NBER WORKING PAPER SERIES

EXTERNALITIES AND INDUSTRIAL DEVELOPMENT

Vernon Henderson

Working Paper No. 4730

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1994

Support of NSF for this research is gratefully acknowledged. This first draft benefited from comments of participants in the RSA and AEA meetings in 1993 and seminars at the Philadelphia Federal Reserve, Harvard, HUD, and a CEPR conference in Vigo, Spain. John Frost, Arindam Mitra, Ari Kuncoro and Mark Bearsell worked as able RA's on this data intensive project. Any opinions expressed in this paper are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper #4730
May 1994

EXTERNALITIES AND INDUSTRIAL DEVELOPMENT

ABSTRACT

Using a panel data set of county-level employment in machinery, electrical machinery, primary metals, transportation, and instruments, this paper analyzes the role of dynamic externalities for individual industries. Key issues examined include the role of externalities from own industry concentration (localization, or MAR externalities) versus the role of externalities from overall diversity of the local environment (urbanization, or Jacobs externalities). In contrast to previous studies, use of panel data allows us to separate these effects out from fixed/random effects influencing industries over time.

Panel data also allow us to estimate a lag structure to externality variables, indicating how long history matters and the time pattern of effects. A particular issue concerns whether conditions from the immediate year or so prior to the current have the biggest impact on current employment, or periods several years prior have the largest impact.

For all industries both localization and urbanization effects are important. For traditional industries most effects die out after four or five years, but for high tech industries effects can persist longer. The biggest effects are typically from conditions of three to four years ago, in the county and metropolitan area.

Vernon Henderson
Department of Economics
Brown University
Providence, RI 02912
and NBER

The traditional industrial location literature assumes that firms move in response to changes in the current comparative advantage offered by different locations (see Herzog and Schlottman 1991 for a review). The existing pattern of location of firms in an industry then depends on current wages, population, industrial composition, utility prices, tax rates, etc. of these different localities. In contrast, a relatively new literature assumes that existing location patterns for an industry are also strongly influenced by "history", in particular the historical industrial environment of cities (Glaeser, Kallal, Scheinkman and Shleifer (1992), Henderson, Kuncoro and Turner (1992) and Miracky (1992)). Historical conditions determine the intangibles of today's economic environment. Such intangibles include the local stock of knowledge relevant for an industry or availability of a labor force with specific acquired skills. Generally, we can't measure these current intangibles, but we can measure their historical determinants and calculate their impact on firm location decisions and industry employment in a locality today.

The importance or non-importance of history is a critical issue. If history is important, the spatial allocation of resources -- the landscape -- we see at a point in time will be largely predetermined. Individual localities will have limited ability to influence or change what they do, particularly in the short run. What firms choose to locate in a particular city will be determined by the current industrial environment of that city, which in turn will largely depend on the locality's specific history.

In considering the nature of history's role, there are also key questions. How long does history matter -- are conditions from six or seven years ago versus one or two years ago important? Do conditions thirty years ago matter? In essence, how tied to the past is a locality? Second, what is the timing of the strongest impacts? Are last year's industrial environment conditions most important to today; or is there an aging, maturing, or transmission mechanism, giving the strongest role to industrial conditions of several years ago? This paper investigates these key questions.

Externalities are central to all discussions of the industrial environment of a locality. Firms cluster together to better receive information spilling over from other firms, to improve search and matches between workers and firms, to enhance the diversity of firms and local products available and to derive benefits from local intra-industry specialization. Apart from the general economics literature, most of these phenomena have been explicitly modeled with a distinctive spatial flavor in the urban-economic geography literature (Fujita and Ogawa 1982, Helsley and Strange 1990, and Abdel-Rahman and Fujita 1991). But there is a long standing debate about whether externalities for a firm derive specifically from other firms in the same industry locally or from the general diversity and scale of the local environment. If the former, externalities are Marshall-Arrow-Romer (MAR) externalities internal to an industry in a city, or economies of locationalization. If the latter they are Jane Jacobs (1969) externalities, or generalized external economies of urbanization for an industry within a city.

Consider information spillovers. The debate concerns whether a firm learns primarily from other firms in its own industry or from firms outside the industry, whether they be suppliers, purchasers, or folks on the cocktail circuit or at the neighborhood bar. Sometimes the lines are fuzzy. Does a firm learn about the best suppliers for a particular item by observing what other firms in the industry do, being told by those other firms what they do, or by being told by the suppliers themselves what other firms do, or, in fact, all three? What about information concerning new technologies? Who is relevant to a firm may vary by industry and stage of product development. R&D activities may be more sensitive to Jacobs externalities, given cross-fertilization of ideas across market sectors, while standardized production activities may be more sensitive to localization economies, and spillovers of industry specific information.

There is now a second debate, which I initially posed, about whether externalities of whatever type are primarily static or dynamic. Does the historical industrial environment matter, or just the current environment? Work by Jaffe et al (1993) suggests that location- industry specific information diffuses slowly over space, so that access to that knowledge binds firms to the same location over time. A larger scale of own industry activity historically means firms today in that locality will operate with greater intangibles, such as accumulated knowledge (about technology, sources of supply of different quality inputs, etc.) than otherwise. Also, the maturity of the social information network of a locality matters in facilitating communications and information spillovers among local firms.

The papers by Glaeser et al. (1992), Henderson et al. (1992) and Miracky (1992) cited earlier assert a role for dynamic externalities based on comparing employment patterns across locations at just two points in time. The time gaps range from 10-30 years. These papers find that industrial environments of 10, 17 or even 30 years ago causally affect location decisions and patterns today. The reasoning is that aspects of the prior industrial structure of a city are a dynamic externality for firms today, giving a strong role to history. The effects are very strong. For example, in Henderson et al. (1992) preliminary regressions of the log of current employment on just the log of base period employment yield coefficients of 0.85 to 0.96 for individual industries indicating strong persistence (although unit roots can be rejected[1]): and, even in the "structural" equations, coefficients remain close to 0.5. But rather than this link between the past and the present representing dynamic externalities, an alternative interpretation is that there is a location fixed/random effect that gives rise to the role of history and high persistence in employment patterns. The fixed/random effect captures relatively time invariant (within a 10-30 year horizon), unmeasured location attributes. For example differences in unmeasured regional resource endowments persist over time, reflecting ongoing regional comparative advantages, so current industrial location patterns are correlated with historical patterns simply because they draw upon the same relative endowments. Additional time invariant, unmeasured attributes would include notions of local culture affecting the local legal, business and institutional climate, as well as attributes of the relatively immobile, specific skill portions of the labor force. These papers do not seriously address this problem. While, for example, Henderson et al. consider issues of endogeneity, there is an inherent problem with cross-sectional work in finding instrumental variables that themselves are uncontaminated by the fixed/random effects.

In this paper I investigate the role of dynamic versus static externalities, as well as MAR versus Jacobs externalities. I utilize an eleven year panel for 1977-1987 of data on county employment levels in different 2-digit manufacturing industries. Eleven years is the length of the panel currently available. That isn't a very long history, but it will turn out to be long enough to allow us to infer the impact of past industrial environments on employment today. By using a

[1] Frost (1994) conducts tests for a variety of industries at the county level used in this paper.

panel, I can specifically model fixed/random effects and sort out whether past industrial environment conditions matters, per se, or whether correlations over time are due to the presence of fixed/random effects. More directly I can estimate a lag structure of the impact of past conditions, and how long these impacts of different externality dimensions persist. This exercise is fraught with many conceptual and statistical problems. Rather than overwhelm the introduction with a list, I will try to deal with them each comprehensively at the appropriate point in the paper.

The Model

The structural model focuses on analyzing either industry employment or the number of firms in a two-digit industry in a location in any year. The model is a reduced form version of a detailed structural form model of industrial location derived and estimated in Henderson (1994) for a single cross section. Conceptually the model is based on standard empirical work on firm location decisions (see Herzog and Schlottmann 1991 for a review; also see Carlton 1983). Firms in location j in time period t have a profit function $\Pi_{jt} = \tilde{\Pi}(\tilde{Y}_{jt}, s_{jt}, \tilde{u}_{jt})$, where $\tilde{Y}_{jt}$ 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. $\tilde{u}_{jt}$ is an error term. The current supply of entrepreneurs to an industry in a locality is a function of current and historical arguments, 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 industry size is determined by the intersection of the $\tilde{\Pi}(\cdot)$ and $\Pi(\cdot)$ functions. Solving for s_{jt} we get a reduced form equation

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

While s_{jt} is labeled total industry employment, this scale measure could be replaced by the total number of firms in the industry in the locality. I experimented with both measures. Y_{jt} combines $\tilde{Y}_t$ and G_{jt} and u_{jt} combines $\tilde{u}_t$ and d_{jt}. In a "stable" equilibrium, sign $[\partial s(\bullet)/\partial \tilde{y}]$ = sign $[\partial \Pi/\partial \tilde{y}]$, so 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 (1) and a breakdown of the arguments to Y_{jt} and e_{jt}. Specifically we hypothesize that

$$s_{jt} = \alpha + \sum_{\ell=1}^{m} \alpha_\ell s_{j,t-\ell} + \sum_{\ell=0}^{m} \delta_\ell X_{j,t-\ell} + \beta Z_j + f_j + e_{jt} + d_t \qquad (2)$$

$s_{\ell,t-1}$ are lagged values of the dependent variable. For lagged values of the dependent variable, the lag structure is specified to start at t-1 and run m periods. $X_{j,t-\ell}$ are lagged values of other variables -- exogenous or endogenous. For variables not directly representing aspects of the lagged dependent variable the lag structure starts today (t-0). Such variables will include wages and measures of locality diversity and size. t-0 and perhaps t-1 are "current" values: for t-1 this corresponds 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. Z_j are time invariant variables (over an 11 year period) that differ across space (e.g. coastal location, state regulations).

