Des Moines, Jackson MS, Little Rock, Memphis, Mobile, Nashville, Oklahoma City, Omaha, Salt Lake City, Shreveport, Syracuse, Richmond, Detroit, Hartford, Milwaukee, Tampa and Pittsburgh.

Cities in the top tier are given a zero for the distance variable. The mileage from a city in the regional nodal center class is judged to the nearest city in the top tier; mileages for cities in subregional class may be to a city in either the top tier or regional class. Mileages to all other cities are judged by the nearest larger city in any of the three classifications above. The only exceptions are Honolulu and Anchorage, which receive 600 simply because mileage is a relatively meaningless measure, and 600 is larger than any other actual mileage. A list of cities in the sample, given in order of size, year of entry, their region, and distance from a "nodal" city follows:

CITY	YEAR	REGION	DISTANCE	FROM NODE
New York City NY	1900	MA	0	New York City
Los Angeles CA	1900	PC	0	Los Angeles
Chicago IL	1900	ENC	0	Chicago
Philadelphia PA	1900	MA	101	New York City
Detroit MI	1900	ENC	62	Cleveland
Washington D.C.	1900	SA	0	Washington D.C.
San Francisco CA	1900	PC	0	San Francisco

Houston TX	1900	WSC	0	Houston
Atlanta GA	1900	SA	0	Atlanta
Boston MA	1900	NE	222	New York City
Nassau NY	1980	MA	20	New York City
Riverside CA	1910	PC	60	Los Angeles
Dallas TX	1900	WSC	246	Houston
San Diego CA	1910	PC	124	Los Angeles
Minneapolis MN	1900	WNC	465	Chicago
St. Louis MO	1900	WNC	299	Chicago
Anaheim CA	1970	PC	30	Los Angeles
Baltimore MD	1900	SA	45	Washington D.C.
Pittsburgh PA	1900	MA	157	Cleveland
Newark NJ	1960	MA	40	New York City
Phoenix AR	1920	MT	369	Los Angeles
Tampa FL	1910	SA	282	Miami
Seattle WA	1900	PC	0	Seattle
Miami FL	1930	SA	0	Miami
Denver CA	1900	MT	0	Denver
Cleveland OH	1900	ENC	360	Chicago
Kansas City MO	1900	WNC	527	Chicago
San Jose CA	1900	PC	44	San Francisco
Sacramento CA	1910	PC	91	San Francisco
Portland OR	1900	PC	174	Seattle
Cincinnati OH	1900	ENC	295	Chicago
Milwaukee WI	1900	ENC	90	Chicago
Fort Worth TX	1900	WSC	28	Dallas
San Antonio TX	1900	WSC	195	Houston
Fort Lauderdale FL	1960	SA	26	Miami
Indianapolis IN	1900	ENC	185	Chicago
Columbus OH	1900	ENC	326	Chicago
New Orleans LA	1900	WSC	352	Houston
Buffalo NY	1900	MA	191	Cleveland
Orlando FL	1940	SA	84	Tampa
Rochester NY	1900	MA	91	Syracuse
Long Branch(Mon) NJ	1980	MA	50	New York City
Nashville TN	1900	ESC	246	Atlanta
Providence RI	1900	NE	49	Boston
Memphis TN	1900	ESC	311	St. Louis
Norfolk VA	1900	SA	93	Richmond
Oklahoma City OK	1910	WSC	209	Dallas
Louisville KY	1900	ESC	103	Cincinnati
Salt Lake City UT	1900	MT	493	Denver
Birmingham AL	1900	ESC	153	Atlanta
Jacksonville FL	1910	SA	348	Atlanta
Albany NY	1900	MA	146	Syracuse
West Palm Beach FL	1960	SA	65	Miami
Honolulu HA	1960	PC	600	Honolulu
Charlotte NC	1900	SA	240	Atlanta
Dayton OH	1900	ENC	55	Cincinnati
Austin TX	1910	WSC	186	Houston
New Brunswick NJ	1980	MA	50	New York City

