Preliminary and Incomplete
Do not cite

Spatial Mobility and Geographic Concentration[1]

Mohammad Arzaghi and James C. Davis[2]

Prepared for the NBER/CRIW Conference:
Producer Dynamics: New Evidence from Micro Data

April 9, 2005

Abstract

This paper uses the Longitudinal Business Database to explore the geographic mobility of three digit SIC industries across cities, and the extent to which this mobility is linked to concentration, scale economies and industry characteristics. We find industries are quite heterogeneous in their degree of mobility, that many industry distributions across cities exhibit a fair degree of churn over time. For example, a city ranked near the top with regard to its share of employment in men's apparel has an estimated mean first passage time to the bottom of the distribution of seven decades. However, a top ranked steel city would be expected to fall to the bottom in twenty four decades. Surprisingly, industry mobility turns out to be independent of either concentration or scale economies. Our main finding is that technologically advancing industries are more mobile. As new technologies replace older ones, centers of agglomerated industry activity are more likely to change location.

[1] The research in this paper was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Boston Census Research Data Center (BRDC). Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed. We are very grateful to Vernon Henderson and Duncan Black for their many contributions, and to Gilles Duranton and Bill Kerr for helpful suggestions. The authors thank participants at the 2004 RSAI and 2005 WRSA conferences.
[2] Respectively Brown University, Providence RI, marzaghi@brown.edu and U.S. Census Bureau, Boston Census Research Data Center, Cambridge, MA jdavis@nber.org.

1. Introduction

It has been widely documented that individual industries are agglomerated.[3] We find these industry concentrations are not necessarily wed to particular city locations, that there exists a degree of mobility of industries over time. There is mobility even for the largest industries. For example, consider the ten largest manufacturing industries by employment in 1963. In Table 1 we list the largest city-employer for each industry, and the city's 1963 share of national employment in the industry. Of these ten largest industries, in only three cases is the largest city employer in 1963 still the largest city employer in 1997. In four cases the decline is only to the second spot, but in two cases it is to the fifth spot. The decline in shares of national employment can be dramatic, for example, from 16% for communications equipment in 1963 in Los Angeles to 4% by 1997. Auto production has dispersed from Detroit; the city keeps its top rank but its share of national production falls from 21% to 12% by 1997. A similar comment applies to aircraft in LA (but not missiles), except that LA loses its top rank to Seattle. LA also drops out of the top running in communications equipment and measuring instruments, with San Jose becoming dominant. The top city employer in 1997 for these industries is also given in the right most column of the table, and some of these top 1997 cities are ranked back in 1963 as low as, for example, 36th or 13th, climbing quite a long way in 34 years.

Beardsell and Henderson (1999) show for the computer industry that though the overall distribution of industry employment across cities is stable, considerable churn exists within the distribution with some cities gaining share and others loosing it. Duranton (2004) further shows that though cities move slowly up and down the size distribution in regard to overall employment, this movement is indeed much quicker for individual industries. He embeds the

[3] See Rosenthal and Strange (2004) for a review.

Grossman and Helpman quality-ladder model of growth into an urban framework and shows new innovation can induce cities to win or loose industries.

The theoretical literature (Arthur 1990 and Rauch (1992)) suggests that concentration and mobility should be linked. While those papers focus on scale economies as driving concentration, their point is that concentrated industries will be immobile. When scale economies are driving concentration, firms can't easily defect from locations with high concentration of own industry employment to those with, for example, no own industry employment. The latter locations offer no existing scale benefits and hence low productivity. An exception (Rauch 1992) is that, if Henderson (1974) type land developers exist that facilitate en masse movement and, if scale economies are static in nature, then mobility is easier and not necessarily linked to concentration. Developers can offer contractual arrangements and timing to facilitate en masse and "instant" scale economies at a new location. If concentration is driven by access to natural resources (Ellison and Glaeser (1999)), mobility is limited to sites of discoveries of new deposits.

This work explores the extent of spatial mobility of industries. Are there significant differences in spatial mobility of industries across cities? How do we measure/characterize the degree of industry mobility? As implied by Arthur (1990) and Rauch (1992), is mobility related to geographic concentration? That is, are geographically concentrated industries, especially those subject to scale economies, less mobile than more ubiquitous industries? Do established locations tend to hold onto industries subject to scale economies?

If we think of mobility and concentration as outcomes, what industry characteristics – scale economies, capital to labor ratio, use of heavy materials, size, volatility, technology use, technology transition, etc. – determine mobility and concentration?

The paper is organized as follows. Section 2 introduces the data, followed by methods in Section 3 for the construction of measures. Section 4 presents results for mobility and concentration as well as for their underlying determinants.

2. Data

The industries examined are 3-digit SIC industries, every Census (fifth) year from 1977 to 1997, giving us five time periods, or four mobility transitions. For manufacturing industries, we are able to extend the time period further back to 1963 to include seven mobility transitions. The data cover 315 PMSA cities[4] and about 740 urban counties.

