6-8-2010

The Role of Geographic Proximity And Industrial Structure In Metropolitan Area Business Cycles

Michael Hollar
U.S. Department of Housing and Urban Development

Anthony Pennington-Cross
Marquette University, anthony.pennington-cross@marquette.edu

Anthony Yezer
George Washington University

The Role of Geographic Proximity And Industrial Structure In Metropolitan Area Business Cycles

Michael Hollar
U.S. Department of Housing and Urban Development, Office of Policy Development and Research; e-mail: michael_k_hollar@hud.gov

Anthony Pennington-Cross
College of Business Administration, Marquette University; e-mail: Anthony.pennington-cross@marquette.edu

Anthony Yezer*
Department of Economics, George Washington University; e-mail: yezer@gwu.edu

* Corresponding author

A previous version of this paper was presented at the Use of Composite Indexes in Regional Economic Analysis conference, Philadelphia Federal Reserve Board (September 2003).

U.S. Department of Housing and Urban Development Working Paper # REP06-01

The opinions expressed in this paper are those of the authors and do not necessarily reflect the opinions of the Department of Housing and Urban Development.

Abstract. Measurement and prediction of aggregate economic fluctuations at the region, state, and metropolitan area level is a major challenge. As data quality and analytical techniques have improved, the analysis of coincident economic cycle indicators (CEI) has progressed from national to regional to state levels. This paper continues the trend of geographic disaggregation by constructing and analyzing CEI at the MSA level. The theoretical advantage of MSA level indexes is that they reflect labor market areas.

Given lack of quarterly economic time series at the MSA level, we construct a new variable, the EPI (export price index). The EPI is an index number constructed to measure changes in the prices of goods produced by major industries located in metropolitan areas. Using non-agricultural employment and the EPI as MSA-specific variables, we are able to estimate following a Stock/Watson type single factor models. We find that, for larger states, with multiple MSAs, there is substantial variation in the amplitude and timing of cycles across MSAs. Further tests group MSAs within states by applying cluster analysis to the state series for the MSAs within a state. The groupings are interesting for two reasons. First, they confirm the differences observed within states. Secondly, and perhaps most important, the groupings of cyclically similar MSAs are not always based on geographic proximity as might be expected. It appears that industrial similarity of the MSA economies is also important for cyclical performance.
I. Introduction

Since the 1930's researchers at the National Bureau of Economic Research have produced leading, coincident, and lagging indicators of the level of economic activity for the entire United States. The number and quality of national time series variables available monthly for incorporation into such indexes have expanded over time and econometric techniques have evolved rapidly. Stock and Watson (1989) suggested the use of a dynamic single factor model of the business cycle and developed procedures for the estimation of weights on observable series to be used in estimating movement of that latent variable. The success of this technique in providing a concrete mathematical framework within which variables and indexes could be evaluated, combined with improved data availability at the regional and state level, has led to efforts to explore geographic disaggregation of business cycle indicators. Clayton-Matthews and Stock (1998/1999) demonstrated the feasibility of disaggregation to the state level for Massachusetts. Recently, Crone (2003) has developed consistent economic indexes for the 50 states and regional business cycle analysis has been the object of a growing literature such as the recent exploration of commonality in cycles by Carlino and Sill (2001).

This paper continues the trend toward geographic disaggregation by extending the common factor approach to the analysis of cyclical activity at the MSA level. There are three major reasons to undertake such an effort. The first is to determine whether the Stock-Watson technique can be extended to the MSA level given the paucity of city-level data. The second begins with the observation that MSA boundaries are set based on empirical observation of work and commuting patterns to reflect a labor market area. Given that the current CEI is determined, to a significant extent by factors, such as non-agricultural employment and personal income less transfers, that reflect the local labor market, the basic level of geographic disaggregation at which fluctuations in these variables are generated should be the MSA. In view of this, we hypothesize that there may be substantial variation in cycles among MSAs within a given state. Third, previous geographic disaggregation has been conducted on the basis that proximity is the basis for cyclical similarity. States are routinely grouped into regions based on geographic proximity. A competing view is that cyclical similarity is based on industrial composition. According to this view, MSAs with similar industrial structure should have cycles that are coordinated even if they are geographically dispersed and adjacent MSAs with different industrial structure might have dissimilar cycles.