Hartford CT	1900	NE	111	Boston/NYC
Las Vegas NV	1960	MT	275	Los Angeles
Richmond VA	1900	SA	105	Washington D.C.
Tulsa OK	1920	WSC	105	Oklahoma City
Grand Rapids MI	1900	ENC	156	Detroit
Allentown PA	1900	MA	68	Philadelphia
Greensboro NC	1910	SA	91	Charlotte
Oxnard-Ventura CA	1970	PC	62	Los Angeles
Tuscon AR	1960	MT	117	Phoenix
Fresno CA	1910	PC	183	San Francisco
Syracuse NY	1900	MA	276	New York City
Akron OH	1900	ENC	39	Cleveland
Greenville SC	1900	SA	110	Charlotte
Omaha NE	1900	WNC	136	Kansas City
Toledo OH	1900	ENC	62	Cleveland
Gary IN	1960	ENC	15	Chicago
Knoxville TN	1900	ESC	178	Nashville
El Paso TX	1910	WSC	437	Phoenix
Harrisburg PA	1900	MA	114	Philadelphia
Tacoma WA	1900	PC	31	Seattle
Wilmington DE	1900	SA	35	Philadelphia
Jersey City NJ	1960	MA	20	New York City
Bakersfield CA	1960	PC	112	Los Angeles
Springfield MA	1900	NE	90	Boston
New Haven CT	1900	NE	40	Hartford
Baton Rouge LA	1930	WSC	80	New Orleans
Little Rock AR	1900	WSC	325	Dallas
Charleston SC	1900	SA	204	Charlotte
Wichita KS	1910	WNC	157	Oklahoma City
Youngstown OH	1900	ENC	77	Cleveland
Albuquerque NM	1940	MT	447	Denver
Stockton CA	1910	PC	79	San Francisco
Mobile AL	1900	ESC	148	New Orleans
Columbia SC	1910	SA	91	Charlotte
Paterson NJ	1960	MA	31	New York City
Vallejo CA	1970	PC	32	San Francisco
Bridgeport CT	1900	NE	57	Hartford
Worcester MA	1900	NE	42	Boston
Johnson City TN	1980	ESC	285	Nashville
Newport News VA	1960	SA	78	Richmond
Chattanooga TN	1900	ESC	117	Atlanta
Lansing MI	1910	ENC	88	Detroit
Flint MI	1910	ENC	64	Detroit
Raleigh NC	1900	SA	153	Charlotte
Lancaster PA	1900	MA	74	Philadelphia
York PA	1900	MA	103	Philadelphia
Scranton PA	1900	MA	129	Philadelphia
Lakeland FL	1980	SA	33	Tampa
Melbourne FL	1980	SA	130	Tampa
Colorado Springs CO	1960	MT	67	Denver
Augusta GA	1900	SA	151	Atlanta