The data are from the Longitudinal Business Database (LBD), the Longitudinal Research Database (LRD) and the Standard Statistical Establishment List (SSEL), all data sets of the U.S. Census Bureau made available through Research Data Centers.[5] The SSEL is the master business register of the economic census, and includes for private firms the employment, industry and location of all U.S. establishments for all sectors of the economy.[6] The LBD links the register longitudinally, and through extensive name and address matching has built or repaired links for establishments across time. The LRD combines the eight Census of Manufacturers conducted since 1963 into a panel data set of plants. These data for the

[4] We use 1990 OMB definitions for Primary Metropolitan Statistical Areas (PMSA), however defining PMSA for New England using counties as opposed to townships, to use a consistent methodology to that of the rest of the country.
[5] See Jarmin and Miranda (2002) for information on the LBD, and McGuckin and Pascoe (1988) and Davis, Haltiwanger and Schuh (1996) for information on the LRD. Information on the Center for Economic Studies and the Census Research Data Centers is available at www.ces.census.gov and on the Boston RDC at www.nber.org/brdc.
[6] The SSEL draws from a variety of sources of data to maintain a universe of establishments in all sectors of the U.S. economy. Examples of these sources include previous economic census and surveys, administrative data from the Internal Revenue Service (who are required to file the firm name and address to receive a federal tax id), from the Bureau of Labor Statistics (providing some industry classification information for small firms), and the Census Company Organization Survey used to break up multi establishment firms into their component establishments. The SSEL was replaced by a new Business Register for the 2002 economic census.

continental U.S. are aggregated into city-industry observations by year using a common denominator set of urban counties and city definitions over time, and mapping years before 1992 from a 1972 SIC basis to the 1987 definitions.[7] We include all 3-digit SIC industries for manufacturing (includes all in 1-digit SIC 2 and 3), communications (in SIC 48), finance and insurance excluding agents (in SIC 60-63, 67), business services (in SIC 73), motion picture production (SIC 781), theatrical (SIC 792), health (in SIC 80), legal (SIC 811), universities (SIC 822) and engineering and management (in SIC 87).[8] We excluded extractive and transportation related industries, as well as industries such as retail, personal services and repair which are local demand driven.

Measures for characteristics of manufacturing industries were taken from a variety of sources. R&D expenditure shares are from the National Science Foundation (average for years of 1972-1980 where sufficient 2-digit SIC detail is presented), scientists and engineers from the Bureau of Labor Statistics (1999 occupations by 3-digit industry), total factor productivity from the CES-NBER Manufacturers Productivity Database (a weighted average to the 3-digit SIC industry, and averaged over years), and the capital to labor ratio from the LRD (an average over the years capital stock is available).

Instruments, used in the estimation of scale elasticities (discussed later), include amenities drawn mostly from the 1970 Census of Population and Housing. Specifically, we include measures of housing availability, share of units in buildings of ten or more units, the ratio of housing units to business establishments and the median residential rent for one bedroom

[7] Considerable attention was devoted to mapping the earlier years of data forward from the 1972 SIC to the 1987 SIC definitions for industries. At the 3-digit SIC level, this mapping is straight forward using the appendices of the OMB Standard Industrial Classification Manual, though some measurement error does remain where subset portions of industries are moved from within one SIC to within another. This issue is significantly mitigated by the use of 3-digit SIC definitions where the occurrence of irreconcilable changes based on available data is minimal. The problem is much more acute at the 4-digit SIC level, and as a result we did not perform our analysis at this lower level of industry definition.

apartments (this last using 1990 data). We also use local government expenditure for 1987 from the City and County Data Books, market potential and retail wages in 1930 from the decennial census (Census (2001)), and proximity (within 80 miles) to the ocean as in Rappaport and Sachs (2003).

3. Mobility and Concentration Measures

To measures mobility, we follow Eaton and Eckstein (1997), Beardsell and Henderson (1999), and Black and Henderson (2003), and characterize the evolution of the city size distribution as a stationary first order Markov process. For geographic concentration, we use a standard measure of the Ellison-Glaeser (1997) index, as well as the Gini index (Krugman (1991)). We discuss each in turn.

Mobility

We use transition matrix to analyze mobility of an industry across cities, imposing a homogeneous stationary first order Markov process (and testing for stationarity). We use a discrete cell division into 4 and 5 cells. The five cell division has cells with following shares, or percent of all cities: smallest 55% of city employers, then the next 25%, 10%, 7%, and 3%, corresponding to 174, 79, 31, 22, and 9 cities. We use a small high-end cell, since the top nine cities account for so much of national employment.[9]

---

[8] Some industries were out of scope of the early censuses, such as communications, finance and insurance. These industries were included in the SSEL at the time, though likely not as well verified by Census operations. We expect sources of potential measurement error to be minimal at the city level, and include them in the analysis.

[9] The 1997 population sums for each cell division group of cities are in order 31.1 for the 174 smallest cities, 41.0, 36.6, 48.9 and 48.8 million for the 9 largest cities.

A problem in constructing the cell division for the size distribution and transition process is that often more than 55% of cities have zero employment in an industry in a year. For those cities, we can't distinguish whether a city is in the bottom, or next to bottom cell. This problem increases dramatically as national employment falls below around 60,000. Therefore we impose criterion to be in the sample that industries are "large", employing over 60,000 in every Census year. This leaves 94 industries that have at least some cities in the bottom cell. That is still a considerable sample loss, so for much of the paper we combine the bottom two cells such that the bottom cell accounts for the 80% of small and zero city-employers in an industry. This gives us a 4-cell distribution, with cell shares of 80%, 10%, 7%, and 3%. Then we can examine transition processes for 104 industries.

