

Real-Time Measurement of Business Conditions

S. Borağan Aruoba*

University of Maryland

Francis X. Diebold

University of Pennsylvania and NBER

Chiara Scotti

Federal Reserve Board

First Draft: March 2007

This Draft: September 11, 2008

Abstract

We construct a framework for measuring economic activity at high frequency, potentially in real time. We use a variety of stock and flow data observed at mixed frequencies (including very high frequencies), and we use a dynamic factor model that permits exact filtering. We illustrate the framework in a prototype empirical example and a simulation study calibrated to the example.

Key Words : Business cycle, Expansion, Recession, State space model, Macroeconomic forecasting, Dynamic factor model, Contraction, Turning point

JEL Codes: E32, E37, C01, C22

*Corresponding author. Department of Economics, University of Maryland, College Park, MD 20742.
aruoba@econ.umd.edu

1 Introduction

Aggregate business conditions are of central importance in the business, finance, and policy communities, worldwide, and huge resources are devoted to assessment of the continuously-evolving state of the real economy. Literally thousands of newspapers, newsletters, television shows, and blogs, not to mention armies of employees in manufacturing and service industries, including the financial services industries, central banks, government and non-government organizations, grapple constantly with the measurement and forecasting of evolving business conditions. Of central importance is the *constant* grappling. Real economic agents, making real decisions, in real time, want accurate and *timely* estimates of the state of real activity. Business cycle chronologies such as the NBER's, which proclaim expansions and contractions long after the fact, are not useful in that regard.

Against this background, we propose and illustrate a framework for high-frequency business conditions assessment in a systematic, replicable, and statistically optimal manner. Our framework has four key ingredients.

Ingredient 1. We work with a dynamic factor model, treating business conditions as an unobserved variable, related to observed indicators. Latency of business conditions is consistent with economic theory (e.g. Lucas, 1977), which emphasizes that the business cycle is not about any single variable, whether GDP, industrial production, sales, employment, or anything else. Rather, the business cycle is about the dynamics and interactions (“co-movements”) of many variables.

Treating business conditions as latent is also a venerable tradition in empirical business cycle analysis, ranging from the earliest work to the most recent, and from the statistically informal to the statistically formal. On the informal side, latency of business conditions is central to many approaches, from the classic early work of Burns and Mitchell (1946) to the recent workings of the NBER business cycle dating committee, as described for example by Hall et al. (2003). On the formal side, latency of business conditions is central to the popular dynamic factor framework, whether from the “small data” perspective of Geweke (1977), Sargent and Sims (1977), Stock and Watson (1989, 1991) and Diebold and Rudebusch (1996), or the more recent “large data” perspective of Forni, Hallin, Lippi and Reichlin (2000), Stock and Watson (2002) and Bai and Ng (2006). (For discussion of small-data vs. large-data dynamic factor modeling, see Diebold (2003).)

Ingredient 2. We explicitly incorporate business conditions indicators measured at different frequencies. Important business conditions indicators do in fact arrive at a variety of frequencies, including quarterly (e.g., GDP), monthly (e.g., industrial production), weekly

(e.g., employment), and continuously (e.g., asset prices), and we want to be able to incorporate all of them, to provide continuously-updated measurements.

Ingredient 3. We explicitly incorporate indicators measured at high frequencies. Given that our goal is to track the high-frequency evolution of real activity, it is important to incorporate (or at least not exclude from the outset) the high-frequency information flow associated with high-frequency indicators.

Ingredient 4. We extract and forecast latent business conditions using linear yet statistically optimal procedures, which involve no approximations. The appeal of exact as opposed to approximate procedures is obvious, but achieving exact optimality is not trivial, due to complications arising from temporal aggregation of stocks vs. flows in systems with mixed-frequency data.

Related to our concerns and framework is a small but nevertheless important literature, including Stock and Watson (1989, 1991), Mariano and Murasawa (2003), Evans (2005) and Proietti and Moauro (2006), as well as Shen (1996), Abeysinghe (2000), Altissimo et al. (2002), Liu and Hall (2001), McGuckin, Ozyildirim and Zarnowitz (2003), Ghysels, Santa Clara and Valkanov (2004), and Giannone, Reichlin and Small (2008).

Our contribution is different in certain respects, and similar in others, and both the differences and similarities are intentional. Let us begin by highlighting some of the differences. First, some authors like Stock and Watson (1989, 1991) work in a dynamic factor framework with exact linear filtering, but they don't consider data at mixed frequencies or at high frequencies.

Second, other authors like Evans (2005) do not use a dynamic factor framework and do not use high-frequency data, instead focusing on estimating high-frequency GDP. Evans (2005), for example, equates business conditions with GDP growth and uses state space methods to estimate daily GDP growth using data on preliminary, advanced and final releases of GDP and other macroeconomic variables.

Third, authors like Mariano and Murasawa (2003) work in a dynamic factor framework and consider data at mixed frequencies, but not high frequencies, and their filtering algorithm is only approximate. Proietti and Moauro (2006) avoid the Mariano-Murasawa approximation at the cost of moving to a non-linear model with a corresponding rather tedious non-linear filtering algorithm.

Ultimately, however, the *similarities* between our work and others' are more important than the differences, as we stand on the shoulders of many earlier authors. Effectively we (1) take a small-data dynamic factor approach to business conditions analysis, (2) recognize

the potential practical value of extending the approach to mixed-frequency data settings involving some high-frequency data, (3) recognize that doing so amounts to a filtering problem with a large amount of missing data, which the Kalman filter is optimally designed to handle, and (4) provide a prototype example of the framework in use. Hence the paper is really a “call to action,” a call to move the state-space dynamic-factor framework closer to its high-frequency limit, and hence to move statistically-rigorous business conditions analysis closer to *its* high-frequency limit.

