REGIONAL ECONOMIC ISSUES

Working Paper Series

Estimating Monthly Regional Value Added by Combining Regional Input with National Production Data

Philip R. Israilevich and Kenneth N. Kuttner

FEDERAL RESERVE BANK
OF CHICAGO

WP - 1992 / 8

Estimating Monthly Regional Value Added by Combining Regional Input with National Production Data

Philip R. Israilevich Kenneth N. Kuttner*

April 17, 1992

1 Introduction

Since 1970, many regional Federal Reserve Banks have published manufacturing production indices to monitor district manufacturing activities. Currently, five of the twelve Banks (Chicago, Cleveland, Philadelphia, Richmond, and Dallas) regularly publish regional manufacturing indices, using them to track fluctuations in current activity as well as long-run trends.

The main challenge in the construction of a regional manufacturing index has always been to combine dissimilar data into a single monthly series. Value added (VA), the most commonly used as a proxy for a region's manufacturing output, is available at an annual frequency, and only with a considerable lag. As a result, a monthly manufacturing index must rely on some form of interpolation to track production fluctuations within the year. Most of the published indices utilize monthly data on regional labor and energy inputs for this purpose.

None of the regional manufacturing indices published to date has found an entirely satisfactory solution to this interpolation problem. Existing methods all begin with an estimate of the *annual* production function, generating the monthly index by inserting monthly energy and labor data.[1] Their specific approaches to estimating the production function fall into two categories: parametric and nonparametric.

The parametric approach is based on an econometric estimate of the production function fitted to annual data, with labor and energy data used in the generation of the monthly output series. Because the model is estimated on annual data, the small number of observations (at most 15) makes the estimates very imprecise. In addition,

*Research Department, Federal Reserve Bank of Chicago. The authors thank David Weiss for his assistance in preparing the data.

[1]Fomby (1986), Wozniak (1990) and Israilevich, *et al.* (1989) provide surveys of the existing measures.

by limiting the number of estimable parameters, this approach effectively forces the factor shares to remain constant over the sample.

The nonparametric approach relaxes the assumption that the production function parameters remain constant over the sample. In this approach, the observed factor shares, calculated from annual data, are used in place of econometric estimates of the production function. To prevent the index from changing abruptly from the end of one year to the beginning of the next, the monthly weights are then interpolated between years. However, as noted by Fomby (1986), because the weights are functions of the prices of labor and capital, the nonparametric approach may make the regional business cycle appear smoother than it actually is.

One serious problem common to both approaches is that neither constrains the estimated monthly VA series to sum to the annual figure. As a result, even though these techniques may produce plausible monthly indices, their annual growth rates may differ significantly from the *observed* growth rate of annual VA. A second shortcoming of these methods is their failure to exploit monthly data in estimating the production function. In addition, both are subject to approximation error from their logarithmic specifications.

This paper proposes a new method for estimating regional VA based on a parametric *monthly* production function. This allows monthly data on factor inputs to be used in the estimation of the production function, and exploits the link between regional and national economic fluctuations.

The key to this method is to model monthly output as a stochastic latent variable that depends on observable energy and labor inputs, and is correlated with national output as measured by the Federal Reserve Board's monthly index of Industrial Production (IP). This does not require that IP serve as a perfect indicator of regional output — only that fluctuations in the two series are correlated. A useful by-product of the technique is a set of measures of the degree and nature of the linkages between regional and national economic activity.

Another important feature of this technique is its imposition of consistency between the estimated monthly index and the observed annual VA series. Annual VA is modeled as the sum of the underlying monthly series; constraining the sum of the monthly log differences to equal the log difference of annual value added produces an index consistent (up to a small approximation error) with the annual VA data.[2]

A third distinguishing feature of this method is the estimation technique, in which the model's parameters are estimated jointly along with the unobserved monthly VA series. Applying the Kalman filter to the model, an example of a "Multiple Indicator Multiple Cause" or MIMIC model, yields an estimate of the monthly VA and facilitates the maximum-likelihood estimation of its parameters. Similar methods are applicable to a variety of other problems in regional economics in which regional

[2]In this respect, the technique resembles Milton Friedman's (1962) technique of interpolation by related series. However, because the annual VA data are sums rather than "snapshots" of the underlying series, the problem is what Friedman called *distribution* rather than interpolation.

data are reported annually, while related national data are reported monthly. These techniques represent an improvement over traditional econometric methods, such as regression models, which do not readily permit the combination of data of differing frequencies.

The results presented below apply this method to 15 two-digit SIC industries from the Seventh Federal Reserve district (which includes all or part of Illinois, Iowa, Indiana, Wisconsin, and Michigan) using data from 1973 through 1989. A subsequent section of the paper modifies the model to estimate monthly VA on the basis of factor input and IP data over the 1990–91 period, in which annual VA data are not yet available.

2 Modeling monthly regional value added

2.1 The production model

The model's foundation is a monthly production function. Following much of the parametric literature, we use a first-differenced logarithmic version of a Cobb-Douglas specification,

$$\Delta x^7_{t,s} = \gamma + \phi \Delta e^7_{t,s} + \theta \Delta l^7_{t,s} + \eta_{t,s}, \tag{1}$$

where ϕ and θ are the (constant) shares of energy and labor, and γ is the (monthly) rate of Hicks-neutral technical change. In the absence of reliable capital stock data, energy usage can be thought of as a proxy for utilized capital. The η disturbance represents a stochastic shock to the production function, or a change in an omitted factor of production.

Because we will later introduce annual data, every variable measured at a monthly frequency is indexed by two subscripts. The first subscript (usually t) denotes the year, and the second (usually s) indexes the month of year t. For example, $\Delta x^7_{3,6}$ refers to June of the third year of the sample.

Without monthly output data, direct estimation of equation 1 is impossible. Fortunately, the availability of indicator variables — series that are linear functions $\Delta x^7_{t,s}$ — allow the indirect estimation of the production model, as described below.