In constructing the transition matrix, the bottom state cell is labeled 1; and the top is 5 (4). p_{jk} is probability of moving from state j to state k in the time interval (5 years). $p_{jk}(t)$ is estimated (MLE) as the total number of cities in state j in time $t-1$ that move to state k in t, divided by total number of cities in state j (which is fixed by the cell division). p_{jk} is estimated (MLE) as total number of cities that start in j in time $t-1$ that move to state j in time t summed over all t and divided by the number of cities in state j times the number of years. To test for stationarity, we test if the set of $p_{jk}(t)$ over all years and states differs from the p_{jk} -- i.e. are the transition probabilities changing over time. The test statistic is

$$\chi^2 = -2\ln\left\{\pi_t \pi_j \pi_k \left[\frac{\hat{p}_{jk}}{\hat{p}_{jk}(t)}\right]^{m_{jk}(t)}\right\} \tag{1}$$

where $m_{jk}(t)$ is the number of cities moving from j to k in t. Degrees of freedom are (T-1) K (K-1) where T is the number of years and K the number of cells.

We measure mobility using the trace of the transition matrix (under regularity conditions that our industries meet). As the trace goes up, mobility goes down. This is because higher diagonal cells mean a given city is less likely to move out of the current state. This has the advantage of providing a single measure to characterize features of the transition matrix. Mobility can also be measured by mean first passage times. The mean first passage time (mp) is the expected time to next visit state k from state j. That is,

$$mp_{jk} = \sum_{t=0}^{\infty} t \phi_{jk}^{\ t} \tag{2}$$

where $\phi_{jk}^{\ t}$ is probability a city in state j visits state k in t periods from now and

$$\phi_{jk}^{\ t} = [M^t]_{jk} - \sum_{s=0}^{t-1} \phi_{jk}^{\ s} [M^{t-s}]_{kk} \ \ \forall t \geq 1$$

$$\phi_{jk}^{\ 0} = 0; \ \phi_{jk}^{\ 1} = p_{jk} \equiv [M_{jk}]$$

where $[M^t]_{jk}$ is the jk element of M raised to power t. So $\phi_{jk}^{\ t}$ are defined recursively starting at t =0. In practice in calculating mp_{jk}, the summation in equation (2) is truncated when the cumulative probability of transiting from k to j equals 0.9999. There are many mp's. So picking any one to measure mobility is difficult. We rely on the trace to measure mobility and give illustrative mean first passage times from state 1 to state 5 and state 5 to state 1.

Concentration

To measure the concentration nationally, we look at the whole USA geography at the county level. For industry j, the Ellison-Glaeser (EG) index is

$$EG_j = \left[\sum_i (S_{ij} - S_i)^2 / (1 - \sum_i S_{ij}{}^2) - H \right] (1 - H) \quad (3)$$

where S_{ij} is county i share of national employment in industry j, S_i is county i share of all national employment, and H is the Hirschman-Herfindahl index of plant concentration in the industry nationally.

The basic part of the index in equation (3) is the $\sum_i (S_{ij} - S_i)^2$, which is the sum of squared shares of county i's share of industry j normalized by county i's share of all employment. This measures how S_{ij} deviates from S_i. If S_{ij} mimics S_i for all i, the industry is perfectly spread, or deconcentrated, just being proportional to county scale everywhere. As S_{ij} starts to deviate from S_i, these deviations are squared and summed across locations. In general because of squaring and because most industries have a few dominant locations with large shares of j, the top 3-4 counties dominate the index. That is EG_j is highly correlated with the statistic, the share of national employment in the top three locations. For example a regression of the time averaged EG_j against the time averaged share of the top three counties has an R^2 of .63. The other parts of the EG index are meant to control or standardize for the coarseness of the geographic units $(1 - \sum_i S_{ij}{}^2)$ and the possibility (in smaller industries) of large plant economies of scale driving concentration (hence the corrections with H).

An alternative way to represent concentration is to use Krugman's way to represent spatial Gini coefficient (normalized to lie between 0 and 1). This measure complements EG by capturing how industries are spread out in the lower part of the distribution. To calculate the Gini,

$$I_{ij} = S_{ij} / S_i$$

n

where I_{ij} is the ratio of county i's share of industry j employment to its share of national employment in all industries. We then rank the counties from highest I_{ij} to lowest I_{ij}. The Gini coefficient for industry j is the sum from the highest to the lowest ranked county of the difference between the cumulative share of industry j by the ranked countries and their cumulative share of all employment multiplied by 2. That is, if counties are ranked from 1 to n, in order of highest to lowest

$$\mathrm{Gini_j} = 2\left[\frac{1}{2}S_iS_{tj} + \sum_{t=2}^{n} S_t\left(\sum_{i=1}^{t-1} S_{ij} + \frac{1}{2}S_{tj}\right) - 1/2\right] \qquad (4)$$

Determinants of Mobility and Concentration

Mobility and concentration measures are outcomes. They are driven by the fundamentals of the industries for which they are measured. The fundamentals include the capital intensity, the heavy raw material intensity, the degree of external economies of scale, the share of output sold directly to consumers, the national industry size, and technology intensity measured by employment in science and engineering occupations, or in information systems occupations, R&D expenditures as a share of sales, and growth in total factor productivity for the industry. All of these are to be measured directly from raw employment and capital stock numbers, taken from input-output tables or other public data, except for the degree of scale economies which are estimated econometrically.