We proceed as follows. In section 2 we provide a detailed statement of our dynamic-factor modeling framework, and in section 3 we represent it as a state space filtering problem with a large amount of missing data. In section 4 we report the results of a four-indicator prototype empirical analysis, using quarterly GDP, monthly employment, weekly initial jobless claims, and the daily yield curve term premium. In section 5 we report the results of a simulation exercise, calibrated to our empirical estimates, which lets us illustrate our methods and assess their efficacy in a controlled environment. In section 6 we conclude and offer directions for future research.

2 The Modeling Framework

Here we propose a dynamic factor model at daily frequency. The model is very simple at daily frequency, but of course the daily data are generally not observed, so most of the data are missing. Hence we explicitly treat missing data and temporal aggregation, and we obtain the measurement equations for observed stock and flow variables. Following that, we enrich the model by allowing for lagged state variables in the measurement equations, and by allowing for trend, both of which are important when fitting the model to macroeconomic and financial indicators.

2.1 The Dynamic Factor Model at Daily Frequency

We assume that the state of the economy evolves at a very high frequency; without loss of generality, call it “daily.” In our subsequent empirical work, we will indeed use a daily base observational frequency, but much higher (intra-day) frequencies could be used if desired. Economic and financial variables, although evolving daily, are of course not generally *observed* daily. For example, an end-of-year wealth variable is observed each December 31, and is unobserved every other day of the year.

Let x_t denote underlying business conditions at day t , which evolve daily with $AR(p)$ dynamics,

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \dots + \rho_p x_{t-p} + e_t, \quad (1)$$

where e_t is a white noise innovation with unit variance. We are interested in tracking and forecasting real activity, so we use a single-factor model; that is, x_t is a scalar, as for example in Stock and Watson (1989). Additional factors could be introduced to track, for example, wage/price developments.

Let y_t^i denote the i -th daily economic or financial variable at day t , which depends linearly on x_t and possibly also on various exogenous variables and/or lags of y_t^i :

$$y_t^i = c_i + \beta_i x_t + \delta_{i1} w_t^1 + \dots + \delta_{ik} w_t^k + \gamma_{i1} y_{t-D_i}^i + \dots + \gamma_{in} y_{t-nD_i}^i + u_t^i, \quad (2)$$

where the w_t are exogenous variables and the u_t^i are contemporaneously and serially uncorrelated innovations. Notice that we introduce lags of the dependent variable y_t^i in multiples of D_i , where $D_i > 1$ is a number linked to the frequency of the observed y_t^i . (We will discuss D_i in detail in the next sub-section.) Modeling persistence only at the daily frequency would be inadequate, as it would decay too quickly.

2.2 Missing Data, Stocks vs. Flows, and Temporal Aggregation

Recall that y_t^i denotes the i -th variable on a daily time scale. But most variables, although *evolving* daily, are not actually *observed* daily. Hence let \tilde{y}_t^i denote the same variable observed at a lower frequency (call it the “tilde frequency”). The relationship between \tilde{y}_t^i and y_t^i depends crucially on whether y_t^i is a stock or flow variable.

If y_t^i is a stock variable measured at a non-daily tilde frequency, then the appropriate treatment is straightforward, because stock variables are simply point-in-time snapshots. At any time t , either y_t^i is observed, in which case $\tilde{y}_t^i = y_t^i$, or it is not, in which case $\tilde{y}_t^i = NA$, where NA denotes missing data (“not available”). Hence we have the stock variable measurement equation:

$$\tilde{y}_t^i = \begin{cases} y_t^i = c_i + \beta_i x_t + \delta_{i1} w_t^1 + \dots + \delta_{ik} w_t^k + \gamma_{i1} y_{t-D_i}^i + \dots + \gamma_{in} y_{t-nD_i}^i + u_t^i & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise.} \end{cases} \quad (3)$$

Now consider flow variables. Flow variables observed at non-daily tilde frequencies are

intra-period sums of the corresponding daily values,

$$\tilde{y}_t^i = \begin{cases} \sum_{j=0}^{D_i-1} y_{t-j}^i & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise,} \end{cases} \quad (4)$$

where D_i is the number of days per observational period (e.g., $D_i=7$ if y_t^i is measured weekly). Combining this fact with equation (2), we arrive at the flow variable measurement equation:

$$\tilde{y}_t^i = \begin{cases} \sum_{j=0}^{D_i-1} c_i + \beta_i \sum_{j=0}^{D_i-1} x_{t-j}^i + \delta_{i1} \sum_{j=0}^{D_i-1} w_{t-j}^1 + \dots + \delta_{ik} \sum_{j=0}^{D_i-1} w_{t-j}^k \\ \quad + \gamma_{i1} \sum_{j=0}^{D_i-1} y_{t-D_i-j}^i + \dots + \gamma_{in} \sum_{j=0}^{D_i-1} y_{t-nD_i-j}^i + u_t^{*i} & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise,} \end{cases} \quad (5)$$

where $\sum_{j=0}^{D_i-1} y_{t-D_i-j}^i$ is by definition the observed flow variable one period ago ($\tilde{y}_{t-D_i}^i$), and u_t^{*i} is the sum of the u_t^i over the tilde period.

Discussion of two subtleties is in order. First, note that in general D_i is time-varying, as for example some months have 28 days, some have 29, some have 30, and some have 31. To simplify the notation above, we ignored this, implicitly assuming that D_i is fixed. In our subsequent empirical implementation, however, we allow for time-varying D_i .

Second, note that although u_t^{*i} follows a moving average process of order $D_i - 1$ at the daily frequency, it nevertheless remains white noise when observed at the tilde frequency, due to the $(D_i - 1)$ -dependence of an $MA(D_i - 1)$ process. Hence we appropriately treat u_t^{*i} as white noise in what follows, where $var(u_t^{*i}) = D_i \cdot var(u_t^i)$.