2.2 Annual value added

While monthly regional output data do not exist, *annual* observations of Seventh-District VA are available. Because these data correspond to the sum of the underlying monthly series, they can be put to use in estimating a production function at a monthly frequency. Specifically, this relationship between the monthly and the annual data implies a linear relation between the annual percentage change in year t and the prior 24 months' percentage changes. Letting upper case letters denote the untransformed Seventh-District VA, the annual and the monthly data are related in

Table 1: Summary of Notation

x_t^7	Log of 7th district value added, annual data
$\Delta^A x_t^7$	$\equiv x_t^7 - x_{t-1}^7$ (annual difference)
$x_{t,s}^7$	Unobserved monthly log of 7th district value added series, month s of year t
$x_{t,s}^N$	Log of FRB industrial production index, month s of year t
$\Delta x_{t,s}^N$	$\equiv x_{t,s}^N - x_{t,s-1}^N$ (monthly difference)
$l_{t,s}^7$	Log of 7th district labor input, month s of year t
$e_{t,s}^7$	Log of 7th district energy input, month s of year t
$\Delta x_{t,s}^7$	$\equiv x_{t,s}^7 - x_{t,s-1}^7$ (monthly difference)
$\Delta l_{t,s}^7$	$\equiv l_{t,s}^7 - l_{t,s-1}^7$ (monthly difference)
$\Delta e_{t,s}^7$	$\equiv e_{t,s}^7 - e_{t,s-1}^7$ (monthly difference)

levels through the summation

$$X_t^7 = \sum_{s=1}^{12} X_{t,s}^7.$$

In terms of percentage changes, the relation is

$$\frac{\Delta^A X_t^7}{\bar{X}^7} = \frac{\Delta^A \sum_{s=1}^{12} X_{t,s}^7}{\bar{X}^7} = \frac{1}{12}\sum_{s=1}^{12} \frac{\Delta^A X_{t,s}^7}{\frac{1}{12}\bar{X}^7},$$

where Δ^A represents the change from a year ago, and $\bar{X}^7$ is an appropriate mean value of X_t^7. The second equality expresses the year-to-year percent change in annual VA as an average of the year-ago changes in monthly VA (subject to an appropriate choice of $\bar{X}^7$).

Although this relationship holds exactly in terms of percentage changes, because the production relation is specified in terms of logarithms, we approximate the percent changes as log differences, yielding

$$\Delta^A x_t^7 = \frac{1}{12}\sum_{s=1}^{12} \Delta^A x_{t,s}^7.$$

Finally, noting that the year-ago change in monthly VA equals the sum of the past twelve months' monthly changes, we can express the log difference of annual VA as a linear combination of lagged differences of the unobserved monthly series:

$$\Delta^A x_t^7 = \frac{1}{12}\sum_{s=1}^{12}\sum_{j=0}^{11} \Delta x_{t,s-j}^7 \tag{2}$$

where $\Delta x^7_{t,s} \equiv x^7_{t,s} - x^7_{t,s-1}$. (Note that for some month $m < 1$, $x^7_{t,m} \equiv x^7_{t-1,m+12}$. When $m = 1$, for example, $\Delta x^7_{t,1} \equiv x^7_{t,1} - x^7_{t-1,12}$.) A detailed derivation of equation 2 appears in Appendix C. In estimating the monthly production model, Equation 2 is imposed with equality. While the logarithmic approximation introduces some error in the relation, in practice, its size is quite small. Statistics on the size of the error appear later in section 4.

2.3 Industrial production indices as indicators

With annual regional VA as the only measure of output, nothing is gained by using monthly data on regional inputs. Lacking any additional information at the monthly frequency, the only feasible estimate would be based on annual input data, as in the traditional parametric approach described above. Fortunately, annual VA data is not the only source of information on regional production. Because month-to-month fluctuations in regional production are correlated with nationwide economic fluctuations, the national aggregate IP index provides a second indirect measure of regional economic activity.

Statistically, national IP can be modeled as a "noisy" indicator of underlying monthly regional VA. A natural specification links the monthly fluctuations of national IP and regional VA, plus a stochastic error:

$$\Delta x^N_{t,s} = \mu + \delta \Delta x^7_{t,s} + \nu_{t,s}. \tag{3}$$

In other words, the monthly change in (log) IP varies with the (log) change in regional VA according to the constant δ. The ν error term captures non-Seventh-District movements in IP. The equation includes a constant term to allow the growth rates of national IP and regional VA to differ; μ is the difference between these growth rates on a monthly basis.

As the inclusion of national IP through equation 3 is one of the more novel aspects of this method, it requires some additional explanation. The most important point to highlight is that unlike the production model in equation 1, it does not describe any fundamental structural relationship between Seventh-District and national output. Nor is national IP in any way a determinant of Seventh-District output.

Instead, the appropriate interpretation of equation 3 is as a description of the covariation between Seventh-District and national output. Together, δ and σ_ν determine the covariance between national and regional fluctuations. The δ parameter is a scale factor, relating the *amplitudes* of regional and national fluctuations. Values of δ less than unity imply that national fluctuations are, on average, smaller than regional fluctuations. In this sense, δ is analogous to the slope coefficient in a regression equation.

The standard deviation of the disturbance in equation 3, σ_ν, determines the *correlation* between regional and national fluctuations, while saying nothing about their relative sizes, much as the standard error of a regression is related to its fit. For the purposes of extracting a monthly VA series, its practical significance is as a measure of the amount of information in national IP data which is relevant to the region. A large standard deviation would indicate that fluctuations in IP yield little information about economic activity in the region, while a small value of σ_ν would allow a more precise estimate to be made of regional VA on the basis of national IP.

Although there are similarities between equation 3 and a regression equation, it is important to stress the key distinction between them: the unobservability of its right-hand-side variable. The fact that $\Delta x_{t,s}^7$ is unobserved, combined with the stochastic error terms in equations 1 and 3, prevents the model from being expressed as a regression. Only if σ_η were zero, rendering the production equation deterministic, could one substitute equation 1 into 3 to turn it into a regression. Similarly, only a correlation coefficient of unity between national IP and regional VA (that is, $\sigma_\nu = 0$) would permit $\Delta x_{t,s}^N$ to be used in place of $\Delta x_{t,s}^7$ in equation 1, allowing it to be estimated as a regression.

A better interpretation of equation 3 is in terms of a factor model of the sort employed in the work of Norrbin and Schlagenhauf (1988), in which the sources of output fluctuations are unobserved industry- and region-specific factors.[3] According to this interpretation, δ is the factor *loading* or weight associated with Seventh-District output fluctuations. The ν disturbance represents those sources of IP fluctuations uncorrelated with the Seventh District's.

Under any of the alternative interpretations, δ and σ_ν measure distinct aspects of the linkage between the regional and national economies. Generally speaking, those industries with large δs and small σ_νs are closely linked to the nation, either because of a common dependence on aggregate activity, or because Seventh District output is a major part of the national aggregate.

2.4 Estimation strategy

Estimation of the model defined by equations 1 through 3 requires it to be recast in state-space form. This representation consists of two matrix equations: a Markovian transition equation describing evolution of the vector of latent state variables, and an the observation equation describing the relationships between the state variables and observable indicator variables. In this framework, equation 1 is the transition equation, while equations 2 and 3 are components of the observation equation. The state variable is the unobserved log difference in monthly VA, while the indicator variables are the observed log differences of Seventh-District annual VA, and the log differences of the monthly national IP series.

The most convenient way to accommodate the combination of annual and monthly data is to include in the period t state vector the log difference of Seventh-District VA for each month in years t and $t-1$, enabling the imposition of the summation condition (equation 2). Similarly, the vector of indicators includes the annual log difference of

[3] For an introduction to factor analysis, see Morrison (1976).

Seventh-District VA, as well as the twelve monthly log differences of overall IP. A complete description of the state space representation appears in Appendix A.

