MIGRATION AND URBAN ECONOMIC GROWTH

John R. MIRON

University of Toronto, West Hill, Ont. M1C 1A4, Canada

Why some cities grow while others decline in population is a relevant question for public policy. Although there are many theories about why metropolitan areas develop, very few of these concepts have been empirically applied. In this paper, several empirical models of urban and regional growth are examined in terms of postulated causes of growth. In particular, the role of migration as a cause or consequence of development is emphasized. It is concluded that these models take a very narrow view of the growth process and that a promising area of future research lies in models of urban labour market dynamics.

1. Introduction

There is much statistical evidence, as in Berry (1973, ch. 1), that many metropolitan regions have been until recently experiencing sustained above-average population growth. This urban size 'ratchet' effect, as initially hypothesized by Thompson (1965), has been part of a conventional wisdom which saw metropolitan areas growing forever larger relative to smaller urban centres. The ratchet effect was seen to arise in part from the political and economic powers which a metropolis could use to ensure its own future growth.

More recent empirical evidence however suggests that some of these larger urban areas have begun periods of either absolute decline or relatively slow growth. If this recent change is more than a short-run phenomenon, some rethinking about the causes of metropolitan growth may be in order. Certainly, there are numerous explanations offered for the decline of large cities and these tend to center around the negative effects on households: pollution, congestion, crime, and placelessness are among the evils mentioned. However, there has been little empirical testing of these hypotheses and of the 'tipping point' if one exists between ratchet and negative effects.

In this paper, a review is undertaken of empirical models of urban and regional economic growth. In general, an emphasis is put on the postulated...
causes of city growth. Since the level of net migration into an urban area is highly correlated with its rate of growth, this review puts a specific emphasis on the relationship between the two. To what extent can migration be viewed as a cause or merely as a consequence of the growth of cities? Although writers such as Thompson have offered a rich variety of hypotheses about why cities do or do not grow, it is a central conclusion of this paper that virtually all empirical models treat the causes of growth and the migration linkage very simplistically.

2. The main issue

The over-riding emphasis in current models of urban economic growth is on the urban labour market. In particular, emphasis is placed on those exogenous forces which come to affect the supply of and demand for labour in a particular city. Those models which emphasize supply factors highlight the role of migration in affecting growth. Those which emphasize demand aspects often ignore migration. Therefore a classification of models as supply- or demand-oriented is also a classification according to the role assigned to migration in affecting urban growth.

2.1. Demand and supply models of urban growth

Engle (1974) identifies two polar cases: the demand- and the supply-oriented model. Pure demand models presume perfectly elastic factor supplies. Labour (and capital) in-migration to a city occurs exactly and only so as to maintain a given real factor price. A city’s growth is limited only by its factor demand at that price. Urban growth is dependent on exogenous expenditures such as local investment, government spending, and export (outside the city) demand.

At the other polar extreme, pure supply models presume that factors are not completely price elastic. The amount of labour is only partly responsive to local wage variations. Migration and natural increase, based partly on a non-wage incentive, affect the total supply. In their simplest form, these models presume that the city’s producers face a perfectly elastic demand for exports at a given output price. The output level of the city is then limited only by the availability of factors.

These two kinds of model are evidently polar extremes. Any one city may have partially inelastic factor supplies and partially elastic demands placed on its outputs. Different cities may be approximated better by models at different points in the spectrum depending on the market conditions facing them. The labour market characteristics of the two polar solutions and their differences are illustrated in fig. 1.
2.2. Conceptual problems with a demand-supply orientation

While the polar demand and supply-oriented models seem to provide an interesting classification scheme, they are not without considerable shortcomings. This is noted immediately when one tries to define what the supply and demand curves for labour in Fig. 1 represent. Consider first the supply curve for which there are at least two alternative definitions. One is that this curve denotes the amount of labour forthcoming from the current residents of a town at any particular wage. The shape of the curve would reflect the residents' work-leisure trade-offs. The second is that the supply curve of labour reflects the total amount of labour which would be made available to employers in a town at any given wage rate. This would be defined as the sum of the labour made available by current residents and prospective in-migrants attracted by that wage less the labour withdrawn by out-migrants repelled by that same wage.

The latter definition is more useful here because it permits a simple definition of growth-inducing migration. Since the labour supply curve under this second concept includes all wage-related migration, shifts in the supply curve represent only those portions of natural increase or migration which are in response to non-wage changes in relative local conditions. Thus, migration which implies a movement along the supply curve is induced by demand variations while migration which shifts the whole supply schedule is independent of demand variations and, therefore, potentially growth-inducing. For this reason, the second definition of the labour supply function is used in the remainder of this paper.

However, this definition commits one to a particular time frame. The first definition is apparently a short-run one. When producers change their wage offer, one might expect the local labour force to make a reasonably quick
adjustment in terms of their labour offer. However, migrants need considerably more time to gain information about job and wage offers, make a decision on relocation, and actually complete the move. Thus, the second involves a longer run time frame.

A similar problem emerges when attempting to define what the demand schedule represents. A short run viewpoint is that the demand schedule for labour is derived from (i) the demand function for the city's output and (ii) the existing stocks of other factors of production (notably fixed capital). A longer run viewpoint is that the demand schedule is derived assuming that these other factors of production can be varied to suit the profit-maximizing firm's requirements. Since a long-run time frame has been adopted for the labour supply curve, the same frame is used with the demand curve.