2.3 Trend

The exogenous variables w_t are the key to handling trend. In particular, in the important special case where the w_t are simply deterministic polynomial trend terms ($w_{t-j}^1 = t - j$, $w_{t-j}^2 = (t - j)^2$ and so on) we have that

$$\sum_{j=0}^{D_i-1} [c_i + \delta_{i1} (t - j) + \dots + \delta_{ik} (t - j)^k] \equiv c_i^* + \delta_{i1}^* t + \dots + \delta_{ik}^* t^k. \quad (6)$$

In the appendix we derive the mapping between the “starred” and “unstarred” c ’s and δ ’s for cubic trends, which are sufficiently flexible for most macroeconomic data and of course include linear and quadratic trends as special cases. Assembling the results, we have the stock variable measurement equation

$$\tilde{y}_t^i = \begin{cases} c_i^* + \beta_i x_t^i + \delta_{i1}^* t + \dots + \delta_{ik}^* t^k + \gamma_{i1} \tilde{y}_{t-D_i}^i + \dots + \gamma_{in} \tilde{y}_{t-nD_i}^i + u_t^{*i} & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise.} \end{cases} \quad (7)$$

and the flow variable measurement equation,

$$\tilde{y}_t^i = \begin{cases} c_i^* + \beta_i \sum_{j=0}^{D_i-1} x_{t-j}^i + \delta_{i1}^* t + \dots + \delta_{ik}^* t^k + \gamma_{i1} \tilde{y}_{t-D_i}^i + \dots + \gamma_{in} \tilde{y}_{t-nD_i}^i + u_t^{*i} & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise.} \end{cases} \quad (8)$$

This completes the specification of our model, which has a natural state space form, to which we now turn.

3 State Space Representation, Signal Extraction, and Estimation

Here we discuss our model from a state-space perspective, including filtering and estimation. We avoid dwelling on standard issues, focusing instead on the nuances specific to our framework, including missing data due to mixed-frequency modeling, high-dimensional state vectors due to the presence of flow variables, and time-varying system matrices due to varying lengths of months.

3.1 State Space Representation

Our model is trivially cast in state-space form as

$$y_t = Z_t \alpha_t + \Gamma_t w_t + \varepsilon_t \quad (9)$$

$$\alpha_{t+1} = T \alpha_t + R \eta_t \quad (10)$$

$$\varepsilon_t \sim (0, H_t), \eta_t \sim (0, Q) \quad (11)$$

$$t = 1, \dots, \mathcal{T},$$

where y_t is an $N \times 1$ vector of observed variables (subject of course to missing observations), α_t is an $m \times 1$ vector of state variables, w_t is a $e \times 1$ vector of predetermined variables containing a constant term (unity), k trend terms and $N \times n$ lagged dependent variables (n for each of the N elements of the y_t vector), ε_t and η_t are vectors of measurement and transition shocks containing the u_t^i and e_t , and \mathcal{T} denotes the last time-series observation.

In general, the observed vector y_t will have a very large number of missing values, reflecting not only missing daily data due to holidays, but also, and much more importantly, the fact that most variables are observed much less often than daily. Interestingly, the missing data *per se* does not pose severe challenges: y_t is simply littered with a large number of *NA* values, and the corresponding system matrices are very sparse, but the Kalman filter remains valid (appropriately modified, as we discuss below), and numerical implementations may indeed be tuned to exploit the sparseness, as we do in our implementation.

In contrast, the presence of flow variables produces more significant complications, and it is hard to imagine a serious business conditions indicator system without flow variables, given that real output is itself a flow variable. Flow variables produce intrinsically high-dimensional state vectors. In particular, as shown in equation (3), the flow variable measurement equation contains x_t and $\max_i\{D_i\} - 1$ lags of x_t , producing a state vector of dimension $\max\{\max_i\{D_i\}, p\}$, in contrast to the p -dimensional state associated with a system involving only stock variables. In realistic systems with data frequencies ranging from, say, daily to quarterly, $\max_i\{D_i\} \approx 90$.

There is a final nuance associated with our state-space system: several of the system matrices are time-varying. In particular, although T , R and Q are constant, Z_t , Γ_t and H_t are not, because of the variation in the number of days across quarters and months (i.e., the variation in D_i across t). Nevertheless the Kalman filter remains valid.

3.2 Signal Extraction

With the model cast in state space form, and for given parameters, we use the Kalman filter and smoother to obtain optimal extractions of the latent state of real activity. As is standard for classical estimation, we initialize the Kalman filter using the unconditional mean and covariance matrix of the state vector. We use the contemporaneous Kalman filter; see Durbin and Koopman (2001) for details.

Let $\mathcal{Y}_t \equiv \{y_1, \dots, y_t\}$, $a_{t|t} \equiv E(\alpha_t|\mathcal{Y}_t)$, $P_{t|t} = \text{var}(\alpha_t|\mathcal{Y}_t)$, $a_t \equiv E(\alpha_t|\mathcal{Y}_{t-1})$, and $P_t =$

$\text{var}(\alpha_t | \mathcal{Y}_{t-1})$. Then the Kalman filter updating and prediction equations are

$$a_{t|t} = a_t + P_t Z_t' F_t^{-1} v_t \quad (12)$$

$$P_{t|t} = P_t - P_t Z_t' F_t^{-1} Z_t P_t' \quad (13)$$

$$a_{t+1} = T a_{t|t} \quad (14)$$

$$P_{t+1} = T P_{t|t} T' + R Q R', \quad (15)$$

where

$$v_t = y_t - Z_t a_t - \Gamma_t w_t \quad (16)$$

$$F_t = Z_t P_t Z_t' + H_t, \quad (17)$$

for $t = 1, \dots, \mathcal{T}$.