Once the model is written in state-space form, the Kalman filter can be applied to extract estimates of the unobserved state vector, conditional on a guess of the unknown parameters. The Kalman filter also generates a sequence of uncorrelated forecast errors, which can be used to evaluate a normal likelihood function. A maximization algorithm applied to this function yields maximum-likelihood estimates of the unknown parameters. For details, see Harvey (1981, 1989) and Watson and Engle (1983).

3 The data

Consistent with other regional manufacturing indices, value added (VA), defined as the difference between the value of shipments and intermediate purchases of goods, energy and some services, is used to measure regional manufacturing activity. This residual includes capital and labor rents and some services.[4] Regional VA is reported annually in the Annual Survey of Manufacturing (ASM), with the exception of the Census years, in which the data are reported by the Census of Manufacturing. These Census data, available every five years, include virtually the entire population, while the ASM data are based on a representative sample. The real VA series used in estimation is constructed by deflating these series with the appropriate Gross State Product deflators.

The model presented above specifies the change in VA as a function of two region-specific inputs: labor and energy. Labor accounts for both production and supervisory workers as reported by the Bureau of Labor Statistics, and is expressed in terms of hours. Energy, measured in kilowatt hours, is collected by the regional Federal Reserve banks, and is widely interpreted as a proxy for utilized capital.[5] Both input series are available monthly.

In addition to VA and factor input data, the model also uses the well-known Federal Reserve Board industrial production (IP) indices as indicators of unobserved regional economic activity. This index is constructed from a number of variables available on a monthly basis, including national aggregates of the energy and labor data discussed above. Depending on the industry, they may also utilize data on the value of shipments, as well as physical measures of output.[6]

[4]Although VA is widely accepted as the best available measure of regional activities, one alternative is the total value of shipments, which corresponds to the total value of output. The drawback to this measure is that it includes the value of materials imported to the region, which may not reflect regional economic activity. These differences complicate the comparison of regional and national economic activity; see Israilevich *et al.* (1989).

[5]Fomby (1986) discusses the rationale for these input measures.

[6]A detailed description of the industrial production indices appears in the Federal Reserve Board publication *Industrial Production, 1986 Edition*.

3.1 SIC change

Because of a change in the way computer manufacturers were assigned to SIC categories 36 (electrical machinery) and SIC38 (instruments), post-1987 data are incompatible with pre-1987 data for those industries. To internalize the changes, the two industries are combined and estimated as a single industry. Their value added, labor, and energy series are added, and a combined industrial production index was formed as a weighted average of the two reported by the Federal Reserve Board.

3.2 Adjustments to labor and energy data

One troubling aspect of working with regional input data is the significant amount of statistical noise present in the series. This noise usually takes one of two forms. The first appears to be occasional, large realizations of measurement error in levels, which, in first differences, become spikes that reverse themselves in the following month. Before using the data for estimation, any readily-identifiable two-month spikes were replaced by an average of the two months' data, thereby smoothing the series.

Another type of error consists of once-and-for-all jumps in the levels of the inputs. The strategy for dealing with these episodes is to include a dummy variable in equation 1, so that the month in question is simply ignored in the estimation. A summary of both types of adjustments appears in Appendix B.

4 Results

This section describes the results of estimating the model defined by equations 1, 2 and 3 on data from 1973 through 1989 for 15 two-digit manufacturing industries. As described above in section 2.4, the parameters were estimated via maximum-likelihood iterations in conjunction with the Kalman filter algorithm.

4.1 Parameter estimates

Parameter estimates for each of the 15 industries appear in Table 2. While the estimates' precision varies significantly between industries, the estimated production function coefficients generally correspond to economically plausible values. In only two cases, SIC 29 (petroleum) and 39 (miscellaneous durables), was the fitted θ negative, probably due to the unusually poor quality of data in those industries. The model was re-estimated for these two industries with θ set equal to zero. In most cases, the sum of ϕ and θ is not far from unity. The overall success in estimating the production function is particularly noteworthy in light of the use of energy as a proxy for capital.

All but three industries show a positive rate of technical progress, γ, although this

Table 2: Unconstrained Parameter Estimates

	Production Equation				IP Indicator			
SIC	γ	ϕ	θ	σ_η	μ	δ	σ_ν	LLF
20	0.24	0.10	0.36	2.88	0.14	0.28	0.43	52.9185
	(0.57)	(0.12)	(0.28)	(1.00)	(0.07)	(0.08)	(0.16)	
24	0.17	0.03	0.64	3.53	0.08	0.36	2.20	39.9850
	(0.29)	(0.13)	(0.45)	(0.80)	(0.24)	(0.21)	(0.51)	
25	0.02	0.32	0.09	2.58	0.07	0.61	1.00	44.5047
	(0.34)	(0.06)	(0.12)	(0.40)	(0.13)	(0.09)	(0.25)	
26	0.23	0.08	1.22	2.08	0.09	0.38	1.66	44.3320
	(0.29)	(0.16)	(0.50)	(0.72)	(0.14)	(0.19)	(0.27)	
27	−0.01	0.16	0.35	1.15	0.27	0.63	0.89	50.6536
	(0.12)	(0.09)	(0.13)	(0.30)	(0.08)	(0.17)	(0.18)	
28	0.17	0.22	0.42	2.34	0.14	0.33	1.16	47.7049
	(0.21)	(0.10)	(0.36)	(0.54)	(0.15)	(0.19)	(0.21)	
29	−0.17	0.11	0.00	9.64	0.06	0.11	1.83	41.3567
	(0.99)	(0.59)	...	(2.78)	(0.35)	(0.08)	(0.47)	
30	0.13	0.55	0.57	2.53	0.09	0.55	1.36	43.7856
	(0.33)	(0.07)	(0.12)	(0.50)	(0.13)	(0.07)	(0.43)	
32	0.01	0.46	0.45	1.08	0.14	0.45	1.68	45.5280
	(0.11)	(0.15)	(0.19)	(0.38)	(0.22)	(0.15)	(0.11)	
33	0.18	0.31	1.23	3.66	−0.06	0.56	2.45	37.3686
	(0.32)	(0.18)	(0.47)	(0.98)	(0.23)	(0.15)	(0.97)	
34	0.03	0.42	0.62	1.74	0.09	0.44	0.80	50.8644
	(0.22)	(0.10)	(0.08)	(0.32)	(0.11)	(0.04)	(0.15)	
35	0.46	0.43	0.44	2.49	0.44	0.29	1.34	46.4758
	(0.34)	(0.20)	(0.30)	(1.01)	(0.19)	(0.18)	(0.32)	
36	0.32	0.11	1.00	1.09	0.31	0.45	0.89	52.1182
	(0.11)	(0.09)	(0.20)	(0.18)	(0.06)	(0.09)	(0.09)	
37	0.23	0.48	0.87	2.73	0.16	0.40	1.60	43.5203
	(0.25)	(0.10)	(0.23)	(0.85)	(0.23)	(0.11)	(0.42)	
39	−0.08	0.24	0.00	4.80	0.11	0.38	0.97	42.7476
	(0.52)	(0.09)	...	(0.78)	(0.10)	(0.06)	(0.23)	

Notes: Standard errors are in parentheses.
See Appendix B for definitions of the SIC designations.

parameter is generally small and imprecisely estimated.[7] Similarly, most estimates of μ, the difference between the region's and the nation's growth rates, are generally positive, signifying slower growth in the Seventh District than in the country as a whole.