Are the conceptual problems resolved completely if a long-run time frame is adopted toward the urban labour market? There is reason to be negative. Research on locational interdependency as typified by the Koopmans-Beckmann problem suggests that the spatial concentration of industry may be similar to a non-stationary stochastic process. If so, the growth of cities may be cumulatively responsive to sudden and short-run changes in local conditions. Short-run changes may accumulate to produce behaviour which in the long run does not approximate either the demand- or supply-oriented model. A study of short-run behaviour within models of long-run growth may well be a necessary complement to this study.³

2.3. The concept of growth-inducing migration

The central question of this paper is how and why, in empirical and theoretical terms, migration is important in causing urban growth. Most of the models to be discussed attempt to examine how migration relates to the urban growth process. Little is said about why growth-inducing migration occurs at all. A brief review of three different theories about migration is in order.

One theory holds that non-wage incentives, in and of themselves, account for this kind of migration. Hirsch (1973, ch. 9), among others, speaks of 'household-initiated' urbanization in which it is the amenities of living in a particular area which attracts migrants. These newcomers are willing to remain unemployed or underpaid for a time in the belief either that the amenities outweigh any wage loss or that new jobs will be created and wages rise in the near future.

A second theory views growth-inducing migration not as a deliberate act

³An additional potential problem is posed by a long-run orientation. In moving from the short to the long-run, the elasticity of both the supply and the demand curve should increase. In the long run, the possibility exists that both curves are perfectly elastic. This extreme case would imply that there is no unique equilibrium city size.
but as the consequence of a sluggish response to wage changes. Richardson's (1973, ch. 4) review suggests the following important reasons for inertia in migration flows: (i) The information channels through which wage and job availability data are passed back to areas of out-migration are at best imperfect. Migration generated on the basis of perceived labour market conditions thus appear to be sluggishly responsive to actual conditions. (ii) Initial movements by out-migrants from an area tend to affect the destination patterns of later out-migrants through a 'friends and neighbours' or 'stem family' process. The availability of temporary lodging, job information and contacts, and spiritual encouragement at the home of friends may encourage more migrants than would have been justified on the basis of an initial wage or job opening difference.

A third reason for non-wage migration has been advanced by David (1974). He suggests that people may move from one area to another even when the local average wages are equal if the local dispersions of wages about these means are different. Some kinds of risk-bearing migrants will be attracted by the possibility of a substantially higher wage in another town even if the average wage there is equal to or less than their old one.

All of these answers rest on implicit notions about the dynamics of the urban labour market. Migrants cannot usually wait forever to get a job. There are psychic, monetary, and time costs involved in moving and searching which the prospective migrant has to weigh against an uncertain gain in income. At the same time, those producers who have a reasonably elastic demand for labour need time to adjust their production schedules to use the increased labour available.

In the remaining sections of this paper, a number of urban growth models are examined to see the role of growth-inducing migration and the accompanying labour market mechanisms. In section 3, we examine several empirically relevant demand-oriented models. In section 4 a similar array of 'mixed' models are considered, models which are neither pure demand nor supply models but have elements of both. In the final section 5, several conclusions and research suggestions are outlined.

3. Demand-oriented models

In this section, some pure demand models of urban growth are examined. These include two export-base models, an income model, and three econometric growth models. The objective here is to show up the similarities among the models in their treatment of urban growth. In addition, one model is extended to lead in to mixed models.

There is no attempt here to undertake a general review of problems associated with empirical urban and regional growth models. The interested reader is referred to Richardson (1973, ch. 8) and Engle (1974).
3.1. A simple export-base model

As is well known, the simple export-base model recognizes two sectors of employment in a city at full employment, $N$; basic- or export-oriented employment, $B$; and non-basic- or local-oriented employment, $S$. The kernel of this theory lies in two hypotheses: that non-basic employment, $S$, is linearly related to the aggregate city population, $P$, and that its aggregate population is itself linearly related to total employment, $N$. Mathematically these familiar conditions are

$$ N = B + S, \quad (1) $$

$$ S = x_0 + x_1 P, \quad 0 \leq x \leq 1, \quad (2) $$

$$ P = \beta_0 + \beta_1 N, \quad \beta \geq 1. \quad (3) $$

These can be shown to assert that

$$ N = \gamma_0 + \gamma_1 B, \quad (4) $$

where

$$ \gamma_0 = (x_0 + x_1 \beta_0)/(1 - x_1 \beta_1) \quad \text{and} \quad \gamma_1 = 1/(1 - x_1 \beta_1). \quad (5) $$

Given that $x_1 \beta_1 \leq 1$, it is seen from (4) that total employment, $N$, is some multiple (greater than unity) of basic employment.$^6$

It is not difficult to imagine a plausible set of assumptions which could form the basis of this model. Suppose first that a city is one of a large number in a region and that workers migrate to any particular city in search of higher wages. In an equilibrium, the capitalized value of the wage difference between any pair of cities would be at most the net cost of migration. Any increase in the wage in one city could bring about a very large (effectively infinite) in-migration of workers. In this case the supply of labour in any one city is infinitely elastic at a certain wage rate. Note that both the local and export sectors must pay this same wage. Assume secondly that the local sector has constant returns to scale in production, is made up of a large number of firms, and can purchase non-labour inputs in competitive markets. Thus, the local sector’s labour requirements increase proportionately with output. Assume finally that each worker has the same

$^5$Exports are generally viewed as those products shipped from the city or region to any other place.