In estimating equation (3), the focus is on isolating impacts of the historical industrial environment, both MAR (localization) and Jacobs (urbanization) effects. Past employment ($s_{j,t-\ell}$) measures persistence in industry employment patterns over time which may reflect current or historical MAR or localization effects. However measures of the concentration of past own industry employment may better represent current or historical MAR effects. Conceptually (see Fujita and Ogawa 1982) intra firm spillovers appear to accelerate with increases in the intensity of local industry activity. Later I will detail issues on how to measure intensity of own industry activity. Measures of diversity of overall local employment external to the own industry are used to represent Jacobs, or urbanization externalities. Diversity is measured by Hirschman-Herfindahl indices of sums of squared employment shares, as defined below.

In equation (2) the error term from (1) has been decomposed into a random/fixed effect, f_j, a time effect, d_t, applying to all localities in time t, and a contemporaneous drawing, e_{jt}. f_j represents the influence of time invariant unmeasured characteristics of the local area which affect RHS variables, in particular $s_{j,\, t-\ell}$ but also potentially certain of the $X_{j,t-\ell}$. The e_{jt} are generally assumed to be i.i.d. across time and space. The absence of serial correlation of the e_{jt} in the employment level equation (2) is a strong assumption, which I investigate.

Estimation of (2) presents the problem that the f_j are correlated with RHS variables. With OLS 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 (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 $s_{j,t-\ell}$. To eliminate the fixed/random effects, rather than following the standard fixed effects procedure, the equations are first differenced to obtain

$$\Delta s_{j,t} = \sum_{\ell=1}^{m} \alpha_\ell \, \Delta s_{j,t-\ell} + \sum_{\ell=0}^{m} \delta_\ell \Delta X_{j,t-\ell} + \Delta e_{jt} + \Delta d_t \qquad (3)$$

Estimating equation (3) means we lose information on the impact of cross-sectional variation in X_j's and must rely on time series variation within localities. Given the equations are estimated by regressing first differences on first differences, it may be surprising how strong the results turn out to be. In equation (3), note $\Delta s_{jt} \equiv s_{jt} - s_{j,t-1}$; $\Delta s_{j,t-2} \equiv s_{j,t-2} - s_{j,t-3}$; $\Delta e_{jt} = e_{j,t} - e_{j,t-1}$; etc. While first differencing eliminates the fixed/random effect, by construction it introduces simultaneity problems and serial correlation, even though the e_{jt} are i.i.d. 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 with $\Delta e_{j,t-1} = e_{j,t-1} - e_{j,t-2}$.

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

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

or

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

For each county j, in year t for equation (3), I include in this row vector all s_{js}, all X_{js}, plus a few other measures of local industrial characteristics for all years $s \le t-2$. This implies the instrument list varies from year to year, growing as time t increases.

In estimation of equation (3), each year is treated as a separate equation with 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 imposed 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. Note the longer the lag structure the more years we lose in estimation.

The model is estimated by a GMM estimator for panel data.[2] Under conditional homoscedasticity, the estimates are 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 estimation procedure also accounts 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 under the maintained assumptions.

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

$s_{jt} - \sum_{\ell=1}^{m} \hat{\alpha}_\ell s_{j,t-\ell} - \sum_{\ell=0}^{m} \hat{\delta}_\ell X_{j,t-\ell} = \alpha + \beta Z_j + f_j + e_{jt}$, for $t=m+1, \ldots T$. I then average

over the T-m-1 years. This gives an estimating equation where

[2] The TSP econometrics package contains such an estimator.

$$B_j \equiv (\bar{s}_{jt} - \sum_{\ell=1}^{m} \hat{\alpha}_\ell \bar{s}_{j,t-\ell} - \sum_{\ell=0}^{m} \hat{\delta}_\ell \bar{X}_{j,t-\ell})/(T-m-1).$$

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

Provided the 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 i.i.d. error term. In actual estimation of equation (3) we have also time fixed effects. These are Δd_t; where d_t is the time dummy for the level equation. The d_t for t=0, ...T are solved for in obtaining B_j in (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 USA, covering the years 1977-87. Counties are not agglomerated into MSA's since that would lose valuable information. Instead some measures are constructed for both the county and, when relevant, for the surrounding metro area. About three quarters of the counties are in a multi-county metro area. (About half of MSA's and PMSA's are single county ones.) Using both county and metropolitan area measures as variables relevant to employment in a county may suggest that the assumption that the e_{jt} in equation (2) are i.i.d. across space is too simple. Rather e_{jt} may be correlated for all counties within a metropolitan area. Some simple diagnostics on the residuals indicated this may not be a problem in the estimation of equation (3).[3]

[3] Specifically for each year, I calculated $\Delta\hat{e}_{jt}$ in equation 3 and then grouped these county-level residuals by metro area. If these error terms are correlated within metro areas, the mean $\Delta\hat{e}_{jt}$ within each metro area should systematically differ from the overall mean in each year. I conducted F-tests for each year for industries 35 and

The basic data set is from County Business Patterns [CBP] which in our version records employment, number of firms and wages for all 1, 2, and 3-digit industries. The data I use 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. These data were kindly given to me by Bill Miracky. Almost all counties at the two digit level in CBP's data report exact numbers; but, where disclosure is an issue they report employment in intervals, using a fairly fine classification. The Northern Illinois State numbers are an improvement over using mid-point values of the intervals in estimation. It accounts for overall state employment and average firm sizes in constructing point estimates. To ensure that use of this data is not a problem, I also estimated the model using the number of firms as the s_{jt} measure in equations (2) and (3), where then s_{jt} is not subject to any censoring by CBP's. The number of firms data do suffer from a small numbers or integer problem, where, for example, the number of plants in a locality for transportation and primary metals averages about a dozen. Nevertheless, the results are qualitatively very similar.

CBP data are supplemented with county data from the 1977 and 1982 City and County Data book, on education, taxes, land area, etc. In addition I have 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 (see below), and annual measures of local economic activity (e.g. county civilian employment, county manufacturing employment, and corresponding numbers for the surrounding metro area).

---

36. In only 1 of 6 cases could I reject the hypothesis that the group means were collectively the same as an overall mean. For that one case, the F-value was just 1.32, allowing us to barely reject the hypothesis at the .047 significance level.

I constructed panels for primary metals, machinery, electrical machinery, transport equipment and instruments. The panels for each industry are balanced, requiring positive employment in all years. Sample sizes are respectively 454, 674, 508, 402 and 509 of the 742 counties, accounting for balancing and missing data (two counties). Of the remaining counties for an industry, some have zero employment in all years, a large portion have either clearly missing values or a scattered year here or there where minimal employment occurs, and another significant portion have entries that appear to represent long term exit and entry or shorter-term episodes of significant entry and exit.[4] Unfortunately the data are not perfect (!) and besides obviously missing values (e.g., a firm recording zero employment in one year and 4,000-6,000 in the other ten years), there are many episodes that appear to involve either temporarily losing track of a firm or firm SIC reclassifications. For machinery and electrical machinery, there appear to be respectively 3 and 55 episodes of sustained entry (starting with zero in 1977 and then initiating sustained employment at some point) and 5 and 11 episodes of sustained exit. In addition, there appear to be roughly 3 and 32 counties which have multiple (non-sustained) entry and exit episodes. For machinery and electrical machinery we thus have 2-13% of counties in total with valid entry-exit episodes that are omitted in the balanced panel. These are not enormous fractions but the phenomenon is exciting. A separate paper will analyze exit and entry for particular three digit industries where there is much more turnover.

[4] For example for electrical machinery, 53 counties have zero employment in all years, 42 have missing data, 38 have recordings of 1 year of minimal activity and 98 have what appear to be valid entry-exit episodes (55 entry, 11 exit and 32 multiple episodes). Missing data entries typically involve counties with (large) employment in all years except one, a recording of the (positive) number of firms for that year, and a blank for employment. Minimal activity is typically a county with zeros in all years except one, and a recording of one firm with 1-7 employees for that one.

Empirical Results

In this section I present the main results, the estimates of equation (3). In the next section, I will present results on equation (4), pertaining to the impact of time invariant variables. This section is broken into three parts, preliminaries reporting some diagnostics, results on the main industrial environment variables, and results on market variables.

