Externalities and Industrial Development

Vernon Henderson
Brown University

The traditional industrial location literature assumes that firms locate in a community in response to changes in the current comparative advantages of different locations (see Herzog and Schlottmann, 1991, for a review). The existing pattern of location of firms in an industry thus depends on such things as current wages, population, industrial composition, utility prices, and tax rates of these localities. In contrast, a relatively new literature assumes that existing location patterns for an industry are strongly influenced by "history," in particular the historical industrial environment of cities (Glaeser, Kallal, Scheinkman, and Schleifer, 1992; Henderson, Kuncoro, and Turner, 1992; Miracky, 1992). A related paper by Blanchard and Katz (1992) argues that State employment patterns follow a unit root process with drift. Thus a local pattern of growth or decline will persist over time, without "reversion to the mean."

Externalities are central to all location analyses. Firms cluster together at various locations to receive information spilling over from other firms, to reduce transport costs of interfirm trade, and to enhance the diversity of firms and local products available. Questions arise as to whether externalities for a firm derive specifically from nearby firms in the same industry or from the general diversity and scale of the local environment. If the former, then externalities are Marshall-Arrow-Romer (MAR) externalities internal to an industry in a city, or economies of locationalization. If the latter, then they are Jane Jacobs (1969) externalities, or generalized external economies of urbanization for an industry within a city.

Further questions arise as to whether externalities of either type are primarily static or dynamic. Does the historical industrial environment matter, or simply the current environment? Work by Jaffe, Trajtenberg, and Henderson (1993) suggests that location-industry specific information diffuses slowly over space, so that access to the knowledge binds firms to the same location over time. A larger scale of own industry activity historically means that firms currently in the locality will operate with greater accumulated knowledge (about such things as technology and sources of supply of various kinds of inputs) than they would otherwise operate. Also the maturity of a locality's social information network is important in facilitating communications and information spillovers among local firms.

The papers by Glaeser et al. (1992), Henderson, Kuncoro, and Turner (1992), and Miracky (1992) cited earlier assert a role for dynamic externalities based on a comparison of locations at just two points in time. The time gaps range from 10 to 30 years, and the papers find that the industrial environments of 10, 17, or 30 years ago affect location today. But an alternative interpretation is that there is a location fixed/random effect that gives rise to the role of history. Such a fixed/random effect could reflect, for example, differences in unmeasured regional resource endowments that persist over time and show ongoing regional comparative advantages, with the result that current industrial location patterns are correlated with historical patterns simply because they draw upon the same relative endowments. Additional time invariant and unmeasured attributes could include notions of local culture affecting the institutional environment, as well as attributes of immobile, specifically skilled portions of the labor force. The papers do not seriously address this problem. While, for example, Henderson, Kuncoro, and Turner consider issues of endogeneity, there is an inherent problem with cross-sectional work in finding instrumental variables that themselves are uncontaminated by local fixed/random effects.

This article investigates the role of dynamic versus static externalities, as well as MAR versus Jacobs externalities. To delineate the issues, an 11-year panel of data on county employment levels in different two-digit manufacturing industries is used. Use of a panel permits specific modelling of fixed/random effects and investigation of the lag structure of externality measures, as well as other variables. This process provides a means for investigating the length of the period during which history matters.

The Model

The structural model focusses on analysis of the level of employment in a two-digit industry in any year. The model is a reduced version of a detailed structural form model of industrial location derived and estimated in Henderson (1993) for a single cross-section. Firms in location j in time period t have a profit function $\Pi_{jt} = \Pi(\hat{Y}_{jt}, s_{jt}, \tilde{u}_{jt})$, where $\hat{Y}_{jt}$ is a vector of arguments depicting current and historical (lagged) conditions such as prices and externality measures, s_{jt} is the current employment level in location j in time t, and $\tilde{u}_{jt}$ is an error term. The current supply of entrepreneurs to an industry in a locality is a function of current and historical arguments (for example, population) G_{jt}. The supply function is $\Pi_{jt} = \Pi(G_{jt}, s_{jt}, d_{jt})$ where as local industry size rises, per firm profits must rise to attract more entrepreneurs. Local employment is the intersection of the $\Pi(\cdot)$ and $\Pi(\cdot)$ functions. Solving for s_{jt}, I get a reduced form equation:

$$s_{jt} = s(Y_{jt}, u_{jt}). \qquad (1)$$

Y_{jt} combines $\hat{Y}_{jt}$ and G_{jt}, and u_{jt} combines $\tilde{u}_{jt}$ and d_{jt}. In a “stable” equilibrium, sign $[\partial s(\cdot)/\partial \tilde{y}]$ = sign $[\partial \Pi/\partial \tilde{y}]$, so that variables favorably affecting per firm profits also favorably affect local employment levels.

At this point, it is important to detail a more specific form to equation (1) and a breakdown of the arguments to Y_{jt} and e_{jt}. Specifically, it is hypothesized that:

$$s_{jt} = \alpha + \sum_{l=1}^{m} \alpha_l s_{j,t-l} + \sum_{l=1}^{m} \delta_l X_{j,t-l} + \beta Z_j + f_j + e_{jt}, \qquad (2)$$

where $s_{j,t-l}$ are lagged values of the dependent variable. The lag structure is specified to start at t - 1 and run m periods. The t - 1 values are “current” values, corresponding to a context where decisions for employment in period t are based on the best information about current production conditions, which are the previous period’s realizations. Values for $m \geq 2$ represent a role for history, where prior realizations contribute to the current environment for production. $X_{j,t-l}$ are lagged values of other variables—exogenous or endogenous. Z_j are time invariant variables that differ across space (for example, climate or regulations).