$^6$The condition that $x_1 \beta_1 \leq 1$ states that with a unit increase in population, the marginal change in service employment, $x_1$, must be less than or equal to the marginal change in total employment, $(1/\beta_1)$. 
demand function for the output of the local sector. These assumptions ensure that (i) the local sector has fixed marginal costs of production, (ii) the price of local sector output remains fixed as city size increases, and (iii) local sector employment increases proportionately with employment in the basic sector.

Under the above assumptions, the export-base model is readily seen to be a demand-oriented model. Each sector has a downward sloping demand curve for labour. Each faces the same horizontal supply curve as is shown in fig. 2. This is structurally equivalent to fig. 1(a). There is no meaningful sense in which the supply of labour can be increased under the assumptions made and the supply of labour thus can play no role in affecting city size.

![Labour market in the export-base model.](image)

What causes a city to grow in this model? It is the growth of exports which causes the growth of employment. However, there is no theory or mechanism by which the level of exports is explained. The city exists at the whim of a variable defined exogenously with respect to itself and has no means, for example of enabling its own growth.

Finally, the model represents a long-run equilibrium under the assumptions made. There is no role for urban labour market dynamics because migration implicitly brings the market into continuous equilibrium.

3.2. Czamanski's Baltimore model

Czamanski (1965) proposed several improvements to the export-base model. He introduced lead-lag-relationships between population and employment. His lead-lag relationships are based on empirical work dealing with the Baltimore SMSA. They are based on qualitative theoretical arguments and no justification is provided for the particular time lengths used.
He also extended the export-base model in that employment is broken into three sectors. A 'geographically-based' sector with employment $G(t)$ is closest to the earlier notion of an export sector. There is, however, also a sector which is complementary to this with employment $C(t)$. The third is the local sector with employment $S(t)$. The model used is of the following form:

$$P(t) = \psi_0 + \psi_1 N(t-2),$$

(6)

$$N(t) = G(t) + C(t) + S(t),$$

(7)

$$S(t) = z_0 + z_1 P(t-1),$$

(8)

$$C(t) = e_0 + e_1 G(t).$$

(9)

This system of equations reduces to the third-order difference equation

$$P(t) = \psi_0 + \psi_1 G(t-2) + \psi_2 P(t-3),$$

(10)

where

$$\psi_0 = \varphi_0 + \varphi_1 (z_0 + e_0), \quad \psi_1 = \varphi_1 (1 + e_1), \quad \psi_2 = \varphi_1 z_1.$$  

(11)

Menchik (1971) terms this a dynamic equilibrium-seeking model if $-1 < \varphi_1 z_1 < 1$. It has an equilibrium solution associated with a particular level of $G(t)$, say $G(t) = G_0$, of

$$\lim_{t \to \infty} P(t) = \psi_0 (1 - \psi_2) + \psi_1 (1 - \psi_2) G.$$

(12)

This equilibrium solution is equivalent in form to the solution (4) of the earlier version of the economic base model.

In effect, (10) is merely a partial-adjustment model in which population growth tends toward an export-base solution but in which the structural parameter $\psi_2$ serves to determine how quickly the population size tends toward that solution.

Casting the export-base model in partial adjustment form is an important extension in several respects. First, this model introduces short-run labour market dynamics into a previously static equilibrium model. Although there is no explicit notion of labour force here, Czamanski's construct permits a flexible relationship between current population and employment and the most reasonable assumption is that this reflects varying degrees of unemployment. Unemployment exists in the model presumably because it takes time...
for people to gather enough information about the local job market to decide whether to out-migrate. Coupled with a sluggish response by workers is delayed response of the local service sector as hypothesized in (8). Thus, although labour force shortages and surpluses can occur, these do not affect the growth rate of employment.

Also, by extending the model slightly, it now is possible to relate basic employment to city size. Suppose we hypothesize that, in any time period, the geographically-based sector makes available a number of jobs which is linearly related to the previous period's population,

\[ G(t) = \delta_0 + \delta_1 P(t - 1). \]  
(13)

Then (10) becomes

\[ P(t) = (\psi_0 + \psi_1 \delta_0) + (\psi_1 \delta_1 + \psi_2)P(t - 3). \]  
(14)

This extended model is a three-period, first-order difference equation which, if \( \psi_1 \delta_1 + \psi_2 > 1 \), indicates that city size will increase at an asymptotically constant, three-year growth rate. The larger are any of \( \varphi_1 > \varepsilon_1 > \delta_1 \), of \( \alpha_1 \), the faster will the city grow. There will also, of course, be no equilibrium size corresponding to (12).

8This extended model is very similar to the macro-urban model of Niedercorn (1963). Niedercorn's model also attempts to relate employment in the export sector, \( C(t) \), and total population, \( P(t) \). Its structure is as follows:

\[ g(t) = \alpha_1 g^*(t - 1) + \alpha_2 g(t - 1), \]

\[ p(t) = \alpha_0 + \alpha_3 g(t), \]

\[ g^*(t) = (\alpha_3 P(t) - G(t))/G(t). \]

where

\[ p(t) = (P(t) - P(t - 1))/P(t - 1), \]

\[ g(t) = (G(t) - G(t - 1))/G(t - 1). \]

Thus, the first hypothesis above is that the export employment growth rate, \( g(t) \), is a function of the previous period's growth rate and the discrepancy, \( g^*(t - 1) \) between actual export employment in the previous period and some fixed proportion of the population. The population growth rate, \( p(t) \), is tied in the second equation to the export employment growth rate.