Crucially for us, the Kalman filter remains valid with missing data. If all elements of y_t are missing, we skip updating and the recursion becomes

$$a_{t+1} = T a_t \quad (18)$$

$$P_{t+1} = T P_t T' + R Q R'. \quad (19)$$

If some but not all elements of y_t are missing, we replace the measurement equation with

$$y_t^* = Z_t^* \alpha_t + \Gamma_t^* w_t + \varepsilon_t^* \quad (20)$$

$$\varepsilon_t^* \sim N(0, H_t^*), \quad (21)$$

where y_t^* is of dimension $N^* < N$, containing the elements of the y_t vector that are observed. The key insight is that y_t^* and y_t are linked by the transformation $y_t^* = W_t y_t$, where W_t is a matrix whose N^* rows are the rows of I_N corresponding to the observed elements of y_t . Similarly, $Z_t^* = W_t Z_t$, $\Gamma_t^* = W_t \Gamma_t$, $\varepsilon_t^* = W_t \varepsilon_t$ and $H_t^* = W_t H_t W_t'$. The Kalman filter works exactly as described above, replacing y_t , Z_t and H with y_t^* , Z_t^* and H_t^* . Similarly, after transformation the Kalman smoother remains valid with missing data.

3.3 Estimation

Thus far we have assumed known system parameters, whereas they are of course unknown in practice. As is well-known, however, the Kalman filter supplies all of the ingredients

needed for evaluating the Gaussian pseudo log likelihood function via the prediction error decomposition,

$$\log L = -\frac{1}{2} \sum_{t=1}^T [N \log 2\pi + (\log |F_t| + v_t' F_t^{-1} v_t)]. \quad (22)$$

In calculating the log likelihood, if all elements of y_t are missing, the contribution of period t to the likelihood is zero. When some elements of y_t are observed, the contribution of period t is $[N^* \log 2\pi + (\log |F_t^*| + v_t^{*'} F_t^{*-1} v_t^*)]$ where N^* is the number of observed variables, and we obtain F_t^* and v_t^* by filtering the transformed y_t^* system.

4 A Prototype Empirical Application

We now present a simple application involving the daily term premium, weekly initial jobless claims, monthly employment and quarterly GDP. We describe in turn the data, the specific variant of the model that we implement, subtleties of our estimation procedure, and our empirical results.

4.1 Business Conditions Indicators

Our analysis covers the period from April 1, 1962 through February 20, 2007, which is 16,397 observations of daily data. (We use a seven-day week.)

We use four indicators. Moving from highest frequency to lowest frequency, the first indicator is the yield curve term premium, defined as the difference between ten-year and three-month Treasury yields. We measure the term premium daily; hence there are no aggregation issues. We treat holidays and weekends as missing.

The second indicator is initial claims for unemployment insurance, a weekly flow variable covering the seven-day period from Sunday to Saturday. We set the Saturday value to the sum of the previous seven daily values, and we treat other days as missing.

The third indicator is employees on non-agricultural payrolls, a monthly stock variable. We set the end-of-month value to the end-of-month daily value, and we treat other days as missing.

The fourth and final indicator is real GDP, a quarterly flow variable. We set the end-of-quarter value to the sum of daily values within the quarter, and we treat other days as missing.

Several considerations guide our choice of variables. First, we want the variables chosen

to illustrate the flexibility of our framework. Hence we choose four variables measured at four different frequencies ranging from very high (daily) to very low (quarterly), and representing both stocks (term premium, payroll employment) and flows (initial claims, GDP). Second, we want our illustrative analysis to be credible, if not definitive. Although reasonable people can (and will) disagree on the number and choice of indicators, our choices are easily defensible. GDP needs no defense. Labor market variables like payroll employment and initial claims are strongly cyclical and feature prominently in many coincident indexes. The Conference Board’s composite coincident index, for example, also uses payroll employment. Finally, the term premium is also strongly cyclical, as studied for example in Diebold, Rudebusch and Aruoba (2006) and several of the papers they cite.

4.2 Model Implementation

In the development thus far we have allowed for general polynomial trend and general $AR(p)$ dynamics. In the prototype model that we now take to the data, we make two simplifying assumptions that reduce the number of parameters to be estimated by numerical likelihood optimization. First, we de-trend prior to fitting the model rather than estimating trend parameters simultaneously with the others, and second, we use simple first-order dynamics throughout. In future work we look forward to incorporating more flexible dynamics but, as we show below, the framework appears quite encouraging even with simple $AR(1)$ dynamics.

Hence latent business conditions x_t follow zero-mean $AR(1)$ process, as do the observed variables at their observational frequencies. For weekly initial claims, monthly employment and quarterly GDP, this simply means that the lagged values of these variables are elements of the w_t vector. We denote these by \tilde{y}_{t-W}^2 , \tilde{y}_{t-M}^3 and \tilde{y}_{t-q}^4 , where W denotes the number of days in a week, M denotes the number of days in a month and q denotes the number of days in a quarter. (Once again, the notation in the paper assumes M and q are constant over time, but in the implementation we adjust them according to the number of days in the relevant month or quarter.)

For the term premium, on the other hand, we model the autocorrelation structure using an $AR(1)$ process for the measurement equation innovation, u_t^1 , instead of adding a lag of the term premium in w_t .

