Journal of Monetary Economics 7 (1981) 151–174. North-Holland Publishing Company

A NEW APPROACH TO DECOMPOSITION OF ECONOMIC TIME SERIES INTO PERMANENT AND TRANSITORY COMPONENTS WITH PARTICULAR ATTENTION TO MEASUREMENT OF THE 'BUSINESS CYCLE'*

Stephen BEVERIDGE

University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Charles R. NELSON

University of Washington, Seattle, WA 93195, USA

This paper introduces a general procedure for decomposition of non-stationary time series into a permanent and a transitory component allowing both components to be stochastic. The permanent component is shown to be a random walk with drift and the transitory or cyclical component is a stationary process with mean zero. The decomposition methodology, which depends only on past data and therefore is computable in 'real time', is applied to the problem of measuring and dating business 'cycles' in the postwar U.S. economy. We find that measured expansions and contractions are of roughly equivalent duration and that our dating of cyclical episodes tends to lead the traditional NBER dating and, to a lesser extent, the 'growth cycle' chronology of Zarnowitz and Boschan (1977).

1. Introduction

The idea that cyclical or transitory movements can be observed in economic time series and can be separated from trend or permanent components is a very old one and has played an important role in shaping our thinking about economic phenomena. The traditional application of the concept is, of course, to the 'business cycle'. In their classic work on the subject, Burns and Mitchell (1946) have as a specific objective the *dating* of cyclical episodes rather than numerical measurement of individual cyclical movements. Their approach emphasized identification of 'turning points', defined to be points in time when a cross-section of economic indicators changed direction from positive to negative, or vice versa. Subsequent dating of business cycles by the National Bureau of Economic Research has followed the methodology of Burns and Mitchell. A number of objections to

*The authors wish to thank Charles Plosser and anonymous referees for helpful suggestions but acknowledge that responsibility for remaining errors is entirely their own. Nelson's participation in this joint research was supported in its later stages by the National Science Foundation through Grant no. SOC-7906948.

the method, both conceptual and practical, can be put forward. On the conceptual side, it is not clear in the context of a growing and perhaps inflationary economy that to declare a cyclical 'downturn' one necessarily should require that indicators show actual declines. To put it another way, if the underlying trend in a time series is strongly positive, then downturns in its cyclical component may occur without any negative change appearing in the series itself. Further, even if a downturn appears in the series it will generally lag the downturn in the cyclical component. From a practical point of view, particularly if one regards counter-cyclical policy to be feasible, a serious shortcoming arises from the fact that turning points are detected only with benefit of hindsight. In this regard it might be more useful to attempt numerical measurement of cyclical movements to provide an on-going record of cyclical movements as they develop.

A number of approaches to numerical measurement of the business cycle have been suggested, almost all of them *ad hoc* in nature. One method which has enjoyed popularity with students of the business cycle assumes that trend is a deterministic function of time, usually a polynomial [for example see Fellner (1956)]. The cycle component emerges in this method as a residual from the trend line. A rather unsatisfactory implication of this approach is that the long-run evolution of the time series is deterministic and therefore perfectly predictable. If in fact the changes in economic series are a random process in a statistical sense, then the deviation of the series from any deterministic path will grow without bound. Further, to impose a deterministic time trend when one is not in fact present may severely distort the apparent statistical properties of the resulting cycle as shown recently by Nelson and Kang (1981).

Friedman's (1957) decomposition of measured income into 'permanent' and 'transitory' components may be interpreted as a trend-cycle decomposition in a behavioural context. For empirical purposes Friedman (1957, pp. 142–144) suggested that the permanent component might be represented as a geometric distributed lag on past incomes which, as Muth (1960) demonstrated, corresponds to the optimal forecast of income if income is generated by a particular stochastic process. An important virtue of such an approach is its freedom from determinism, although to interpret 'permanent' as 'expected' requires rather strong prior assumptions about the stochastic structure of the series in question. Another virtue is the fact that the computation of the components depends only on one-sided filtering of the data, that is, only *past* observations are needed. There have been, to our knowledge, no studies of the utility of exponential smoothing techniques (of which Friedman's procedure is the simplest case) for measurement of business cycles. Our own technique, however, will be seen to proceed rather in that general direction though in a much more general framework which includes exponential smoothing as a special case.

Nerlove (1967) has approached the problem of decomposition as one of signal extraction, making use of a theorem due to Whittle (1963). The fundamental difficulty with the signal extraction approach is that it requires prior knowledge of the particular stochastic processes generating the unobserved components, and tractable solutions appear to be limited to rather simple cases.

Working in the traditional setting of business cycle analysis, Mintz (1969, 1972) has considered the problem of measuring and dating growth cycles: cyclical episodes for an economy where absolute downturns are rare and recession is marked only by a letup in the rate of expansion. How can one make the 'invisible' cycles in such an economy 'visible' by statistical procedures? Two alternative measures of cycle in this setting are suggested by Mintz. One is referred to as the 'deviation cycle' defined to be the residual from a centered, seventy-five month moving average of the data. The same weights are used to form the moving average regardless of the series being analyzed. At the ends of the series the missing thirty-seven observations required by the centered moving average are supplied by extrapolation of the rate of change observed during the adjoining twenty-four months. The second definition offered by Mintz, referred to as 'step cycles', focuses on fluctuations in rates of change. A 'downturn' is defined to be the endpoint of a period of relatively rapid growth and an 'upturn' as the endpoint of a period of relatively low growth. Exact dating of the step cycles proceeds via maximization of the *variance* of mean changes over each tentative breaking point between 'high' growth and 'low' growth. For German data there was strong coincidence between the alternative measures of cycle, although Mintz concluded that for final dating of cycles the deviation measure was preferred. Applying the same methods to U.S. data, Mintz found that discrepancies relative to the traditional NBER dates were all due to their resulting in somewhat earlier turning points at cyclical peaks. On a theoretical level, we might be concerned about using the same centered moving average as used to compute trend for all series — ideally the procedure for extraction of trend should be appropriately tailored to the stochastic properties of each series considered. On a practical level, the centered moving average trend presents a serious problem for studying on-going developments in indicator series since future observations are unavailable for inclusion in the average. To simply extrapolate recent past rates of change into the future to fill the gap may well result in missing turning points at the time they occur, even though they become apparent with hindsight.