As Mills (1972, pp. 65-66) points out, the dynamic properties of this model are not easily established. The model, like the extended version of Czamanski's, does not possess a static equilibrium solution. Unlike Czamanski's model, however, this model also rules out the possibility that export employment and total population could increase at the same average rate. The responsiveness of \( P(t) \) to \( g(t) \), the accelerator mechanism, is the key to understanding these dynamic properties. The larger is \( \alpha_1 \) relative to \( \alpha_0 \), the more rapid is the divergence between \( g(t) \) and \( p(t) \) with time.
This extension of Czamanski's model is no longer an export-base model in the usual sense because there is no meaningful long-run sense in which a city's size is limited by its level of export demand. Since the key to this extension is (13), what set of plausible assumptions would be consistent with such a model? As in the local sector, assume that the export sector has constant returns to scale and can purchase non-labour inputs in competitive markets. Assume further that the city's exporters operate in a competitive national market where they are again price takers. Then, the producers have an infinitely elastic demand for labour at a given wage rate. It is this combination of infinitely elastic demand for and supply of labour in the long-run which makes a notion of equilibrium city size indeterminate.

What causes urban growth in Czamanski's model? In the original version without (13), urban growth is equated in the long-run with export growth. As in the initial export-base model, no hypotheses are offered to explain why a particular level of exports occur. The augmented version with (13) is different. A city also grows in this version because its level of exports increases but export growth in turn is tied to city size. In this case, labour force growth in anticipation of employment becomes a self-fulfilling prophecy. Significantly, this occurs without an economies-of-scale or other force as often popularly argued. This augmented version leaves at least one main issue untouched. The model implies the significance of anticipations. Job-seekers in (6) anticipate the number of new jobs to be created in a time period. Employers, particularly in (8) and (13) anticipate the number of workers required and available. If both groups were able to foresee the other's response, the growth of a city would be indeterminate here. However, the model does not indicate how these anticipations (i) come to be formed or (ii) evolve with experience. If one is to pursue models such as this, explaining the formation of anticipations is a major research issue.

3.3. Income models

Pure Keynesian models of urban growth are another common kind of demand-oriented model. Empirical models of this type are relatively scarce, however, because regional income and product accounts are difficult to obtain. In this section, the model by Anderson (1970) is used as an example.

---

9 This infinitely elastic labour demand is presumed only for the longer run (more than one year). In the short run (less than one year), the existence of unemployment implies that the firms require time to adjust their production levels.

10 This model can be viewed as quite similar to a non-stationary stochastic representation of the locational interdependency problem.
while recognizing that there are many variants on the basic Keynesian model.\textsuperscript{11}

Keynesian models emphasize the determinants of different components of total regional income or expenditure. Anderson’s model takes the following form:

\begin{align*}
Y(t) & = C(t) + I(t) + G(t) + X(t) - M(t), \\
YDP(t) & = Y(t) - D(t), \\
C(t) & = \alpha_{10} + \alpha_{11} YPD(t) + \alpha_{12} YPD(t - 1), \\
YL(t) & = \alpha_{20} + \alpha_{21} Y(t) + \alpha_{22} t, \\
YPD(t) & = YL(t) + YNL(t), \\
I(t) & = \alpha_{30} + \alpha_{31} YNL(t) + \alpha_{32} YNL(t - 1) + A(t) + \alpha_{33} A(t - 1), \\
M(t) & = \alpha + \alpha Y(t).
\end{align*}

where

\begin{itemize}
  \item $Y$ = Regional total income.
  \item $C$ = Regional consumption.
  \item $I$ = Regional gross investment.
  \item $G$ = Regional government expenditures.
  \item $X$ = Regional exports.
  \item $M$ = Regional imports.
  \item $YPD$ = Regional personal disposable income.
  \item $D$ = Difference between $Y$ and $YPO$.
  \item $YL$ = Regional disposable wage income.
  \item $YNL$ = Regional disposable non-wage income.
  \item $A$ = Autonomous investment.
\end{itemize}

It is not difficult to establish a solution to this model. In solving for $Y(t)$ judicious substitutions yield the following form:

\begin{align*}
Y(t) & = \beta_0 + \beta_1 D(t) + \beta_2 t + \beta_3 A(t) + G(t) + X(t) \\
& + \beta_4 D(t - 1) + \beta_5 A(t - 1) + \beta_6 Y(t - 1).
\end{align*}

The $\beta$ coefficients are functions of the original $\alpha$ parameters. As is the case

\textsuperscript{11}An empirically tested model for Ohio of similar (though more disaggregated) structure is discussed in L’Espérance et al. (1969). Also Moody et al. (1970) use a similar model to study inter-regional linkages in the U.S.
with each other endogenous variable, \( Y(t) \) is the solution to a first order difference equation involving seven current and lagged exogenous variables. If \( -1 < \beta_k < 1 \) this is a dynamic equilibrium-seeking model.

What causes the growth of income in this model? If (22) is a dynamic equilibrium-seeking model, the levels of \( D(t), D(t-1), A(t), A(t-1), G(t), X(t), \) and time are the critical determinants. With the exception of 't', these represent either exogenous demands for the output of or leakages from the local economy. Whereas in the export base model the sole source of growth was exports, the income model allows for autonomous investment, government spending, taxes, and capital outflows as well as exports to affect growth. In spite of this, the model also has little to say about why growth occurs. It does not suggest how these important exogenous variables come to change.