We examine these three questions in this paper by constructing MSA-specific single-factor coincident indexes for large, multi-MSA states. In the course of the research, we propose and
implement a new index number to measure export price shocks to the MSA economy. This effort is justified both by economic theory suggesting that such price shocks should be a leading or coincident indicator of cyclical fluctuations and by the lack of other high-quality measures of MSA economic activity. We also construct a new index of MSA economic dissimilarity that allows us to examine the hypothesis that the variation in movements of the common factor across MSAs may be due to industrial composition in addition to geographic proximity.

II. The Search for MSA-level Measurement Series

There are few indicators of economic activity at the metropolitan level available monthly or even quarterly for a significant time for cities across the country. Most time series of metropolitan area economic activity are measured annually and likely have substantial error, if they are even measured at all. Very few are measured monthly or even quarterly and an even smaller number have sufficiently long time series to allow for statistical evaluation. In addition, changes in metropolitan area boundaries and in sampling methods have disrupted some possible candidates for use as economic indicators. Accordingly, the vast majority of variables available for consideration as leading, coincident, or lagging variables nationally, are simply not available at the state, and particularly at the metropolitan area level.

In addition to difficulties arising from data limitations research on business cycles at the metropolitan level is complicated by the diversity of behavior. Fluctuations in urban economies are may appear rather different in character than national cycles. Using the classic NBER definition of cycle, some cities have not experienced a downturn in the past 30 years while others have experienced far more cycles than the aggregate economy, suggesting that it may be difficult to estimate dynamic factor models for some cities.

Given the very short list of possible indicator variables currently available for characterizing cycles at the MSA level, we have undertaken to construct a new variable, the export price index (EPI). While the EPI has its roots in classic regional economics it is also consistent with the new economic geography. It is constructed by first identifying the major export industries in each metropolitan area. Such export industries are differentially concentrated in those cities and therefore must sell a significant proportion of their output outside the area. Secondly, a transformation matrix relates the export industries to their characteristic products for which national price series are available on monthly basis.

---

1Consider, for example, that fundamental variables such as metropolitan area population are generated as projections between decennial census periods. BLS earnings data cover manufacturing sectors only.
Once this is done, the EPI can be constructed for each city based on application of the export weights of each industry to the vector of product prices. In effect, the EPI measures demand shocks to the metropolitan economy. To the extent that local output and employment fluctuate due to these demand shocks, the EPI should be a leading indicator of local economic activity.

The next section of this report discusses the rational for and construction of the EPI. This is followed by an evaluation of the EPI as a leading indicator using metropolitan area employment as a coincident indicator. In doing this we are effectively assuming consistency with the NBER's use of nonagricultural employment as a coincident indicator. Overall, we find the expectation that the EPI measures exogenous demand shocks to the local economy and functions as a leading indicator of employment changes appears to be justified in the data.

III. Rational for and Construction of the Export Price Index

The rational for believing that the EPI, as constructed here, can function as a leading of coincident indicator of urban or regional economic activity is consistent with both the "old" regional economics and the "new" economic geography. The former relies on various demand driven models ranging from the simple economic base model to regional input-output models.\(^\text{2}\) Demand shocks drive changes in output and employment as labor and capital are supplied perfectly elastically to the regional economy. The new economic geography relies on a variation of this demand driven process in which positive demand shocks allow further product differentiation that stimulates local economic growth.\(^\text{3}\) It also allows for the role of positive productivity shocks which lower export prices. Thus, both old and new approaches to regional economic development suggest that sector-specific price shocks have a role in cyclical fluctuations of a city or regional economy.