The methodology for measuring cyclical movements which we propose is based on the fact (proven in section 2) that any time series which exhibits the kind of homogeneous non-stationarity typical of economic time series can be decomposed into two additive components, a stationary series and a pure random walk. The stationary part, which we call the cyclical component, is defined to be the forecastable momentum in the series at each point in time. The random walk is simply the mid-point of the predictive distribution for the future path of the original series. Application of this technique begins with investigation of the stochastic structure of each series and then exploits the particular structure of each to arrive at the appropriate filters. The associated computational procedures are completely operational. Since our measurement of the cyclical component depends only on *past* data it may be performed in 'real time' to monitor business developments. We show how our cyclical component may be used to date cyclical turning points and to judge the severity of economic contractions by reference to postwar U.S. experience. We also compare our cyclical chronology with those of the NBER and the 'growth cycle' chronology of Zarnowitz and Boschan (1977).

2. Decomposition of ARIMA process into permanent and transitory components

A large number of studies over the last decade have shown that many economic time series are well represented by the class of homogeneous non-stationary 'ARIMA' processes for which the first differences are a stationary process of autoregressive-moving average form. Such processes may appear to exhibit 'trend' when they drift persistently upward, but they are in fact the accumulation of changes which in general may be autocorrelated and have a positive mean value. We will assume that the data in hand are non-seasonal since seasonal adjustment logically precedes business cycle analysis. The applicability of our conceptual framework to the problem of seasonal adjustment will be explored in a future paper.

We denote the observations on a non-stationary series by z_t and its first differences by w_t so that $w_t = z_t - z_{t-1}$. If the w's are stationary in the sense of fluctuating around a fixed mean with stable autocovariance structure then the decomposition theorem due to Wold (1938) implies that w_t may be expressed as

$$w_t = \mu + \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \dots, \tag{1}$$

where μ is the long-run mean of the w series, the λ_i are constants, and the ε's are uncorrelated random disturbances with mean zero and variance σ^2. The ε's are often referred to as 'innovations' since they are the part of w_t and z_t which is not predictable from the past. We note that many economic time series require transformation to natural logs before the first differences exhibit stationarity, so the w's then are continuous rates of change.

Our concept of the decomposition of z is motivated by considering the relation of the current value z_t to the forecast profile for future z's. In our framework, the forecast profile takes the place of a deterministic trend as the benchmark for the location of the series and therefore for measuring the cyclical component. The expectation of z_{t+k} conditional on data for z through time t is denoted by $\hat{z}_t(k)$ and is given by

$$\begin{aligned}\hat{z}_t(k) &= E(z_{t+k} \mid \ldots, z_{t-1}, z_t) \\ &= z_t + E(w_{t+1} + \ldots + w_{t+k} \mid \ldots, w_{t-1}, w_t) \\ &= z_t + \hat{w}_t(1) + \ldots + \hat{w}_t(k),\end{aligned} \tag{2}$$

since the z's can be expressed as accumulations of the w's. Now from (1) it is easy to see that the forecast of w_{t+i} at time t is

$$\begin{aligned}\hat{w}_t(i) &= \mu + \lambda_i \varepsilon_t + \lambda_{i+1}\varepsilon_{t-1} + \ldots \\ &= \mu + \sum_{j=1}^{\infty} \lambda_j \varepsilon_{t+i-j},\end{aligned} \tag{3}$$

since future disturbances ε_{t+1} are unknown but have expectation zero. We are assured of convergence of summations $\sum \lambda_i$ by the stationarity of w [see Box and Jenkins (1976, pp. 49–50)]. Substituting (3) into (2) and gathering terms in each ε_t we have

$$\hat{z}_t(k) = k\mu + z_t + \left(\sum_1^k \lambda_i\right)\varepsilon_t + \left(\sum_2^{k+1} \lambda_i\right)\varepsilon_{t-1} + \ldots . \tag{4}$$

If we now consider very long forecast horizons we have approximately

$$\hat{z}_t(k) \simeq k\mu + z_t + \left(\sum_1^\infty \lambda_i\right)\varepsilon_t + \left(\sum_2^\infty \lambda_i\right)\varepsilon_{t-1} + \ldots \tag{5}$$

by virtue of the convergence of $\sum \lambda_i$. It is now apparent that the forecast profile is asymptotic to a linear function of forecast horizon k with slope equal to μ, the rate of drift of the series, and a 'level' (algebraically the intercept) which itself is a stochastic process. It is natural to interpret this level as the *permanent* or *trend component* of z_t. Denoting this level by $\bar{z}_t$ we have

$$\bar{z}_t = z_t + \left(\sum_1^\infty \lambda_i\right)\varepsilon_t + \left(\sum_2^\infty \lambda_i\right)\varepsilon_{t-1} + \ldots . \tag{6}$$

To prove that $\bar{z}_t$ is a random walk with rate of drift μ, we need only demonstrate that the first difference of $\bar{z}_t$ is μ plus a non-autocorrelated noise. Upon taking first differences in (6) we have

$$\bar{z}_t-\bar{z}_{t-1}=w_t+\left(\sum_1^{\infty}\lambda_i\right)\varepsilon_t-(\lambda_1\varepsilon_{t-1}+\lambda_2\varepsilon_{t-2}+\ldots), \tag{7}$$

which in view of eq. (1) for w_t reduces to

$$\bar{z}_t-\bar{z}_{t-1}=\mu+\left(\sum_0^{\infty}\lambda_i\right)\varepsilon_t, \qquad \lambda_0\equiv 1. \tag{8}$$

Since ε_t is serially random by definition, the permanent component $\bar{z}_t$ is indeed a random walk with rate of drift equal to μ and a non-autocorrelated innovation equal to $(\sum_0^{\infty}\lambda_i)\varepsilon_t$. To summarize, we have found that the forecast profile approaches a linear path as it is extended into the indefinite future and we have defined the permanent component of a series as the value the series would have if it were on that long-run path in the current time period. The permanent component is then the long-run forecast of the series adjusted for its mean rate of change and we have shown that it follows a random walk.