In estimating equation (3), the focus is on isolating impacts of the historical industrial environment. Past employment ($s_{j,t-l}$) measures persistence in industry employment patterns over time. Past own industry concentration measures capture current or historical MAR effects. Concentration is measured by the ratio of local own industry to total local employment. Measures of diversity of local employment and local industry scale are used to represent Jacobs, or urbanization externalities. Diversity is measured by Hirschman-Herfindahl indices (HHI) of sums of squared shares, as defined below.

In equation (2), the error term from (1) has been decomposed into a random/fixed effect, f_j and contemporaneous drawing e_{jt}, where f_j represents the influence of time invariant unmeasured characteristics of the local area, which affect the right-hand side variables, in particular $s_{j,t-l}$ but also potentially certain of the $X_{j,t-l}$. The e_{jt} are generally assumed to be identically and independently distributed across time and space. The absence of serial correlation of the e_{jt} in the employment level equation (2) is a strong assumption, which is investigated.

Estimation of equation (2) presents the problem that the f_j are correlated with right-hand side variables. With ordinary least squares estimation, coefficients will be biased. For example, a high degree of persistence in own industry employment may be estimated, not because past employment directly influences present employment, but because of persistence in unmeasured (essentially time invariant) regional endowments determining employments in both years. Estimation of equation (2), where f_j are treated in standard fixed effects estimation procedures, still results in biased estimates because the contemporaneous error term, e_{jt}, is correlated with any time average of $s_{j,t-l}$. To eliminate the fixed/random effects, rather than following the standard fixed effects procedure, the equations are first differenced to obtain:

$$\Delta s_{jt} = \sum_{l=1}^{m} \alpha_l \Delta s_{j,t-l} + \sum_{l=1}^{m} \delta_l \Delta X_{j,t-l} + \Delta e_{jt}. \qquad (3)$$

Note that $\Delta s_{jt} \equiv s_{jt} - s_{j,t-1}$; $\Delta s_{j,t-2} \equiv s_{j,t-2} - s_{j,t-3}$; $\Delta e_{jt} = e_{jt} - e_{j,t-1}$; and so forth. While first differencing eliminates the fixed/random effect, by construction it introduces simultaneity problems and serial correlation, even though the e_{jt} are identically and independently distributed. In particular, $\Delta s_{j,t-1} \equiv s_{j,t-1} - s_{j,t-2}$ is correlated with $\Delta e_{jt} \equiv e_{jt} - e_{j,t-1}$, since $e_{j,t-1}$ affects $s_{j,t-1}$. In fact, it is reasonable to assume also that many of the $X_{j,t-1}$ are affected by $e_{j,t-1}$. Second, $\Delta e_{jt} = e_{jt} - e_{j,t-1}$ is correlated by $\Delta e_{j,t-1} = e_{j,t-1} - e_{j,t-2}$.

To obtain consistent estimates of the parameters requires the use of instrumental variables. I assume that there are no strictly exogenous variables but merely predetermined ones. That is, there is a row vector Z_{jt}, where

$$E\,[e_{jt}\, Z_{js}] = 0 \qquad s = 1, 2, \ \ldots \ t-1$$

$$t = 1, \ \ldots \ , T.$$

For each county j, in year t, I include in this row vector all s_{js}, and all X_{js}, plus a few other measures of local industrial characteristics for all years $s \leq t - 1$. This implies the instrument list varies from year to year.

In estimation of equation (3), each year is treated as a separate equation with a sample size equal to the number of localities and cross-equation constraints imposed on all coefficients other than any constant term (differenced time dummies). The number of equations is the length of the panel, T, minus the length of the observed lag structure m minus 2, or

T - m - 2. Of the minus 2, one is lost in differencing and the other is lost from instrumenting. Thus the longer the lag structure the more years are lost in estimation.

The model is estimated by the generalized methods of moments estimator for panel data in the time series processor econometrics package. Under conditional homoscedasticity, the estimates reduce to a generalization of three-stage least squares (Hayashi, 1992), or full information instrumental variables estimation of Brundy and Jorgenson (1971), which allows for a variable instrument list by year and accounts for the serial correlation across years in the error terms. The generalized methods of moments procedure in time series processor also allows for heteroscedasticity through a White-type correction of the variance-covariance matrix. The estimates are efficient in terms of use of instruments, and coefficients and standard errors are consistently estimated.

First, differencing the level equations eliminates not just the fixed effects but also the constant term and time invariant variables. To recover these, I insert the estimated coefficients from equation (3) into equation (2) to obtain

$$s_{jt} - \sum_{l=1}^{m} \hat{\alpha}_l s_{j,t-l} - \sum_{l=1}^{m} \hat{\delta}_l X_{j,t-l} = \alpha + \beta Z_j + f_j + e_{jt}, \text{ for } t = m + 1, \ldots T.$$

I then average over the T - m - 1 years to get an estimating equation:

$$B_j = \alpha + \beta Z_j + f_j + \bar{e}_j, \qquad (4)$$

where

$$B_j \equiv (\bar{s}_{jt} - \sum_{l=1}^{m} \hat{\alpha}_l \bar{s}_{j,t-l} - \sum_{l=1}^{m} \hat{\delta}_l \bar{X}_{j,t-l})/(T - m - 1).$$

Provided that f_j and Z_j are orthogonal to each other, equation (4) may be estimated by ordinary least squares to obtain estimates of α and β, treating f_j as a random effect and $f_j + \bar{e}_j$ as a composite identically and independently distributed error term.