Jackson MS	1900	ESC	180	New Orleans
Lawrence MA	1960	NE	30	Boston
Canton OH	1900	ENC	58	Cleveland
Des Moines IA	1900	WNC	345	Chicago
Santa Rosa CA	1970	PC	58	San Francisco
McAllen TX	1970	WSC	346	Houston
Daytona Beach FL	1980	SA	90	Jacksonville FL
Modesto CA	1970	PC	92	San Francisco
Santa Barbara CA	1960	PC	92	Los Angeles
Madison WI	1900	ENC	77	Milwaukee
Fort Wayne IN	1900	ENC	132	Indianapolis
Beaumont TX	1920	WSC	88	Houston
Spokane WA	1900	PC	326	Seattle
New London CT	1960	NE	49	Hartford
Salinas CA	1970	PC	94	San Francisco
Davenport IA	1900	WNC	177	Chicago
Corpus Christi TX	1930	WSC	211	Houston
Lexington KY	1920	ESC	82	Cincinnati
Pensacola FL	1960	SA	200	New Orleans
Peoria IL	1900	ENC	170	Chicago
Reading PA	1900	MA	65	Philadelphia
Fort Myers FL	1980	SA	130	Tampa
Shreveport LA	1910	WSC	188	Dallas
Wilkes-Barre PA	1900	MA	115	Philadelphia
Trenton NJ	1900	MA	38	Philadelphia
Atlantic City NJ	1910	MA	60	Philadelphia
Utica NY	1900	MA	56	Syracuse
Appleton WI	1970	ENC	107	Milwaukee
Huntington WV	1900	SA	240	Cincinnati
Durham NC	1930	SA	145	Richmond
Visalia CA	1980	PC	183	Los Angeles
Orange Co. NY	1980	MA	70	New York City
Montgomery AL	1900	ESC	167	Atlanta
Hamilton OH	1900	ENC	22	Cincinnati
Saginaw MI	1900	ENC	101	Detroit
Rockford IL	1910	ENC	84	Chicago
Eugene OR	1960	PC	109	Portland OR
Ann Arbor MI	1960	ENC	41	Detroit
Macon GA	1900	SA	81	Atlanta
Evansville IN	1900	ENC	172	St. Louis
Salem OR	1970	PC	47	Portland OR
Sarasota FL	1980	SA	57	Tampa
Erie PA	1900	MA	106	Cleveland
Fayetteville NC	1970	SA	139	Charlotte
Lowell MA	1960	NE	32	Boston
Lorain OH	1900	ENC	20	Cleveland
Winston-Salem NC	1920	SA	81	Charlotte
Provo UT	1960	MT	43	Salt Lake City
Binghamton NY	1900	MA	75	Syracuse
Brownsville TX	1960	WSC	352	Houston
Poughkeepsie NY	1980	MA	89	New York City

Killeen TX	1980	WSC	153	Dallas
Reno NV	1960	MT	228	San Francisco
Fort Pierce FL	1990	SA	120	Miami
Charleston WV	1900	SA	211	Columbus OH
South Bend IN	1900	ENC	76	Chicago
Columbus GA	1900	SA	104	Atlanta
Savannah GA	1900	SA	140	Jacksonville FL
Springfield MO	1900	WNC	169	Kansas City
Johnstown PA	1900	MA	68	Philadelphia
Duluth MN	1900	WNC	156	Minneapolis
Huntsville AL	1960	ESC	203	Atlanta
Tallahassee FL	1970	SA	163	Jacksonville FL
Santa Cruz CA	1980	PC	83	San Francisco
Anchorage AK	1980	PC	600	Anchorage
Portsmouth NH	1980	NE	63	Boston
Roanoke VA	1910	SA	186	Richmond
Kalamazoo MI	1910	ENC	139	Detroit
Lubbock TX	1940	WSC	351	Dallas
Hickory NC	1980	SA	74	Charlotte
Waterbury CT	1960	NE	32	Hartford
Galveston TX	1920	WSC	49	Houston
Portland ME	1900	NE	115	Boston
Lincoln NE	1900	WNC	193	Des Moines
Bradenton FL	1980	SA	36	Tampa
Salisbury NC	1980	SA	40	Charlotte
Lafayette LA	1970	WSC	134	New Orleans
Boise City, ID	1970	MT	349	Spokane
Gainesville FL	1970	SA	73	Jacksonville
Stamford CT	1960	NE	34	New York City
Biloxi MS	1970	ESC	80	New Orleans
Ocala FL	1980	SA	94	Tampa
Green Bay WI	1910	ENC	117	Milwaukee
St. Cloud MN	1980	WNC	72	Minneapolis
Bremerton WA	1980	PC	63	Seattle
Springfield IL	1900	ENC	100	St. Louis
Yakima WA	1980	PC	140	Seattle
Waco TX	1900	WSC	96	Dallas
Brockton MA	1900	NE	24	Boston
Danbury CT	1970	NE	61	Hartford
Amarillo TX	1930	WSC	368	Dallas
Fort Collins CO	1980	MT	62	Denver
Houma LA	1990	WSC	58	New Orleans
Chico CA	1980	PC	177	San Francisco
Nashua NH	1970	NE	46	Boston
Merced CA	1990	PC	117	San Francisco
New Bedford MA	1960	NE	56	Boston
Fort Smith AR	1960	WSC	160	Little Rock
Asheville NC	1920	SA	143	Charlotte
Racine WI	1910	ENC	30	Milwaukee
Champaign ILL	1960	ENC	137	Chicago
Cedar Rapids IA	1900	WNC	128	Des Moines