Note that the variance of the innovation in the permanent component is $(\sum_0^{\infty}\lambda_i)^2\sigma^2$ which may be larger or smaller than σ^2, the variance of the innovation ε_t of the observed data z_t, depending on the signs and pattern of the λ's. In particular, the innovations in the permanent component will be 'noisier' in this sense than those of the observed data if the λ_i are positive which would typically be the case if the changes in z, the w's, are positively autocorrelated. For example, if the w's were first-order autoregressive with coefficient 0.5, then we would have $\lambda_i=0.5^i$ and $(\sum_0^{\infty}\lambda_i)^2=4$ so that the variance of innovations in the permanent component would be four times as large as the variance of the innovations in the observed data. Note also that $\bar{z}$ shifts from period to period in response only to the *current* innovation while past events have no effect on $\bar{z}$. Thus it is only 'new information' that triggers a revision in our measure of the permanent part of z.

The permanent component as we have defined it may be interpreted as the current observed value of z plus all forecastable future changes in the series beyond the mean rate of drift. To see this we rewrite (6) in the equivalent form

$$\bar{z}_t=z_t+\lim_{k\to\infty}\{[\hat{w}_t(1)+\hat{w}_t(2)+\ldots+\hat{w}_t(k)]-k\mu\}, \tag{9}$$

which sums all forecasted future changes and subtracts the portion due to drift. The second term on the right-hand side of (9) is the difference between

z's permanent component and its current value, in effect the *momentum* contained in z_t at time t. It is natural to regard this second component as the *transitory* or *cyclical* portion of z_t. Denoting the cyclical portion by c_t we have the definition

$$c_t = \lim_{k \to \infty} \{[\hat{w}_t(1) + \ldots + \hat{w}_t(k)] - k\mu\}$$
$$= \left(\sum_1^\infty \lambda_i\right)\varepsilon_t + \left(\sum_2^\infty \lambda_i\right)\varepsilon_{t-1} + \ldots, \tag{10}$$

the equivalence being apparent from eq. (6). Proof that c_t is a stationary process is immediate in the case that w_t is a finite order moving average process ($\lambda_i = 0$, $i > q$) and follows in the autoregressive or mixed ARIMA case from expansion of the inverse of the AR polynomial using partial fractions [Box and Jenkins (1976, p. 54)].

Our definition of c_t as the sum of forecastable future changes in z at time t implies that c will generally be positive when z is rising more rapidly than average and negative when z is rising less rapidly (or falling) since first differences of economic time series are predominantly positively autocorrelated. To illustrate, suppose that the first differences of a particular series have a representation as a first-order moving average process with mean μ. In that case we would have

$$w_t = z_t - z_{t-1} = \mu + \varepsilon_t + \theta\varepsilon_{t-1} \tag{11}$$

so that in terms of the previous notation $\lambda_1 = \theta$ and $\lambda_i = 0$ for all $i > 1$. The parameter θ is bounded $|\theta| < 1$ in general and will be positive if w_t is positively autocorrelated as is the case for most economic time series. From (8) the first difference of the permanent part of z_t is given by

$$\bar{z}_t - \bar{z}_{t-1} = \mu + [(1+\theta)\varepsilon_t], \tag{12}$$

which is a random walk with rate of drift μ and innovation $[(1+\theta)\varepsilon_t]$. The variance of this innovation is $(1+\theta)^2\sigma^2$ which will be larger than σ^2 if the changes in z are positively autocorrelated. The cyclical component of z_t is given by eq. (10) which becomes

$$c_t = \theta\varepsilon_t. \tag{13}$$

In this example, the cyclical part of z is serially random and is simply proportional to the current innovation in z and therefore also to the current innovation in the permanent part of z.

As a check on the reasonableness of our results, consider the special case of the random walk in z which is well known to characterize speculative prices and corresponds to (11) with θ set equal to zero. From (9) or (12) it is apparent that $\bar{z}_t = z_t$ and from (13) that $c_t \equiv 0$ in this special case. These results reflect the fact that a random walk contains no forecastable momentum, in other words, there are no meaningful cycles in the stock market. All price movements are permanent and current prices reflect the best estimate of the 'trend'.

In summary, the permanent/transitory or trend/cycle decomposition proposed in this paper is tailored to the stochastic structure of each time series. The permanent component is invariably a random walk with the same rate of drift as the original data and an innovation which is proportional to that of the original data. The transitory component is a stationary process which represents the forecastable momentum present at each time period but which is expected to be dissipated as the series tends to its permanent level. The remainder of the paper deals with application of this framework to measurement of cyclical components in series traditionally monitored by the NBER and comparison of the results with the postwar business cycle identified and dated by the NBER.

3. Practical implementation of the technique and application to U.S. business cycle indicators

In order to apply the results of section 2 to economic series we must have at hand a procedure for representing such series as linear stochastic processes and a general procedure for numerical computation of c_t and $\bar{z}_t$ given the data and the linear stochastic model. Operationality requires, of course, that we be able to write the linear process in terms of a finite number of parameters, a requirement which is satisfied if we confine our attention to linear processes of rational form; that is, where

$$w_t = \mu + (1 - \lambda_1 L - \lambda_2 L^2 - \dots)\varepsilon_t, \tag{14}$$

may be written in the form

$$w_t = \mu + \frac{(1 - \theta_1 L - \dots - \theta_q L^q)\varepsilon_t}{(1 - \phi_1 L - \dots - \phi_p L^p)}, \tag{15}$$

or equivalently

$$\begin{aligned} w_t = {} & \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \mu(1 - \phi_1 - \dots - \phi_p) \\ & + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \end{aligned} \tag{16}$$

in which case $\{w_t\}$ is referred to as a mixed ARIMA process of orders p and q. Stationarity requires that the roots of $(1-\phi_1 L-\ldots-\phi_p L^p)=0$ lie outside the unit circle. Since theory will, in general, not provide the appropriate values of p and q these must be supplied by analysis of the data. Box and Jenkins have emphasized use of the sample autocorrelations of a stationary series for model 'identification'. Estimation of the ϕ and θ parameters is in general non-linear, involving minimization of a non-quadratic sum of squares function if the ε_t are assumed to be normal. Rather than reviewing this material here, we refer the reader to Box and Jenkins (1976), or Nelson (1973) for the relevant details.

Given an estimated model for a particular series how may we proceed to compute c_t? We note first of all that forecasts $\hat{w}_t(1)$, $\hat{w}_t(2),\ldots$ from any given origin date are readily computed from past observations on w by direct evaluation of the conditional expectations of successive future observations implied by the estimated model [again see Box and Jenkins (1976), or Nelson (1973)]. Given that sequence of predicted changes we may apply the formula implied by (10). In practice the limit in (10) can be replaced by a large number of forecasts: we used 100 forecasts and then checked the value of $\hat{w}_t(100)-\hat{\mu}$ which in all cases was trivially small.