The equations that define the model are

$$\begin{aligned}
\underbrace{\begin{bmatrix} \tilde{y}_t^1 \\ \tilde{y}_t^2 \\ \tilde{y}_t^3 \\ \tilde{y}_t^4 \end{bmatrix}}_{y_t} &= \underbrace{\begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 0 & \beta_2 & 0 & \beta_4 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & 0 & \beta_4 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \beta_4 \text{ or } 0 \\ 0 & 0 & 0 & \beta_4 \text{ or } 0 \\ 0 & 0 & 0 & \beta_4 \text{ or } 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}'}_{Z_t} \underbrace{\begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-\bar{q}-1} \\ x_{t-\bar{q}} \\ u_t^1 \end{bmatrix}}_{\alpha_t} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ \gamma_2 & 0 & 0 \\ 0 & \gamma_3 & 0 \\ 0 & 0 & \gamma_4 \end{bmatrix}}_{\Gamma_t} \underbrace{\begin{bmatrix} \tilde{y}_{t-W}^2 \\ \tilde{y}_{t-M}^3 \\ \tilde{y}_{t-q}^4 \end{bmatrix}}_{w_t} + \underbrace{\begin{bmatrix} 0 \\ u_t^{*2} \\ u_t^3 \\ u_t^{*4} \end{bmatrix}}_{\varepsilon_t} \\
\underbrace{\begin{bmatrix} x_{t+1} \\ x_t \\ \vdots \\ x_{t-\bar{q}} \\ x_{t-\bar{q}+1} \\ u_{t+1}^1 \end{bmatrix}}_{\alpha_{t+1}} &= \underbrace{\begin{bmatrix} \rho & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & \gamma_1 \end{bmatrix}}_T \underbrace{\begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-\bar{q}-1} \\ x_{t-\bar{q}} \\ u_t^1 \end{bmatrix}}_{\alpha_t} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}}_R \underbrace{\begin{bmatrix} e_t \\ \zeta_t \end{bmatrix}}_{\eta_t} \quad (23) \\
\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} &\sim N \left(\begin{bmatrix} 0_{4 \times 1} \\ 0_{2 \times 1} \end{bmatrix}, \begin{bmatrix} H_t & 0 \\ 0 & Q \end{bmatrix} \right), \quad H_t = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_{2t}^{*2} & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_{4t}^{*2} \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}
\end{aligned}$$

where the notation corresponds to the system in Section 3 with $N = 4$, $k = 3$, $m = 93$, $p = 1$ and $r = 2$. We use the current factor and 91 lags in our state vector because the maximum possible number of days in a quarter is 92, which we denote by \bar{q} . (If there are q days in a quarter, then on the last day of the quarter we need the current value and $q - 1$ lags.) Also, in every quarter we adjust the number of non-zero elements in the fourth row of the Z_t matrix to reflect the number of days in that quarter.

4.3 Estimation

The size of our estimation problem is substantial. We have 16,397 daily observations, and even with de-trended data and first-order dynamics we have 93 state variables and 13 coefficients. Using a Kalman filter routine programmed in MATLAB, one evaluation of the likelihood takes about 20 seconds. Maximization of the likelihood, however, may involve a very large number of likelihood evaluations, so it's crucial to explore the parameter space in a sophisticated way. Throughout, we use a quasi-Newton algorithm with BFGS updating of the inverse Hessian, using accurate start-up values for iteration. (Perhaps the methods of Jungbacker and Koopman (2008) could be used to improve the speed of our gradient evaluation, although we have not yet explored that avenue. Our real problem is the huge sample size.)

We obtain our start-up values in two steps, as follows. In the first step, we use only daily and stock variables, which drastically reduces the dimension of the state vector, resulting in very fast estimation. This yields preliminary estimates of all measurement equation parameters for the daily and stock variables, and all transition equation parameters, as well as a preliminary extraction of the factor, \hat{x}_t (via a pass of the Kalman smoother).

In the second step, we use the results of the first step to obtain startup values for the remaining parameters, i.e., those in the flow variable measurement equations. We simply regress the flow variables on the smoothed state extracted in the first step and take the coefficients as our startup values.

Obviously in the model that we use in this paper, the variables that we use in the first step are the daily term premium and monthly employment, and the variables that we use in the second step are weekly initial claims and quarterly GDP. (Note that the same simple two-step method could be applied equally easily in much larger models.) At the conclusion of the second step we have startup values for all parameters, and we proceed to estimate the full model's parameters jointly, after which we obtain our "final" extraction of the latent factor.

4.4 Results

Here we discuss a variety of aspects of our empirical results. To facilitate the discussion, we first define some nomenclature to help us distinguish among models. We call our full four-variable model GEIS ("GDP, Employment, Initial Claims, Slope"). Similarly, proceeding to drop progressively more of the high-frequency indicators, we might consider GEI or GE

models.

4.4.1 The Smoothed GEIS Real Activity Indicator

We start with our centerpiece, the extracted real activity indicator (factor). In Figure 1 we plot the smoothed GEIS factor with NBER recessions shaded. (Because the NBER provides only months of the turning points, we assume recessions start on the first day of the month and end on the last day of the month.)

Several observations are in order. First, our real activity indicator broadly coheres with the NBER chronology. There are, for example, no NBER recessions that are not also reflected in our indicator. Of course there is nothing sacred about the NBER chronology, but it nevertheless seems comforting that the two cohere. The single broad divergence is the mid 1960s episode, which the NBER views as a growth slowdown but we would view as a recession.

Second, if our real activity indicator broadly coheres with the NBER chronology, it nevertheless differs somewhat. In particular, it tends to indicate earlier turning points, especially peaks. That is, when entering recessions our indicator tends to reach its peak and start falling several weeks before the corresponding NBER peak. Similarly, when emerging from recessions, our indicator tends to reach its trough and start rising before the corresponding NBER trough. In the last two recessions, however, our indicator matches the NBER trough very closely.

One can interpret our indicator's tendency toward earlier-than-NBER peaks in at least two ways. The NBER chronology may of course simply be inferior, tending to lag turning points whereas ours does not. Alternatively, the NBER chronology may be accurate whereas our index may actually have some lead, particularly as one of our component indicators is the daily term premium, which is not only cyclical but may actually lead the cycle (e.g., Diebold, Rudebusch and Aruoba, 2006).

Third, our real activity indicator makes clear that there are important differences in entering and exiting recessions, whereas the "0-1" NBER recession indicator can not. In particular, our indicator consistently plunges at the onset of recessions, whereas its growth when exiting recessions is sometimes brisk (e.g., 1973-75, 1982) and sometimes anemic (e.g., the well-known "jobless recoveries" of 1990-91 and 2001).