Clarksville TN	1980	ESC	47	Nashville
Lake Charles LA	1960	WSC	146	Houston
Longview TX	1980	WSC	130	Dallas
Topeka KA	1900	WNC	61	Kansas City
Benton Harbor MI	1980	ENC	205	Chicago
Olympia WA	1980	PC	59	Seattle
Muskegon MI	1970	ENC	186	Detroit
Ogden UT	1930	MT	34	Salt Lake City
Fall River MA	1900	NE	45	Boston
Athens GA	1980	SA	77	Atlanta
Elkhart IN	1980	ENC	100	Chicago
Lima OH	1910	ENC	162	Cleveland
Fargo ND	1960	WNC	240	Minneapolis
Naples FL	1990	SA	116	Miami
Tyler Tx	1960	WSC	98	Dallas
Tuscaloosa AL	1960	ESC	59	Birmingham
Jackson MI	1910	ENC	77	Detroit
Jacksonville NC	1980	SA	239	Charlotte
Wheeling WV	1900	SA	54	Pittsburgh
Richland WA	1980	PC	201	Portland OR
Parkersburg WV	1980	SA	109	Columbus OH
Springfield OH	1900	ENC	84	Cincinnati
Manchester NH	1900	NE	64	Boston
New Britain CT	1960	NE	16	Hartford
Waterloo IA	1920	WNC	108	Des Moines
Redding CA	1980	PC	215	San Francisco
Medford OR	1980	PC	272	Portland OR
Anderson SC	1980	SA	120	Atlanta
Fort Walton Beach FL	1980	SA	234	New Orleans
Steubenville OH	1960	ENC	133	Cleveland
Monroe LA	1960	WSC	102	Shreveport
Jamestown NY	1990	MA	140	Cleveland
Lynchburg VA	1960	SA	117	Richmond
Janesville WI	1980	ENC	77	Milwaukee
Vineland NJ	1970	MA	37	Philadelphia
Eau Claire WI	1980	ENC	79	Minneapolis
Battle Creek MI	1980	ENC	120	Detroit
Las Cruces NM	1980	MT	457	Phoenix
Joplin MO	1980	WNC	150	Kansas City
Laredo TX	1950	WSC	312	Houston
Greeley CO	1980	MT	55	Denver
Alexandria LA	1980	WSC	125	Shreveport
Decatur AL	1990	ESC	80	Birmingham
Rock Hill SC	1980	SA	20	Charlotte
Charlottesville VA	1980	SA	74	Richmond
Lafayette IN	1970	ENC	62	Indianapolis
Florence AL	1980	ESC	128	Birmingham
Dothan AL	1990	ESC	200	Birmingham
Terre Haute IN	1900	ENC	78	Indianapolis
Anderson IN	1970	ENC	48	Indianapolis
Burlington VT	1980	NE	256	Boston