To summarize, we have shown that the procedure for cycle measurement is completely operational, involving the two steps: (1) identification and estimation of an ARIMA model for the first differences of the non-stationary series of interest, and (2) numerical evaluation of c_t using a practical equivalent of (10). At any given time the computed value of c_t will involve only *past* values of the observed series, avoiding the extrapolation problems associated with 'two-sided' filtering techniques such as centered moving averages.

To explore the implications of our technique we chose a set of indicators which the NBER uses in their composite indices of cyclical indicators or were components before the Zarnowitz and Boschan (1975a, b) evaluation and revision of the indices, or can be considered 'important' economic series. The 38 indicators and their NBER classifications are listed in table 1.[1] The series are observed monthly except for the five labelled quarterly.

Indicators (1) through (20) are components of the current NBER composite indices, thus, of the 22 component series, only two were not analyzed. The Stock Price Index is a leading indicator but is a pure random walk process, hence does not have a cyclical component — which perhaps underlies Paul Samuelson's quip that, 'stock prices have accurately predicted nine of the last five recessions'. And the Average Prime Rate Charged by Banks does not have a stochastic structure suitable for ARIMA modeling as

[1] All data for the historical series were obtained from various issues of the *Business Conditions Digest*.

Table 1

Classification of economic indicators.

Number	Indicator	NBER classification[a]	Stochastic cycle classification[b]
(1)	Average Workweek of Production Workers, Manufacturing	L	Lg(2)
(2)	Layoff Rate, Manufacturing	L	C
(3)	Manufacturer's New Orders for Consumer Goods and Materials, Constant Dollars	L	C
(4)	Index of Net Business Formation	L	L(3)
(5)	Contracts and Orders for Plant and Equipment, Constant Dollars	L	L(3)
(6)	Index of New Private Housing Units Authorized by Local Building Permits	L	Lg(6)
(7)	Vendor Performance, Percent of Companies Reporting Slower Deliveries	L	C
(8)	Inventories on Hand and on Order, Constant Dollars	L	C
(9)	Sensitive Price Index	L	C
(10)	Total Liquid Assets	L	Lg(9)
(11)	Money Supply (M1), Constant Dollars	L	L(8)
(12)	Number of Employees on Non-agricultural Payrolls, Establishment Survey	C	C
(13)	Index of Industrial Production, Total	C	C
(14)	Personal Income, Less Transfer Payments, Constant Dollars	C	C
(15)	Manufacturing and Trade Sales, Constant Dollars	C	C
(16)	Index of Labor Cost per Unit of Output, Total Manufacturing	Lg	Lg(10)
(17)	Manufacturing and Trade Inventories, Total Book Value, Constant Dollars	Lg	Lg(3)
(18)	Commercial and Industrial Loans Outstanding	Lg	Lg(3)
(19)	Average Duration of Unemployment	Lg	Lg(4)
(20)	Ratio Consumer Installment Debt to Personal Income	Lg	Lg(3)
(21)	Average Weekly Initial Claims for Unemployment Insurance, State Programs	L	L(2)
(22)	Manufacturer's New Orders, Durable Goods Industries	L	C
(23)	Contracts and Orders for Plant and Equipment, Current Dollars	L	L(3)
(24)	Corporate Profits after Taxes, Quarterly	L	C
(25)	Index of Price per Unit of Labor Cost, Manufacturing	L	L(2)
(26)	Index of Industrial Material Prices	C	C
(27)	Manufacturing and Trade Sales, Current Dollars	C	C
(28)	Unemployment Rate, 15 Weeks and Over	Lg	Lg(9)
(29)	Business Expenditures, New Plant and Equipment, Quarterly	Lg	Lg(2)
(30)	Bank Rates on Short-Term Business Loans, Quarterly	Lg	Lg(3)
(31)	Manufacturing and Trade Inventories, Total Book Value, Current Dollars	Lg	Lg(3)

Table 1—continued

Number	Indicator	NBER classification[a]	Stochastic cycle classification[b]
(32)	Unemployment Rate, Total	U	Lg(7)
(33)	Personal Income	U	C
(34)	Consumer Installment Debt	Lg	C
(35)	Gross National Product, Constant Dollars, Quarterly	C	C
(36)	Gross National Product, Current Dollars, Quarterly	U	C
(37)	Number of Persons Unemployed	U	C
(38)	Total Civilian Employment	U	Lg(2)

[a]L = Leading, C = Roughly coincident, Lg = Lagging, U = Unclassified.
[b]The Stochastic Cycle Classification is developed at a later point in section 3 of the paper using the proposed methodology of decomposition.

it remains unchanged for long periods of time. Indicators (21) through (34) were components before the revision of the indices and the remaining series possess economic importance.[2]

Since our analysis subsumes taking first differences, we analyze the levels of some series which are differenced by the NBER. We study Consumer Installment Debt whereas the NBER indicator is the change in this series (see footnote 2), Inventories on Hand and on Order instead of its net change, Total Liquid Assets rather than its percent change, and the Index of Sensitive Prices in place of Percent Change in Sensitive Prices. Because the last three series are not readily available, they are constructed by setting the first value equal to 100 and generating the remaining values from the change or percent change data.

Five of the indicators are stationary and therefore, by definition, devoid of the random walk component: Average Workweek of Production Workers, Manufacturing; Vendor Performance, Percent of Companies Reporting Slower Deliveries; Layoff Rate, Manufacturing; Unemployment Rate, 15 Weeks and over; and the Total Unemployment Rate. Since these series are already a pure 'cycle', their cyclical components are defined as deviations from their respective means.

ARIMA models for the first differences were obtained for the 33 non-stationary indicators.[3] The data base extends from 1947 through 1974 for 21

[2]Indicators (32) and (33) were components of the index of roughly coincident indicators but are now unclassified by the NBER. Indicator (34), which is classified as lagging, was not part of an index but the change in the series was in the index of leading indicators.