With one exception, the estimated values of δ range from 0.28 for SIC 20 (food) to 0.63 for SIC 27 (printing and publishing), and are statistically significant at the 10% level or better.[8] The coefficients' variation reflects Seventh-District industries' differing degrees of sensitivity to aggregate fluctuations.

In part, these coefficients depend on the size of Seventh-District production relative to the nation as whole. Although there is no theoretical reason to restrict δ to the unit interval, estimates falling in this range are implying a positive, but attenuated, relation between the region and the nation. This result is consistent with the observation that national IP comes in part from Seventh-District production, but is also subject to partially offsetting fluctuations from other regions.

Besides δ, the other parameter describing the "closeness" of the regional–national link is the standard deviation of the disturbance to equation 3, σ_ν, which is inversely related to the correlation between regional and national output. The relatively large estimated σ_ν illustrates the quantitative importance of regional factors in the national aggregate, implying that national IP is a poor substitute for the Seventh-District regional manufacturing index.

Figure 1 summarizes the overall degree of linkage between the regional and the national economies at the monthly frequency. For each industry, the horizontal axis measures the standard deviation of the disturbance associated with the national IP indicator, while the vertical axis is the ratio of δ to the region's average share of the nation's output. Industries falling in the northwestern corner of the plot are therefore those most closely linked to the national economy, controlling for the effect of the size of the region in the overall economy.

4.2 The estimated monthly series

Extracting an estimate of monthly Seventh-District real VA through 1989 on the basis of the estimated model is straightforward. "Fitted" values of the unobserved variable are delivered directly by the Kalman Filter algorithm, using monthly factor input and IP data, and annual value added data. Essentially, the Kalman Filter uses the three monthly series to construct an optimal interpolation of the annual data using the parameters estimated earlier.

[7] In none of the three industries with negative point estimates of γ is it statistically different from zero at the 20% level.

[8] The one outlier is SIC 29 (petroleum), whose small and statistically insignificant δ of 0.11 reflects that industry's small size (in the Seventh District), and its poor quality data quality.

4.2.1 Extending through 1991

Due to the two- to three-year lag in publishing regional VA data, the estimates summarized above used data only through 1989. In the absence of current VA data, generating contemporaneous estimates of monthly regional VA requires a small change to the model. This modification turns out to be straightforward, requiring only the deletion of equation 2. The parameters estimated over the 1973–89 period are inserted in the remaining two equations; conditional on those coefficients, applying the Kalman Filter to the modified model extracts an optimal estimate of $\Delta x_{t,s}^{7}$ for the sample period in which only IP and input data are available.

The top panels in Figures 2 and 3 plot the estimated Seventh-District monthly real VA series for two industries: SIC20 (food) and SIC33 (primary metals). As measured by δ and σ_{ν}, SIC33 is rather closely linked to the industry on a national level, while SIC20 is much more loosely connected. For comparison, the dashed line plots a "naïve" monthly real VA series, constructed by simply dividing the annual figure by 12.

The relative importance of national fluctuations can be seen by comparing the estimated series' deviations from the interpolated series to fluctuations in IP. For SIC33, the national index is much more informative about monthly movements in Seventh-District steel production than it is for SIC20.

4.2.2 The overall Seventh-District MMI

Because the monthly real VA series constructed above are expressed in terms of 1982 dollars, aggregation is straightforward. Figure 4 plots the sum of the monthly series, for the 15 industries analyzed. Again, the dashed line is the "naïve" series. The last observation is for October 1991.

4.2.3 Approximation error

The adding-up condition in equation 2 is only approximate, and depends on the substitution of log differences for exact percentage changes. An important question to consider is the size of the approximation error introduced by this simplification.

For the years in which annual VA data are available, computing a measure of the approximation error is straightforward. The relevant measure is based on a comparison between the year-to-year percent change in the sum of each year's estimated monthly VA, and the *actual* percent change in VA.

Table 3 summarizes the root mean squared and the mean absolute errors for each of the 15 industries analyzed. The figures are expressed in percentages, so that, in the case of SIC 33, for example, the difference between the estimated and actual annual changes is on the order of 0.36%. The largest such error is 0.56%, in the case of SIC 29. Clearly, the estimates' precision suffers little from the logarithmic approximation.