What assumptions permit this model to operate? Nothing is indicated about wages and price although one might expect several of the exogenous variables to respond to them. One plausible route is to assume, as done earlier for the export base model, an infinitely elastic supply of all inputs within the urban region as well as constant returns to scale. By making these assumptions, fixed wages and prices can be assumed. The city's factor markets would then be equivalent to those in the export base model as illustrated in fig. 2. In this case, the income model is demand-oriented. It has no role for growth-inducing migration or labour market dynamics since it disregards the labour market.

3.4. Bell's model of Massachusetts

Three models based on the principles of both export base and income models are considered next. All of these income, population, and employment models have been empirically estimated. The first is the model of a region somewhat larger than a city (Massachusetts). However, the predominance of the Boston SMSA within this region and a number of unique features of this model make it an appropriate starting point. It is considered in more detail than the subsequent models because its structure is illustrative.

Bell’s Model takes the following block-recursive form:

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]

\[ \ldots \]
Block I: Income

\[ X(t) = x_{10} + x_{11} GNP(t), \]  

(23)

\[ S(t) = x_{20} + x_{21} Y(t), \]  

(24)

\[ Y(t) = X(t) + S(t). \]  

(25)

\[ GRP(t) = x_{31} Y(t). \]  

(26)

Block II: Capital stocks

\[ K_m(t)/K_m(t-1) = (K_m^*(t)/K_m(t-1))^z_{41}. \]  

(27)

\[ K_m^*(t) = x_{50}(x_{51})X(t)^{z_{42}}. \]  

(28)

\[ K_n(t) = x_{60}S(t)^{z_{61}}. \]  

(29)

\[ K(t) = K_n(t) + K_m(t). \]  

(30)

Block III: Labour

\[ L(t) = x_{70}x_{71}^{-1}(1 - \beta)K(t)^{1 - \beta}(1 - \beta)GRP(t)^{1 - \beta}. \]  

(31)

\[ P_c(t) = \gamma P(t-1). \]  

(32)

\[ P(t) = P_c(t) + M(t). \]  

(33)

\[ N_c(t) = \delta P_c(t). \]  

(34)

\[ M(t) = x_{80} + x_{81}(L(t-1) - N_c(t-1)). \]  

(35)

where

\[ X = \text{Received income from export sector.} \]

\[ S = \text{Received income from local sector.} \]

\[ Y = \text{Regional received income.} \]

\[ GRP = \text{Regional produced income.} \]

\[ GNP = \text{Gross national product (U.S.A.)} \]

\[ K_m = \text{Capital stock - Manufacturing.} \]

\[ K_n = \text{Capital stock - Non-manufacturing.} \]

\[ K = \text{Regional capital stock.} \]

\[ L = \text{Regional labour demand.} \]

\[ M = \text{Net regional in-migration.} \]
This model is block-recursive because the solution to any block depends only on the solution to preceding blocks and/or exogenous variables. This aspect of the causal ordering is illustrated in fig. 3.

Block I of this model is very similar to an export-base model. Export demand, as represented by GNP, is the sole cause of growth in this block. Depending on the values of $a_0$ and $a_1$, there may be, for example, a sustained net in-migration of workers. However, any surplus labour force will in the future merely reduce net in-migration without any effect on the income or capital blocks. In the long run, the aggregate employment, income and capital stock levels are determined by a time trend and by the demand for regional exports via GNP. There is no possibility of growth-inducing migration here.

3.5. The Glickman and Hall–Licari models

Two other annual econometric models to date have been constructed specifically for metropolitan regions. These are the models by Glickman (1971) of Philadelphia and by Hall–Licari (1974) of Los Angeles. These models are very similar in structure and it is helpful to discuss their structures simultaneously while making comparisons with Bell.
Each model can be divided into a number of recursive blocks of variables as was done for the Bell version. There are three blocks in the Glickman model and two in the Hall–Licari prototype. In both cases, the first block relates to the manufacturing sector.

The manufacturing sector, as a proxy for all export industry, is strongly tied to national conditions in both models.

\[ Q(t) = \alpha_{10} + \alpha_{11} GNP(t), \]  
\[ I(t) = \alpha_{20} + \alpha_{21} Q(t) + \alpha_{22} K(t-1) + \{x_{23} I(t-1)\}, \]  
\[ K(t) = (1-d)K(t-1) + I(t), \]  
\[ E(t) = \alpha_{30} + \alpha_{31} Q(t) + \{x_{32} t\}, \]

where

- \(Q\) = Manufacturing value added.
- \(I\) = Manufacturing gross investment.
- \(K\) = Manufacturing capital stock.
- \(E\) = Manufacturing employment.

The above equations summarize this block for each model.\(^{13}\) Again, export demand as measured by \(GNP\) and a time trend are the two variables causing growth in this sector. This is illustrated diagrammatically in fig. 4 in the case

![Diagram of Glickman's model of Philadelphia](image-url)

**Fig. 4.** Glickman's model of Philadelphia.

\(^{13}\)The \(\{\) brackets enclose terms used by Glickman but not by Hall–Licari.
of Glickman's model. Here, there is no possibility for exogenous increases in the labour force to affect the manufacturing sector whatsoever.

The remaining blocks of the Glickman and Hall–Licari models are also quite similar. The Glickman model has two blocks of variables (an omnibus second block feeding a third government block) in addition to the initial manufacturing block. In the Hall–Licari model, the government-sector variables feed back into the second block so that no separable, recursive sub-blocking is possible. Hall–Licari emphasize this difference as it allows government policy variables in their model to affect the growth of non-manufacturing output, income, and employment variables in the second block. The only exogenous policy variable in the Glickman model is the property tax rate which helps to determine local public expenditures and revenues, but has no effect on any other variables in the model. Even more extreme is the Bell model which has no direct policy variables at all.