[3]Details of the estimated models are available from the authors.

series and 1948 through 1974 for the remaining indicators. As one would expect from the nature of output series, the first differences of many of our indicators display departures from spatial homogeneity in raw form. This tendency disappears after transformation to natural logs and thus the majority of the models are for logged series. In many cases the models are of multiplicative form, that is the moving average polynomial is written as the product of two polynomials in L, and one of the polynomials will have non-zero coefficients only at seasonal lags, that is 11, 12, or 13 months or 3, 4, or 5 quarters. Other models are not multiplicative but have coefficients at those lags. Such models are essentially seasonal models, in apparent contradiction to the indicators being the standard 'seasonally adjusted' series in general use. Usually the seasonality which appears in the models is 'negative seasonality' which is manifest in negative serial correlation at the seasonal lags. Clearly, the adjustment procedures in use by the Department of Commerce and the Bureau of Labor Statistics may not succeed in the extraction of seasonality from these series but rather tend to leave some correlation at lag 12 (4 for quarterly) or at adjoining lags or to introduce negative correlation at those lags. This phenomenon has been noted previously and commented on in Nelson (1972) and is explained by Cleveland and Tiao (1976). Because we want the cycles to be free of any spurious seasonal variation, the seasonal coefficients were disregarded for the purposes of calculating the cycle components.

A few of the non-stationary indicator series are plotted along with their cycle components in figs. 1 through 4. Also shown on the plots are the NBER reference cycles, shaded from peak to trough. The time period covered ends at the third month or first quarter of 1977, thus up to 1974 the cycle components utilize later data in the sense that the data are used to estimate the coefficients, and post-1974 cycle components are computed on a real time basis. The plotted indicators are representative of the series in general in that the cycle components not only tend to anticipate movements in the original series, reflecting the fact that the cycle components measure the forecastable momentum of each series, but also typically lead the traditional NBER turning points. Comparison of figs. 3 and 4 also reveals that the cycle component is as observable in nominal series as in real measures; in fact, nominal and real GNP have very nearly the same cycle components.

Descriptive comparison of the individual indicator cycles is greatly facilitated by construction of a composite index of movements in general business conditions which can serve as a frame of reference for discussing the timing of individual indicators. Since differences in the scale of measurement of indicators is irrelevant to their importance in an index, all of the cycle series were divided by their standard deviations prior to computation of indices. Denoting the standardized cycle component of the ith series at time t

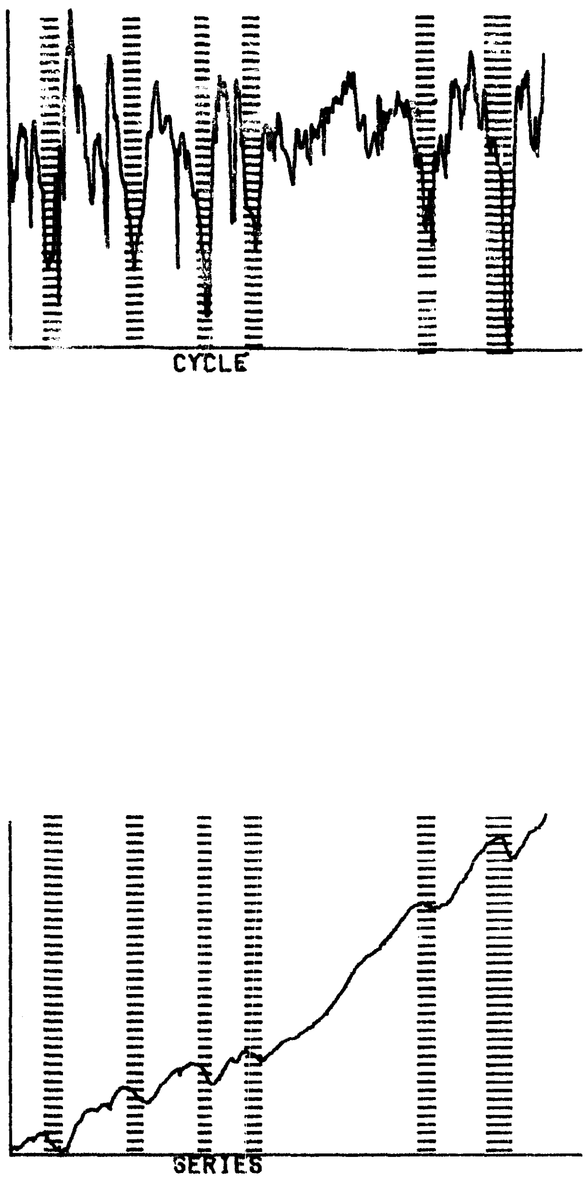
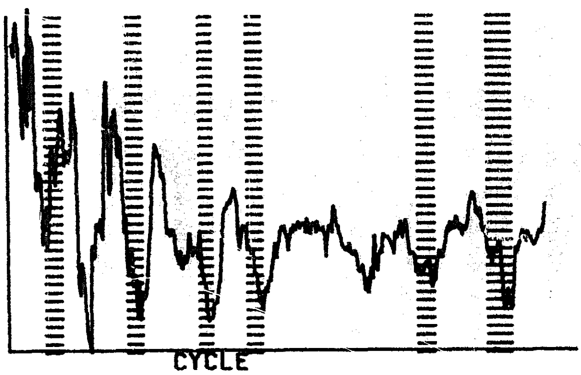


Fig. 1. Employees in non-agricultural establishments, monthly 1947:02–1977:03.



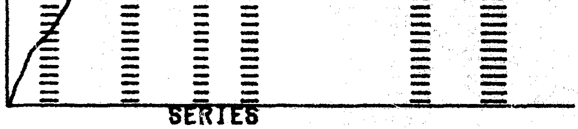
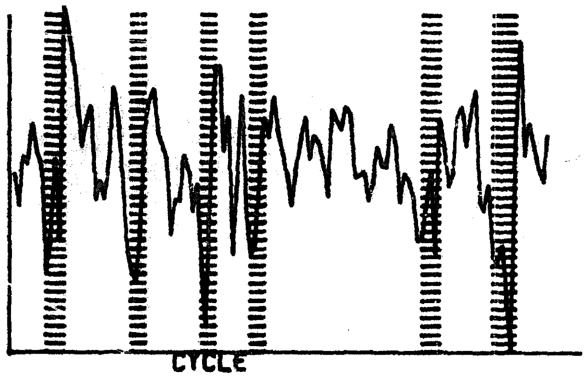


Fig. 2. Consumer installment debt, monthly (natural logs) 1947:02–1977:03.