Fourth, and of crucial importance, our indicator is of course available at high frequency, whereas the NBER chronology is available only monthly and with *very* long lags (often several years). Hence our indicator is a useful "nowcast," whereas the NBER chronology is not.

4.4.2 Gains From High-Frequency Data I: Comparison of GE and GEI Factors

Typically, analyses similar to ours are done using monthly and/or quarterly data, as would be the case in a two-variable GE (GDP, employment) model. To see what is gained by inclusion of higher-frequency data, we now compare the real activity factors extracted from a GE model and a GEI model (which incorporates weekly initial claims).

In Figure 2 we show the smoothed GEI factor, and for comparison we show a shaded interval corresponding to the smoothed GE factor ± 1 s.e. The GEI factor is quite different, often violating the ± 1 s.e. band, and indeed not infrequently violating a ± 2 s.e. band (not shown) as well.

In Figure 3 we dig deeper, focusing on the times around the six NBER recessions: December 1969 - November 1970, November 1973 - March 1975, January 1980 - July 1980, July 1981 - November 1982, July 1990 - March 1991 and March 2001 - November 2001. We consider windows that start twelve months before peaks and end twelve months after troughs. Within each window, we again show the smoothed GEI factor and a shaded interval corresponding to the smoothed GE factor ± 1 s.e. Large differences are apparent.

In Figure 4 we move from smoothed to filtered real activity factors, again highlighting the six NBER recessions. The filtered version is the one relevant in real time, and it highlights another key contribution of the high-frequency information embedded in the GEI factor. In particular, the filtered GEI factor evolves quite smoothly with the weekly information on which it is in part based, whereas the filtered GE factor has much more of a discontinuous “step function” look. Looking at the factors closely, the GE factor jumps at the end of every month and then reverts towards to mean (of zero) while the GEI factor jumps every week with the arrival of new Initial Claims data.

Finally, what of a comparison between the GEI factor and the GEIS factor, which incorporates the daily term structure slope? In this instance it turns out that, although incorporating weekly data (moving from GE to GEI) was evidently very helpful, incorporating daily data (moving from GEI to GEIS) was not. That is, the GEI and GEIS factors are almost identical. It is important to note, however, that we still need a daily state space setup even though the highest-frequency data of value were weekly, to accommodate the variation in weeks per month and weeks per quarter.

4.4.3 Gains From High-Frequency Data II: A Calibrated Simulation

Here we illustrate our methods in a simulation calibrated to the empirical results above. This allows us to assess the efficacy of our framework in a controlled environment. In particular,

in a simulation we know the true factor, so we can immediately determine whether and how much we gain by incorporating high-frequency data, in terms of reduced factor extraction error. In contrast, in empirical work such as that above, although we can see that the extracted GE and GEI factors differ, we can not be certain that the GEI factor extraction is more accurate, because we can never see the true factor, even *ex post*.

In our simulation we use the system and the parameters estimated previously. Using those parameters, we generate forty years of “daily” data on all four variables, and then we transform them to obtain the observed data. Specifically, we delete the weekends from the daily variable and aggregate the daily observations over the week to obtain the observed weekly (flow) variable. We also delete all the observations for the third (stock) variable except for the end-of-the-month observations and sum the daily observations over the quarter to get the fourth (flow) variable. Finally, using the simulated data we estimate the coefficients and extract the factor, precisely as we did with the real data.

In the top panel of Figure 5 we show the true factor together with the smoothed factor from the GE model. The two are of course related, but they often diverge noticeably and systematically, for long periods. The correlation between the two is 0.72 and the mean squared extraction error is 0.45. In the bottom panel of Figure 5 we show the true factor together with the smoothed factor from the GEI model. The two are much more closely related and indeed hard to distinguish. The correlation between the two is 0.98 and the mean squared extraction error is 0.07. This exercise quite convincingly shows that incorporating high-frequency data improves the accuracy of the extracted factor.

4.4.4 Real Time Performance

At any point \mathcal{T} in real time, we simply use the time- \mathcal{T} data vintage to extract the real activity factor at time \mathcal{T} and earlier, as in Corradi, Fernandez and Swanson (2007). As time progresses, we re-estimate the system each period (or less frequently if desired for convenience), always using the latest-vintage data to extract the real activity factor.

In Figure 6 we show what we call a “tentacle plot” for part of 2008, that is, a plot of several series of real activity factors extracted using sequential vintages of 2008 data. The tentacle plot contains six paths, extracted on April 4, June 12, June 28, July 19, August 9, and August 30. The six paths show clearly that newly-arrived data can produce substantial changes in optimal assessments of real activity. It is interesting to note, for example, that the two assessments using August data vintages produce August values of the real activity factor below the levels seen at the onset of both the 1990-91 and 2001 recessions, but still

far from the levels seen at those recessions' troughs, indicating not only recession, but also that conditions would likely worsen before improving.

5 Summary and Concluding Remarks

We view this paper as providing both (1) a “call to action” for measuring macroeconomic activity in real time, using a variety of stock and flow data observed at mixed frequencies, potentially also including very high frequencies, and (2) a prototype empirical application, illustrating the gains achieved by moving beyond the customary monthly data frequency. Specifically, we have proposed a dynamic factor model that permits exactly optimal extraction of the latent state of macroeconomic activity, and we have illustrated it in a four-variable empirical application with a daily base frequency, and in a parallel calibrated simulation.

We look forward to a variety of variations and extensions of our basic theme, including but not limited to:

(1) Incorporation of indicators beyond macroeconomic and financial data. In particular, it will be of interest to attempt inclusion of qualitative information such as headline news.

(2) Construction of a real time composite leading index (CLI). Thus far we have focused only on construction of a composite coincident index (CCI), which is the more fundamental problem, because a CLI is simply a forecast of a CCI. Explicit construction of a leading index will nevertheless be of interest.