Altoona PA	1900	MA	95	Pittsburgh
Bloomington IL	1980	ENC	136	Chicago
Newark OH	1980	ENC	126	Cleveland
Kenosha WI	1920	ENC	38	Milwaukee
Panama City FL	1980	SA	258	New Orleans
Norwalk CT	1960	NE	43	New York City
Bellingham WA	1980	PC	88	Seattle
Mansfield OH	1970	ENC	80	Cleveland
Petersburg VA	1970	SA	23	Richmond
Sioux Falls SD	1930	WNC	285	Des Moines
State College PA	1980	MA	199	Philadelphia
Yuba City CA	1980	PC	130	San Francisco
Pueblo CO	1910	MT	109	Denver
Wichita Falls TX	1920	WSC	137	Dallas
Bryan TX	1970	WSC	102	Dallas
Hagerstown MD	1980	SA	74	Baltimore
Sharon PA	1980	MA	56	Pittsburgh
Abilene TX	1960	WSC	185	Dallas
Muncie IN	1910	ENC	62	Indianapolis
Wilmington NC	1970	SA	204	Charlotte
Texarkana AR	1960	WSC	71	Shreveport
Glens Falls NY	1980	MA	165	Syracuse
Odessa TX	1960	WSC	354	Dallas
Williamsport PA	1980	MA	201	Pittsburgh
Santa Fe NM	1990	MT	386	Denver
Decatur IL	1910	ENC	120	St. Louis
Anniston AL	1980	ESC	66	Birmingham
Pascagoula MS	1980	ESC	118	New Orleans
Wausau WI	1980	ENC	180	Minneapolis
Sioux City IA	1900	WNC	200	Des Moines
Florence SC	1980	SA	100	Charlotte
Albany GA	1960	SA	182	Atlanta
Fayetteville AR	1980	WSC	195	Little Rock
Billings MT	1960	MT	538	Spokane
Bay City MI	1900	ENC	115	Detroit
Columbia MO	1970	WNC	127	St. Louis
Lawton OK	1960	WSC	87	Oklahoma City
Danville VA	1980	SA	149	Richmond
Bloomington IN	1980	ENC	50	Indianapolis
Burlington NC	1980	SA	118	Charlotte
Yuma AR	1990	MT	178	Phoenix
Midland TX	1960	WSC	332	Dallas
Rochester MN	1970	WNC	90	Minneapolis
Sheboygan WI	1980	ENC	57	Milwaukee
Fitchburg MA	1960	NE	51	Boston
Cumberland MD	1980	SA	140	Baltimore
Gadsden AL	1930	ESC	61	Birmingham
San Angelo TX	1950	WSC	255	Dallas
LaCrosse WI	1970	ENC	160	Minneapolis
Kokomo IN	1980	ENC	53	Indianapolis
Iowa City IA	1980	WNC	112	Des Moines

Kankakee IL	1980	ENC	56	Chicago
Sherman TX	1970	WSC	65	Dallas
Elmira NY	1980	MA	93	Syracuse
Bangor ME	1980	NE	249	Boston
Lewiston ME	1960	NE	146	Boston
Owensboro KY	1970	ESC	233	Cincinnati
Dubuque IA	1900	WNC	177	Chicago
Pine Bluff AR	1970	WSC	42	Little Rock
Bismark ND	1980	WNC	438	Minneapolis
St. Joseph MO	1900	WNC	54	Kansas City
Lawrence KS	1980	WNC	38	Kansas City
Rapid City SD	1990	WNC	413	Denver
Pittsfield MA	1900	NE	134	Boston
Bristol CT	1970	NE	22	Hartford
Great Falls MT	1960	MT	401	Spokane
Jackson TN	1990	ESC	86	Memphis
Victoria TX	1980	WSC	125	Houston
Cheyenne WY	1990	MT	100	Denver
Grand Forks ND	1980	WNC	318	Minneapolis
Casper WY	1980	MT	281	Denver
Enid OK	1980	WSC	98	Oklahoma City

APPENDIX B
TRANSITION MATRICES, DECADE BY DECADE

Each cell in these matrices represents the number of cities in the respective category in year *t* (rows) compared to year *t+1* (columns), and the associated probability . For example, in the first matrix, the probability of a city being the smallest category in 1900 and remaining there in 1910 is 68.57 %; 24 cities did so. The probability of a city being in the smallest category in 1900 and moving to the next category (between .3 and .5 of the mean) in 1910 was 31.43 %; 11 cities did so. As explained in the text, these matrices do not pick up "new" cities. The first matrix picks up only the 112 cities that meet our criteria for 1900.