Figure 1

Measures of regional-national linkage

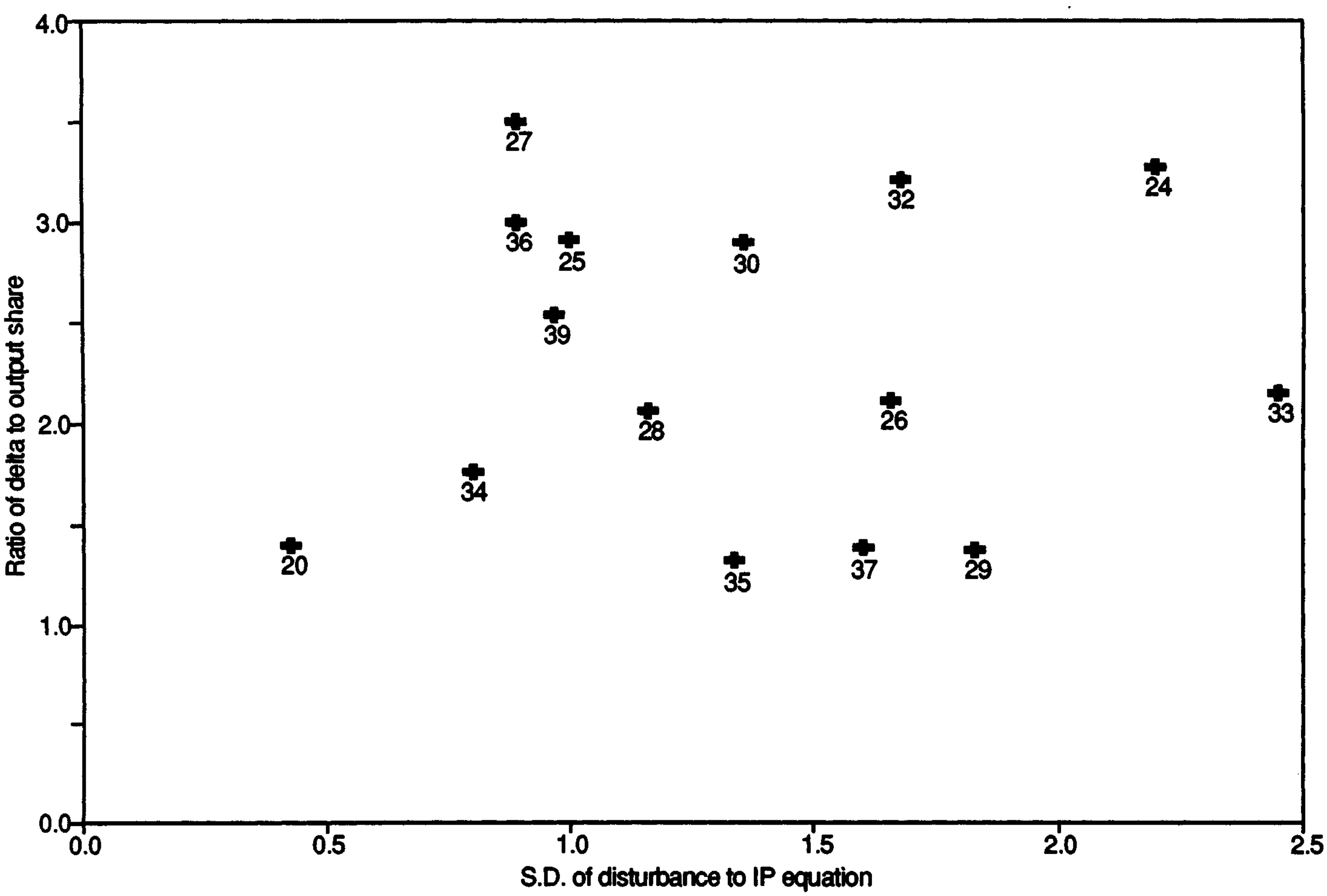

Figure 2

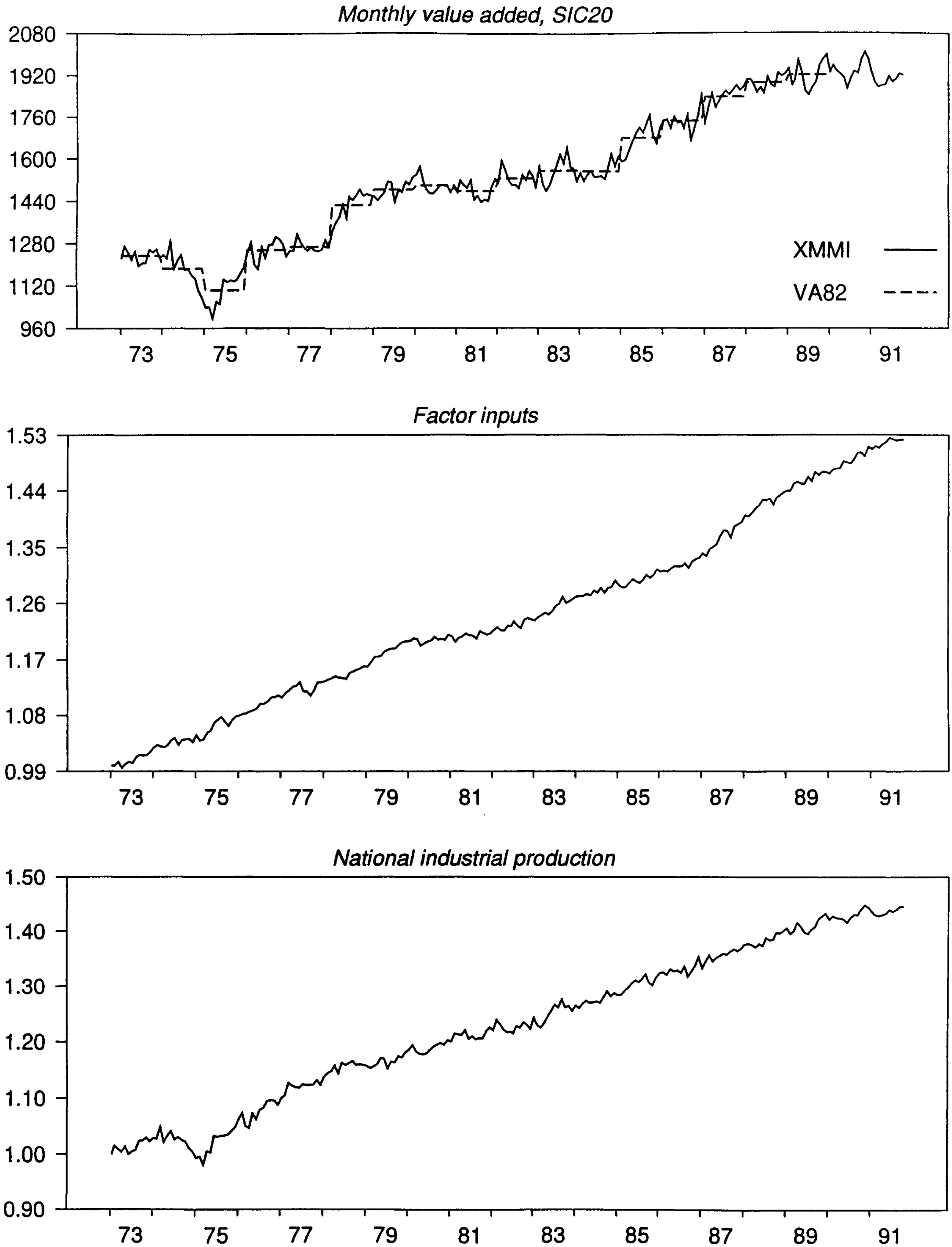
Monthly value added, SIC20
2080
1920
1760
1600
1440
1280
1120
960
73
75
77
79
81
83
85
87
89
91
XMMI
VA82
Factor inputs
1.53
1.44
1.35
1.26
1.17
1.08
0.99
National industrial production
1.50
1.40
1.30
1.20
1.10
1.00
0.90