In both models, changes in the labour force do come to affect variables in the models. Both models use the same kind of non-manufacturing mechanism. Initially, the effect of an increase in labour force is to increase the level of unemployment. This increased unemployment lowers money wages in the non-manufacturing industry sectors. Since there is no direct feedback from wage to employment levels, employment initially remains fixed so that the total real wage bill falls. This reduces personal income lowering the output of the non-manufacturing sectors and reduces employment levels in these sectors. As a first round effect then, an increase in labour force leads to a decline in non-manufacturing activity and thereby to a decline in total output, income, and employment. Because of the internal complexity of the second block in each model, it is difficult to evaluate the full, as opposed to initial, impact of an increase in labour force without numerical solutions. Hall–Licari estimate full impact elasticities, as illustrated in table 1, using 1970 values for their model. These bear out the predominance of these initial effects.

The inability of a labour force increase to generate an increase in employment or output may seem to be surprising. As Engle (1974) argues, however, it is a consequence of the structural assumptions made. In both the Hall-Licari and Glickman models, the effect of a labour force increase is, through an increase in unemployment, to decrease the wage rates in certain industry sectors. There is no means by which wage rate changes encourage employers to make more jobs available. Thus, the three models are similar in their inability to generate growth-inducing migration.

---

14In the Hall-Licari and Glickman models, an exogenous increase in labour force also leads to a decrease in the regional price deflator used in converting from money wages to the real wage bill. This lessens the decline in the real wage bill that would occur if the deflator remained constant. This secondary effect is of a small enough magnitude in empirical work to omit in this discussion.
Table 1

<table>
<thead>
<tr>
<th>Endogenous variable</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of unemployed</td>
<td>23.618</td>
</tr>
<tr>
<td>Total real wage bill</td>
<td>-0.106</td>
</tr>
<tr>
<td>Total money wage bill</td>
<td>-0.200</td>
</tr>
<tr>
<td>Total employment</td>
<td>-0.032</td>
</tr>
<tr>
<td>Average money wage</td>
<td>-0.174</td>
</tr>
<tr>
<td>Regional price deflator</td>
<td>-0.106</td>
</tr>
<tr>
<td>Personal income</td>
<td>-0.097</td>
</tr>
<tr>
<td>Consumer expenditures</td>
<td>-0.083</td>
</tr>
<tr>
<td>Gross regional product</td>
<td>-0.069</td>
</tr>
</tbody>
</table>

Let us summarize how the last two econometric models structuralize the process of urban growth and how they handle the possibility of growth-inducing migration. Both models have a manufacturing sector in which output and employment growth depends strictly on national growth and a time trend. No other variable, endogenous or exogenous, affects the growth of this sector. In both models, growth in the non-manufacturing sectors is at least partly determined by this manufacturing sector. Both models permit other exogenous variables, notably national wages, to affect the non-manufacturing sectors although the Hall-Licari model additionally asserts a specific role for government variables such as federal revenue sharing. Both models, however, permit no direct role for growth-inducing migration and even suggest, through the effect of unemployment on wages, that migration will lead to a decline in total employment.

4. Mixed demand-supply-oriented models

To this point, a number of pure demand models of urban growth have been investigated. We now turn to a series of models which do permit some role for growth-inducing migration. These are termed ‘mixed’ models in that while they permit migration to affect urban growth rates, they also have roles for other sources of growth. An untested, theoretical model by Klein and the simple empirical models of Muth and Greenwood are considered first. Then, the models by Bell and Hall-Licari are partially reestimated using Muth’s formulation.

4.1. The Klein model

Klein’s (1969) model of a regional economy, although never implemented, is sufficiently different in its structure from the others to make it valuable for
study here.\textsuperscript{15} In particular, it emphasizes a role for prices not found in the previous models.

The model is built around a form of the Keynesian identity linking regional product and expenditures. For present purposes, the most relevant portions of his model are the following:

\begin{align*}
p\text{GRP} &= P_cC + P_iI + G_{sl} + G_f + p_cX - p_M M, \\
C &= f_1(PDI/P_c), \\
PDI &= p\text{GRP} - T - D, \\
l &= f_2(\text{GRP}, P_k, K(t - 1)), \\
M &= f_3(\text{GRP}, p/p_M), \\
X &= f_4(GNP, p/p_M), \\
p &= f_5(P, w), \\
w &= f_6(U, u, P_c),
\end{align*}

where

\begin{itemize}
  \item \text{GRP} = \text{Real gross regional product.}
  \item \text{p} = \text{GRP price deflator (regional).}
  \item \text{C} = \text{Real consumption (regional).}
  \item \text{P}_c = \text{Consumption price deflation (national).}
  \item \text{I} = \text{Real investment (regional).}
  \item \text{P}_i = \text{Investment price deflation (national).}
  \item \text{G}_{sl} = \text{State-local government expenditure (regional).}
  \item \text{G}_f = \text{Federal government expenditure.}
  \item \text{X} = \text{Real exports (regional).}
  \item \text{p}_v = \text{X price deflator (regional).}
  \item \text{M} = \text{Real imports (regional).}
  \item \text{p}_M = \text{M price deflator (exogenous, regional).}
  \item \text{PDI} = \text{Personal disposable income (regional).}
  \item \text{T} = \text{Total income tax paid (regional).}
  \item \text{D} = \text{Depreciation (regional).}
  \item \text{p}_k = \text{Rate of return on capital (national).}
  \item \text{K} = \text{Capital stock (regional).}
\end{itemize}

\textsuperscript{15}\text{Note that, in this paper, we shall discuss only 8 of the 20 equations making up the Klein model. This reduction simplifies our exposition without doing great damage.}
Although Klein's model does not explicitly consider migration, it does permit a growth-inducing role for it indirectly. An exogenous increase in immigration might initially be expected to swell the labour force and increase the local unemployment rate. This would decrease the local wage rate and GRP price deflator through (47) and (46). In turn, the real demand for exports would rise (45) and the demand for imports fall in (44). Thus, GRP should increase as a consequence of migration.