(3) Allowance for nonlinear regime-switching dynamics. The linear methods used in this paper provide only a partial (linear) statistical distillation of the rich business cycle literature. A more complete approach would incorporate the insight that expansions and contractions may be probabilistically different regimes, separated by the “turning points” corresponding to peaks and troughs, as emphasized for many decades in the business cycle literature and rigorously embodied Hamilton’s (1989) Markov-switching model. Diebold and Rudebusch (1996) and Kim and Nelson (1998) show that the linear and nonlinear traditions can be naturally joined via dynamic factor modeling with a regime-switching factor. Such an approach could be productively implemented in the present context, particularly if interest centers on turning points, which are intrinsically well-defined only in regime-switching environments.

(4) Comparative assessment of experiences and results from “small data” approaches, such as ours, vs. “big data” approaches. Although much professional attention has recently turned to big data approaches, as for example in Forni, Hallin, Lippi and Reichlin (2000) and Stock and Watson (2002), recent theoretical work by Boivin and Ng (2006) shows that bigger

is not necessarily better. The matter is ultimately empirical, requiring detailed comparative assessment. It would be of great interest, for example, to compare results from our approach to those from the Altissimo et al. (2002) EuroCOIN approach, for the same economy and time period. Such comparisons are very difficult, of course, because the “true” state of the economy is never known, even *ex post*.

(5) Complete real-time analysis, recognizing that at any time \mathcal{T} we have not only the time- \mathcal{T} data vintage, but also all earlier data vintages. That would permit incorporation of the stochastic process of data revisions, as was attempted (in different contexts) in early work such as Conrad and Corrado (1979) and Howrey (1984) and recent work such as Aruoba (2008). Doing so in the rich dynamic multivariate environment of this paper is presently infeasible, however, due to the large additional estimation burden that it would entail.

(6) Exploration of direct indicators of daily activity, such as debit card transactions data, as in Galbraith and Tkacz (2007).

Indeed progress is already being made in work done subsequently to earlier drafts of this paper, such as Camacho and Perez-Quiros (2008).

Acknowledgments:

For helpful guidance we thank the Editor, Associate Editor, and referees, as well as seminar and conference participants at the Board of Governors of the Federal Reserve System, the Federal Reserve Bank of Philadelphia, the European Central Bank, the University of Pennsylvania, SUNY Albany, SCE Cyprus, and American University. We are especially grateful to David Armstrong, Carlos Capistran, Dean Croushore, Martin Evans, Jon Faust, John Galbraith, Eric Ghysels, Sharon Kozicki, Steve Kreider, Simon van Norden, Alexi Onatski, Simon Potter, Scott Richard, Frank Schorfheide, Roberto Sella and Jonathan Wright. For research support we thank the National Science Foundation and the Real-Time Data Research Center at the Federal Reserve Bank of Philadelphia. The views expressed here are solely those of the authors and do not necessarily reflect those of the Federal Reserve System.

References

- [1] Abeyasinghe, T. (2000), “Modeling Variables of Different Frequencies,” *International Journal of Forecasting*, 16, 117-119.
- [2] Altissimo, F., Bassanetti, A., Cristadoro, R., Forni, M., Hallin, M., Lippi, M., Reichlin, L. and Veronese, G. (2001), “Eurocoin: A Real Time Coincident Indicator of the Euro Area Business Cycle,” CEPR Discussion Paper No. 3108.
- [3] Aruoba, S.B. (2008), “Data Revisions are not Well-Behaved,” *Journal of Money, Credit and Banking*, 40, 319-340.
- [4] Bai, J. and Ng, S. (2006), “Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-augmented Regressions,” *Econometrica*, 74, 1133-1150.
- [5] Boivin, J. and Ng, S. (2006), “Are More Data Always Better for Factor Analysis?,” *Journal of Econometrics*, 127, 169-194.
- [6] Burns, A.F. and Mitchell, W.C. (1946), *Measuring Business Cycles*. New York: NBER.
- [7] Camacho, M. and Perez-Quiros, G. (2008), “Introducing the Euro-STING: Short Term INdicator of Euro Area Growth,” Manuscript, Bank of Spain.
- [8] Conrad, W. and Corrado, C. (1979), “Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates,” *Journal of Economic Dynamics and Control*, 1, 177-198.
- [9] Corradi, V., Fernandez, A. and Swanson, N.R. (2007), “Information in the Revision Process of Real-Time Datasets,” Manuscript, University of Warwick and Rutgers University.
- [10] Diebold, F.X. (2003), “‘Big Data’ Dynamic Factor Models for Macroeconomic Measurement and Forecasting” (Discussion of Reichlin and Watson papers), in M. Dewatripont, L.P. Hansen and S. Turnovsky (Eds.), *Advances in Economics and Econometrics, Eighth World Congress of the Econometric Society*. Cambridge: Cambridge University Press, 115-122.
- [11] Diebold, F.X., Rudebusch, G.D. and Aruoba, B. (2006), “The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach,” *Journal of Econometrics*, 131, 309-338.

- [12] Diebold, F.X. and Rudebusch, G. (1996), "Measuring Business Cycles: A Modern Perspective," *Review of Economics and Statistics*, 78, 67-77.
- [13] Durbin, J. and Koopman, S.J. (2001), *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
- [14] Evans, M.D.D. (2005), "Where Are We Now?: Real Time Estimates of the Macro Economy," *The International Journal of Central Banking*, September, 127-175.
- [15] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), "The Generalized Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82, 540-554.
- [16] Galbraith, J. and Tkacz, G. (2007), "Electronic Transactions as High-Frequency Indicators of Economic Activity," Manuscript, McGill University and Bank of Canada.
- [17] Geweke, J.F. (1977), "The Dynamic Factor Analysis of Economic Time Series Models," in D. Aigner and A. Goldberger (eds.), *Latent Variables in Socioeconomic Models*. Amsterdam: North Holland, 365-383.
- [18] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004), "The MIDAS Touch: Mixed Data Sampling Regression Models," Manuscript, University of North Carolina.
- [19] Giannone, D., Reichlin, L. and Small, D. (2008), "Nowcasting: The Real Time Informational Content of Macroeconomic Data," *Journal of Monetary Economics*, in press.
- [20] Hall, R.E., et al. (2003), "The NBER's Recession Dating Procedure," Available at <http://www.nber.org/cycles/recessions.html>.
- [21] Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357-384.
- [22] Howrey, E.P. (1984), "Data Revision, Reconstruction, and Prediction: An Application to Inventory Investment," *Review of Economics and Statistics*, 66, 386-393.
- [23] Jungbacker, B. and Koopman, S.J. (2008), "Likelihood -Based Analysis of Dynamic Factor Models," Manuscript, Free University of Amsterdam.
- [24] Kim, C.-J. and Nelson, C.R. (1998), *State Space Models with Regime Switching: Classical and Gibbs Sampling Approaches with Applications*. Cambridge, Mass.: MIT Press.