Preliminaries

I estimate equations (3) and (4) for five 2-digit capital goods industries: primary metals, machinery, electrical machinery, transportation, and instruments. These industries, in contrast to, say, non-metallic minerals and metallurgy, are the capital good industries which products are widely traded across cities and in which localities absolutely and relatively specialize. All level variables (employment, population, and price measures) are in logs and hence equation (3) is in growth rates (differences in logs). Diversity measures (see below) are not transformed to logs.

Hausman-type tests (Hayashi, 1992) were carried out on different instrument lists to test for the time when variables become predetermined. In particular in (2), I assume e_{jt} are i.i.d. over time, so that in (3) Δe_{jt} are only correlated with $\Delta s_{j,t-1}$ and possibly $\Delta X_{j,t-1}$. Thus all $s_{j,t-\ell}$ and $X_{j,t-\ell}$ for $\ell \geq 2$ may be used as instruments. But if the e_{it} are themselves correlated over time, then Δe_{jt} will be correlated with earlier $\Delta s_{j,t-\ell}$ or $\Delta X_{j,t-\ell}$, in which case for example, $s_{j,t-2}$ or $X_{j,t-2}$ would be inappropriate instruments for all industries.[5] In Hausman-type tests I did not

[5] While the Hausman test suggests that Δe_{jt} is not correlated with $\Delta s_{j,t-2}$, (implying that $e_{j,t-1}$ is not correlated with $e_{j,t-2}$ and hence $s_{j,t-2}$), the estimated covariance matrices are not fully supportive of a simple i.i.d. process for the e_{jt}. In theory, the simple correlation coefficient between $\Delta e_{j,t} = e_{j,t} - e_{j-t}$ and $\Delta e_{j,t-1} = e_{j,t-1} - e_{j,t-2}$ should be -0.5, if the variances of $e_{j,t}$ are the same across years. Typically the coefficients are -0.1. But the estimation does not assume equal variances across years, or (given White-corrections) even across counties.

come near to rejecting 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. Thus for efficiency reasons we include $s_{j,t-2}$ and $X_{j,t-2}$ as instruments.

In terms of a lag structure we set m=6, or look back seven years from the current. Then for an eleven year panel T-m-2=3, or, in estimation of equation (3), three years are covered. Evidence for moving the lag structure beyond four or five years is not that strong. Pseudo-F tests on the value of the objective function under different lag structures suggest for two of five industries that adding a sixth year onto a five year structure does not improve the value of the objective function. I used a lag structure of six years because I wanted to explicitly look at as long a structure as possible in the data. I was uncomfortable about losing further degrees of freedom by going to seven years. As we will see, with a few notable exceptions for particular industries, most effects tend to peak by $\ell=4$ and disappear for $\ell=5$ or 6. Further, the lag pattern seems to differ noticeably across variables, making it difficult to impose exogenous uniformity to the lag structure, such as geometric or Pascal, so as to estimate infinite lag structures.[6]

Industrial Environment Variables

The set of tables in Appendix A contains the coefficients for equation (3). The variables are broken into two groups, although the separation is not strict, as discussion will reveal. The first group contains variables measuring the condition of the industrial environment, current or

[6] I note that in estimation the variables are first differenced. Variables in the vectors $\Delta X_{j,t}$ and $\Delta X_{j,t-1}$ have simple correlation coefficients which always have absolute values less than 0.30 and $\Delta X_{j,t}$ and $\Delta X_{j,t-2}$ have correlation coefficients less than 0.13 and typically around 0.05. That is, there is fairly low multicollinearity among lagged regressors. Thus shutting down the lag structure at m=6, may result in limited bias to the coefficients $\delta_1, \ldots \delta_6$, even if lagged effects persist beyond m=6.

past, and includes lagged own industry employment and Hirschman-Herfindahl indices of diversity of the industrial environment external and also internal to the industry. The second group contains variables that control for market conditions, such as wages and scale of activity in the metropolitan area or state, as measured by total employment or population, reflecting local demand for the product. We start the discussion by looking at the industrial environment variables.

Past Own Industry Employment. Previous work suggests that the level of current employment in an industry is strongly affected by past own industry employment levels and the concentration of that activity (Henderson et al. 1992 and Miracky 1992). Two questions arise from that work. Does the association between past and current employment levels reflect causation, perhaps through dynamic externalities; or does it reflect, say, the presence of a fixed/random effect? Second, how can we measure concentration?

In Table 1, I present some representative results for electrical machinery that demonstrate an initial puzzle. On the left hand side of that Table are the coefficients for lagged own industry employment in the county going back seven years, and the coefficients for a typical concentration measure, the share of total county employment devoted to the own industry. The coefficients on lagged own industry employment are little affected by the presence or absence of the concentration measure.

Table 1 suggests that, after differencing out the fixed/random effect, the coefficients on lagged own industry employment don't simply diminish further from 1 or 0.5. They actually become negative, although modest in magnitude. Why is this? From work on plant turnover (Dunne, Roberts, and Samuelson 1990) and from raw simple correlation coefficients in the data, there is a ready answer. The $\Delta s_{j,t-\ell}$ are negatively correlated over time. A positive shock in t-1 generates firm births in that period, but most new firms die out almost immediately in the next

Table 1

Electrical Machinery

	ln (own ind. emp.) (β_{1l})	concentration own ind./total employ (β_{2l})	ln (own ind. employ) γ_{1l})	ln (own ind.emp))^2 (γ_{2l})
lag 1	-.130*	6.129*	-.605*	.053*
	(.025)	(1.054)	(.075)	(.0089)
lag 2	-.068*	2.508*	-.378*	.033*
	(.012)	(.670)	(.038)	(.0040)
lag 3	-.035*	.121	-.149*	.0097*
	(.012)	(.450)	(.037)	(.0035)
lag 4	-.023*	1.920*	-.102*	.0085*
	.011	(.548)	(.031)	(.0029)
lag 5	.006	-.573	-.071*	.0059*
	(.012)	(.431)	(.029)	(.0026)
lag 6	-.069*	2.354*	-.114*	.0070*
	(.011)	(.563)	(.025)	(.0029)

mean concentration ≈ .031
mean employ ≈ 3000 (ln (3000)=8)
mean ln (employ) ≈ 6.6

$\frac{\partial \ln (emp_t)}{\partial \ln (emp_{t-l})} = \beta_{1l} + \beta_{2l} conc_{t-l}$ $\qquad$ $\frac{\partial \ln (emp_t)}{\partial \ln (emp_{t-l})} = \gamma_{1l} + 2\gamma_{2l} \ln emp_{t-l}$

for l1: = .06 at conc=.031 $\qquad$ = .25 at lnemp = 8 few periods, as predicted by a Jovanovic (1982) model of learning. So a high birth rate in t-1 is followed by a high death rate in t. Note the negative correlation applies to changes in s, not levels (i.e., some births survive, so levels in t are higher because of births in t-1). Our estimated coefficients have a levels (equation (2)) interpretation. However, to estimate the level coefficients by first differencing out the fixed effects and de-emphasizing the cross-sectional variation in the data, we accentuate the Jovanovic phenomenon (which is not explicitly part of the model), complicating the interpretation of the results.

There is a further result in the left hand panel of Table 1 that higher concentration has a positive ameliorating effect on this Jovanovic process. The coefficients of past concentration are generally positive, large, and significant. So increased employment also increases concentration and at high concentration levels that generates benefits more than sufficient to offset the Jovanovic effect. For electrical machinery for $\ell=1$, a 1% increase in last year's employment nets a .06% increase in employment today at average concentration levels (.031 share). At very high concentration levels such as 0.2 to 0.25 shares, a 1% increase in own industry employment last year would result in a 1.1 to 1.4% increase in employment today.

Upon consideration, I decided that concentration in counties of fixed geographic size is better measured by a non-linear representation of the level of own industry employment. A concentration of 50% of a labor force of 2,000 doesn't represent much opportunity to accumulate local industry knowledge relative to a concentration of 5% of a labor force of 200,000. By focusing just on own industry employment, the impact of total county employment can then be separated out and evaluated on its own. Also, I measure MAR effects at the level of the county, not the overall metro area, since your "immediate" neighbors are the ones whose information spills over onto you. Experimentally, own industry employment in the rest of the metro area controlling for all other variables (see below) has little impact on own industry county employment for all industries and I dropped it as a variable in the main results, but footnote its impact here.[7]

On the right hand side of Table 1 for electrical machinery, I reformulate the overall impact of past own industry employment as a quadratic. Inspection reveals that at low own industry employment levels the negative Jovanovic effect dominates while at higher levels, what I interpret to be positive MAR (localization) externalities dominate. For electrical machinery the turning point (depending on the lag point) appears to be about a log employment of 5-6.5, or actual employment of 150-700. At mean employment of about 3,000 per county in electrical machinery at $\ell=1$ a 1% increase in employment last year generates a .24% increase in employment today.