Is this an appropriate structural mechanism for generating growth-inducing migration? Should the effect of a labour surplus be through variations in the wage rate? The answer put forward by some is that increasingly institutional factors such as national labour unions make wage changes uniform among all cities. This seems to be consistent with the empirical finding by Bell (1967, pp. 114–116) that local labour unemployment has no effect on wage levels. This argument is also supported by Freund (1973, pp. 284–288) who could find no significant relationship with annual data covering manufacturing in 35 American SMSA’s from 1961 to 1967. Further, the Hall–Licari wage equations were found to have insignificant coefficients for the unemployment rate. In view of this, Glickman’s solitary significant results should be viewed with some apprehension.

Further, the Hall–Licari wage equations were found to have insignificant coefficients for the unemployment rate. In view of this, Glickman’s solitary significant results should be viewed with some apprehension.

The other alternative is to think of a labour surplus as acting directly on the equilibrium demand for labour rather than on the wage rate. Such a hypothesis is consistent with the polar-oriented growth model of fig. 1(b). It seems surprising to find that these hypotheses have not been included for testing in any of the models discussed to this point.

4.2. The models of Muth and Greenwood

Muth (1968, 1971) was among the first to suggest the structure of a formal theory of supply-oriented urban growth. Together with his theoretical model, he presents a simple two-equation model which he empirically estimates using data for some 78 urban areas in the United States. It is to this model that we now turn our attention.

Muth suggests that the growth in employment and the level of immigration are, in fact, interdependent. He argues that each is partly determined by the other as well as by other variables. One of the more successful forms used by him is as follows:16

---

16 Figs. (48) and (49) are adapted from column 4 of tables IV and V in Muth (1971, p. 304), omitting variables whose coefficients are statistically insignificant.
\[ (E(60)/E(50))^* = \alpha_{10} + \alpha_{11}(1 + M(60)/L(50))^* \]
\[ + \alpha_{12}(1 + N(60)/L(50))^* \]
\[ + \alpha_{13}(Y(60)/Y(50))^*, \]
\[ (1 + M(60)/L(50))^* = \alpha_{20} + \alpha_{21}(E(60)/E(50))^* \]
\[ + \alpha_{22}(1 + N(60)/L(50))^* \]
\[ + \alpha_{23}(1 + \Delta ML/L(50))^* \]
\[ + \alpha_{24}(U(50)/L(50))^*, \]

where

- \( E \) = Total civilian employment.
- \( M \) = Net in-migration over the previous decade.
- \( L \) = Total civilian labour force.
- \( N \) = Natural population increase over the previous decade.
- \( Y \) = Median family income.
- \( \Delta ML \) = The decade change in military personnel.
- \( U \) = The number of unemployed persons.
- \( 50 \) = Subscript referring to 1950.
- \( 60 \) = Subscript referring to 1960.
- * = Superscript denoting a natural logarithm.

As in Klein's model, the effect of an exogenous increase in migration is to change the employment levels. However, unlike Klein, Muth does not have this change occur via an apparent wage rate variation.

Greenwood (1973) develops a similar kind of model in which there are five endogenous variables; (i) the level of out-migration (\( OM \)), (ii) the level of in-migration (\( IM \)), (iii) the growth rate of median personal income (\( \Delta INC \)), (iv) the growth rate of employment (\( \Delta EMP \)), and (v) the growth rate of unemployment (\( \Delta UN \)). He used data for the 100 largest SMSA's in the United States with the migration data based on the 1955–60 period and growth rates referring to the decade change from 1950 to 1960. Each endogenous variable is hypothesized to be a function of some subset of other endogenous variables as well as a set of exogenous variables as illustrated in fig. 5.\(^{17}\) Note that, like Muth, Greenwood sees a direct link between migration and employment without any reference to a wage change linkage.

Both of these models see the sources of urban growth as being manifold. In Muth's model, the growth in jobs and in-migrants feed on one another

\(^{17}\)Note that the depicted structure is a subset of the structure original version hypothesized by Greenwood. For simplicity, non-significant statistical relationships have been omitted as in the case of Muth.
completely simultaneously. In Greenwood's model, however, there are three recursive blocks; the first contains just $OM$, the second $EMP$, $IN$, and $INC$, and the final one just $UNEMP$. Thus, an exogenous change in $UNEMP$ has no effect on the remainder of his model, while a change in $OM$ affects all other variables. An exogenous increase in in-migration, then, brings about increases in employment, income, and unemployment. Note, however, that among the exogenous forces affecting the endogenous variables in either model, there is no longer a measure of export demand present as found in all previous models.