Figure 3

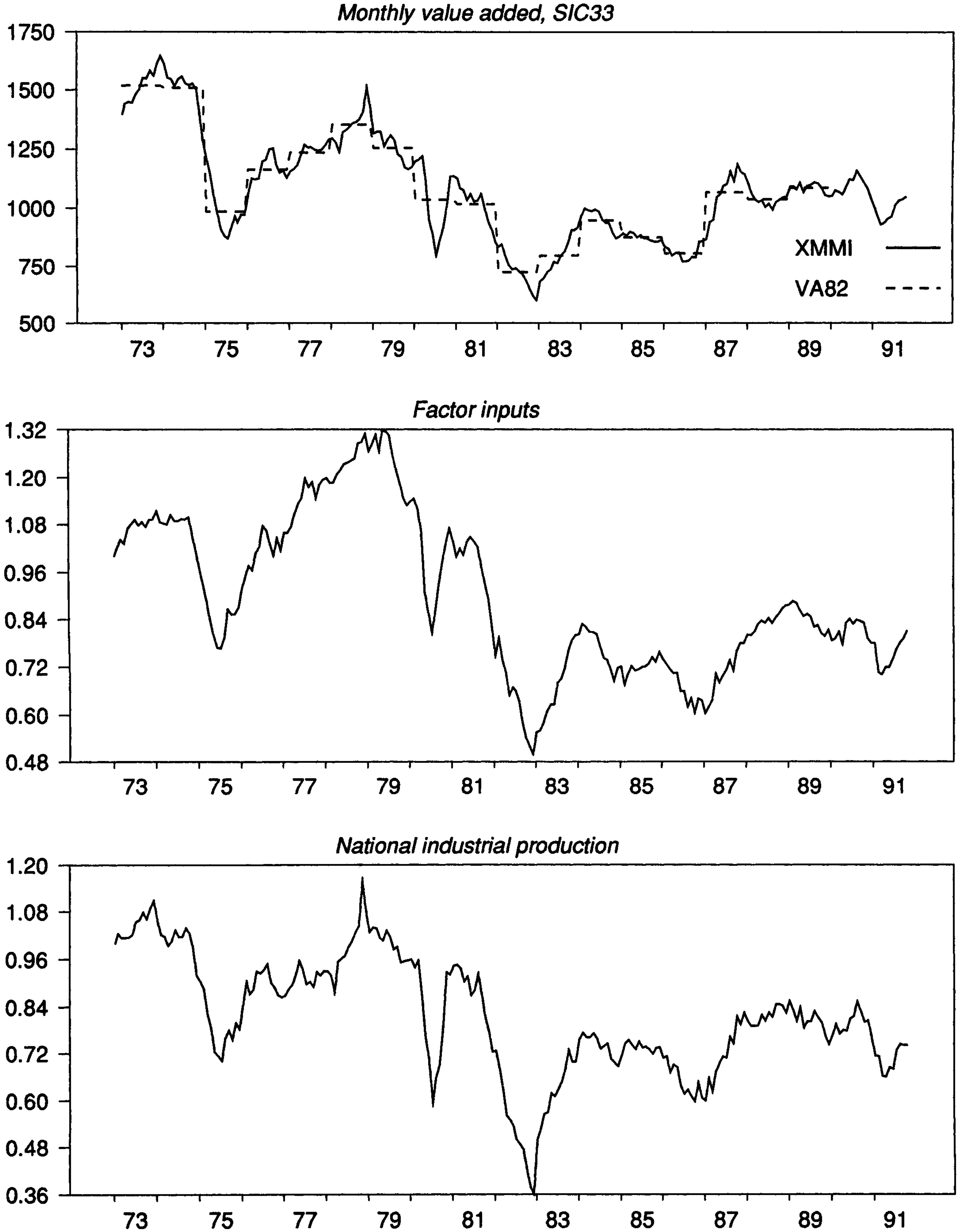
Monthly value added, SIC33
1750
1500
1250
1000
750
500
73
75
77
79
81
83
85
87
89
91
XMMI
VA82
Factor inputs
1.32
1.20
1.08
0.96
0.84
0.72
0.60
0.48
73
75
77
79
81
83
85
87
89
91
National industrial production
1.20
1.08
0.96
0.84
0.72
0.60
0.48
0.36
73
75
77
79
81
83
85
87
89
91

Figure 4

Estimated seventh district manufacturing real value added

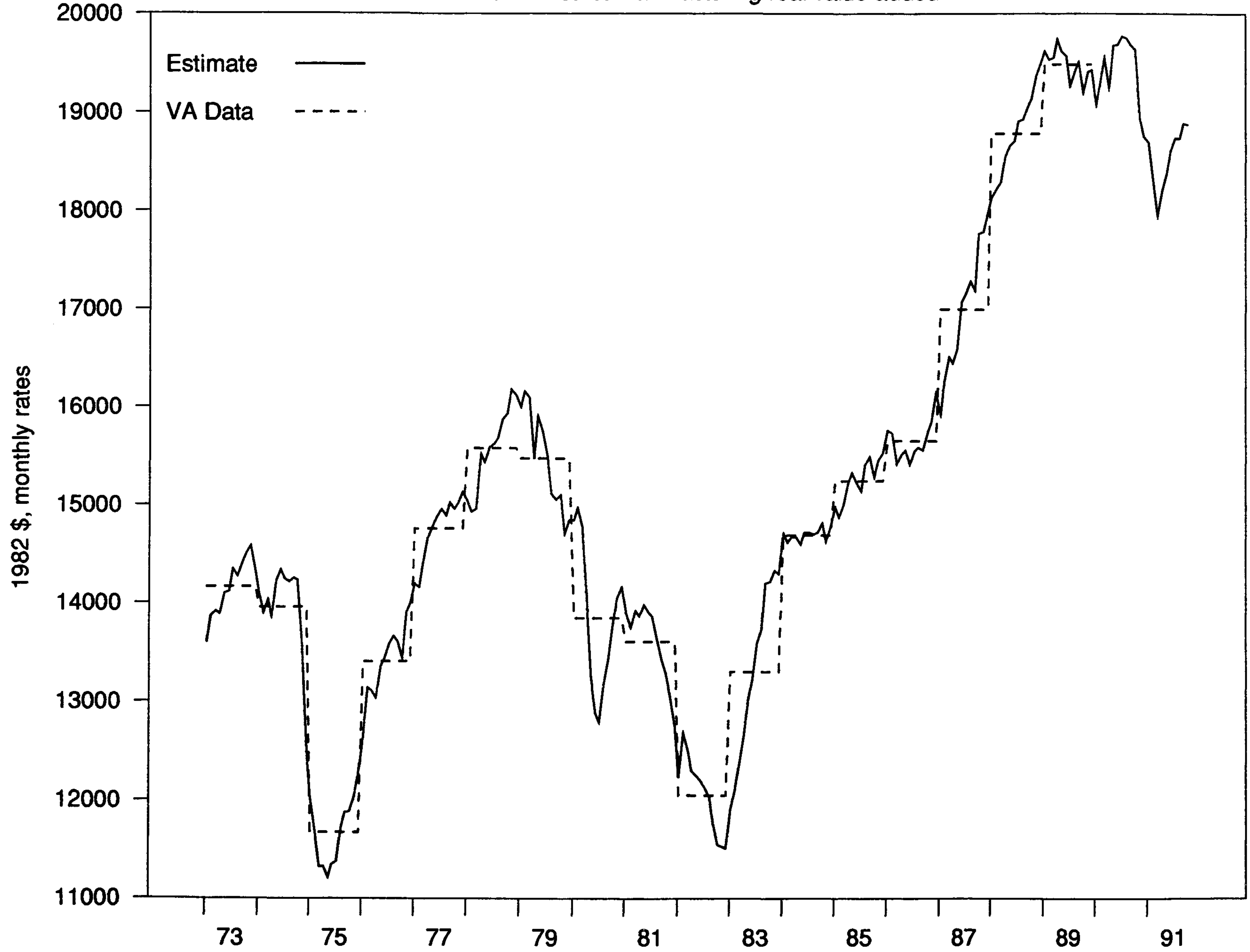

Table 3: Approximation error

SIC	Root mean squared error, %	Mean absolute error, %
20	0.04	0.03
24	0.14	0.11
25	0.08	0.06
26	0.06	0.04
27	0.02	0.01
28	0.06	0.04
29	0.56	0.35
30	0.19	0.15
32	0.05	0.04
33	0.36	0.25
34	0.13	0.09
45	0.21	0.13
36	0.04	0.04
37	0.20	0.16
39	0.11	0.09

4.3 Testing CRTS restrictions

Do the estimated production functions for these industries obey constant returns to scale? Table 4 presents parameter estimates in which the coefficients on labor and energy in equation 1 have been constrained to sum to one. The final column of the table reports likelihood-ratio test statistics for the $\theta = 1 - \phi$ constraint, asymptotically distributed as a χ_1^2 random variable.