Three steps are required to construct a price index that is specific for individual MSAs that is capable of measuring price shocks for goods and services produced for export. First the export industries for each area is identified and separated from locally produced and consumed goods. Second, each industry is weighted by its importance to the MSA. Third, using the weights, a MSA unique price index of locally produced goods for export called the Export Price Index (EPI) is created.\(^\text{4}\)

\(^\text{2}\) See, for example, the discussion in Tiebout (1962) or Richardson (1985).
\(^\text{3}\) See the discussion in Chapters 3, 6, and 11 of Fugita, Krugman, and Venables (2001).
\(^\text{4}\) For a discussion of an earlier version of the EPI, see Pennington-Cross (1997). The sector weights are taken from output patterns from the 1999 Census of Business and hence can be thought of as a Paasche index.
Identification of Export Industries for Each Metropolitan Area

To identify the export industries in any metropolitan area, location quotients ($LQ$) are used. There is a substantial literature including Greytak (1969), Leigh (1970), Gibson and Warder (1981), and Isserman (1977) on the use of location quotients to identify export industries for cities, states, and regions. The $LQ$ is the share of employment of industry $i$ in region $r$ divided by the same industry employment share for the nation. Any $LQ$ greater than one identifies an industry that exports at least part of its product out of the MSA. $LQ$s are computed for the year 1987 using data collected from the U.S. Army Corps of Engineers Construction Engineering Research Laboratory (CERL) and CERL's Economic Impact Forecast System (EIFS). All data used to calculate $LQ$s are organized by SIC code at the four digit level. This level of detail is desirable because it creates more homogeneous product categories and reduces problems associated with cross hauling.

Location Quotients

Location quotients are typically used to measure the economic base of a region when surveys or inter-regional trade flow data are not available. The $LQ$ is defined as follows:

$$LQ_{ir} = \frac{e_{ir}/e}{(e_i/e_r)} \quad (1)$$

where $i$ is the industry, $r$ is the region or MSA, $LQ_{ir}$ is the location quotient for industry $i$ in region $r$, $e_{ir}$ is employment in industry $i$ region $r$, $e_r$ is total employment in region $r$, $e_i$ is national employment in industry $i$, and $e$ is national employment. Richardson (1985) identifies the following assumptions needed to estimate equation (1) accurately: productivity and consumption per employee in the region must be the same as national levels for each industry $i$, the nation must be neither a net exporter nor importer of $i$, and if region $r$ exports $i$ then all of $i$ consumed locally must be produced locally. As a result of these stringent requirements traditional estimates can greatly underestimate exports if aggregated definitions of industries (two digit SIC) are used. This level of aggregation creates cross hauling problems and violates last assumption noted above. Bloomquist (1988) uses four digit SIC data from EIFS to help alleviate the cross hauling problem. We follow his approach in computing the EPI.

Industry Weights

The ideal weight to use for each export industry would be the fraction of all export value added for the metropolitan area produced in each industry. Because value added is not readily
available, we have applied employment weights. The weight assigned to each identified export
industry in each MSA is the estimated export employment for industry \( i \) divided by the total export
employment in the whole MSA or the industry's share of MSA export employment.\(^6\)

\[
\begin{align*}
    w_{ir} &= \frac{ex_i}{ex_r} = \frac{ex_i}{S_e x_i} \\
    &= \frac{1-1/LQ_i}{e_i} = \left( \frac{e_i}{e_i} - \frac{e_r}{e_i} \right) e_i .
\end{align*}
\]

where: \( i \) and \( r \) are the industry and MSA, \( ex \) is the estimated export employment, \( w \) is the weight and
\( ex_{ir} \)

\( LQ_i \) is less than or equal to one, \( w_{ir} \) and \( ex_{ir} \) equal zero.

Construction of the Export Price Index (EPI)

Even the largest cities are specialized in the production of a moderate number of products for
export to the rest of the world. This section describes what we believe to be the first successful
attempt to construct an index of export prices, the EPI. Our approach to this task emphasizes
simplicity, transparency, periodicity, and comprehensive coverage over elegance and complexity.
The EPI is designed to be available for counties, metropolitan areas, and states on a monthly or
quarterly basis for a substantial period of time and to be easily updated.

The EPI is calculated by transforming commodity-based price indexes into industrial
categories and then multiplying by a weighting scheme that identifies specific export industries in
each geographic area. Specifically, the EPI can be expressed as: \( CTW = x \); where \( C \) is a 1xm
dimensional vector of \( m \) commodity prices for a particular year, \( T \) is an \( m \times i \) matrix transforming the
\( m \) commodity prices into \( i \) four-digit SIC categories, \( W \) is an \( i \times r \) matrix of export intensity weights
for \( i \) industries in \( r \) regions, and \( x \) is a 1xr dimensional vector of export price index values for the \( R \)
regions at a particular time. The previous paragraphs explain the construction of \( W \), the matrix of
weights associated with each with each industry. The remainder of this section details the
construction of \( C \), the vector of commodity prices and the transformation matrix \( T \), which takes
commodity prices into industrial categories.