In actual estimation of equation (3), I have time-fixed effects, or dummies, for the T years of estimates. These are Δd_t; where d_t is the time dummy for the level equation. The d_t for $t = 0, \ldots T$ are solved for in obtaining B_j in equation (4) by imposing the normalization that

$$\sum_{t=0}^{T} d_t = 0.$$

The Data

The sample consists of 11 years of complete data for the 742 urban counties of the United States, covering the years 1977 to 1987. Counties are not agglomerated into metropolitan statistical areas (MSAs) since that would lose valuable information. Instead, some measures are constructed for both the county and, when relevant, for the surrounding metropolitan area. About three-quarters of the counties are in a multicounty metropolitan area. (About half of the MSAs and primary metropolitan statistical areas are single-county areas.)

The basic data set is from County Business Patterns (CBP), which in this version records employment, number of firms, and wages for all one, two, and two-digit industries. The data used have been treated by the Center for Governmental Studies, Northern Illinois University (Gardocki and Baj, 1985), to give point estimates for employment in those cases where, for disclosure reasons, employment is reported in interval form. Most counties at the two-digit level in CBP's data report exact numbers, but, when disclosure is an issue, they report employment in intervals, using a fairly fine classification. The Northern Illinois State numbers are an improvement over the use of midpoint values of the intervals, in estimation, thus accounting for overall State employment and average firm sizes in constructing point estimates.

CBP data are supplemented with county data on education, taxes, land area, and so on from the 1977 and 1982 *City and County Data* book. In addition there are data on State right-to-work laws, coastal location, annual State population, and annual average State electricity prices for industrial users. The CBP data are used to construct annual county wage rates (in all other industries than the own industry), annual measures of concentration and diversity for the county and metropolitan area in various dimensions (discussed later), and annual measures of local economic activity (for example, county civilian employment, county manufacturing employment, and corresponding numbers for the surrounding metropolitan area).

The panels for each industry are balanced, requiring positive employment in all years. Sample sizes average about 550 of the 742 U.S. urban counties (as of 1990). For the industries that are examined, only about 10 percent of the counties not in the balanced panel have some employment during some (but not all) of the 11 years. The remainder have zero employment in these industries in all years.

Empirical Results

Preliminaries

I estimate equations (3) and (4) for four two-digit industries: chemicals, primary metals, machinery, and electrical machinery. All level variables (employment, population, and price measures) are in logs and therefore in differences in logs (annual growth rates) in equation (4). Concentration and diversity measures (discussed below) are in linear form.

Hausman-type tests (Hayashi, 1992) were carried out on various instrument lists to test for the time when variables become predetermined. In particular in equation (2), I assume e_{jt} are identically and independently distributed over time, so that in equation (3) Δe_{jt} are only correlated with $\Delta s_{j,t-1}$ and possibly $\Delta X_{j,t-1}$. But if the e_{it} are themselves correlated over time, then Δe_{jt} will be correlated with earlier $\Delta s_{j,t-l}$ or $\Delta X_{j,t-l}$, in which case, for example, $s_{j,t-2}$ or $X_{j,t-2}$ would be inappropriate instruments. In Hausman-type tests, I could not reject the hypotheses for all industries that either including $s_{j,t-2}$ as an instrument or including $s_{j,t-2}$ and all $X_{j,t-2}$ as instruments yield the same results (by χ^2 tests) as excluding them as instruments. (I also broke variables into two groups, those that pertain to own industry measures and those that do not, and conducted the same types of tests with the same results.) Moreover, I could reject the hypothesis that including $X_{j,t-1}$ as instruments yields the same results as excluding them. In short I selected an instrument list that, for year t, excludes all t - 1 variables as instruments and includes all $t - l$ variables for $2 \leq l \leq t - 1$ as instruments.

Regarding a lag structure, I set m = 6, or regressed 7 years from the present. Then for an 11-year panel T - m - 2 = 3, or, in estimation of equation (3), 3 years are covered. For machinery, primary metals, electrical machinery, and chemicals, the numbers of counties, N, are, respectively, (for balanced panels) 674, 454, 508, and 548. Pseudo-F tests on the value of the objective function (with instrumental variables) under different lag structures accepted the hypotheses that adding a 5th year onto a 4-year lag structure or adding a 6th year onto a 5-year structure significantly improves the results except for chemicals, where a 4-year lag structure dominates longer lag structures. However, for primary metals and electrical machinery (but not machinery), adding a 7th year also modestly improves over a 6-year lag structure but loses another year in terms of equations and degrees of freedom.

I compromised on a lag structure of 6 years. As will be seen, the lag pattern seems to differ noticeably across variables, making it difficult to impose exogenous uniformity, such as geometric or Pascal, to the lag structure so as to estimate infinite lag structures. It is noted, however, that the variables are first differenced. Variables in the vectors $\Delta X_{j,t}$ and $\Delta X_{j,t-1}$ have simple correlation coefficients that always have absolute values less than 0.15 and typically less than 0.025; that is, there is almost no multicollinearity among lagged regressors. Thus shutting down the lag structure at m = 6 may result in little bias to the coefficients $\delta_1, \ldots \delta_6$, even if lagged effects persist beyond m = 6. To aid in presentation, I also estimated a quadratic Almon lag structure for s = 0, . . . 5 (that is, $l = s + 1$) where:

$$\delta_s = \alpha_0 + \alpha_1 s + \alpha_2 s^2 \quad s = 0, \ldots 5; \; l = s + 1.$$

The coefficients α_0 and δ_5 are unconstrained.