Coefficients for all industries for the quadratic formulation are reported in Table A1 in the Appendix for all industries. The general pattern is the same for all industries as for electrical machinery. However for machinery, primary metals, and transport equipment coefficients are generally insignificant at $\ell=5$ and $\ell=6$, so no significant effects persist beyond $\ell=4$. Table 2 in the text presents a summary of the results. On the LHS of Table 2 the impact of last year's ($\ell=1$) employment on this year is calculated for different employment levels. As we know from the quadratic, effects increase (become more positive or else less negative) as county own industry past employment rises from 300 through 3,000 to 30,000. For most industries maximum county own industry employment is about 60,000.

On the RHS of Table 2, the time pattern or lag structure to past employment is presented. Only the direct effects are presented; the indirect effects of increased employment at $\ell=6$ include

[7] For example a 1% permanent increase in own industry employment in surrounding counties (see Table 3 and discussion below for details on this conceptual experiment) for machinery, electrical machinery, primary metals, transport equipment and instruments, results over a six year period in percentage changes in own industry county employment of 0, -.02, -.01, -.04, and -.12, respectively. Over the years initial small negative impacts (competition?) offset small later positive impacts (externalities), with net minimal negative impacts. In terms of statistical significance, electrical machinery and primary metals are strongest, but those net impacts (-.02 and -.01) are truly minimal!

Table 2

Prior Own Industry Employment

	direct impact (I1) at:*			direct lagged impacts (at 3000)					
	300	3000	30000	I1	I2	I3	I4	I5	I6
machinery	-.06	.04	.13	.04	.02	.08	.09	(-.10)	(.01)
electrical machinery	0	.24	.49	.24	.15	.01	.03	.02	0
primary metals	-.15	-.01	.14	-.01	.05	.01	.11	(0)	.03
transport equip.	-.08	0	.08	(0)	-.02	.03	.07	(-.01)	(.02)
instruments*	-.06	.06	.17	(-.11)	-.06	-.01	.01	.10	.10

* Since one of the coefficients at I1 for instruments is not close to being statistically significant, I present the results for I2.

the impacts on employment levels in intervening years and their own impact on today. All impacts are evaluated at the typical mean employment of 3,000 (or 8 in logs). Given we have a quadratic estimated by regressing first differences on first differences, while there may be some considerable numerical imprecision, it is surprising how strong the results are. Table 2 indicates that (1) for all industries significant effects persist to ℓ=4; (ii) for only one industry is the largest effect at ℓ=1; (iii) for all industries but one effects are either insignificant or small at ℓ=5 and ℓ=6; (iv) for the high-tech instruments the strongest effects persist at ℓ=5 and ℓ=6; and (v) for the other partially high-tech industry, electrical machinery, while effects at ℓ=5 and ℓ=6 are small (respectively .02 and 0 at employment of 3,000) they are statistically significant. For higher employment levels such as 30,000 for electrical machinery at ℓ=5 and ℓ=6, the direct lagged impacts would be more noticeable (.05 and .04 respectively).

In summary, for standard industries, MAR effects persist from conditions up to five years ago, with typically a strong effect remaining at ℓ=4. For more high tech industries significant effects persist through to the maximum length of our lag structure, ℓ=6, or seven years prior. Even after differencing out fixed/random effects, industrial environment conditions from seven years ago affect employment directly today in high tech industries.

The lag structure itself is of particular interest. Because there is a quadratic formulation (in contrast to other variables analyzed below), it is not easy to directly summarize the lag structure for past own industry employment, beyond what I have already stated. However there is one critical result. For most industries initial impacts are smaller than impacts from four or five years ago (ℓ=3 and ℓ=4). How could this be the case? For example, suppose we utilize an accumulation of (local) knowledge interpretation. Presumably knowledge depreciates. Thus we might think that an increase in knowledge (represented by an increase in own industry employment) from four years ago ought then to be less beneficial than a ceteris paribus increase from one year ago. The lag structures in Table 2 suggest otherwise. That means there is a lag process within the county and metro area in terms of either the transmission of knowledge across firms or the maturing of ideas before firms act upon information. Not only may it take an entrepreneur some months to learn of a specific new "development" locally (information received or created and revealed by another local firm), but it may take many more months before a firm acts upon new information, while it observes and determines that such information is useful. Note this is all in the realm of dynamic externalities. Firms in locations without this historical accumulation of knowledge have no such information to filter and act upon and are disadvantaged at those locations.

Finally I note that direct plus indirect effects are typically larger than direct effects. An increase in employment at, say, ℓ=5 not only directly impacts employment today, but also indirectly impacts today's employment by stimulating employment levels at ℓ=1, ... 4, which in turn each affect employment at intervening time periods.[8] For example, the total impacts of an increase in employment from 3,000 at ℓ=5 in electrical machinery and in instruments respectively are <u>.06</u> (vs. a direct effect of .02) and <u>.09</u> (vs. a direct effect of .10). For instruments indirect effects are minimal and negative, (given the negative coefficients at ℓ=1 and ℓ=3), whereas for machinery they are relatively large and positive.

<u>The External Environment</u>. Jacobs effects, or related economies of urbanization, derive from the diversity and related scale of the urban environment which surrounds an industry. Diversity in that environment, for example, enhances knowledge accumulation as producers in an industry can draw upon a greater diversity of ideas from other industries, through interacting with a greater diversity of local suppliers (including suppliers of labor, entrepreneurial and R&D

8 If we denote by a_1, ... a_5 as the direct net (in a quadratic) effects on employment today of increases in past employment evaluated at 3,000, the total impact of an employment increase at l=5 is $a_5 + 2a_1{*}a_4 + a_3(2a_2 + 3a_1^2) + a_2(4a_1^3 + 3a_1a_2) + a_1^5$. This is an approximation given everywhere we evaluate employment effects at 3,000 workers.

services) and purchasers. To measure diversity of the surrounding environment, I calculated a variety of Hirschman-Herfindahl indices (HHI) for employment in other 2-digit industries. While there is no completely satisfactory way to measure diversity, the HHI has the virtue of being a recognized, standard measure. In terms of different HHI's, I calculated indices for all-manufacturing and for all industries, for both the county and the MSA/PMSA the county is part of. For example, for manufacturing in the metro area, the HHI is the sum of squared shares of each 2-digit manufacturing industry (other than the own industry) in total (all other) manufacturing employment in the metro area. An HHI index measures lack of diversity, in the environment surrounding the own industry. For 19 other 2-digit manufacturing industries, HHI takes a maximum of 1 if remaining employment is concentrated in just 1 other industry and a minimum of .0526 if it is uniformly distributed across all 19 other industries.

Given that the different HHI measures are strongly correlated, that they only imperfectly measure what we have in mind conceptually, and that results are based on annual changes in HHI (which are typically quite small), there is a limit to how well I can sort out what is the relevant diversity measure. Ultimately I settled on diversity in all other industries for the MSA the county is part of. This corresponds most closely to Jacobs notions of urbanization -- the benefits of greater diversity of the entire metro area for an industry. Results with an HHI just for manufacturing for the MSA are very similar. Results for HHI measured at the MSA level did seem to dominate measures at the county level (in terms of what generally survives as significant coefficients especially beyond $\ell=3$). Results with HHI's for both manufacturing and all industries or for both MSA and county levels generally yielded one set of signs for one measure and the opposite set for the other, but no consistent pattern across industries and suggested significant multicollinearity.

For the HHI for all other industries in the MSA, in Table A2 in the Appendix, two thirds of all coefficients have the expected negative sign. An increase in HHI is a decrease in diversity which under a Jacobs hypothesis hurts own industry employment. It is clear from Table A2 that coefficients on our first differenced measure bounce around. To get a better view of overall effects and the lag structure I report two other results. First in Table 3 I sum coefficients over the lag structure, which tells us the direct impact on current own industry employment of a permanent increase in HHI. The effects are large. A one standard deviation permanent increase in HHI decreases current own industry employment by 24-74%, depending on the industry. Diversity appears most important for high tech electrical machinery and for primary metals where surviving employment has been focused in newer non-ferrous, ("space age") alloys.