The annual index of prices of tradeable commodities, \( x \), is taken from three sources: the
Producer Price Index (PPI), the Consumer Price Index (CPI), and various sector prices (SP). These
three indexes are produced by the Bureau of Labor Statistics (BLS). The BLS provides a series of
translation tables designed to transform commodity prices into SIC industrial categories. However,
codes from the various price categories sometimes overlap or may be missing, and hence an element
of judgment is required to form the translation matrix. The elements of \( T \) were selected by imposing
the following priority: match in the standard BLS PPI/SIC translation table, match in the BLS SP/SIC translation table, match in a manual search through the SIC 1987 Manual and the PPI code descriptions, and match in a manual search through the SIC 1987 Manual and the CPI code descriptions. Imposing this order of search results in a T matrix constructed from 326 PPI, 54 SP, and 49 CPI price series. If this sequence of searches was unable to locate even one match for a particular SIC category, then the industry was dropped from the analysis. Some industries have been determined to be for local consumption only (e.g., construction sectors, public utilities, second hand stores) and have therefore been removed from export analysis. For other industries a price could not be found. The most commonly excluded four-digit SICs for this reason are (1) printing and publishing, (2) military hardware, (3) metal, coal, and oil and gas services, and (4) very vague general categories. Approximately 84 industries in total are dropped from the analysis. In cases where a given source provides more than one price series per SIC category, the price series are averaged. The elements of W, or the weights associated with each industry in each MSA, are constructed as described in the previous subsection.

III. Estimating Indexes of Coincident Economic Indicators

In this section, we report on the estimation indexes of coincident economic indicators constructed for the largest states, specifically California, New York, Florida, and Texas, and for the larger MSAs within those states. The Stock/Watson model estimates are essentially a dynamic factor model in which the unobserved common factor or “state” series, Δc_t, is expressed as a weighted combination of the indicator variables or “measurement” series, Δx_t. Formally the model structure may be expresses as:

\[ Δx_t = β + γ(L) Δc_t + μ_t \]  
\[ D(L) μ_t = P(L) ε_t \]  
\[ φ(L) Δc_t = δ + q(L) η_t \]

The measurement series include logarithms of: quarterly estimates of the EPI, monthly employment (from the Bureau of Labor Statistics); monthly unemployment rate estimates (from the Bureau of Labor Statistics, based on the Current Population Survey); monthly average hours worked in manufacturing (from the Bureau of Labor Statistics) and quarterly real wages from the Bureau of Economic Analysis. The last three variables were taken from Crone’s (2003) state indexes. The EPI and monthly
employment are observed at the MSA level. All variables were seasonally adjusted, either previously by
the BLS or using the X-11 procedure. Estimation of the system represented by equations 3-5 is
accomplished using a (quasi) maximum likelihood procedure developed by Clayton-Matthews (2001)
that accommodates both quarterly and monthly frequencies in a model producing estimates of the state
series with monthly frequency. The scale of the \( \gamma(L) \) coefficients is determined by setting the variance
of \( \eta \) equal to unity and the timing of the coincident index is fixed by setting all but the current element
of \( \gamma(L) \) equal to zero in equation (1) for employment following the approach in Crone (2003).

The model was estimated for four large states and for the large MSAs within those states in order
to evaluate the effect of substate diversity in location and economic structure. The results are displayed
in a series of tables in Appendix A. Stock and Watson (1989) suggest a test of the assumption of a
single state series in the form of a test of the ability of past values of the measurement series or past
values of the errors from the measurement equations are related to disturbances in the measurement
equations. For each model estimated in Appendix A, a table reports the F-statistics for rejecting the
hypothesis that the coefficients obtained by regressing the errors from each measurement equation on a
constant, six lags of the errors from each measurement equation and six lags of each measurement series
are jointly equal to zero. Using a critical value of the F-statistic at the 0.05 level, we find that only 16 of
200 F-statistics for the four states and 79 of 950 F-statistics reported for the 19 MSAs are significant.
Overall, these results support our assumption that cyclical fluctuations in these areas can be well
characterized by a single state series.