Results on Time Variant Variables

Tables 3 and 4 contain the basic results for equation (3) estimation. Table 3 starts with key industrial environment variables, and Figure 1 presents the corresponding results for an Almon lag structure for machinery that is representative.

Lagged own industry employment is a necessary control, measuring the prior record of the industry in the county. After differencing out the fixed effect, coefficients on lagged employment are zero or even negative. With level equations, even under attempted instrumental variables estimation, coefficients on prior employment levels are typically about 0.5 (implying mean reversion). Thus, a 1 percent higher employment level in the past results in a 0.5 percent higher employment level today. Table 3 and Figure 1 part (a) suggest that this association is spurious and arises simply because of the existence of fixed/random effects. Here a 1 percent higher employment level last year, or separately a year earlier and beyond, appear to modestly reduce employment today. Any effects disappear after 6 years.

To more fully explore the role of own industry historical activity, I experimented with a variety of other measures of historical own industry activity. A common way to measure MAR (own) industry externalities is to look at concentration; in this case, I compared the ratio of own industry to total employment in the county. Higher concentrations of own industry employment imply higher levels of spatial interaction among firms because they are more clustered together, as opposed to being spatially dispersed. Clustering improves information spillovers among firms.

Table 3 and Figure 1 part (b) report the results for the concentration measure. First, it is noted that including the concentration measure, as opposed to excluding it completely, has virtually no impact on other variables. In particular, if I remove the measure for all industries and all lags, the coefficients on lagged own industry employment experience no sign changes, nor do they change appreciably in magnitude. This is also true for other variables, such as metropolitan area employment in Table 4.

Concentration has a positive significant impact on current employment. Last year's concentration (a "static" externality) presents a positive impact for machinery. But historical ($l \geq 2$) concentrations also affect current employment. In fact, in Figure 1, part (b), the concentration coefficient rises and peaks around $l = 3$ and then declines to zero at $l = 6$. A cubic Almon lag indicates that δ starts off at $\delta_1 \approx 1$, drops to $\delta_2 \approx .5$, and then rises to peak at about $\delta_l = 1.4$ before phasing out at $l = 6$. These results suggest the presence of a dynamic externality where historical own industry concentration positively affects productivity and employment today.

In Table 1, I report the long-term impacts of permanently raising concentration from a given year on. These impacts are obtained by summing coefficients in Table 3 on concentration over the 6-year lag, ignoring the small impacts of the lagged dependent variables. (Of course, concentration contains an own industry employment measure, making the experiment for this variable questionable.) For primary metals and electrical machinery, the sums are near 12, meaning that a one standard deviation increase in concentration would raise own industry employment over time by about 50 percent. At least 50 percent of that amount is based on the historical (beyond one lag) impacts.

In addition to concentration, other significant measures of own industry activity are subsequently discussed. In the following paragraphs, I examine the rest of the industrial environment.

To examine the impact of the remainder of the industrial environment, I used diversity measures. I calculated a variety of Hirschman-Herfindahl indices for two-digit industries; that is, indices for all manufacturing and all industries, for both the county and the MSA/primary metropolitan statistical area of which the county is a part. For example, for manufacturing in the metropolitan area, the HHI is the sum of squared shares of each two-digit manufacturing industry (other than the own industry) in total (all other) manufacturing employment in the metropolitan area. An HHI measures lack of diversity in the environment surrounding the own industry. For 19 other two-digit manufacturing industries, HHI takes a maximum of 1 if remaining employment is concentrated in just one other industry, and a minimum of 0.0526 if it is uniformly distributed across all 19 industries.

In a static context, diversity improves efficiency for a firm through Dixit-Stiglitz production impacts of being able to purchase a greater diversity of intermediate inputs locally (see Abdel-Rahman and Fujita, 1991), and also through Jacobs-type effects of being in a richer information environment. This latter effect is thought to persist to form a dynamic externality—a base of diverse location-specific historical knowledge.

In empirical estimation of HHI impacts, I rely on year-by-year composition *changes* to provide information on diversity impacts. Although these changes are typically very small, Table 3 shows that increases in HHI, or decreases in past diversity, generally hurt current employment. Moreover the first year effects (static externalities) are often smaller than second or some latter year effects. Even though, for machinery, electrical machinery, and primary metals, all but one of the significant coefficients in the lag structure are negative, there are some positive coefficients or insignificant negative coefficients or both. The Almon lag structure in Figure 1 indicates that the impact of HHI persists over the 6 years. For long-run impacts within the 6-year horizon, in Table 1, I look at the impact of a one-time permanent change in the HHI. The sums range up to 2, suggesting that a one standard deviation *decline* in HHI can raise employment by up to approximately 20 percent over the long term in machinery or primary metals.

Taken together the results on concentration and diversity suggest that both MAR and Jacobs effects are present for major two-digit manufacturing industries. An industry benefits both from concentrated own industry employment and from diverse all other industry employment. If all industries are similarly structured, there is a real tension for localities taken as a whole. Ideally the goal is to achieve concentrated, yet diverse employment bases. The tension may be resolved by having concentrated employment in a county's major export industries but diversity among other supporting industries. For large metropolitan areas that can be attained if various counties concentrate in different activities. For small, single-county metropolitan areas, modest employment levels in a variety of supporting activities are required.