For the lag structure from Table A2, for machinery, primary metals, transport equipment and instruments, the negative HHI effects appear to peak at ℓ=3 or 4, often starting positive and ending positive or small. Only for the high tech industries of instruments and electrical machinery are there strong negative effects at both ℓ=5 and ℓ=6. For other industries effects at ℓ=5 or ℓ=6 are small, of perverse sign, and/or statistically insignificant. Given the lack of smooth movement from one lag to the next in coefficients, to get a better sense of the lag structure, I also imposed and estimated an Almon lag structure for this variable. Results are reported in Table A6, for either a cubic or quadratic lag structure for the industry. A cubic was chosen if adding the cubic term improved the results. For a cubic the impact at any lag is

$$\delta_{\ell} = \alpha_0 + \alpha_1 S + \alpha_2 S^2 + \alpha_3 S^3 \quad \ell, S = 0, \dots 6. \tag{5}$$

The α's are reported in Table A6[9], part a. Of the 18 reported coefficients, all are either

---

[9] For an Almon lag the RHS variables take the form

$$\alpha_0 \sum_{\ell=1}^{m} z_{\ell} + \alpha_1 \sum_{\ell=0}^{m} \ell z_{\ell} + \alpha_2 \sum_{\ell=0}^{m} \ell^2 z_{\ell} + \alpha_3 \sum_{\ell=0}^{m} \ell^3 z_{\ell}$$

Given RHS variables are first differences, the explanatory variables corresponding to this Almon lag structure for

Table 3

IMPACT OF A ONE-TIME PERMANENT INCREASE

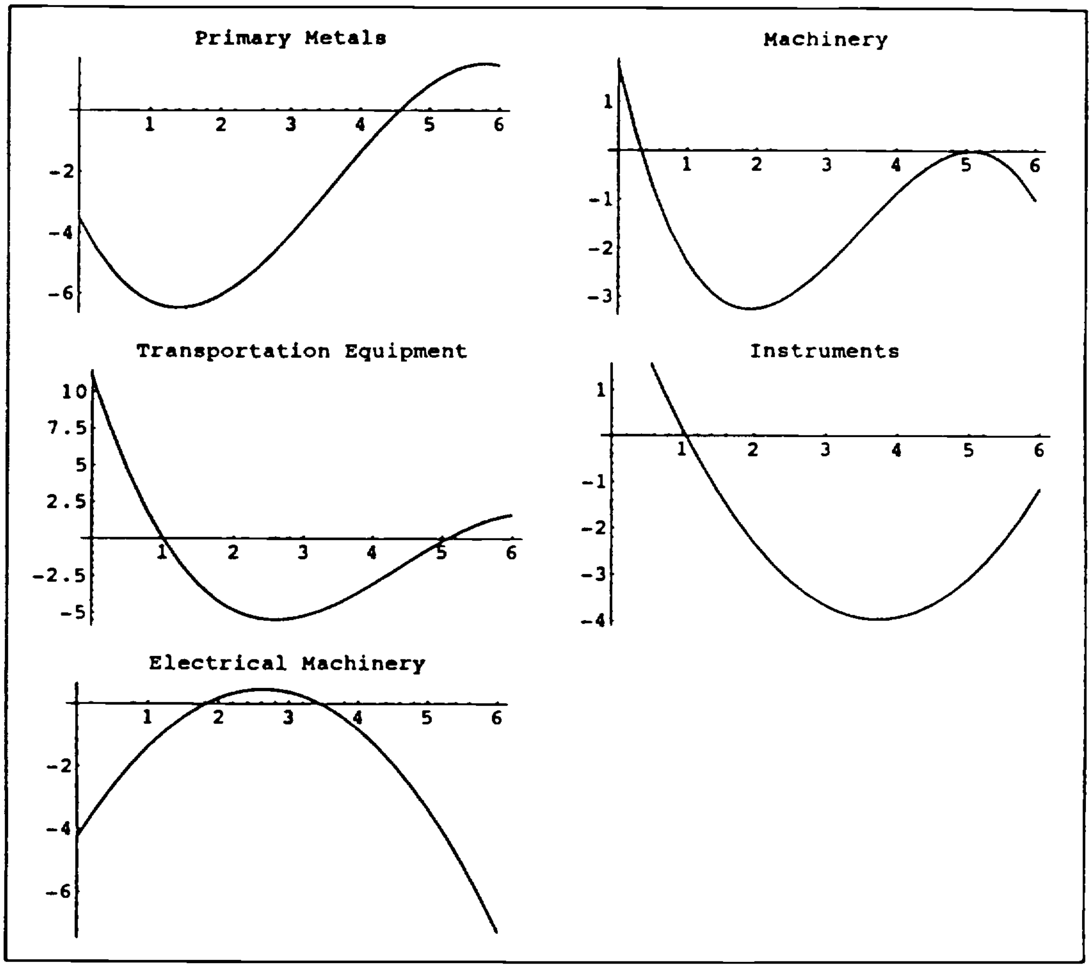
$$\sum_{l=0}^{6} \delta_l$$

	HHI (external)		HHI (internal)	
	Σδ	1 s.d. increase	Σδ	1 s.d. increase
machinery	-6.5	-.27	-.59	-.20
electrical mach.	-16.5	-.66	-1.0	-.45
primary metals	-18.4	-.74	-.08	-.04
transportation equip.	-6.0	-.24	-.66	-.46
instruments	-8.0	-.32	.08	.04

	Total (other) MSA employ	Wages
machinery	.90	-.21
elect. mach.	.38	-.63
primary metals	1.83	-.95
transportation equip.	.83	-.76
instruments	1.02	(1.30)

Figure 1 HHI: EXTERNAL

(δ_ℓ for time ℓ=0, ... 6 from Almon Lags)



statistically significant or have t-statistics over 1.9. These coefficients give a very clear lag structure to this measure of diversity external to the industry. This lag structure is graphed in Figure 1.

Figure 1 plots the δ_ℓ's from ℓ=0 to 6 in equation (5) for the α's reported in Table A6. I start with machinery, primary metals, transport equipment, and instruments. The pattern is similar for these industries and is U-shaped. Starting today (ℓ=0) HHI effects are either small negative or positive; they rapidly become (more) negative peaking in negative values somewhere between ℓ=1 and ℓ=4; then the effects diminish (become less negative) so that at ℓ=5 and ℓ=6 they typically have relatively small positive or negative values. In Figure 1, for these four industries the biggest negative value at ℓ=6 is for high tech instruments (-1.24), the other negative value is for machinery which fluctuates around zero after ℓ=3. If we sum the lagged effects at the discrete time points 0, 1 ... 6, we get sums similar to those reported in Table 3. In summary, in Figure 1, for those four industries diversity matters, starting small, peaking somewhere between ℓ=1 and ℓ=4, and then diminishing at six or seven years out. Only for instruments does a strong negative effect seem to remain at ℓ=6.

For electrical machinery, consistent with the original lagged coefficients in Table 2A, the pattern is different - an inverted-U. Effects start negative, diminish, and then accelerate (become more negative) with time. As noted earlier, seven years out, for the two high tech industries only, the historical external industrial environment significant affects today.

Similar to the effects of own industry concentration, the effects of diversity start small or even positive, and then peak at years further back. The reasoning is also similar, as for own industry concentration. An increase in local diversity (decline in HHI) improves the local stock

---

1985 are for example $\alpha_0 A + \alpha_1 B + \alpha_2 C + \alpha_3 D$ where
A= HHI85 - HHI78, B= HHI84 + ... + HHI79 - 6HHI78
C= HHI84 + 3HHI83 + 5HHI82 + ... + 11HHI79 - 36HHI78, and
D= HHI84 + 7 HHI83 + 19HHI82 + ... + 91 HHI79 - 216HHI78

of knowledge but there is a diffusion process where it takes time for knowledge to spread and an "aging" process where firms want to observe if new information is good. So the effect on current employment of knowledge increases is greatest from knowledge changes of several years ago.

In looking at Figure 1, it is also natural to ask why there are positive effects (for transportation, instruments and machinery) of HHI increases (diversity declines) at $\ell=0$. I think this occurs because at $\ell=0$ (only) we are inadvertently picking up competitive effects of other industries on the own industry. An increase in today's HHI probably occurs because some other industries contract in the MSA, perhaps helping own industry expansion.

<u>**Internal Diversity**</u>. Apart from diversity external to an industry, internal diversity may also matter. For any sub-sector, interaction with a diversity of other sub-sectors may be informationally important and greater diversity within an industry indicates a greater variety of intra-industry specialized functions are being performed. The inclusion or emission of this variable has little impact on coefficients of other variables, so its inclusion creates no problems. I experimented with a measure of internal diversity at the metro area level which is the sum of squared employment shares of the 3-digit sub-categories in corresponding 2-digit employment.[10] Since it is own industry employment, I start the lag structure at $l=1$, not today ($l=0$).

Results are reported in Table A5. If greater diversity enhances productivity, an increase in the internal HHI will decrease it. Indeed in Table A5, three quarters of the coefficients are negative. In Table 3, the coefficients summed over the lag structure are negative except for instruments and quite large for machinery, electrical machinery, primary metals, and transportation. This sum indicates the direct impact on current employment of a permanent

[10] Again, consistent with Jacbos, I measure HHI at the metro area level. Also measures at the county level were problematical in some counties because three-digit employments were not reported. Moreover reporting at the metro area level tends to be more accurate (fewer numbers reported in interval form and then massaged to get point estimates).

increase in HHI.

This internal HHI measure is the only variable in equation (3) where expected (in this case negative) effects persist for all industries at both $\ell=5$ and $\ell=6$, although most coefficients at that lag time are not statistically significant. Eyeballing the coefficients in Table A5 indicates a variable lag structure across industries. To get a better sense of the lag structure I also estimated an Almon lag structure for internal diversity. The results are reported in Table A6, part b. Results in Tables A5 and A6 are reasonably similar and Almon lagged coefficients (δ_ℓ from equation (5)) sum up to similar numbers as those in Table 3. Note however the coefficients of the Almon lag structure are often statistically weak. Only for the high tech electrical machinery and instruments, does the Almon lag structure seem strong enough to graph. In Figure 2 there is a convex relationship where negative effects fizzle out by $\ell=5$ or $\ell=6$.