Estimating external economies of scale.

We can use an error component model for profits of an establishment as an underlying motivation,

$$\pi_{ik} = \Sigma_j \beta_j X_{kj} + \delta_k + \varepsilon_{ik}$$

and use the observed location choices by new establishments (births) to estimate the parameters of the profit function as well as firms' responses to location attributes (Carleton (1983), see Rosenthal and Strange (2003), Arzaghi (2005) and Davis and Henderson (2004)). We presume there exist unobserved location attributes δ_k which may be correlated with observed attributes. As such, we estimate a discrete choice model for births using a non-linear generalized instrumental variables estimator for a general exponential model based on the following set of moments (Windmeijer and Santos (1997), Mullahy (1997))

$$\Sigma_k (n_k - \exp(\Sigma_j \beta_j X_{kj})) w_k = 0$$

Based on Henderson (2003), we assume the best measure of own industry scale economies is the number of plants in the own industry locally, where the local spatial extent of scale benefits is the county not the MSA. The choice is critical; scale economies measured by local employment scale are not statistically significant in as many industries and are not highly correlated with scale economies under plant counts. Using MSA measures results in even weaker scale effects, consistent with recent evidence on the very local nature of such effects (Rosenthal and Strange (2001, 2003), Arzaghi and Henderson (2005)). As such, using all urban counties, we estimate

$$\begin{aligned}\text{births} = {} & \beta_0 + \beta_1 \ln(\text{own industry establishment count}) \\ & + \beta_2 \ln(\text{own industry wage}) + \beta_3 \ln(\text{county employment}) \\ & + \beta_4 (\text{central business district indicator})\end{aligned}$$

for each three digit industry, using the instruments described earlier in the data discussion (see Arzaghi (2005) for details). As an alternative, we estimated this model using only central business district counties, a specification that can be better identified for some business services industries.

In a discrete choice model setting as this one, it is the ratio of coefficients that has economic meaning (see Train (2003)), and we interpret the ratio of β_2 / β_1 as the willingness by the firm to pay wages in return for the externality received from greater local own industry scale.[10] The willingness to pay ratios of the wage coefficient to the scale coefficient are then comparable across industries, and are captured for use as a covariate later.

We also estimate external economies of scale with a common approach in the literature of using aggregated level variables in a city production function specification. The basic production function for any industry j

$$\ln(\text{value of production})_i(t) = \alpha + \beta \ln X_i(t) + \gamma_0 \ln(\text{number of plants in ind } j)_i(t)$$

$$+\gamma_1 \ln(\text{population MSA})_i(t) + \delta(t) + f_i + \varepsilon_i(t) \quad (5)$$

The X_i are industry-county inputs of labor (production workers' hours plus 1800 times the number of non-production workers), materials used in production, and capital (book value). Value of production is sales adjusted for inventory changes and materials are also adjusted for inventory changes. The $\delta(t)$ control for price changes over time and productivity shocks, the f_i are county fixed effects, and the $\varepsilon_i(t)$ are the contemporaneous error terms. The γ_0 is the degree of localization (own industry) economies and γ_1 the degree of urbanization economies (which spatial scope is assumed to be the whole metro area).

Given the large number of industries involved, the estimation is for aggregated county-industry data.[11] Results are fairly standard, with weak coefficients (under fixed effects) on the poorly measured capital stock variable. While always positive, the capital coefficient is only statistically

---

[10] This can be seen by taking the total derivative of the profit function with respect to wages and scale and set it to zero such that $d\pi = \beta_1 \partial \ln(\text{count}) + \beta_2 \partial \ln(\text{wage}) = 0$ and solve for the change in wage that keeps profit constant for a change in scale, or $\partial \ln(\text{count}) / \partial \ln(\text{wage}) = - \beta_2 / \beta_1$.

[11] Aggregation "avoids" cleaning the data plant by plant (and thus involves imputed numbers for many small plants accounting for a small fraction of industry output). Coefficients on inputs sum to less than one, a problem for aggregation, however this could reflect attenuation bias on the capital coefficient.

significant in 28 cases. The estimation assumes inputs are chosen before the contemporaneous output shock $(\varepsilon_i(t))$ occurs and f_i controls for unmeasured county characteristics which affect the value of production and input levels.

<u>Determinants of mobility</u>.

The industry primitives we considered for manufacturing include (1) the capital to labor ratio -- the industry national book value deflated by the CPI and divided by national industry employment; (2) average plant employment; (3) share of heavy industries in output -- the sum of input-output share coefficients for each industry for forestry, iron and non-ferrous ores, coal, stone and clay quarrying, chemical and fertilizer mineral mining, primary non-ferrous and iron and steel products and stone and clay products; (4) the share of output sold to consumers from input-output tables; (5) the point estimate for the willingness to pay wages for the degree of scale economies; (6) the national industry size (total employment); (7) the employment share of scientists and engineers; (8) R&D expenditures as a share of sales; and (9) total factor productivity growth.