In interpreting static versus dynamic externalities, I have distinguished between effects at one lag versus latter lags. Unfortunately there are some problems with that interpretation. I examine industry employment responses to changes in the industrial environment and assume that an improvement in the environment acts as a static externality. Firms respond to the improved environment by expanding employment and by entering the locality. Because that response in scale of industry operation may not be instantaneous, there may be a lagged adjustment. Thus latter lagged impacts of changes in the industrial environment might not merely reflect dynamic externalities but could be simply delayed responses to static improvements.

To clarify this possibility, one can look at adjustment times for variables that do not represent externalities but simply reflect market conditions. If adjustments to changes in market conditions are very quick, then the longer lags for industrial environment variables more plausibly imply that dynamic externalities rather than adjustment lags are at work.

Table 4 and Figure 1, parts (d) and (e), present results on market variables such as wages and measures of regional demand for a county's output, which affect firm employment. Regional demand is measured by the total metropolitan area scale as represented by metropolitan area employment. Using State civilian population, rather than metropolitan area employment, yields very similar types of results, with coefficients of other variables essentially unaffected.

The scale variable, whether measured by metropolitan area employment or State population, has a consistent lag pattern across industries. After one or two periods, the effects diminish and often show up as negative for longer lags, or with signs switching between positive and negative. The Almon lag structure confirms this pattern for all industries. With a cubic structure, effects go to zero within two periods and stay near zero for the duration. Thus the adjustment to changes in demand conditions, at least the scale of demand, occurs quickly. This rapid adjustment to market scale variables is in contrast to the longer lags for the HHI and concentration measures, and these may be plausibly interpreted as dynamic externalities rather than adjustment time.

Regarding wages I have generally been able to get the correct negative sign for all industries by instrumenting. Clearly, higher wages reduce employment. Wages may rise because of industrial conditions in a city or because of a worsening of amenities and the resulting increased compensating wage differentials. How do firms react to higher wages? Wage effects seem to persist for four or five periods in Table 4 or Figure 1, part (d), a condition that cannot be attributed to dynamic externalities. Rather, it must reflect an adjustment process in which adjustment to increased wages as firms are adjusting technologies and dealing with unions requires a number of periods. Wages can be viewed as affecting firm factor proportions which may require some time to adjust, given the purchase of new machines and technologies, whereas industry scale adjustments (number of firms and level of firm scale)—from, for example, metropolitan area employment changes—occur quickly.

In Table 1, I summarize the impacts of a one-time permanent increase in wages or metropolitan area scale. The wage effects vary considerably across industries, as do the scale effects. Except for primary metals, the long-run elasticities of both wage and scale impacts are less than 1.

Finally, in Table 5, I consider other variables depicting the own industry industrial environment. I calculated an HHI for shares of the three-digit components of each own two-digit industry for the metropolitan area. Greater intraindustry diversity implies greater intraindustry specialization and a potentially richer range of activities for encouraging intraindustry advancement within the metropolitan area. For electrical machinery, primary metals, and machinery, increased HHI—or reduced diversity—in the past has hurt current industry employment. Effects for primary metals persist in strength through $l = 5$ and 6 but diminish for $l \geq 5$ for the other two industries. Table 1 presents the impact of a permanent change in this internal industry HHI. A one standard deviation permanent *decrease* in HHI can lead to a 15- to 20-percent increase in long-run employment for primary metals and electrical machinery.

The other measure in Table 5 concerns the level of own industry employment in other counties in the metropolitan area. A greater scale of own industry activity in the rest of the metropolitan area may promote a richer environment for a county's industry, but it also represents a demand base for intraindustry trade. For primary metals and electrical machinery, such effects appear to exit, over a short horizon. However, the coefficients and long-term impacts (in Table 1) are small.

A final comment on the estimation of the model in equations (2) and (3) is that the model is a location model for which history matters. For individual industry employment levels, this model has been chosen over a growth from a base model as in Barro and Sala-i-Martin (1991) or Glaeser et al. (1992). Such a model would regress the employment growth rate between 1987 and 1977 on 1977 base variables. In this model I could reject this formulation outright if base-period variables in equation (3) had no impact. Base-period own industry employment is consistently negative and significant, but its coefficient is tiny (-.012 to -.025). The other variables (in particular concentration and HHI for all metropolitan manufacturing) have mixed sign patterns and significance.

Results on Time Invariant Variables

Thus far I have examined the results for time variant variables based on estimating first differenced equations. The first differencing eliminates the time invariant variables in equation (2). To recover these I estimate equation (4), whereby the dependent variable is B_j, the estimated average "residual" from the reinstated level equation (2). Results are reported in Table 2, for residuals from simple 2SLS estimates for balanced panels.

In Table 2 the top panel forms the basic estimating equation. In general being in a county having a more educated adult population (in 1980) boosts industry employment by increasing the general quality of the labor force and perhaps encouraging better public decisionmaking. Simply being on the coast also helps, indicating easy access to shipping and perhaps more desirable climate and air quality conditions. However, being part of a larger metropolitan area with multiple counties detracts from industrial employment. Larger metropolitan areas are more service oriented, more congested, and have land prices of greater relative disadvantage to manufacturing.

The dummy for State right-to-work laws indicates a potentially perverse effect. In theory, although State right-to-work laws should create an environment unattractive to unions and more attractive to firms, the net result seems to be the reverse. This may be a problem of separating out fixed/random effects from time invariant variables. States with poor inherent conditions for manufacturing (low draws of f_i) may choose to pass State right-to-work laws to try (unsuccessfully on average) to compensate. A similar problem arises for tax rate variables. Insertion of county property tax or State income tax measures in Table 2 results in positive signs for the variables. That may reflect greater f_i's, resulting in greater demand for and realization of per capita infrastructure, and hence higher taxes.