__Interpretation and Summary for Industrial Environment Variables__. Past own industry employment and locational diversity matter for capital goods industries. Employment composition and intensity from 6 or 7 years ago appear to matter most for the more high tech industries of instruments and electrical machinery. Only for these industries do strong effects of past own industry employment persist at both $\ell=5$ and $\ell=6$. These are also the only two industries where external HHI remains strongly negative and significant for both $\ell=5$ and $\ell=6$ (Table A2). For other industries these two effects tend to evaporate beyond $\ell=4$.

The results generally indicate that both own industry concentration in a county and diversity in employment of the surrounding metro area are important for an industry. That suggests a tension for an MSA in maintaining a strong industrial environment. Cities, counties and metro areas are all highly specialized in manufacturing activities, reflecting the benefits of concentration. But having an otherwise diversified metro area employment base is also important.

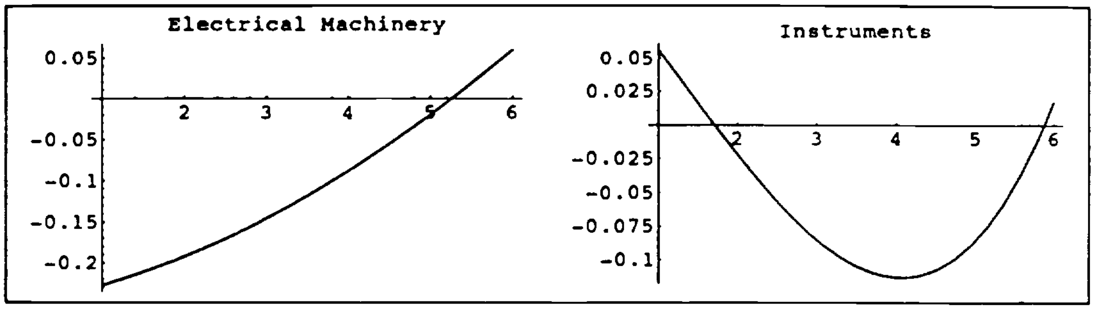
In terms of lag structures my basic interpretation of these results is that history matters. Industrial conditions from 6 or 7 years ago can effect strongly productivity and hence employment today, even after removing persistence in relationships due to fixed/random effects. By raising productivity and profits today, history dictates a larger scale of industry operation (employment) today in the locality. Moreover direct effects from four or five years ago are more important than effects from last year, suggesting an aging, maturing and/or transmission mechanism.

Overall, the interpretation of the results is that there is strong evidence of dynamic externalities. However there is a caveat. Suppose history doesn't matter and there are no dynamic externalities. We still might get a lag structure in equation (2), because of expectations formation or because of delays in firms responding (by expanding employment) to improved static conditions. Now such lag structures typically look very different than those in Figures 1 and 2. For lagged responses to static conditions, we typically expect big initial impacts which rapidly taper off. We can test for this by picking current market variables where we expect no dynamic externalities and investigating whether they have any lag to their response and, if so, whether it is markedly different from Figures 1 and 2.[11] If it is, we may reasonably infer that Figures 1 and 2 deal with non-static phenomena. I turn to this next.

[11] Data on productivity might allow us to directly sort this out, since lagged variables then would only matter if they reflected dynamic externalities, but such data are unavailable in CBP data and strongly censored in the Census of Manufacturing.

Figure 2 **HHI: INTERNAL**

(δ_ℓ for time ℓ=1, ... 6 from Almon Lags)



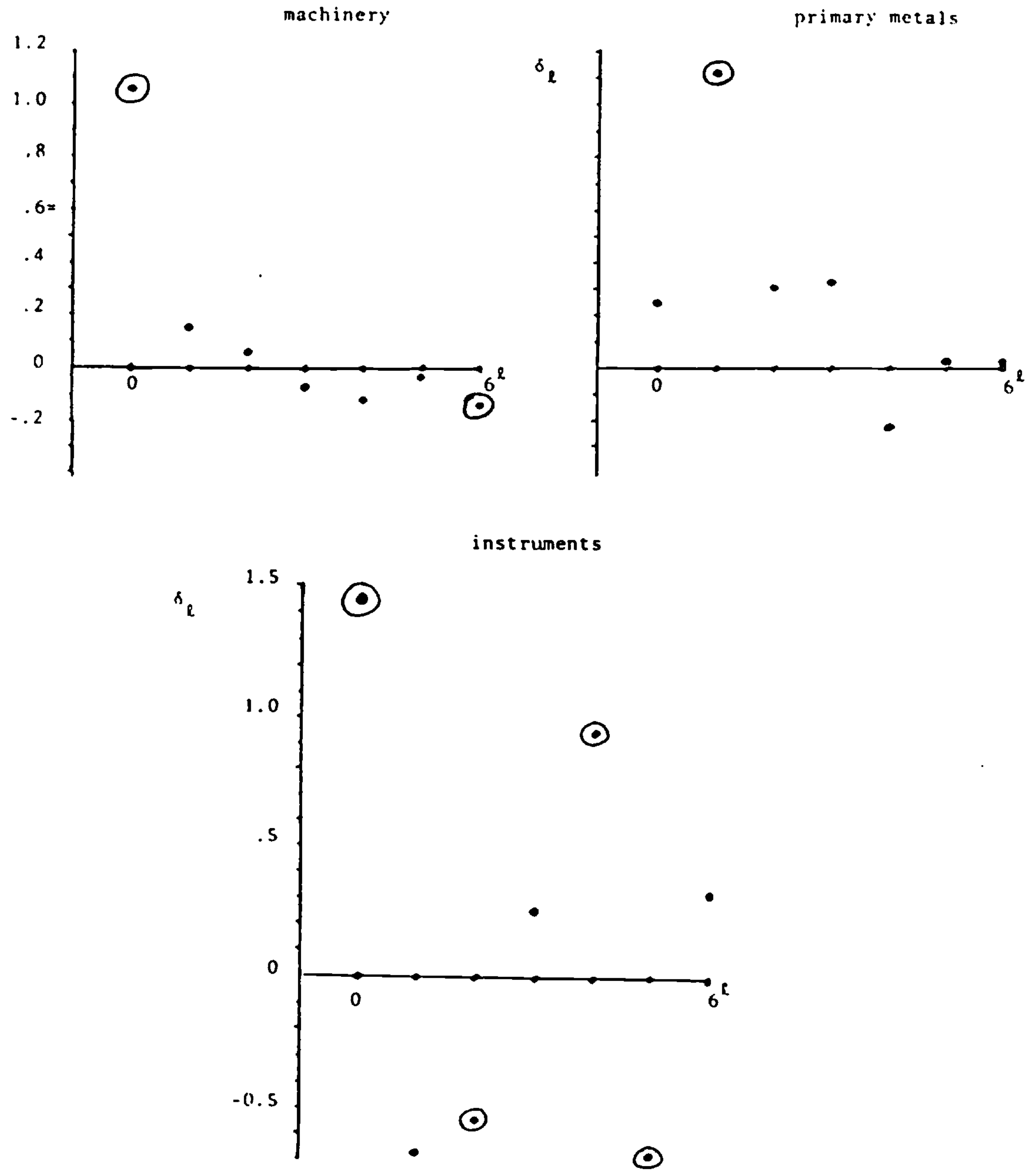
Market Conditions

I have two primary variables representing market conditions -- the scale of the local metro area market representing local demand for these industrial products and the wage rate. We start with the scale of demand measure.

Market Scale. I measured market scale variously by state population, employment in the metro area, and employment in the county. From experimentation it seemed that including more than one measure wasn't helpful and metro area scale dominated the other measures in terms of consistency. State population is poorly measured on a year to year basis and our qualitative results are not sensitive to the choice of metro versus county level employment.

The results for metro area employment in Table A3 are quite striking. For machinery, electrical machinery, primary metals, and instruments, at either $\ell=0$ or $\ell=1$, there is a large initial scale effect, where a 1% increase in last year's metro area employment (other than the own industry) increases own industry employment today by 1.05 to 1.45%. Thereafter for machinery and primary metals, any remaining effects are small. We graph the actual coefficients in Figure 3 for machinery, primary metals, and instruments. As Figure 3 strongly suggests the reaction to an increase in scale of local market demand for machinery and primary metals is essentially immediate (at $\ell=0$ or $\ell=1$) and thereafter basically disappears. There is no long adjustment lag to current market conditions. For instruments some individual effects beyond $\ell=0$ or 1 are large but they fluctuate in sign, with no consistency to the pattern and they are dominated by the initial impact. The pattern for electrical machinery is similar to that for instruments in Figure 3. The final industry, transportation, doesn't contradict the others. It's just that metro area scale effects for it are smaller, which is not surprising since its products tend to be mostly exported from cities.