This illustrates a central difficulty with Muth and Greenwood in that they fail to present a formal theory which might underly the specific models they have chosen to estimate. While both models contain a growth-inducing role for migration, it is not clear why the growth of jobs is limited by anything more than the growth of labour force. What process is being modelled here? Muth and Greenwood are really modelling the short run dynamics of the urban labour market where entrepreneurs try to anticipate the change in labour supply and migrants in turn try to anticipate the number of new jobs to be created. However, the five to ten year time frames of the two studies is too long to realistically estimate such a model.\textsuperscript{18} Is there any evidence to

\textsuperscript{18}The extended version of Czamanski's Baltimore model seems to be on a more appropriate, one-year time frame.
suggest that, within a shorter time frame, these models would still be empirically valid?

4.3. Bell and Hall–Licari revisited

Some reworking of the Bell model is possible because the original data are presented in his article. The data for the Hall–Licari model was graciously provided by one of the authors. This enables a reexamination of the growth mechanisms underlying both models.

The hypothesis is forwarded that the level of exports (or manufacturing) value-added is responsive to the growth in labour force. The decision to tie in export growth in this way is related in part to the recursive structure of these models. If the level of migration affects the export level, in other words, the whole set of endogenous variables is affected. Certain complexities in the estimation of non-linear simultaneous models are posed by such a relationship and we do not consider them at present. The Ordinary Least Squares method is used in both cases to derive the following equations to replace (23) and (36):

\[ e(t) = -0.01348 + 1.00310 g(t) + 1.50272 l(t), \]
\[ R^2 = 0.678, \quad DW = 2.690, \quad N = 15, \] (4.705) (1.785)

\[ q(t) = -0.11396 + 0.99539 g(t) + 3.75270 l(t), \]
\[ R^2 = 0.568, \quad DW = 1.952, \quad N = 11. \] (1.615) (2.018)

The two equations share some similarities. The first, using Bell's data, relates the growth rate of received export income, \( e \), to the growth rate of GNP, \( g \), and the growth rate of the regional labour force, \( l \). The second equation poses the same model, using the data of Hall–Licari, where the endogenous variable, \( q \), is the growth rate of manufacturing output. Both equations show a near-unity relationship between the manufacturing-export sector growth and national growth. The difference between the coefficients of \( 't' \) is somewhat larger but this may be in part due to the many differences between 'manufacturing output' and 'received export income'. Additionally, in both models the slope coefficients are either significant or near-significant at the 5\% level.

These two equations are in the spirit of the work of Muth and Greenwood. They suggest in addition that the simultaneity between econ-
5. Conclusions

What has been learned from all of the above discussions? Several interesting models of urban growth have been examined. All of these models have attributed urban growth to one of two sources: either external demand for the city's output (usually related to GNP) or a black box called labour market dynamics. Is that all there is to city growth? The answer to that, by any serious student of urban form, must be no. So, in concluding, it is appropriate to speculate on where research on urban growth should be headed. Three areas seem to offer special promise.

The first area of interest involves an examination of the concept of short-run labour market dynamics, especially on the supply side. This would involve the construction of models which hypothesize the behaviour of the urban job seekers. What processes are involved in his search for a job? How does he decide where he will look for a job and how long he will look? How does he weigh housing and other factors in making his decision to in- (or out-) migrate to (or from) an area? What is the relationship between migration and population growth both in terms of fertility-mortality changes and in terms of subsequent movements of dependents and others? On the general topic of search theory, starting points here might come from the work of David (1974), Phelps (1970), and Zarembka (1972, pp. 54-62). One operational model of this type is being developed by Cordey Hayes and Gleave (1974).

The second area of interest concerns some spatial aspects of urban growth. There are a number of questions which could be usefully raised here. One set of questions would be concerned with the two-way linkage between urban growth and a city's hinterland or market area. To what extent, for instance, is migration to a city in any period limited by the nature of the urban system? How do improvements in technology alter the relationship between a city and its hinterland so that hinterland labour is freed to migrate to the city? Answers to such questions may prove to be an important determinant of the growth of particular cities.

Other spatial aspects which might affect urban growth have to do with the interior spatial structure of the city. Simple Alonso-type models suggest, for instance, that there are several kinds of costs which increase with city size. Hoch (1972) finds similar empirical evidence. None of the models examined, however, introduce such costs at all. Variables such as average commuting costs, rent levels, environmental quality, the dispersion of job opportunities,
and recreational access costs could be introduced explicitly and endogenously into aggregate growth models.

This issue of rising costs with city size brings into perspective the third major area of interest. If we view the city simply as a concentration of production activities, one principal raison d'être lies in indivisibilities of production which generate economies of scale. In the theory of urban economic growth, an equilibrium size is reached when further increases to city size increase the costs of such size faster than these economies of scale. This trade-off between economies and costs has not entered any of the models discussed here.

This issue of why cities exist has more implications than just the introduction of a more refined or plausible growth model. It brings out a central conflict with respect to the theory behind a pure supply model of urban growth. As was mentioned earlier, the pure supply model assumes constant returns to scale in production. This would seem to deny a major reason for the existence of cities at all. A careful reconciliation of these two assumptions is required as part of the theoretical structure underlying pure supply models. One might try to argue, for example, that there are economies of scale, but that these are exhausted at the present scale of large cities. If so, then why and to whom does such a city export its goods? If each city services only its hinterland then in what sense can it be viewed as a price taker; another assumption of the pure supply model. What prevents the city from behaving monopolistically within such a market area? Some careful rethinking has to be done on this whole conflict.

References
Berry, B.J., 1974, Growth centers in the American urban system, Vol. 1 (Ballinger, Cambridge, MA).