The bottom panel of Table 2 presents results for an industrial environment variable that may be arguably independent of the fixed/random effect. The weighted average size of firms in all other industries in 1977 is intended as a proxy for an age measure. Larger firms are typically old firms that generally are less dynamic and represent a more stagnant environment. In Table 3 the coefficient is generally positive and insignificant. Only for electrical machinery does it have the expected effect.

Conclusions

Today there are pressures for a city to formulate an effective industrial environment. Cities desire concentrations of employment in particular industries, but also want the advantages of a diverse industrial base. Diversity tends to raise productivity and hence employment in a city's particular concentration of production and export activity. Cities, metropolitan areas, and counties are all highly specialized in terms of manufacturing, reflecting the benefits of concentration. Yet diversity at some level is also important.

These effects appear to persist over time. Increased concentrations of own industry activity appear to affect employment levels for the next 5 to 6 years. For diversity measures, effects appear to persist beyond the 6-year horizon examined in this article. Given the rapid adjustment to market-scale or demand variables, these long lags suggest a presence of dynamic externalities.

The interpretation here is that the past industrial environment affects future employment, but there is a broader causation issue. One needs to know if an industry needs a diverse environment to initiate production in a locality, or can it count on such an environment to generate itself? For emerging high technology industries, Henderson, Kuncoro, and Turner (1992) suggest that a preexisting diverse environment is important in attracting the industries to a locality. Here I suggest that it is important in retaining industries, as well.

Author

Vernon Henderson is Eastman Professor of Political Economy and professor of economics and urban studies at Brown University, as well as research associate in the National Bureau of Economic Research. *The author gratefully acknowledges the support of the National Science Foundation and the U.S. Department of Housing and Urban Development (HUD). The paper benefitted from comments made by James R. Follain, Norman J. Glickman, and participants of the 1993 HUD conference on Regional Growth and Community Development. The author also thanks research assistants Arindam Mitra, John Frost, Ari Kuncoro, and Mark Beardsell.*

Figure 1

Almon Lag: Machinery

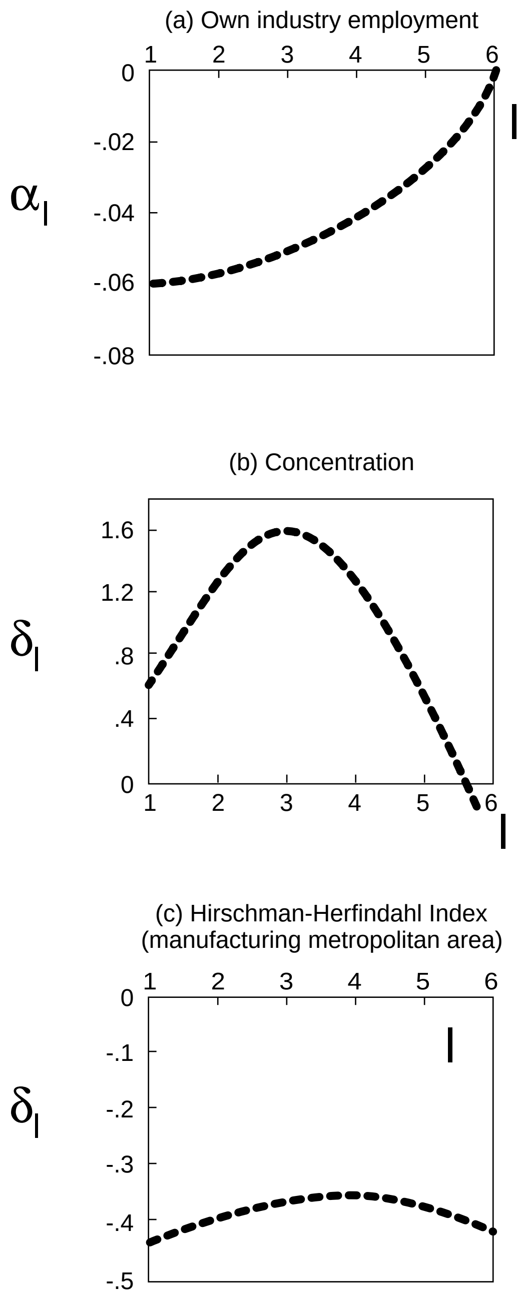

Figure 1 (continued)