The relations among movements of the state series for each state and for larger MSAs within
each state are displayed in Figures I through VII. In California, the frequency and timing of cycles across
cities is similar. There is an obvious difference in amplitude associated with regions of the state.
Northern California, particularly San Francisco and Riverside but also Oakland, has large cycles
compared to the Southern California MSAs, Los Angeles, Orange County, and San Diego. Florida cities
display the most consistent frequency and amplitude. Fort Lauderdale has the largest amplitude but
differences across cities and between city and state are small. There is a division in behavior of cities in
New York that appears to be based on proximity. Nassau-Suffolk and New York City are closely related,
while Syracuse and Rochester are generally similar. Note that Rochester has a recession in 1994 that
does not occur in any other cities. There is substantial diversity among cities in Texas. The cycle in
Dallas has particularly large amplitude, although its periodicity is similar to its Fort Worth neighbor.
The state series for Corpus Christi, Houston, and San Antonio are nearly identical, with very shallow
recessions. Indeed, Houston and San Antonio have been in a continuous expansion since 1987.

IV. Is Cyclical Similarity Based on Proximity or Industrial Structure

The previous section illustrated that there may be significant diversity in the timing and/or amplitude of cycles for MSAs within a given state. In this section, we consider the extent to which these within-state differences are associated with geographic proximity or similar industrial structure. Classification of MSAs in terms of proximity is trivial but finding a metric for identifying similarity of industrial structure is far from obvious.

After searching unsuccessfully for a measure of industrial similarity that could be used to distinguish MSAs such as those in our sample, we have decided to use the correlation between the EPIs for pairs of cities as a measure of industrial similarity. Given that EPI movements are the product of weights based on employment shares of city export industries and price movements for the products of those industries, high rates of correlation between EPI measures over time are unlikely unless cities have similar export industrial structure. Table I shows the matrix of correlation coefficients between EPIs for cities in each state and groups MSAs with high pairwise correlations. In the case of California and, particularly, New York, geographic proximity and industrial similarity tend to produce similar groupings of MSAs. The three Southern California MSAs, Los Angeles, Orange County, and San Diego have EPIs that are highly correlated. San Francisco’s industrial structure is similar to that of the south but its two northern neighbors, Oakland and Riverside are dissimilar to any other MSAs. In New York, Nassau-Suffolk and New York city have similar industrial structure. Rochester and Syracuse, while not close industrially based on correlation of their EPIs are closer to one another than they are to Nassau-Suffolk or New York City. In Texas, Dallas, Fort Worth, and San Antonio all have EPI correlations above 0.8 with one another as do Houston and Corpus Christi. These groupings are somewhat different than pairs associated with geographic proximity. Finally, Florida MSAs show a pattern of industrial similarity that is definitely in conflict with geographic proximity. Jacksonville and Fort Lauderdale form a close industrial pair as do Orlando and Tampa.

Having identified MSAs within each state that are industrially similar and geographic proximity being obvious, we determine which cities within each state are cyclically similar by applying the k-means clustering technique used by Crone (2003) to cluster state indexes into cyclically similar regions. The measure of dissimilarity of observations in a cluster is the squared Euclidean distance from the mean of the cluster. The number of clusters in each state was set equal to two in order to partition MSAs based
on the most fundamental separation criterion. Finally, the initial selection of clusters was randomized and the clustering problem solved 100 times. In the case of this simple clustering exercise, the solutions were robust to initial selection.

Geographic proximity was of primary importance in two states, California and New York. For California, the three northern MSAs, Oakland, Riverside, and San Francisco, formed one cluster while Los Angeles, Orange County, and San Diego formed the other. Industrial similarity would have grouped San Francisco with the southern MSAs. For New York, industrial similarity would also have grouped Nassau-Suffolk and New York City, leaving Rochester and Syracuse as the other cluster. Industrial similarity was dominant in the Florida clusters where Jacksonville and Fort Lauderdale were grouped and Orlando and Tampa formed the other cluster. Clearly geographic proximity played no role in cyclical similarity in Florida. For Texas Dallas consistently formed a singleton cluster and the other four cities, Fort Worth, Corpus Christi, Houston and San Antonio, constituted the other cluster. Recall that Texas was a state where industrial similarity was rather inconsistent with proximity but Dallas and Fort Worth were identified as similar by either criterion. In the case of Texas MSAs neither geography nor industrial similarity was predictive of cyclical similarity. This may be taken either as a commentary on the precision with which the state series was estimated in Texas or on the importance of other determinants of cyclical similarity.