For service industries, data on industry characteristics are less available. We considered (1) the national industry size (total employment); (2) the point estimate for the willingness to pay wages for the degree of scale economies; and (3) the log of employment in computer and mathematical occupations.[12]

4. Results

[12] This includes occupations such as computer and information scientists, computer software engineers, computer systems analysts, database, network and systems administrators, actuaries, mathematicians and statisticians.

In the following, we generally present results that include both manufacturing and business service industries, however we sometimes limit the discussion to the subset of industries that include manufacturers, using LRD data for the longer span of 1963 to 1997. Most previous empirical work has focused on manufacturing, so in some cases we find presenting the discussion for manufacturing industries can help offer information in a comparable context.

Table 2A gives the average across the 94 industries of 5 cell transition probabilities. It also gives the standard deviation of the transition probabilities across those industries. Table 2B does this for the 104 industries under a 4-cell matrix. So in Table 2A, .912 is the probability of staying in state 1 from $t-1$ to t, while .084 is the probability of moving from state 1 to state 2, or .0015 is the probability of moving from the top state (5) to the bottom (1). Under either a 4 or 5 cell distribution, every industry passes stationarity except for the 5-cell industry 822 (colleges and universities) and for the 4-cell 373 (ship and boat building). Most p-values exceed 0.9, resoundingly passing stationarity. For the 5-cell distribution, 80% of the industries have p-values that exceed 0.9, each resoundingly passing stationarity, and for the 4-cell it is 70% of the industries. It is evident from the table that there exists a very high probability of staying in the bottom state if a city starts there. Getting a new industry is difficult for a city. Note also that the standard deviations are not small; there is a lot of variation across industries in transition probabilities.

Table 3 gives the trace and mean first passage times for the eight fastest and the eight slowest industries, as defined by the eight lowest and highest values of the trace. The R^2 for the five cell distribution in regressions of the trace on the mean first passage times of going from state 1 to 5 and state 5 to 1 are respectively .60 and .84. In column 2 for the 5-cell, a trace of 2.69 is associated with a mean first passage time of 44 decades to go from the bottom to top state for holding offices, while a trace of 4.62 is associated with a corresponding mean first passage time of 325.2 decades for legal services.

Two issues are worth noting. Although we use the trace to examine mobility across MSA's for different industries, we could have used the county as the geographic unit of observation.[13] The traces defined for 5 cell distributions using counties versus using MSA's are strongly correlated. For 51 common manufacturing industries, a regression of one on the other yields a slope coefficient of 1 and an R^2 of .82. Secondly, note in Table 3 that the slow move from the bottom to top state, compared to the quick move from the top to bottom state, in terms of mean first passage times, is a function of our cell sizes. The bottom cell is huge and the top cell tiny. What is relevant is the comparative speeds between fast and slow industries.

In Table 4, we list those among our 104 large industries that have the eight highest and eight lowest EG concentration, where the index is calculated for each year using counties and then averaged over time. Some of the most concentrated industries are motion pictures, securities, and advertising books, as well as aircraft and photo equipment. The least concentrated are ubiquitous industries such as health related, telephone communications and commercial printing.[14] The average EG over years and industries is .0108.

Focusing on manufacturing, the average EG is .0096, and declines from .0117 in 1963 to .0079 in 1997. Note that our 68 large manufacturing industries do not include the most concentrated industries which are very small industries such as fur goods (sic 237) with an EG of .623 and only 4,000 employees nationally on average per year. The average EG for all 140

[13] The problem for counties is that out of 2918 common denominator counties typically 85-90% have zero employment in any one industry. The bottom state contains many places that essentially don't "count" -- have no economic significance (unlike MSA's with zero industry employment).
[14] Note that this is a sampling, in the sense that we did not include in our SIC list many ubiquitous industries such as retail, personal services, and so on.

manufacturing industries over all years is .0167. Results for the Gini are different, in that the Gini only decreases from 1963 to 1997 in 32 of 68 cases, versus 54 cases for EG, and the average actually rises from .746 to .762 over the time period. Regressing EG on the Gini gives a positive and significant coefficient, yet with an R^2 of only .082.

Turning to the relationship between concentration and mobility, for industries that are concentrated due to scale economies, theory suggests a location without the industry has difficulty attracting firms since it can't offer the productivity advantage of built-up scale. The scatter plot in Figure 1 between the trace of the transition matrix and the time-averaged EG for each industry does not indicate a systematic relationship between the two measures. A regression shows a statistically insignificant relationship as well. This is a surprising result, and led us to explore other possibilities. First, time averaging is not the issue. Regressions of the trace for manufacturing for the transition matrix in 1963 against the EG for 1963, and the same for the trace and EG measures for 1997, also yield insignificant relationships. Second, the EG index tends to be dominated by the largest three or four sites. For example, the change in the EG index from 1963 to 1997 for the 68 manufacturing industries has an R^2 of .550 when regressed on the change in the share of the top three locators, while the change in EG with respect to the change in the share of the bottom ranked locations (below the top 40) has an R^2 of only .043. As a complement to the EG index, the Gini index gives a better view of concentration in the lower part of the distribution. A regression of mobility on the Gini yields an R2 of .097, and the trace declines as the Gini rises with the coefficient significant. These preliminary results appear to indicate mobility could be larger for more concentrated industries.