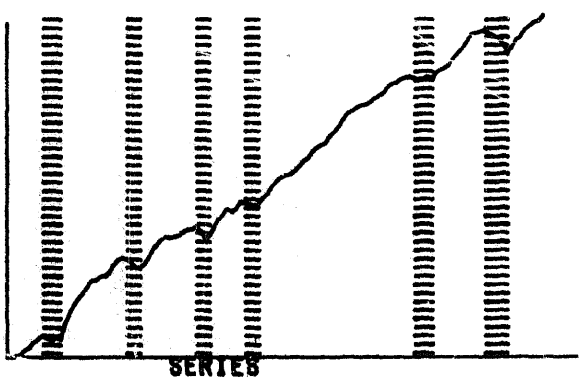


Fig. 3. Gross national product, constant dollars, quarterly (natural logs) 1947:02–1977:01.

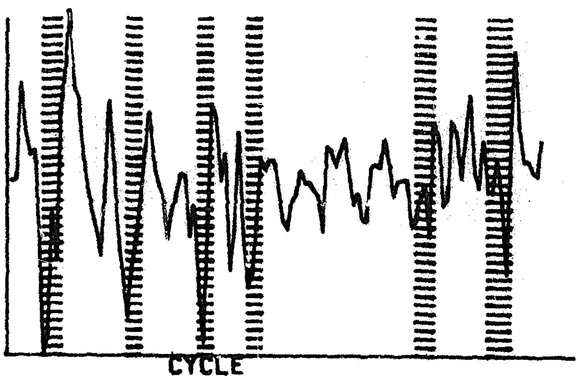


Fig. 4. Gross national product, current dollars, quarterly (natural logs) 1947:02–1977:01.

by $c_{i,t}$ then all linear composite indices, I_t, are of the form

$$I_t = \sum_{i=1}^{M} \beta_i c_{i,t}, \tag{17}$$

where the weight applied to the ith component is denoted β_i and there are M components making up the index.[4]

The method used to choose the weights conforms to Burns and Mitchell's classical definition of a business cycle. They state: 'Our definition presents business cycles as a consensus among expansions in 'many' economic activities, followed by 'similarly general' recessions, contractions, and revival.'[5] Hence, it is necessary to select weights which extract a common source of variation from the different series. Since principal component analysis attempts to capture common movement in a set of observations, an index which meets the criterion is the first principal component vector of the constituent cycle series. Consequently the weights chosen maximize the variance of the index and, because their inner product is unity, the variance is independent of the weights.

The first principal component of the indicator cycles, which accounts for 32.8% of their total variation, is called the 'Index of Business Conditions' and pictured in fig. 5. The index captures clearly the leading nature of the cycle components compared to the reference dates. Most of the indicator cycles have positive weights in the index but eleven weights are negative.[6] For six of the indicator cycles, layoffs (2), duration of unemployment (19), initial unemployment claims (21), unemployment rates (28, 32), and number of unemployed (37), we expect the weights to be negative because the NBER inverts those series, thus it is the negative of the cycle which will move with business conditions. Indicator cycles for manufactures' new orders for consumer goods in constant dollars (3) and for new plant and equipment orders in constant (5) and current dollars (23) move counter to the original series because of negatively autocorrelated data. When a series displays negative serial correlation, a positive change in the series will frequently be associated with negative momentum. The inverse behavior of indicator cycles for housing permits (6) and unit labor costs (16) is explained below when the indicators are classified with respect to timing.

Sample correlations between the index and the individual indicator cycles are useful in describing the coherence of the indicators without computing all

[4]For the purpose of index construction, in the five cases of quarterly series the cycle component measure was assumed to have occurred in the middle month of the quarter and the cycle of the intervening months estimated by interpolation.

[5]Burns and Mitchell (1946, p. 6).

[6]The exact weights are available from the authors.

pair-wise correlations within the indicator set. By considering leading and lagging correlations, that is

$$\operatorname{corr}[I_t, c_{i,t+k}], \qquad k=0, \pm 1, \pm 2, \ldots, \tag{18}$$

we may hope to characterize each of the indicator cycles as a leading, lagging, or coincident indicator.

It may be of interest to note that the correlations are quite symmetric around their maximum absolute value. The intensity of the correlation dampens with increasing lag in both directions. This dampening of

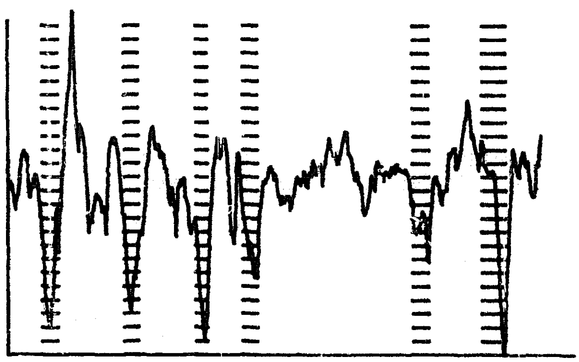


Fig. 5. Index of business conditions, monthly.

correlation together with the cyclical pattern of correlation is indicative of the pseudo-periodic nature of indicator cycles and the index. If we were in fact dealing with true periodic functions of time then correlation would fail to dampen out. The contemporaneous correlations have signs appropriate in view of the weights given the indicator cycles in the composite index.

From the cross-correlation results, it was evident that the indicator cycle for unit labor costs (16) has a negative weight in the index because the indicator cycle lags behind business conditions to the extent that its contemporaneous movement reflects the previous cycle phase. On the other hand, we find that the indicator cycle for housing permits (6) has a dominant 'negative-lagging' relationship with movements in the economy rather than the NBER's 'positive-leading' classification. This suggests that the housing industry exhibits stronger (faster and more widespread) recoveries from recessions than the deterioration it suffers going into a cyclical downturn.

Indicator cycles are defined as roughly coincident if the maximum correlation (in absolute value) occurs within one month of the contemporaneous relationship and the rest are classified as leading or lagging depending on whether the maximum correlation occurs at a negative or positive lag respectively. Our suggested classifications and the number of months by which an indicator cycle leads or lags the index (shown in parentheses) are presented in the last column of table 1.