In ten of the 15 industries, the CRTS constraint is not rejected at even the 10% level. However, in five of the 15 industries, the restriction is rejected at the 5% level or better. In three cases — 20 (food), 25 (furniture) and 27 (printing) — CRTS is rejected in favor of declining returns to scale. In the other two — 33 (primary metals) and 37 (transportation equipment) — it is rejected in favor if *increasing* returns to scale. As expected, this restriction generally yields somewhat more precise estimates of the production function parameters for those industries in which the restriction is consistent with the data.

5 Conclusion

This paper presents a new approach to the measurement of industrial production on the regional level. The main advantage to this mixed-frequency state-space approach relative to traditional parametric and nonparametric methods is its superior ability to integrate information from monthly and annual sources. At the same time, its reasonable estimates of the underlying monthly production function are consistent with economic theory and intuition. In addition, it provides useful information on relationships between regional and national economic fluctuations. As regional indices become more prominent in the economic and business literature (see Bechter *et al.* (1991)), refinements in the measurement of regional economic activity become more important.

Although the work presented here is a significant first step in the construction of an improved regional manufacturing index, at least three important tasks remain. The first is to examine its out-of-sample forecasting performance, and to evaluate the CRTS restrictions on this criterion. Second, the production model can be extended to include richer specifications of technology with time-varying factor shares. Another interesting extension is an investigation of the dynamics and lead-lag relationships between national and regional economic activity. All three topics are promising avenues for future research.

Table 4: CRTS Constrained Parameter Estimates

	Production Equation				IP Indicator				χ^2 Test
SIC	γ	ϕ	θ	σ_η	μ	δ	σ_ν	LLF	for CRTS
20	0.26	0.15	0.85	2.18	0.17	0.18	0.83	52.6736	8.33***
	(0.32)	(0.17)		(0.85)	(0.18)	(0.06)	(0.14)		
24	0.16	0.03	0.97	3.88	0.10	0.29	2.27	39.9193	2.23
	(0.31)	(0.16)		(0.89)	(0.25)	(0.09)	(0.29)		
25	−0.11	0.60	0.40	2.46	0.12	0.34	1.65	44.1544	11.91***
	(0.25)	(0.13)		(0.44)	(0.29)	(0.08)	(0.18)		
26	0.21	0.06	0.94	2.20	0.08	0.46	1.55	44.2984	1.14
	(0.27)	(0.14)		(0.56)	(0.12)	(0.10)	(0.20)		
27	−0.08	0.27	0.73	1.24	0.29	0.37	1.05	50.4915	5.51**
	(0.11)	(0.08)		(0.53)	(0.11)	(0.08)	(0.11)		
28	0.16	0.26	0.74	2.44	0.15	0.27	1.21	47.6673	1.28
	(0.22)	(0.12)		(0.48)	(0.11)	(0.08)	(0.11)		
29	−0.39	1.00	0.00	8.60	0.06	0.01	2.10	41.2918	2.21
	(0.82)	...		(1.84)	(0.44)	(0.05)	(0.10)		
30	0.16	0.50	0.50	2.44	0.07	0.58	1.33	43.7626	0.78
	(0.26)	(0.04)		(0.43)	(0.12)	(0.04)	(0.38)		
32	0.02	0.49	0.51	1.03	0.14	0.41	1.70	45.5150	0.44
	(0.11)	(0.16)		(0.34)	(0.19)	(0.13)	(0.09)		
33	0.05	0.27	0.73	3.77	−0.05	0.62	2.29	37.1531	7.33***
	(0.26)	(0.15)		(0.42)	(0.17)	(0.06)	(0.39)		
34	0.03	0.41	0.59	1.76	0.09	0.45	0.78	50.8614	0.10
	(0.23)	(0.05)		(0.28)	(0.09)	(0.04)	(0.13)		
35	0.47	0.47	0.53	2.45	0.45	0.26	1.38	46.4657	0.34
	(0.41)	(0.13)		(0.96)	(0.17)	(0.09)	(0.28)		
36	0.31	0.08	0.91	1.14	0.30	0.49	0.85	52.0947	0.80
	(0.11)	(0.05)		(0.18)	(0.06)	(0.03)	(0.08)		
37	0.19	0.35	0.65	2.65	0.15	0.48	1.53	43.3494	5.81**
	(0.29)	(0.06)		(0.70)	(0.21)	(0.06)	(0.38)		
39	−0.23	1.00	0.00	5.03	0.11	0.13	1.95	42.1715	0.58
	(0.55)	...		(1.57)	(0.21)	(0.03)	(0.13)		

Notes: Standard errors are in parentheses.
** Statistically significant at the 5% level.
*** Statistically significant at the 1% level.
See Appendix B for definitions of the SIC designations.

Appendix A: State space representation

In the transition equation, $z_t = Az_{t-1} + Gw_t + Bu_t$, z_t is the 24-dimensional state vector comprised of two years' of differenced log monthly VA. The vector of exogenous variables w_t includes a constant and the differenced logarithms of energy and labor in year t. The 12-dimensional vector of disturbances, u_t, represents shocks to the production function.

$$\underset{(24\times 1)}{z_t} \equiv \begin{bmatrix} \Delta x^7_{t,12} \\ \vdots \\ \Delta x^7_{t,1} \\ \Delta x^7_{t-1,12} \\ \vdots \\ \Delta x^7_{t-1,1} \end{bmatrix}, \quad \underset{(25\times 1)}{w_t} \equiv \begin{bmatrix} 1 \\ \Delta e_{t,12} \\ \vdots \\ \Delta e_{t,1} \\ \Delta l_{t,12} \\ \vdots \\ \Delta l_{t,1} \end{bmatrix} \text{ and } \underset{(12\times 1)}{u_t} \equiv \begin{bmatrix} \eta_{t,12} \\ \vdots \\ \eta_{t,1} \end{bmatrix}.$$