V. Summary and Conclusions

We began with three objectives: to demonstrate that the Stock/Watson single dynamic factor approach could be extended to MSAs to determine if MSAs with a given state had dissimilar cycles; and, if within-state differences were found, to evaluate geographic proximity versus industrial similarity as a basis for similarity of cycles across MSAs. Finding MSA-level time series measures was a struggle and involved creation of a new index, the EPI. Nevertheless reasonable estimates of state series for larger MSAs were achieved. Differences in amplitude and periodicity of cycles across MSAs were observed. Using an index of industrial similarity based on pairwise correlation of EPIs across MSAs, we were able to determine the extent to which clusters of MSAs within each state were associated with proximity versus industrial similarity. In some instances (California), it was apparent that geographic proximity was more predictive but in other cases (Florida) industrial similarity was more important. For New York, both effects were reinforcing and, finally, in Texas, neither proximity nor industrial structure was predictive of cyclical similarity.
Clearly, the study of MSA-level cycles will benefit from coming data improvements. Over time, some series that did not go back far enough to be useful in this paper will become available. We are confident that geographic disaggregation beyond the state level will both produce results that are hidden by state-level aggregation and allow researchers to test interesting hypothesis about determinants of cyclical variability using panels of MSAs.
References


Table I: Pairwise correlation Matrix: MSA Export Price Index

<table>
<thead>
<tr>
<th></th>
<th>California</th>
<th>Florida</th>
<th>New York</th>
<th>Texas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Los Angeles</td>
<td>Oakland</td>
<td>Orange County</td>
<td>Riverside</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>3,450</td>
<td>2,857</td>
<td>4,349</td>
<td>3,323</td>
</tr>
<tr>
<td>Oakland</td>
<td>3,450</td>
<td>2,412</td>
<td>3,827</td>
<td>2,534</td>
</tr>
<tr>
<td>Orange County</td>
<td>2,857</td>
<td>2,412</td>
<td>3,148</td>
<td>1,873</td>
</tr>
<tr>
<td>Riverside</td>
<td>4,349</td>
<td>3,827</td>
<td>3,148</td>
<td>3,222</td>
</tr>
<tr>
<td>San Diego</td>
<td>3,323</td>
<td>2,534</td>
<td>1,873</td>
<td>3,222</td>
</tr>
<tr>
<td>San Francisco</td>
<td>2,623</td>
<td>2,656</td>
<td>2,436</td>
<td>3,902</td>
</tr>
<tr>
<td>San Jose</td>
<td>7,764</td>
<td>6,011</td>
<td>5,197</td>
<td>7,874</td>
</tr>
</tbody>
</table>
Figure I

California CEI Indexes

% Change in CEI Index

-3 -2 -1 0 1 2 3


Los Angeles
Oakland
Orange County
Riverside
San Diego
San Francisco
CA-State
Figure II

Northern California CEI Indexes

% Change in CEI Index

Oakland
Riverside
San Francisco
CA-State
Figure III

Southern California CEIIndexes

% Change in CEI Index

Los Angeles
Orange County
San Diego
CA-State
Figure IV

Florida CEI Indexes

- Ft. Lauderdale
- Jacksonville
- Orlando
- Tampa
- FL-State
Figure V

New York CEI Indexes

% Change in CEI Index

Nassau
New York City
Rochester
Syracuse
NY-State
Figure VI

Texas CEI Indexes

% Change in CEI Index

Corpus Christi
Dallas
Ft. Worth
Houston
San Antonio
TX-state
Figure VII

Texas CEI Indexes
without Dallas or Ft. Worth

% Change in CEI Index

Corpus Christi
Houston
San Antonio
TX-state