Figure 3. MARKET SCALE: METRO EMPLOYMENT



Further evaluating the lag structure, Table 4 part a, row two indicates, the sums of the remaining coefficients (apart from the largest one, given in the first row) are collectively significantly smaller than the initial major impact for all industries. Further there is no consistency in the sign of the residual impact. But perhaps most telling is the fact that, in strong contrast to the industrial environment variables, for market scale, for four of the industries, no industry has two significant coefficients of the same sign for the four years from $\ell=3$ to $\ell=6$. Moreover for three of the five industries, pseudo F-tests indicate that the value of the objective function doesn't change significantly if we truncate the lag structure for this one variable at $\ell=1$.

In summary, there appears to be little evidence of a consistent lag structure of firms adjusting to the (static) market condition of altered market size. There is only consistent evidence of a large initial ($\ell=0$, 1) positive impact. This is in contrast to the impact of industrial environment conditions, where conditions from $\ell=3$ to $\ell=6$ matter in a very strong consistent fashion, suggesting the presence of dynamic externalities.

Finally if we sum coefficients in Table 3 we get the impact of a permanent increase in metro area size. There are sizable long run impacts for transport equipment, as well as machinery, primary metals and instruments, with elasticities averaging over one.

Wages. Wages are the wages in the county for workers in all other industries -- total payroll divided by total workers. Despite the fact that we are regressing first differences in employment on first differences in wages, I do indeed find fairly consistent negative wage impacts in four of the five industries. Only instruments represents a perverse case, perhaps reflecting the absence of a separate wage variable for highly skilled workers. Elasticities of the direct effect of a permanent increase in wages on employment for primary metals, electrical machinery, and transport equipment in Table 3 are reasonably near -1, ranging from -.63 to -.95. For machinery this magnitude is much smaller and individual coefficients are also small.

Table 4

(a) Metro Area Employment

	Machinery	Electrical Machinery	Primary Metals	Transportation Equipment	Instruments
highest coefficient in l0 or l1	1.05	1.05	1.11	(.51)	1 .45
sum of remaining coefficients	-.15	-.67	.72	.32	-.43

(b) Wages

	Machinery	Electrical Machinery	Primary Metals	Transportation Equipment
$\sum_{l=0}^{2}\delta_l$: early sum	.02	-.17	-.29	.09
$\sum_{l=4}^{6}\delta_l$: late sum	-.21	-.35	-.57	-.83

For wages, in contrast to metro area employment (market scale), the lag structure is different. As Table 4, part b reveals immediate impacts are minimal and significant employment adjustments today can remain from wages changes six or seven years ago. Impacts of wage changes within the last three years are much smaller than impacts from four to seven years ago. Wage, like market scale is a contemporaneous market condition; and, if there is a lag structure to its impact, couldn't that be evidence of long lags in market adjustment to (static) conditions? However there is a fundamental difference with wages, suggesting that these effects are institutionally, not market driven.

Responses to changes in market scale or static industrial environment conditions can be immediate as new firms come into business or existing firms hire more workers. Responses to changes in local market wages are delayed by union contracts, which fix wages facing firms within the own industry for a discrete time period. That is, for heavily unionized industries, effective wages do not respond to changes in the going wage rate in a locality. The response comes when contracts are renegotiated and those negotiations are themselves based on the prior recent history of wages. So a firm's effective wages may be based on a contract negotiated three years ago, and those negotiations will have been based on local wage conditions for the several years prior to that time. Moreover part of the response to changes in effective wages is to adjust capital-labor ratios, which again is delayed by irreversible capital investments and technology turnover. It is instructive to note that by far the largest delayed impacts in Tables A4 and 4 are for primary metals and transport equipment. Relative to our other industries, these are industries which are traditionally highly unionized and whose production is characterized by large investments in fixed plant.

Results for Time Invariant Variables

In the previous section I presented results for equation (3), based on estimating first differenced equations. The first differencing eliminates the time invariant variables in equation (2), which do vary cross-sectionally. To recover these, we estimate equation (4) by OLS, where the dependent variable is B_i, the estimated average residual from the reinstated level equation (2). Results are reported in Table 5.

Two versions of equation (4) are presented in Table 5. The first column for each industry reports truly exogenous variables -- land area, regional dummies, and a dummy for coastal (ocean or Great Lakes) location. The second column adds in other explanatory variables such as a dummy for being in a multi-county MSA, 1980 college educational attainment, and median value of owner-occupied housing in 1980 (as a proxy for land prices). These additional variables could be correlated with the random effect f_j now in the error term of equation (5), if such a term is county or MSA-specific, rather than just industry-county specific.[12] To conserve on space, I discuss only the second column results.

The results are similar across industries. For the truly exogenous variables, in general being in a coastal MSA with access to shipping and different climatic conditions helps employment, as does having a larger land area and hence less congested conditions. Compared to the West (the basic constant term), the Northeast still has higher employment levels in machinery and electrical machinery, and the mid-West also has higher machinery employment. Despite relative growth in the 1970's and 1980's the South remains significantly less

[12] Again, there is an issue in evaluating whether standard error estimates are unbiased, given possible intra-metro area correlation in the random effects. Following the procedure in footnote 2, there appears to be more of a problem for our random effects than for the differenced contemporaneous error terms. For industry 35, a hypothesis of equal means for the error terms grouped by metro area is decisively rejected. For industry 36, we accept the hypothesis of equal means, but just.

Table 5

	Machinery		Electrical Machinery		Primary Metals		Transport Equipment		Instruments	
Constant	-3.92* (.908)	4.32 (5.56)	11.04* (.730)	10.01* (4.62)	-13.77* (1.55)	34.63* (7.62)	-3.07* (.997)	25.91 (6.26)	-3.94* (1.04)	7.99 (7.38)
Dummy: MSA on coast	.328* (.155)	.283 (.154)	.336* (.122)	.290* (.127)	-.330 (.206)	.164 (.197)	.541* (.164)	.800* (.171)	.408* (.186)	.540* (.198)
Region NE	.710 (.400)	1.13* (.428)	.340 (.309)	.691* (.344)	.706 (.556)	-.578 (.556)	-.059 (.414)	-.787 (.452)	.580 (.442)	.619 (.511)
Region Mid-West	.588 (.363)	1.16* (.392)	-.464 (.282)	-.028 (.325)	1.10* (.518)	-.124 (.538)	.404 (.371)	-.256 (.427)	-.368 (.419)	-.304 (.496)
Region South	-.903* (.352)	-.342 (.407)	-.596* (.277)	-.169 (.336)	-.073 (.509)	-1.38* (.566)	-.246 (.354)	-1.12* (.437)	-.745 (.413)	-.815 (.521)
ln (land area)	.674* (.117)	.470* (.116)	-.120 (.094)	-.165 (.097)	.693* (.174)	.174 (.161)	.322* (.129)	.125 (.132)	.427* (.132)	.400* (.139)
dummy: multi-county MSA		-1.77* (.234)		-.653* (.187)		-2.65* (.302)		-.829* (.246)		-.794* (.285)
% adults w/ at least college education		.095* (.019)		.055* (.016)		.051 (.028)		.076* (.023)		.094* (.025)
ln (median housing value)		-.705 (.512)		.050 (.425)		-4.01* (.703)		-2.59* (.579)		-1.20 (.682)
dummy: state right-to-work law	-.128 (.306)		.133 (.251)		-.295 (.496)		.042 (.350)		.131 (.400)	
1983 state corporate tax rate	.038 (.065)		-.103* (.049)		-.036 (.091)		-.024 (.070)		-.024 (.080)	
R^2	.14	.26	.06	.13	.06	.27	.05	.12	.09	.14
N	674	674	508	508	454	454	509	509	402	402

industrialized in capital goods employment. These regional dummies may proxy for regional resource endowment considerations, regional public infrastructure investments geared to specific industries, and cultural and institutional factors interacting with long term labor force characteristics.

For the other three variables, results were very similar across industries. Multi-county MSA's are less involved in capital goods production, being more oriented to service activities. Counties with higher quality labor forces are more attractive to industries, and higher land prices hurt capital goods employment (which are relatively land intensive).

Finally we experimented in the bottom panel in separate regressions with two state level variables. The 1983 state corporate tax rate has generally the expected negative sign but is only significant for one industry. The dummy for state right-to-work laws has no significant nor consistent impact. This is surprising since state right-to-work laws create a labor market environment favorable to firms. But states may only pass such laws if they have otherwise poor inherent conditions for attracting manufacturing.

Conclusions

To maintain strength in a particular industry a county wants concentrations of employment in that industry, yet it also wants a surrounding diverse industrial base. Diversity tends to raise productivity and hence employment in a city's particular concentration of production and export activity.

In this process history is critical. Increased concentrations of own industry activity appear to affect employment levels for five years afterwards and longer for high tech industries. For diversity measures, affects appear to persist beyond the seven year horizon examined in this paper for high tech industries. Given the rapid adjustment to contemporaneous market scale or demand, these long lags and the lag pattern suggest a presence of dynamic externalities. Conditions four or more years ago typically have a greater direct impact than conditions last year, suggesting the presence of an aging, maturing, and/or transmission mechanism..