Table 5a presents basic information on results for the willingness of producers to pay wages for an increase in own industry scale economies, from the estimation of establishment births as described earlier. Using the sample of all urban counties for the estimation, we see that the industries with the strongest localization externalities include manufacturers that produce nonferrous rolling and drawing, household furniture, aircraft and general industrial machinery. Repeating the estimation using only central counties of each city, we get similar results. For 85% of industries in the sample, the localization coefficient is of expected sign and significant at the 5% level.

Table 5b presents basic information on results for the own industry scale economies from the estimation of aggregated city production functions. The coefficient, γ_0 is almost always positive and often significant. Coefficients are of expected magnitude. The average of .057 suggests that a 1% increase in the number of plants yields a .057% increase in output for the same inputs. Going from 10 to 150 plants locally raises output by 17%. Urbanization economies don't really exist in standardized manufacturing. There are two industries with significant positive coefficients, two with significant negative ones, and the rest are insignificant.

Given Ciccone and Hall's objections to the form utilized, we also estimated for manufacturing an equation for value added per production worker as a function of the capital to production worker ratio. As with the basic equation, we find little evidence for urbanization economies. Localization economies are about 2.7 times larger on average than in the basic equations and the R^2 between the two measures is .292.

Table 6 shows for manufacturing industries the relationship of mobility to some fundamentals of the underlying industries. We examine the trace as a function of time averaged industry variables. We find the share of heavy raw materials, plant size, and the share sold to consumers have no effect on mobility, and were thus excluded from the results presented. From the table, increases in the capital-labor ratio increase the trace, or reduce mobility, a traditional fixed capital stock impediment to mobility. Industry size nationally reduces mobility. This is consistent with urban theory, which predicts an optimal scale in equilibrium where greater overall industry size would yield replication of urban sites. Thus, a larger industry would be expected to be present in more cities, making it less likely to have reordering in the urban ranking.

The coefficient on the increase in the willingness to pay for scale economies has an insignificant effect on mobility for manufacturers. From Table 7 for service industries however, we see those firms that benefit from localized scale economies tend to be more mobile. This finding is unexpected, though may in part reflect heterogeneity in service industries' ability to export their services across cities. In addition, lower fixed capital intensity in service production may facilitate movement of footlose industries. This warrants further exploration.

Our main results show technologically oriented or advancing industries are more mobile. For manufacturing in Table 6 we use a variety of different measures of technology, share of employment in science and engineering, R&D intensity and increasing total factor productivity over time, and all of these measures enter with significant coefficients and demonstrate this basic relationship. Likewise in Table 7 for services industries, we use the log of employment in computer and mathematical occupations, and find a similar positive and statistically significant effect on mobility.

As outlined in the introduction, Duranton (2004) introduces a theoretical model of cities where the effect of technological change can induce mobility. We find the empirical evidence presented here lend support to the ideas introduced in this work.

5. Conclusion

Audretch and Feldman (1996) show concentration is higher with R&D intensity, and Henderson (2003) for a selection of industries shows those that are high tech exhibit higher productivity from localization economies, as captured by the scale of like industry plants nearby. The contribution of this paper is to show that these agglomerations of industrial activity are not necessarily fixed over time, and that technologically active industries in fact tend to be more mobile.

We offer a stylized example, one that is consistent with empirical evidence of a highly mobile U.S. computer industry (Beardsell and Henderson (1999)). Computer production was originally concentrated in New York state, with the production of mainframes. Later, new agglomerations emerged, notably that of the route 128 beltway around Boston, home to many successful "mid-range" computer manufacturers in the 1980's. Finally with the advent of microprocessor computing, the emergence of the "Silicon Valley" agglomeration is both well known and widely documented.

As a final note, one not explored in this paper, perhaps this relationship of technology and industry mobility is in part explained by technological experimentation, where attempts at new technologies are uncertain and may not necessarily occur in existing agglomerated locations due to their high cost or a need for other or diverse inputs as in Duranton and Puga (2002). Dumais, Ellison and Glaeser (1999) find that new births reduce concentration, yet plant closures reinforce it. To the extent that technology successes accelerate the depreciation of physical and human capital stocks in existing agglomerations, they can facilitate spatial shift to newly emerging sources of knowledge in the industry. Technological failures on the other hand, as reflected in plant closures, would leave existing agglomerations intact.

References

Arthur, B. (1990), “Silicon Valley location clusters: when do increasing returns imply monopoly.” Mathematical Social Sciences, 19(3), 235-251.

Arzaghi, M. (2005), “Quality Sorting and Networking: Evidence from the Advertising Agency Industry,” Brown University mimeo.

Arzaghi, M. and J.V. Henderson (2005), “Networking Off Madison Avenue“, Brown University mimeo.

Audretsch, D. and M. Feldman (1996), “R&D spillovers and the geography of innovation and production,” American Economic Review, 86(3), 630-640.

Beardsell, M. and J.V. Henderson (1999), “Spatial evolution of the computer industry in the U.S.A.,” European Economic Review, 43, 431-456.