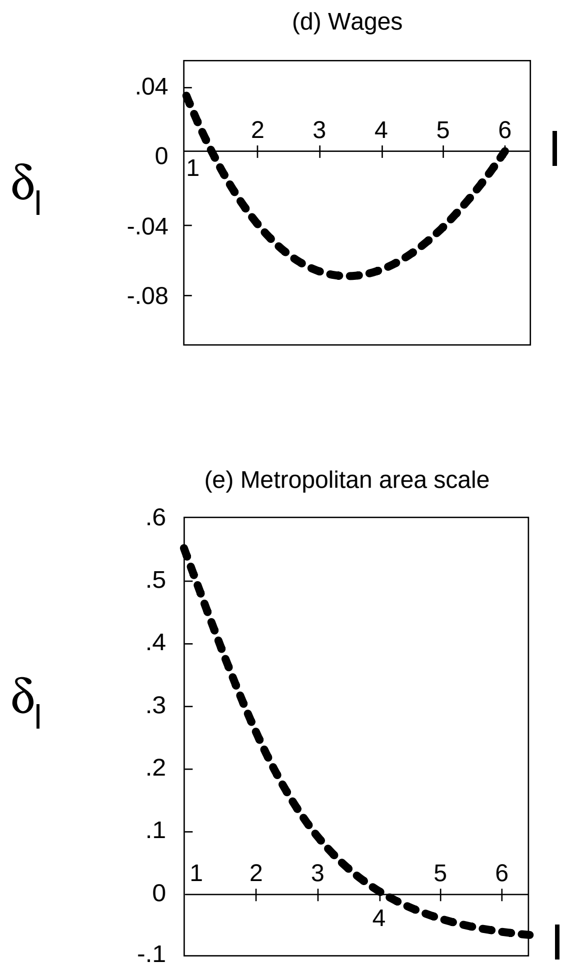

Table1

Impact of a one-time permanent increase

$$\sum_{l=1}^{6} \delta_l$$

	Machinery		Primary Metals	
	$\Sigma\delta$	**1 standard deviation increase**	$\Sigma\delta$	**1 standard deviation increase**
Concentration	3.29	.13	11.6	.46
HHI (manufacturing industry)	-2.08	-.19	-1.96	-.15
l_n (wages)	-.198		-1.33	
l_n (metropolitan area employment)	.079		1.71	
HHI (internal own industry)	-.306	-.06	-.86	-.19
l_n (own industry employment, surrounding counties)	-.057		.171	

	Electrical Machinery		Chemicals	
	$\Sigma\delta$	**1 standard deviation increase**	$\Sigma\delta$	**1 standard deviation increase**
Concentration	12.3	.54	4.68	.21
HHI (manufacturing industry)	-.509	.04	0	0
l_n (wages)	-.321		-.627	
l_n (metropolitan area employment)	.813		.856	
HHI (internal own industry)	-.743	-.16	-.085	-.02
l_n (own industry employment, surrounding counties)	.239		0	

Note:

HHI stands for the Hirschman-Herfindahl indices.

Table 2

Time invariant effects[1]

	Machinery	**Primary metals**	**Electrical machinery**	**Chemicals**
Constant	-2.361* (.336)	1.237* (.573)	-1.564* (.299)	1.719* (.317)
Percent of adults with at least a high school education	.021* (.006)	.011 (.010)	-.003 (.005)	-.006 (.006)
Percent of adults with at least a college education	.023* (.008)	.003 (.015)	.036* (.007)	.018* (.008)
Dummy: MSA/ PMSA on coast[2]	-.056 (.083)	.049 (.131)	.144* (.069)	.318* (.074)
Dummy: county in multicounty MSA/PMSA	-.368* (.082)	-.456* (.124)	-.066 (.065)	.031 (.072)
Dummy: State has right-to-work law	-.154 (.079)	-.314* (.127)	-.157* (.068)	-.158* (.073)
Weighted average of firm sizes in other industries in county, 1977	.018 (.016)	.048 (.033)	-.070* (.019)	.003 (.017)
N	676	456	510	550

Notes:

[1]Standard errors are in parentheses.

[2]MSA is metropolitan statistical area, and PMSA is primary metropolitan statistical area.

*Significance at the 95-percent level.

Table 3

Industrial environment variables[1]

County own industry employment				
	Machinery	**Primary metals**	**Electrical machinery**	**Chemicals**
Lag 1	-.086 (.047)	-.324* (.024)	-.159* (.023)	-.288* (.020)
Lag 2	-.031* (.145)	-.201* (.015)	-.075* (.011)	-.116* (.018)
Lag 3	-.059* (.010)	-.047* (.012)*	-.037* (.011)	-.070 (.014)
Lag 4	-.059* (.011)	-.084* (.013)	-.024* (.011)	-.010 (.011)
Lag 5	-.013 (.012)	.015 (.011)	.008 (.012)	.021 (.012)
Lag 6	.012 (.012)	-.031* (.011)	-.058* (.011)	.027* (.013)

Table 3 continued on next page.

Table 3 (continued)

Concentration in county (county own industry employment/county total employment)				
	Machinery	**Primary metals**	**Electrical machinery**	**Chemicals**
Lag 1	2.30* (1.08)	4.42* (1.10)	6.12* (1.08)	2.90* (.683)
Lag 2	.270 (.395)	2.85* (.576)	2.41* (.675)	2.31* (.637)
Lag 3	1.29* (.326)	2.24* (.50)	.241 (.452)	.428 (.474)
Lag 4	1.66* (.374)	1.83* (.517)	1.55* (.522)	-.443 (.469)
Lag 5	-1.22* (.379)	-1.17 (.621)	-.395 (.423)	-.921 (.514)
Lag 6	-1.01* (.373)	1.39* (.421)	2.33* (.545)	.408 (.868)

Metropolitan area diversity (HHI[2] for *all* two-digit manufacturing industries in metropolitan area)				
Lag 1	-.278 (.357)	-.968 (.588)	.405 (.499)	-.272 (.488)
Lag 2	-.408* (.202)	-1.437* (.277)	-.599* (.280)	.006 (.240)
Lag 3	-.565* (.194)	.907* (.276)	.009 (.54)	-.738* (.255)
Lag 4	-.404* (.177)	.458 (.317)	-.213 (.265)	.606* (.279)
Lag 5	.314 (.171)	-.429 (.276)	-.066 (.266)	-.101 (2.78)
Lag 6	-.737* (.182)	-.495 (.270)	-.045 (.221)	.479* (.222)

Notes:

[1] Standard errors are in parentheses.
[2] HHI stands for Hirschman-Herfindahl indices.
* Significance at 95-percent level.