The matrices A, B and G are then defined as:

$$\underset{(24\times 24)}{A} \equiv \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \begin{matrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{matrix} & \mathbf{0} \end{bmatrix}, \quad \underset{(24\times 12)}{B} \equiv \begin{bmatrix} \boldsymbol{I} \\ \mathbf{0} \end{bmatrix} \text{ and}$$

$$\underset{(24\times 25)}{G} \equiv \begin{bmatrix} \gamma & \phi & 0 & \dots & 0 & \theta & 0 & \dots & 0 \\ \gamma & 0 & \ddots & \ddots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ \gamma & 0 & \dots & 0 & \phi & 0 & \dots & 0 & \theta \\ \mathbf{0} & & \mathbf{0} & & & & \mathbf{0} & & \end{bmatrix}.$$

The observable annual VA and national IP depend on unobservable state vector via the measurement equation, $y_t = Cz_t + Hw_t + Dv_t$, where

$$\underset{(13\times 1)}{y_t} \equiv \begin{bmatrix} \Delta^A x^7_t \\ \Delta x^N_{t,12} \\ \vdots \\ \Delta x^N_{t,1} \end{bmatrix} \text{ and } \underset{(12\times 1)}{v_t} \equiv \begin{bmatrix} \nu_{t,12} \\ \vdots \\ \nu_{t,1} \end{bmatrix}.$$

The matrices C, D and H are then defined as:

$$\underset{(13\times 24)}{C} \equiv \begin{bmatrix} 1/12 & 2/12 & \dots & 12/12 & 11/12 & 10/12 & \dots & 0 \\ \delta & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \delta & 0 & \dots & 0 & 0 \end{bmatrix},$$

$$\underset{(13\times 12)}{D} \equiv \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix} \text{ and } \underset{(13\times 25)}{H} \equiv \begin{bmatrix} 0 & 0 & \dots & 0 \\ \mu & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mu & 0 & \dots & 0 \end{bmatrix}.$$

Appendix B: Data

Table 5: Seventh-District output shares and SIC codes

SIC	Industry	Average share
20	foods	0.20
24	lumber and wood products	0.11
25	furniture and fixtures	0.21
26	paper and paper products	0.18
27	printing and publishing	0.18
28	chemicals and products	0.16
29	petroleum products	0.08
30	rubber and plastic products	0.19
32	clay, glass & stone products	0.14
33	primary metals	0.26
34	fabricated metal products	0.25
35	nonelectrical machinery	0.22
36*	electrical machinery & instruments	0.18
37	transportation equipment	0.29
39	miscellaneous	0.15

Note: SIC 36* combines SICs 36 and 38; see text.

Table 6: Summary of data adjustments

SIC	Dummy variables	Data replaced by average
26	January 1981 October 1987	
27	April 1980 September 1985 June 1987	
28		Labor: April–May 1979
29	February 1980 January 1984 January 1988	Energy: January–February 1973 Labor: January–February 1975 Energy: September–October 1981 Energy: May–June 1985 Labor: June–July 1985 Energy: March–May 1987 Energy: April–June 1989
32	June 1987 January 1988	
33		Labor: August–September 1974 Energy: August–September 1976 Energy: December–January 1984–85
34		Labor: April–May 1979 Energy: September–October 1987
35		Energy: December–January 1979–80 Energy: December–January 1984–85
37		Energy: October–November 1976 Labor: April–May 1979 Energy: April–May 1984 Labor: January-February 1990 Energy: December–January 1990–91
39	October 1987	Labor: October–November 1973 Labor: August–September 1976 Energy: January–February 1983

Appendix C: Deriving the adding-up condition

This appendix sketches a demonstration of equation 2, and derivation of the arithmetic sequence of coefficients in the first row of matrix C in Appendix A.

Neglecting the 1/12 coefficient, setting $t = 2$ and omitting the 7 superscript, the right-hand side of equation 2,

$$\sum_{s=1}^{12}\sum_{j=0}^{11} \Delta x_{t,s-j},$$

can be expanded in tabular form (starting from the end of the series) as:

Repeated terms				
1	$x_{2,12} - x_{2,11}$			
2	$x_{2,11} - x_{2,10}$	$x_{2,11} - x_{2,10}$		
3	$\vdots$	$x_{2,10} - x_{2,9}$	$\cdots$	
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
12	$x_{2,1} - x_{1,12}$	$\vdots$	$\ddots$	$x_{2,1} - x_{1,12}$
11		$x_{1,12} - x_{1,11}$	$\cdots$	$x_{1,12} - x_{1,11}$
$\vdots$			$\ddots$	$\vdots$
1				$x_{1,2} - x_{1,1}$
Sum:	$x_{2,12} - x_{1,12}$	$x_{2,11} - x_{2,10}$	$\cdots$	$x_{2,1} - x_{1,1}$

The key simplification comes from the cancellation of all the intermediate terms in the column sums, leaving only year-ago differences. The last row of the table can be written in terms of simple summations:

$$= \sum_{s=1}^{12} x_{2,s} - \sum_{s=1}^{12} x_{1,s} = 12\Delta^A x_2,$$

which equals (up to the 1/12 factor) the left-hand-side of equation 2, verifying the equality.

Summing the table *across columns* and collecting terms produces an expression entirely in terms of monthly differences, whose coefficients (divided by 12) appear in the first row of matrix C.

References

Bechter, D.M., Z. Kenessey, F. Siegmund, and R.D. Whitman (1991), "A Total Production Index for Washington, D.C.," Federal Reserve Bank of Richmond *Economic Review* **77**, May/June, pp.13–22.

Fomby, T.M. (1986), "A Comparison of Forecasting Accuracies of Alternative Regional Production Index Methodologies," *Journal of Business and Economic Statistics* **4**, pp. 177–86.

Friedman, Milton (1962), "The Interpolation of Time Series by Related Series," *Journal of the American Statistical Association* **57**, December, pp. ?–?.

Harvey, A.C. (1981), *Time Series Methods.* New York: Wiley.

Harvey, A.C. (1989), *Forecasting, Structural Time Series Models, and the Kalman Filter.* Cambridge: Cambridge University Press.

Industrial Production, 1986 Edition. Washington, D.C.: Board of Governors of the Federal Reserve System.

Israilevich, P.R., R.H. Schnorbus, and P.R. Schneider (1989), "Reconsidering the Regional Manufacturing Indexes," Federal Reserve Bank of Chicago *Economic Perspectives* **XIII**, July/August, pp. 13–21.

Morrison, D.F. (1976), *Multivariate Statistical Methods.* New York: McGraw-Hill.

Norrbin, S.C. and D.E. Schlagenhauf (1988), "An Inquiry in the Sources of Macroeconomic Fluctuations," *Journal of Monetary Economics* **22**, pp. 43–70.

Watson, M.W. and R.F. Engle (1983), "Alternative Algorithms for the Estimation of Dynamic Factor, MIMIC, and Varying Coefficient Regression Models," *Journal of Econometrics* **23**, pp. 385–400.

Wozniak, G.D. (1990), "Manufacturing Output Indexes for a Metropolitan Area: A Look at New Economic Indicators," *Growth and Change* , Summer, pp. 61–76.