Black, D. and J.V. Henderson (2003), “Urban evolution in the USA,” Journal of Economic Geography, 3, 343-372.

Carleton, D. (1983), “The location and employment choices of new firms: An econometric model with discrete and continuous endogenous variables,” Review of Economics and Statistics, 65, 440-449.

Davis, J. and J.V. Henderson (2004), “The Agglomeration of Headquarters,” Working Paper 04-02, U.S. Census Bureau, Center for Economic Studies.

Dumais, G., G. Ellison and E. Glaeser (2002), “Geographic concentration as a dynamic process,” Review of Economics and Statistics, 84(2), 193-204.

Duranton, G. (2004), “Urban Evolutions: The Still, the Slow, and the Fast,” London School of Economics mimeo.

Duranton, G. and D. Puga (2001), “Nursery cities: Urban diversity, process innovation, and the life cycle of products,” American Economic Review, 91(5), 1454-1477.

Eaton, J. and Z. Eckstein (1997), “Cities and growth: Theory and evidence from France and Japan,” Regional Science and Urban Economics, 27, 443-474.

Ellison, G. and E.L. Glaeser (1997), “Geographic concentration in U.S. manufacturing industries: A dartboard approach,” Journal of Political Economy, 105, 889-927.

Ellison, G. and E.L. Glaeser (1999), “The geographic concentration of industry: Does natural advantage explain agglomeration?,” American Economic Review, 89, 311-316.

Henderson, J.V. (1974), “The sizes and types of cities,” American Economic Review, 64(4), 640-656.

Henderson, J.V. (2003), “Marshall’s scale economies,” Journal of Urban Economics, 53, 1-28.

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

Miranda, J. and R. Jarmin (2002), "The Longitudinal Business Database," Working Paper 02-17, U.S. Census Bureau, Center for Economic Studies.

Mullahy, J. (1997), "Instrumental-variable estimation of count data models: Application to a model of cigarette smoking behavior," The Review of Economics and Statistics, 79(4), 586-593.

Rauch, J. (1993), "Does history matter only when it matters little? The case of city-industry location," Quarterly Journal of Economics, 108(3), 843-867.

Rosenthal, S. and W. Strange (2001), "The determinants of agglomeration," Journal of Urban Economics, 50, 191-229.

Rosenthal, S. and W. Strange (2003), "Geography, industrial organization, and agglomeration," Review of Economics and Statistics, 85(2), 377-393.

Rosenthal, S. and W. Strange (2004), "Evidence on the nature and sources of agglomeration economies," in Henderson, J.V. and J.-F. Thisse (Eds.), Handbook of Regional and Urban Economics, vol. 4. Elsevier, Amsterdam, 2119-2171.

Windmeijer, F. and J. Santos (1997), "Endogeneity in Count Data Models: An application to demand for health care," Journal of Applied Econometrics, 12(3), 281-294.

Table 1. Mobility of Largest City-Employers for 10 Largest Industries in 1963

[SIC]	MSA top in 1963	1963 shares of national employment	1997 rank of top 1963 city	1997 share of national employment of top 1963 city	MSA top in 1997 [1963 rank, 1997 share]
autos [371]	Detroit	.211	1	.116	
steel [331]	Pittsburgh	n.a.*	2	.057	Gary [2, .136]
aircraft [372]	Los Angeles	.175	2	.115	Seattle [6, .178]
men's [232] furnishings	New York	.071	3	.030	Los Angeles [2, .065]
women's [233] outer wear	New York	.296	2	.219	Los Angeles [2, .266]
communications [366] equipment	Los Angeles	.155	5	.043	San Jose [36, .121]
missiles [376]	Los Angeles	.260	1	.251	
fab. metals [344] (structures)	Los Angeles	.049	1	.032	
measuring [382] instruments	Los Angeles	.088	5	.028	San Jose [13, .076]
newspapers [271]	New York	.066	2	.026	Los Angeles [2, .031]

* Cannot be disclosed.

Table 2A. Five Cell Transition Probabilities: Average Across Industries (standard deviation across industries of transition probabilities)*

		state in t				
		1	**2**	**3**	**4**	**5**
state in t-1	1	.912 (.026)	.084 (.023)	.0032 (.0033)	.0008 (.0015)	.00008 (.00032)
	2	.184 (.050)	.723 (.074)	.085 (.026)	.0070 (.0075)	.00071 (.00155)
	3	.021 (.021)	.213 (.065)	.641 (.101)	.122 (.036)	.0037 (.0065)
	4	.0080 (.0139)	.025 (.030)	.170 (.048)	.721 (.088)	.075 (.026)
	5	.0015 (.0063)	.0068 (.0181)	.011 (.021)	.183 (.065)	.800 (.079)

B. 4 Cell Transition Probabilities

		state in t			
		1	**2**	**3**	**4**
state in t-1	1	.967 (.012)	.029 (.0095)	.0029 (.0029)	.00029 (.00053)
	2	.237 (.074)	.638 (.100)	.121 (.037)	.0036 (.0064)
	3	.037 (.042)	.166 (.051)	.720 (.086)	.076 (.027)
	4	.0091 (.0210)	.012 (.021)	.185 (.064)	.794 (.078)