TRANSITIONS: 1900 TO 1910

1910/1900	0.3	0.5	0.75	1	2	20	TOTAL
0.3	24 68.57	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	24
0.5	11 31.43	20 76.92	0 0.00	0 0.00	0 0.00	0 0.00	31
0.75	0 0.00	5 19.23	10 45.45	0 0.00	0 0.00	0 0.00	15
1	0 0.00	1 3.85	10 45.45	3 42.86	0 0.00	0 0.00	14
2	0 0.00	0 0.00	2 9.09	4 57.14	9 69.23	0 0.00	15
20	0 0.00	0 0.00	0 0.00	0 0.00	4 30.77	9 100.00	13
Total	35 31.25	26 23.21	22 19.64	7 6.25	13 11.61	9 8.04	112 100.00

TRANSITIONS: 1910 TO 1920

1920/1910	0.3	0.5	0.75	1	2	20	TOTAL
0.3	42 82.35	4 12.90	0 0.00	0 0.00	0 0.00	0 0.00	46
0.5	9 17.65	21 67.74	2 13.33	0 0.00	0 0.00	0 0.00	32
0.75	0 0.00	5 16.13	11 73.33	0 0.00	1 6.67	0 0.00	17
1	0 0.00	1 3.23	2 13.33	10 71.43	0 0.00	0 0.00	13
2	0 0.00	0 0.00	0 0.00	4 28.57	13 86.67	1 7.69	18
20	0 0.00	0 0.00	0 0.00	0 0.00	1 6.67	12 92.31	13
Total	51 36.69	31 22.30	15 10.79	14 10.07	15 10.79	13 9.35	139 100.00

TRANSITIONS: 1920 TO 1930

1930/1920	0.3	0.5	0.75	1	2	20	Total
0.3	45 81.82	3 9.09	0 0.00	0 0.00	0 0.00	0 0.00	48
0.5	10 18.18	23 69.70	2 11.76	0 0.00	0 0.00	0 0.00	35
0.75	0 0.00	7 21.21	12 70.59	2 15.38	0 0.00	0 0.00	21
1	0 0.00	0 0.00	3 17.65	8 61.54	2 11.11	0 0.00	13
2	0 0.00	0 0.00	0 0.00	3 23.08	16 88.89	0 0.00	19
20	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	13 100.00	13
Total	55 36.91	33 22.15	17 11.41	13 8.72	18 12.08	13 8.72	149 100.00

TRANSITIONS: 1930 TO 1940

1940/1930	0.3	0.5	0.75	1	2	20	Total
0.3	51 92.73	2 5.56	0 0.00	0 0.00	0 0.00	0 0.00	53
0.5	4 7.27	31 86.11	0 0.00	0 0.00	0 0.00	0 0.00	35
0.75	0 0.00	3 8.33	18 85.71	2 15.38	0 0.00	0 0.00	23
1	0 0.00	0 0.00	3 14.29	10 76.92	1 5.26	0 0.00	14
2	0 0.00	0 0.00	0 0.00	1 7.69	17 89.47	0 0.00	18
20	0 0.00	0 0.00	0 0.00	0 0.00	1 5.26	13 100.00	14
Total	55 35.03	36 22.93	21 13.38	13 8.28	19 12.10	13 8.28	157 100.00

TRANSITIONS: 1940 TO 1950

1950/1940	0.3	0.5	0.75	1	2	20	Total
0.3	48 85.71	2 5.71	0 0.00	0 0.00	0 0.00	0 0.00	50
0.5	8 14.29	25 71.43	4 17.39	0 0.00	0 0.00	0 0.00	37
0.75	0 0.00	8 22.86	15 65.22	3 21.43	1 5.56	0 0.00	27
1	0 0.00	0 0.00	3 13.04	9 64.29	1 5.56	0 0.00	13
2	0 0.00	0 0.00	1 4.35	2 14.29	16 88.89	0 0.00	19
20	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	14 100.00	14
Total	56 35.00	35 21.88	23 14.38	14 8.75	18 11.25	14 8.75	160 100.00