Eleven of the classifications have changed, furthermore, the five indicators which are unclassified by the NBER have been classified.[7] Ten of the eleven reclassifications are for indicators which the NBER has categorized as leading, a finding not at odds with the results of those who have attempted to use leading indicators to forecast economic activity.[8] The one exception is the consumer debt (34) indicator which the NBER classifies as lagging and its changes as leading, thus, our roughly coincident classification falls midway between the two. For the indicators which the NBER does not classify, our lagging classification for the unemployment rate (32) indicator cycle is in agreement with the classification for the fifteen weeks and over unemployment rate (28) indicator cycle, and we would expect nominal GNP (36) and Personal Income (33) to be coincident indicators. Total Civilian Employment which we have classified as lagging could equally well retain its unclassified status since its maximum correlation with business conditions occurs in both the coincident and lagging categories.

4. The dating of U.S. growth and business cycles

The indicator cycles classified as roughly coincident can provide the basis for a redating of U.S. growth and business cycles. A composite 'Index of Roughly Coincident Indicator Cycles' was formed from the first principal component vector of the coincident cycles. The first component accounted for 51.1% of the variation and the weights changed very little from those used in the index of business conditions. The index is plotted in fig. 6 and like its component cycles, it seems to lead traditional cycle dates.

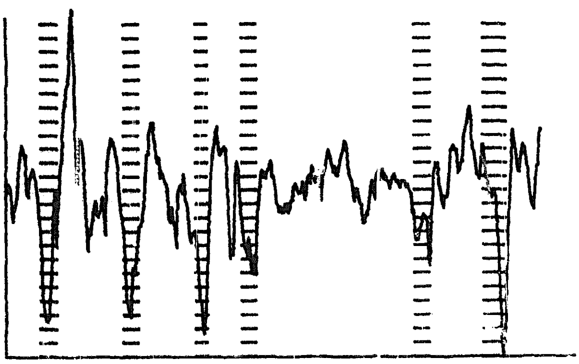


Fig. 6. Index of roughly coincident indicator cycles, monthly.

[7]There may be no real disagreement in the reclassification of indicators constant dollar inventories (8), sensitive prices (9), and liquid assets (10) for, as is discussed above, it is the differences of these series that are utilized by the NBER.

[8]See, for example, Hymans (1973).

While cycle component decomposition is done on a real time basis, the identification of peaks and troughs can only be apparent in retrospect and necessarily depend upon definitions which are to some extent arbitrary. The rules we have used for identifying peaks and troughs are patterned after those applied by Mintz in her analysis of U.S. data. First, reject a peak or trough when there are like turns without an intervening opposite turn. In deciding which of the pair to reject, we retain the largest (smallest) peak (trough). Second, the minimum duration of a cycle phase is five months; and the last criterion is that like turning points cannot be less than twelve months apart. The remaining turning points are our estimates of postwar growth cycles, that is, retardations in the rate of growth in the economy. The turning points of the stochastic growth cycles are reported in column (1) of table 2 and, for comparison, Zarnowitz and Boschan's (1977) growth cycle dates are presented in column (2).[9]

In general, our datings are similar to Zarnowitz and Boschan's (ZB); however, the turning points in our index lead theirs with the exception of the concurrent dating of the Feb. '61 turn.[10] Notable differences between the two sets of datings are the following: the July '51 turning point is one year before ZB's date, the Dec. '62 recovery precedes the turn in ZB's index by twenty-two months, and the Dec. '67 peak leads by eighteen months. We also found growth cycles at Mar. '55–July '56, Nov. '58–Aug. '59, and Aug. '75–Oct. '76, which were not detected by ZB.

The composite index can also be utilized to date classical business cycles which Burns and Mitchell have defined as absolute declines in economic activity. Thus we add the further rules that to qualify as a business cycle turning point, the index trough (peak) must be negative (positive) and the momentum must be negative (positive) for at least five consecutive months in the neighborhood of the trough (peak) date. This leads to the rejection of five turning points and yields the business cycle datings given in column (3) of table 2. The reference cycle dates established by the NBER are presented in column (4) of the table.

Like the growth cycle analysis, the turning points in our coincident index lead the classical turning points. Further, four distinct episodes are present in our index which were not dated at all by classical methods. Taking the NBER-dated episodes in order we have

Nov. '48 (peak)–Oct. '49 (trough): Our index records a loss in momentum beginning in Nov. '47, a brief recovery in '48, and then falls steadily until the April '49 trough. If the Nov. '47 peak is an artifact of the data (37 percent of

[9]The Zarnowitz and Boschan growth cycle datings are almost identical to Mintz's but have the advantage of covering a longer time period.

[10]The first turning point in our index occurs in 1947 and therefore must be considered tentative since many of the component cycles do not begin until 1948.

Table 2

Peaks(P) and troughs(T) in growth cycles and classical business cycles, 1947:01–1977:03.

(1) Stochastic growth cycle turning points	(2) Zarnowitz–Boschan growth cycle turning points	(3) Stochastic business cycle turning points	(4) NBER reference dates
P-Nov. '47	July '48	P-Nov. '47	Nov. '48
T-April '49	Oct. '49	T-April '49	Oct. '49
P-Aug. '50	Mar. '51	P-Aug. '50	
T-July '51	July '52	T-July '51	
P-Nov. '52	Mar. '53	P-Nov. '52	July '53
T-Dec. '53	Aug. '54	T-Dec. '53	May '54
P-Mar. '55		P-Mar. '55	
T-July '56		T-July '56	
P-Dec. '56	Feb. '57	P-Dec. '56	Aug. '57
T-Feb. '58	April '58	T-Feb. '58	April '58
P-Nov. '58		P-Nov. '58	April '60
T-Aug. '59		T-Feb. '61	Feb. '61
P-Jan. '60	Feb. '60	P-Dec. '61	
T-Feb. '61	Feb. '61	T-Dec. '62	
P-Dec. '61	May '62	P-Feb. '66	
T-Dec. '62	Oct. '64	T-Feb. '67	
P-Feb. '66	June '66	P-Dec. '67	Dec. '69
T-Feb. '67	Oct. '67	T-Nov. '70	Nov. '70
P-Dec. '67	Mar. '69	P-Feb. '73	Nov. '73
T-Nov. '70	Nov. '70	T-Jan. '75	Mar. '75
P-Feb. '73	Mar. '73		
T-Jan. '75	Mar. '75		
P-Aug. '75	[a]		
T-Oct. '76			

[a]Time period not examined by Zarnowitz–Boschan.

the component cycles do not begin until 1948), then we would date the peak at June '48.

July '53 (peak)–May '54 (trough): This recession shows up in our index as a peak in Nov. '52 and a continuous loss in momentum until Dec. '53.