Table 4
Market variables[1]

	Machinery	Primary metals	Electrical machinery	Chemicals
Wages				
Lag 1	.078 (.083)	-.133 (.075)	-.269* (.081)	-.218* (.103)
Lag 2	-.048 (.045)	-.140* (.068)	.243* (.075)	.047 (.044)
Lag 3	-.017 (.037)	-.092 (.070)	-.075* (.036)	-.039 -.039
Lag 4	-.021 (.064)	-.199* (.071)	-.225* (.077)	-.207* (.092)
Lag 5	-.173 (.095)	-.394* (.169)	-.042 (.092)	-.267* (.126)
Lag 6	-.017 (.436)	-.371* (.141)	.047 (.058)	.057 (.076)
Metropolitan area employment				
Lag 1	.215 (.240)	1.55* (.352)	.576 (.312)	.475 (.305)
Lag 2	.414* (.128)	.320 (.199)	.084 (.163)	.293 (.192)
Lag 3	.237* (.107)	.172 (.155)	.270* (.136)	-.113 (.149)
Lag 4	-.061 (.099)	-.376* (.142)	-.266* (.122)	.141 (.148)
Lag 5	-.064* (.021)	.054 (.086)	.624* (.160)	.069* (.027)
Lag 6	-.143* (.027)	-.010 (.210)	-.475* (.154)	-.009 (.038)

Notes:

[1] Standard errors in parentheses.
*Significance at 95-percent level.

Table 5

Other own industry[1]

	Machinery	Primary metals	Electrical machinery	Chemicals
Internal (three-digit) HHI[2]				
Lag 1	-.036 (.138)	-.406* (.136)	-.220* (.102)	-.015 (.114)
Lag 2	.096 (.073)	-.237* (.074)	-.130* (.062)	-.027 (.064)
Lag 3	-.136* (.053)	.197* (.055)	-.241* (.049)	-.070 (.073)
Lag 4	-.121* (.056)	.055 (.062)	-.110 (.064)	-.017 (.063)
Lag 5	-.099 (.071)	-.251* (.068)	-.019 (.055)	-.061 (.063)
Lag 6	-.010 (.050)	-.218* (.064)	-.023 (.058)	-.071 (.057)
Own industries employment, surrounding counties				
Lag 1	-.008 (.020)	.018 (.027)	.031 (.016)	.018 (.021)
Lag 2	-.012 (.015)	.024 (.018)	.046* (.010)	.009 (.012)
Lag 3	.012 (.012)	.027 (.014)	.037* (.008)	-.055* (.015)
Lag 4	-.029* (.010)	.026* (.012)	.029* (.008)	.010 (.008)
Lag 5	-.015 (.012)	.049* (.009)	.050* (.007)	(.002) (.011)
Lag 6	-.005 (.009)	.027* (.011)	.046 (.010)	.019 (.010)

Notes:

[1] Standard errors in parentheses.
[2] HHI stands for the Hirschman-Herfindahl indices.
* Significance at 95-percent level.

References

Abdel-Rahman, H.M., and M. Fujita. 1991. "Specialization and Diversification in a System of Cities," University of Pennsylvania, mimeo.

Barro, R., and X. Sala-i-Martin. 1991. "Convergence Across States and Regions," *Brookings Papers on Economic Activity*, 1.

Blanchard, O., and L. Katz. 1992. "Regional Evolutions," *Brookings Papers on Economic Activity*, 1, 1–75.

Brundy, J., and D.W. Jorgenson. 1971. "Efficient Estimation of Simultaneous Equations Systems Using Instrumental Variables," *Review of Economics and Statistics*, 53, 207–224.

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

Gardocki, B.C., and J. Baj. 1985. "Methodology for Estimating Non-Disclosure in County Business Patterns Data," Center for Governmental Studies, Northern Illinois University, mimeo.

Glaeser, E., H. Kallal, J. Scheinkman, and A. Schleifer. 1992. "Growth in Cities," *Journal of Political Economy*, 100, 1126–1152.

Hayashi, F. 1992. "Comment," *Journal of Business and Economic Statistics*, 10, 15–17.

Henderson, J.V. 1993. "An Econometric Model of Industrial Location," *Journal of Urban Economics*, forthcoming.

Henderson, J.V., A. Kuncoro, and M. Turner. 1992. "Industrial Development in Cities," NBER working paper no. 4178.

Herzog, H.W., and A.M. Schlottmann, eds. 1991. *Industrial Location and Public Policy*, Knoxville: University of Tennessee Press.

Holtz-Eakin, D., W. Newey, and H. Rosen. 1988. "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56, 1371–1395.

Hsaio, C. 1986. *Analysis of Panel Data*, Cambridge University Press.

Jacobs, Jane. 1969. *The Economy of Cities*, New York: Random House.

Jaffe, A.B., M. Trajtenberg, and R. Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, 577–598.

Keane, M., and D. Runkle. 1992. "On the Estimation of Panel Data Models with Serial Correlation When Instruments Are Not Strictly Endogenous," *Journal of Business and Economic Statistics*, 10, 1–9.

Levin, A., and C.F. Lin. 1992. "Unit Root Tests in Panel Data," San Diego discussion paper no. 92–23.

Miracky, W.F. 1992. "Technological Spillovers, the Product Cycle and Regional Growth," Massachusetts Institute of Technology, mimeo (November).