- Variance of any element is approximated by $p_{jk}(1-p_{jk})/(T*$ number of cities in state $j)$

Table 3. Fastest and Slowest Industries

Fastest Industries (SIC)	4-cell Trace	5-cell Trace	5-cell mfp up (decades)	5-cell mfp down (decades)
Holding Offices (671)	2.34	2.69	43.5	4.63
Pension, Health and Welfare Funds (637)	2.51	2.98	63.8	6.2
Misc. Electrical Equipment (369)	2.65	3.19	70.7	6.4
Misc. Wood Products (249)	2.67	3.23	67.6	6.8
Men's Furnishings (232)	2.73	3.26	80.9	8.0
Toys and Sporting Goods (394)	2.69	3.28	73.4	7.6
Logging (241)	2.85	3.38	121.6	8.3
Communications Equipment (366)	2.84	3.41	83.1	7.8
Farm and Garden Machinery (352)	2.87	3.45	90.4	8.4
Meat Products (201)	2.80	3.46	81.7	8.9
Slowest Industries				
Legal Services (811)	3.72	4.62	325.2	37.6
Doctors Offices and Clinics (801)	3.49	4.35	214.7	33.4
Fire, Marine and Casualty Insurance (633)	3.46	4.32	251.0	30.7
Commercial Printing (275)	3.52	4.29	253.9	24.3
Hospitals (806)	3.51	4.26	256.9	24.8
Newspapers (271)	3.45	4.25	178.1	24.4

Table 4. Concentration

Most Concentrated Industries	Ellison-Glaeser Index (average over years)
Motion Pictures (781)	.281
Security Brokers and Dealers (621)	.114
Photographic Equipment (386)	.109
Periodicals (272)	.057
Women's Outerwear (233)	.036
Advertising (731)	.033
Producers, Orchestras, Entertainers (792)	.026
Aircraft and Parts (372)	.023
Least Concentrated Industries	
Hospitals (806)	.00023
Colleges and Universities (822)	.00027
Medical Service and Health Insurance (632)	.00040
Telephone Communications (481)	.00044
Health and Allied Services nec. (809)	.00058
Doctors Offices and Clinics (801)	.00072
Commercial Printing (275)	.00075
Other Health Practitioners (804)	.00094

E.G. average over years and industries .0108

Table 5a. Birth Model Estimation of Scale Economies

Degree of willingness to pay for localization economies across industries

Highest ranked industries Urban counties sample	
Industry (SIC)	Willingness to pay
Nonferrous rolling & drawing (335)	2.94
Household furniture (251)	2.93
Metal Services, NEC (347)	2.51
Aircraft & parts (372)	2.27
Soap, cleaners & toilet goods (284)	1.96
General industrial machinery (356)	1.91
Misc. Manufacture (399)	1.85

Highest ranked industries CBD counties sample	
Industry (SIC)	Willingness to pay
Metal services, NEC (347)	2.90
Household furniture (251)	2.83
Soap, cleaners & toilet goods (284)	2.40
Meat products (201)	2.34
Misc. plastic products (308)	2.19
Aircraft & parts (372)	1.94
Kitting mills (225)	1.77

* Industries included are those with statistically significant estimates of the willingness to pay ratio.

Number of times localization coefficient (urban counties sample) is:
- positive and significant at 5% level: 87
- positive and significant at 10% level: 3

Rest insignificant with 9 of 12 having positive coefficients.

Table 5b. City Production Function Estimation of Scale Economies

Degree of localization economies across manufacturing industries

average,	**standard deviation**,	**maximum**
.057	.036	.149

Number of times localization coefficient is:
- positive and significant at 5% level: 37
- positive and significant at 10% level: 4

Rest insignificant with 24 of 27 having positive coefficients

Table 6. Manufacturing Determinants of Mobility

Dependent Variable: log 4-cell trace	(1)	(2)	(3)	(4)	(5)
Capital to labor ratio	.000492** (.000138)	.000488** (.000176)	.000569** (.000139)	.000474** (.000075)	.000510** (.000128)
Log industry total employment	.0356** (.0111)	.0358** (.0118)	.0404** (.0111)	.0383** (.0118)	.0425** (.0105)
Willingness to pay for scale elasticity		-.000159 (.00525)			
Share of scientists and engineers			-.465** (.217)		
R&D expenditure share				-.641** (.272)	
Total factor productivity					-1.45** (.42)
Number of industries	68	66	68	68	68
R-squared	.219	.216	.271	.278	.342

** indicates significant at 5%, and * at 10%. Standard errors are in parenthesis.

Table 7. Services Determinants of Mobility

Dependent Variable: log 4-cell trace	(1)	(2)	(3)	(4)
Log industry total employment	.0363** (.0129)	.0172 (.0130)	.0548** (.0155)	.0386** (.0158)
Willingness to pay for scale elasticity		-.0155** (.0069)		-.0160** (.0065)
Log information systems occupations			-.0171* (.0086)	-.0180** (.0084)
Number of industries	35	32	35	32
R-squared	.193	.240	.281	.348

** indicates significant at 5%, and * at 10%. Standard errors are in parenthesis.

Figure 1. Mobility and Concentration