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.

Becker, G.S. and K.M. Murphy (1992), "The Division of Labor, Coordination Costs and Knowledge," Quarterly Journal of Economics, 107, 1137-1160.

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.

Carlton, D. (1983), "The Location and Employment of New Firms," Review of Economics and Statistics, 65, 440-449.

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

Dunne, T., M.J. Roberts and L. Samuelson (1990), "The Growth and Failure of U.S. Manufacturing Plants," Quarterly Journal of Economics.

Frost, J. (1994), "**An Examination of Levin and Lin's Unit Root Test for Panel Data,**" Brown University mimeo.

Fujita, M. and H. Ogawa (1982), "Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations," Regional Science and Urban Economics, 16, 161-196.

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.

Hayashi, F. (1992), "Comment," Journal of Business and Economic Statistics, 10, 15-17.

Helsley, R.W. and W.C. Strange (1990), "Matching and Agglomeration Economies in a System of Cities," Regional Science and Urban Economics, 20, 189-212.

Henderson, J.V. (1994), "An Econometric Model of Industrial Location," Journal of Urban Economics.

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 (1991), (eds.) 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, 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.

Jovanovic, B. (1982), "Selection and the Evolution of Industry," Econometrica. 649-670.

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 (92-23).

Miracky, W.F. (1992), "Technological Spillovers, the Product Cycle and Regional Growth," MIT mimeo (November).

Table A1

Lagged Own Industry Employment

	Machinery		Electrical Machinery		Primary Metals	
	ln (emp)	(ln (emp))^2	ln (emp)	(ln (emp))^2	ln (emp)	(ln (emp))^2
I1	-.301* (.090)	.021* (.0083)	-.605* (.075)	.053* (.0089)	-.501* (.046)	.031* (.0062)
I2	-.145* (.030)	.010* (.0028)	-.378* (.038)	.033* (.0040)	-.394 (.038)	.028* (.0039)
I3	-.226* (.027)	.019* (.0027)	-.149* (.037)	.0097* (.0035)	-.093 (.028)	.0064* (.0030)
I4	-.196* (.030)	.018* (.0028)	-.102* (.031)	.0085* (.0029)	-.276* (.028)*	.024* (.0029)
I5	.034 (.027)	-.0086* (.0027)	-.071* (.029)	.0059* (.0026)	-.037 (.027)	.0022 (.0032)
I6	.0036 (.019)	.00045 (.0022)	-.114* (.025)	.0070* (.0027)	-.059* (.020)	.0058* (.0026)

Table A1 (Continued)

Lagged Own Industry Employment

	Transport Equipment		Instruments	
	ln (emp)	(ln(emp))^2	ln (emp)	(ln (emp))^2
l1	-.288*	.018	-.260*	.094
	(.077)	(.0095)	(.062)	(.0079)
l2	-.136*	.0073*	-.361*	.026*
	(.029)	(.0033)	(.028)	(.0031)
l3	-.104*	.0086*	-.129*	.0072*
	(.027)	(.0029)	(.024)	(.0026)
l4	-.142*	.013*	-.132*	.0090*
	(.027)	(.0030)	(.022)	(.0027)
l5	.022	-.0017	-.203*	.019*
	(.025)	(.0028)	(.026)	(.0029)
l6	-.018	.0022	-.158*	.016*
	(.024)	(.0025)	(.024)	(.0028)

Table A2

External Hirschman-Herfindahl Index

	Machinery	Electrical Machinery	Primary Metals
I0	.685	-2.28	-9.32*
	(1.28)	(2.50)	(2.21)
I1	-1.39	-2.67	2 .18
	(2.00)	(2.54)	(2.54)
I2	-.829	-2.73	-7.03*
	(.969)	(1.55)	(1.39)
I3	-4.86*	2.52*	-10.26*
	(1.14)	(1.24)	(1.47)
I4	-.622	.980	-.117
	(.950)	(1.27)	(1.70)
I5	1.79	-5.55*	6.10*
	(.992)	(1.88)	(1.43)
I6	-1.28*	-6.78*	.031
	(.504)	(1.52)	(.559)

Table A2 (Continued)

External Hirschman-Herfindahl Index

	Transport Equipment	Instruments
I0	8.97*	-.669
	(3.84)	(2.50)
I1	-7.41	4.18*
	(4.19)	(1.99)
I2	-2.08	.514
	(2.00)	(1.29)
I3	-8.34*	-2.76
	(2.31)	(1.49)
I4	-2.20	-5.04*
	(1.90)	(1.33)
I5	3.48	-3.26*
	(2.30)	(1.32)
I6	1.59	-.942*
	(1.98)	(.761)

Table A3

Metro Area Employ (all other industries)

	Machinery	Electrical Machinery	Primary Metals
I0	1.05* (.251)	-.070 (.324)	.268 (.271)
I1	.160 (.280)	1.05* (.438)	1.11* (.345)
I2	.056 (.135)	-.638* (.206)	.304 (.199)
I3	-.081 (.124)	.040 (.171)	.333 (.186)
I4	-.126 (.110)	-.389* (.160)	-.215 (.166)
I5	-.022 (.025)	.882* (.193)	.019 (.071)
I6	-.139* (.025)	-.498* (.176)	.0079 (.026)

Table A3 (Continued)

Metro Area Employ (all other industries)

	Transport Equipment	Instruments
I0	.509	.145*
	(.428)	(.364)
I1	.168	-.669
	(.482)	(.396)
I2	.227	-.567*
	(.275)	(.199)
I3	.411*	.266
	(.185)	(.170)
I4	-.133	.938*
	(.202)	(.194)
I5	-.486*	-.695*
	(.217)	(.174)
I6	.129	.301
	(.238)	(.207)

Table A4

Wages

	Machinery	Electrical Machinery	Primary Metals
I0	.012	-.125	.011
	(.042)	(.073)	(.047)
I1	.070	-.214*	-.135*
	(.079)	(.075)	(.058)
I2	-.061	.174*	-.162
	(.042)	(.072)	(.086)
I3	-.016	-.113	-.083
	(.046)	(.065)	(.070)
I4	-.057	-.265*	-.113
	(.063)	(.074)	(.075)
I5	-.074	-.155	-.400*
	(.097)	(.005)	(.162)
I6	-.078*	.073	-.059
	(.018)	(.082)	(.157)

Table A4 (Continued)

Wages

	Transport Equipment	Instruments
I0	.054 (.079)	.031 (.095)
I1	.121 (.126)	.111 (.098)
I2	-.087 (.075)	.073 (.075)
I3	-.018 (.086)	-.072 (.096)
I4	-.041 (.112)	-.159 (.088)
I5	-.329 (.171)	1.16* (.14)
I6	-.462* (.146)	.155 (.088)

Table A5

Internal Hirschman-Herfindahl Index

	Machinery	Electrical Machinery	Primary Metals
I1	-.041* (.124)	-.266* (.092)	-.066 (.148)
I2	.080 (.074)	-.233* (.059)	-.026 (.081)
I3	-.159* (.055)	-.237* (.049)	.147* (.056)
I4	-.142* (.055)	-.160* (.062)	.025 (.058)
I5	-.220* (.074)	-.080 (.061)	-.099 (.068)
I6	-.106 (.061)	-.035 (.056)	-.056 (.065)

Table A5 (Continued)

Internal Hirschman-Herfindahl Index

	Transport Equipment	Instruments
I1	-.333* (.170)	.216* (.108)
I2	-.085 (.065)	.033 (.062)
I3	.093 (.070)	.098* (.048)
I4	-.090 (.070)	-.038 (.045)
I5	-.095 (.065)	-.172* (.051)
I6	-.147* (.069)	-.055 (.049)

Table A6

Almon Lag Structure for Diversity Indices

(a) External HHI

	Primary Metals	Machinery	Electrical Machinery	Transport Equipment	Instruments
α_0	-3.52 (1.84)	1.75* (1.63)	-4.28* (1.54)	11.54 (3.20)	3.68 (1.89)
α_1	-4.58 (2.35)	-5.98* (1.64)	3.59* (1.18)	-14.75 (4.36)	-4.09* (1.05)
α_2	2.03* (.839)	2.15* (.593)	-.681* (.186)	4.02* (1.61)	.547* (.125)
α_3	-.188* (.083)	-.205* (.060)		-.304 (1.66)	

(b) Internal HHI

	Primary Metals	Machinery	Electrical Machinery	Transport Equipment	Instruments
α_0	-.117 (.115)	-.059 (.111)	-.252* (.064)	.076 (.098)	.129 (.091)
α_1	.097 (.057)	-.086 (.052)	.019 (.020)	-.051 (.030)	-.063 (.041)
α_2	-.0011 (.0024)	-.0011 (.0027)	.0055* (.0023)	.0046 (.0035)	-.013* (.0028)
α_3	-.0037* (.0016)	.0025 (.0015)			.0034* (.0011)