TRANSITIONS: 1950 TO 1960

1960/1950	0.3	0.5	0.75	1	2	20	Total
0.3	36 69.23	2 5.41	1 3.70	0 0.00	0 0.00	0 0.00	39
0.5	15 28.85	20 54.05	2 7.41	0 0.00	0 0.00	0 0.00	37
0.75	1 1.92	14 37.84	13 48.15	1 7.69	2 10.53	0 0.00	31
1	0 0.00	1 2.70	6 22.22	3 23.08	2 10.53	0 0.00	12
2	0 0.00	0 0.00	5 18.52	9 69.23	11 57.89	0 0.00	25
20	0 0.00	0 0.00	0 0.00	0 0.00	4 21.05	14 100.00	18
Total	52 32.10	37 22.84	27 16.67	13 8.02	19 11.73	14 8.64	162 100.00

TRANSITIONS: 1960 TO 1970

1970/1960	0.3	0.5	0.75	1	2	20	Total
0.3	59 84.29	1 2.08	0 0.00	0 0.00	0 0.00	0 0.00	60
0.5	11 15.71	32 66.67	2 6.25	0 0.00	0 0.00	0 0.00	45
0.75	0 0.00	15 31.25	24 75.00	0 0.00	0 0.00	0 0.00	39
1	0 0.00	0 0.00	5 15.63	7 53.85	0 0.00	0 0.00	12
2	0 0.00	0 0.00	1 3.13	6 46.15	21 77.78	0 0.00	28
20	0 0.00	0 0.00	0 0.00	0 0.00	6 22.22	20 100.00	26
Total	70 33.33	48 22.86	32 15.24	13 6.19	27 12.86	20 9.52	210 100.00

TRANSITIONS: 1970 TO 1980

1980/1970	0.3	0.5	0.75	1	2	20	Total
0.3	54 65.85	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	54
0.5	28 34.15	19 35.85	0 0.00	0 0.00	0 0.00	0 0.00	47
0.75	0 0.00	32 60.38	17 41.46	0 0.00	0 0.00	0 0.00	49
1	0 0.00	2 3.77	16 39.02	2 16.67	0 0.00	1 3.70	21
2	0 0.00	0 0.00	8 19.51	10 83.33	19 67.86	0 0.00	37
20	0 0.00	0 0.00	0 0.00	0 0.00	9 32.14	26 96.30	35
Total	82 33.74	53 21.81	41 16.87	12 4.94	28 11.52	27 11.11	243 100.00

TRANSITIONS: 1980 TO 1990

1990/1980	0.3	0.5	0.75	1	2	20	Total
0.3	100 90.91	15 23.44	0 0.00	0 0.00	0 0.00	0 0.00	115
0.5	10 9.09	41 64.06	13 25.49	0 0.00	0 0.00	0 0.00	64
0.75	0 0.00	8 12.50	33 64.71	2 8.70	0 0.00	0 0.00	43
1	0 0.00	0 0.00	5 9.80	13 56.52	3 7.89	0 0.00	21
2	0 0.00	0 0.00	0 0.00	8 34.78	32 84.21	0 0.00	40
20	0 0.00	0 0.00	0 0.00	0 0.00	3 7.89	36 100.00	39
Total	110 34.16	64 19.88	51 15.84	23 7.14	38 11.80	36 11.18	322 100.00

VITA

Linda Harris Dobkins holds an associate of arts degree in journalism (1968), and a bachelors degree (*magna cum laude*, 1971) in political science from Missouri Southern State College in Joplin, Missouri, her hometown. She holds a master of science degree in economics (highest honors, 1973) from Pittsburg State University in Pittsburg, Kansas, and a master of arts degree in economics (1993) from Virginia Polytechnic Institute and State University. The Ph.D. in economics was awarded to her in June of 1995 by Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Previous careers include eight years as a reporter and photographer for newspaper and television; and eight years as a math teacher in a private, college-prep high school.