Aug. '57 (peak)–April '58 (trough): We date the beginning of this recession at Dec. '56 when the momentum in our index begins to decline until Aug. '57 and then falls rapidly until the Feb. '58 recovery. Thus the index detects the weakening economy eight months before it is recognized by the NBER.

April '60 (peak)–Feb. '61 (trough): The loss in momentum in our index shows up nearly a year and one-half prior to the reference date, although the subsequent recovery is in the same month. The index peaks in Nov. '58, declines until Oct. '59, and then reveals a short recovery which peaks in Jan. '60 (the momentum at this point is less than half of its value on Nov. '58). This behavior suggests that the economy had already entered a recession when the steel strike occurred (July '59 to Nov. '59) and the peak detected by the NBER is the temporary stimulant of the strike settlement.

Dec. '69 (peak)–Nov. '70 (trough): Our index does not show a pronounced peak but instead gradually decreases beginning in Dec. '67. The momentum gives a picture of the economy sliding rather than falling into a recession. The trough dates, however, are the same.

Nov. '73 (peak)–March '75 (trough): This cycle phase shows up as a distinct loss of momentum in Feb. '73 and a recovery in Jan. '75. From Aug. '74 to Jan. '75 there is a substantial loss in momentum indicating that over this period the economy suffered a rapid and severe contraction.

Four 'mini-recessions' which our index detects do not appear in the classical chronology. The stochastic cycle decomposition employed in this study, however, also allows one to gauge the severity of a recession. The values of the coincident index at the ten recorded troughs are, in ascending order of magnitude, as follows: Mar. '75 (−10.0), Feb. '58 (−8.6), April '49 (−7.9), Dec. '53 (−7.6), Feb. '61 (−5.0), Nov. '70 (−4.5), July '51 (−3.5), July '56 (−3.3), Feb. '67 (−1.8), and Dec. '62 (−1.1). The ordering is in general accord with other economic measures of the depth of a contraction, the most recent recession being the most severe in the postwar period. Thus it appears that the momentum in the economy has to fall to about −4.0 before a classical recession is declared.

5. Summary and conclusions

In this paper we have introduced a general procedure for decomposition of non-stationary time series into permanent and transitory components. The theoretical basis for the decomposition does not require that the time series follow a deterministic trend but rather begins with the assumption that the successive changes in the series (or its natural logarithm) are stationary with a representation as an ARMA process. We show that the long-run forecast profile for such a series at any point in time is asymptotic to a linear function of time which we define as the permanent component when it is evaluated at the current time period. The permanent component is shown to be always a random walk with drift. The difference between the permanent component and the actual value of the series is then the momentum contained in the series at that point in time and is a natural measure of its transitory or cyclical component. The transitory component is itself a stationary process with mean zero. The innovations in the permanent component are shown to be larger or smaller than those in observed data depending on the autocorrelation structure of the particular series in question. The decomposition depends only on past data and therefore is computable in 'real time'.

The decomposition methodology was applied to the problem of measuring and dating business 'cycles' in the postwar U.S. economy, using thirty-eight indicator series monitored by the NBER. While our results should only be regarded as exploratory, we feel that they warrant some tentative conclusions based on comparison with the traditional NBER chronology as well as with the recent work of Zarnowitz and Boschan (1977). A composite index of indicator cycles formed by taking their first principal component allowed a classification of individual indicators into leading, coincident, and lagging categories which differed from that of the NBER in about one-third of the cases. Since the index is a linear combination of stationary series with mean zero it exhibits fluctuations above and below the mean of roughly equivalent duration. Thus the measured expansions and contractions are of roughly equivalent duration, in contrast with the traditional NBER chronology which results in contractions which are very brief relative to intervening expansions. Reflecting our definition of the cyclical component as the forecastable momentum in a series at a given time, our dating of cyclical episodes tends to lead the traditional NBER dating and, to a lesser extent, the 'growth cycle' chronology of Zarnowitz and Boschan (1977).

References

Box, G.E.P. and G.M. Jenkins, 1976, Time series analysis forecasting and control, revised ed. (Holden-Day, San Francisco, CA).

Burns, A.F. and W.C. Mitchell, 1946, Measuring business cycles (National Bureau of Economic Research, New York).

Cleveland, W.P. and G.C. Tiao, 1976, Decomposition of seasonal time series: A model for the census X-11 program, Journal of the American Statistical Association 71, 581–587.

Fellner, W., 1956, Trends and cycles in economic activity (Henry Holt, New York).

Friedman, M., 1957, A theory of the consumption function (Princeton University Press, Princeton, NJ).

Hymans, S.H., 1973, On the use of leading indicators to predict cyclical turning points, Brookings Papers on Economic Activity 2, 339–375.

Mintz, I., 1969, Dating postwar business cycles (National Bureau of Economic Research, New York).

Mintz, I., 1972, Dating American growth cycles, in: V. Zarnowitz, ed., The business cycle today (National Bureau of Economic Research, New York).

Muth, J.F., 1960, Optimal properties of exponentially weighted forecasts, Journal of the American Statistical Association 55, 299–306.

Nelson, C.R., 1972, The prediction performance of the FRB-MIT-Penn model of the U.S. economy, American Economic Review 62, 902–917.

Nelson, C.R., 1973, Applied time series analysis for managerial forecasting (Holden-Day, San Francisco, CA).

Nelson, C.R. and H. Kang, 1981, Spurious periodicity in inappropriately detrended time series, Econometrica, forthcoming.

Nerlove, M., 1967, Distributed lags and unobserved components in economic time series, in: W. Fellner, ed., Ten economic studies in the tradition of Irving Fisher (Wiley, New York).
Whittle, P., 1963, Prediction and regulation by linear least-squares methods (English Universities Press, London).
Wold, H., 1938, A study in the analysis of stationary time series (Almquist and Wiksell, Stockholm).
Zarnowitz, V. and C. Boschan, 1975a, Cyclical indicators: An evaluation and new leading indexes, Business Conditions Digest, May, v–xxii.
Zarnowitz, V. and C. Boschan, 1975b, New composite indexes of coincident and lagging indicators, Business Conditions Digest, Nov., v–xxiv.
Zarnowitz, V. and C. Boschan, 1977, Cyclical indicators, National Bureau of Economic Research 57th Annual Report, 34–38.