- [25] Liu, H. and Hall, S.G. (2001), "Creating High-frequency National Accounts with State-Space Modelling: A Monte Carlo Experiment," *Journal of Forecasting*, 20, 441-449.
- [26] Lucas, R.E. (1977), "Understanding Business Cycles," *Carnegie Rochester Conference Series on Public Policy*, 5, 7-29.
- [27] Mariano, R.S. and Murasawa, Y. (2003), "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18, 427-443.
- [28] McGuckin, R.H., Ozyildirim, A. and Zarnowitz, V. (2003), "A More Timely and Useful Index of Leading Indicators," Manuscript, The Conference Board.
- [29] Proietti, T. and Moauro, F. (2006), "Dynamic Factor Analysis with Non Linear Temporal Aggregation Constraints," *Applied Statistics*, 55, 281-300.
- [30] Sargent, T.J. and Sims, C.A. (1977), "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory," in C. Sims (ed.), *New Methods in Business Research*. Minneapolis: Federal Reserve Bank of Minneapolis.
- [31] Shen, C.-H. (1996), "Forecasting Macroeconomic Variables Using Data of Different Periodicities," *International Journal of Forecasting*, 12, 269-282.
- [32] Stock, J.H. and Watson, M.W. (1989), "New Indexes of Coincident and Leading Economic Indicators," *NBER Macro Annual*, Volume 4. Cambridge, Mass.: MIT Press.
- [33] Stock, J.H. and Watson, M.W. (1991), "A Probability Model of the Coincident Economic Indicators." In K. Lahiri and G. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge: Cambridge University Press, 63-89.
- [34] Stock, J.H. and Watson, M.W. (2002), "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20, 147-162.

Appendix: Trend Representations

Here we provide the mapping between the “unstarred” and “starred” c 's and δ 's in equation (4) of the text, for cubic polynomial trend. Cubic trends are sufficiently flexible for most macroeconomic data and of course include linear and quadratic trend as special cases.

The “unstarred” representation of third-order polynomial trend is

$$\sum_{j=0}^{D-1} [c + \delta_1 (t - j) + \delta_2 (t - j)^2 + \delta_3 (t - j)^3],$$

and the “starred” representation is $c^* + \delta_1^*(t) + \delta_2^*(t)^2 + \delta_3^*(t)^3$. We seek the mapping between $(c, \delta_1, \delta_2, \delta_3)$ and $(c^*, \delta_1^*, \delta_2^*, \delta_3^*)$.

The requisite calculation is tedious but straightforward. The “unstarred” expression can be expanded as

$$\begin{aligned} & \sum_{j=0}^{D-1} c + \delta_1 \sum_{j=0}^{D-1} (t - j) + \delta_2 \sum_{j=0}^{D-1} (t - j)^2 + \delta_3 \sum_{j=0}^{D-1} (t - j)^3 \\ = & Dc + \delta_1 \sum_{j=0}^{D-1} t - \delta_1 \sum_{j=0}^{D-1} j + \delta_2 \sum_{j=0}^{D-1} t^2 + \delta_2 \sum_{j=0}^{D-1} j^2 - 2\delta_2 \sum_{j=0}^{D-1} tj \\ & + \delta_3 \sum_{j=0}^{D-1} t^3 - \delta_3 \sum_{j=0}^{D-1} j^3 - 3\delta_3 \sum_{j=0}^{D-1} t^2 j + 3\delta_3 \sum_{j=0}^{D-1} tj^2 \\ = & Dc - \delta_1 \sum_{j=0}^{D-1} j + \delta_2 \sum_{j=0}^{D-1} j^2 - \delta_3 \sum_{j=0}^{D-1} j^3 + t \left[D\delta_1 - 2\delta_2 \sum_{j=0}^{D-1} j + 3\delta_3 \sum_{j=0}^{D-1} j^2 \right] \\ & + t^2 \left[D\delta_2 - 3\delta_3 \sum_{j=0}^{D-1} j \right] + t^3 (D\delta_3). \end{aligned}$$

But note that

$$\sum_{j=0}^{D-1} j = \frac{D(D-1)}{2} \quad \sum_{j=0}^{D-1} j^2 = \frac{D(D-1)(2D-1)}{6} \quad \sum_{j=0}^{D-1} j^3 = \left[\frac{D(D-1)}{2} \right]^2$$

which yields

$$\begin{aligned}c^* &= Dc - \frac{\delta_1 D(D-1)}{2} + \frac{\delta_2 D(D-1)(2D-1)}{6} - \frac{\delta_3 [D(D-1)]^2}{4} \\ \delta_1^* &= D\delta_1 - \delta_2 D(D-1) + \frac{\delta_3 D(D-1)(2D-1)}{2} \\ \delta_2^* &= D\delta_2 - \frac{3\delta_3 D(D-1)}{2} \\ \delta_3^* &= D\delta_3.\end{aligned}$$

Figure 1
Smoothed Real Activity Factor, Full Model (GEIS)

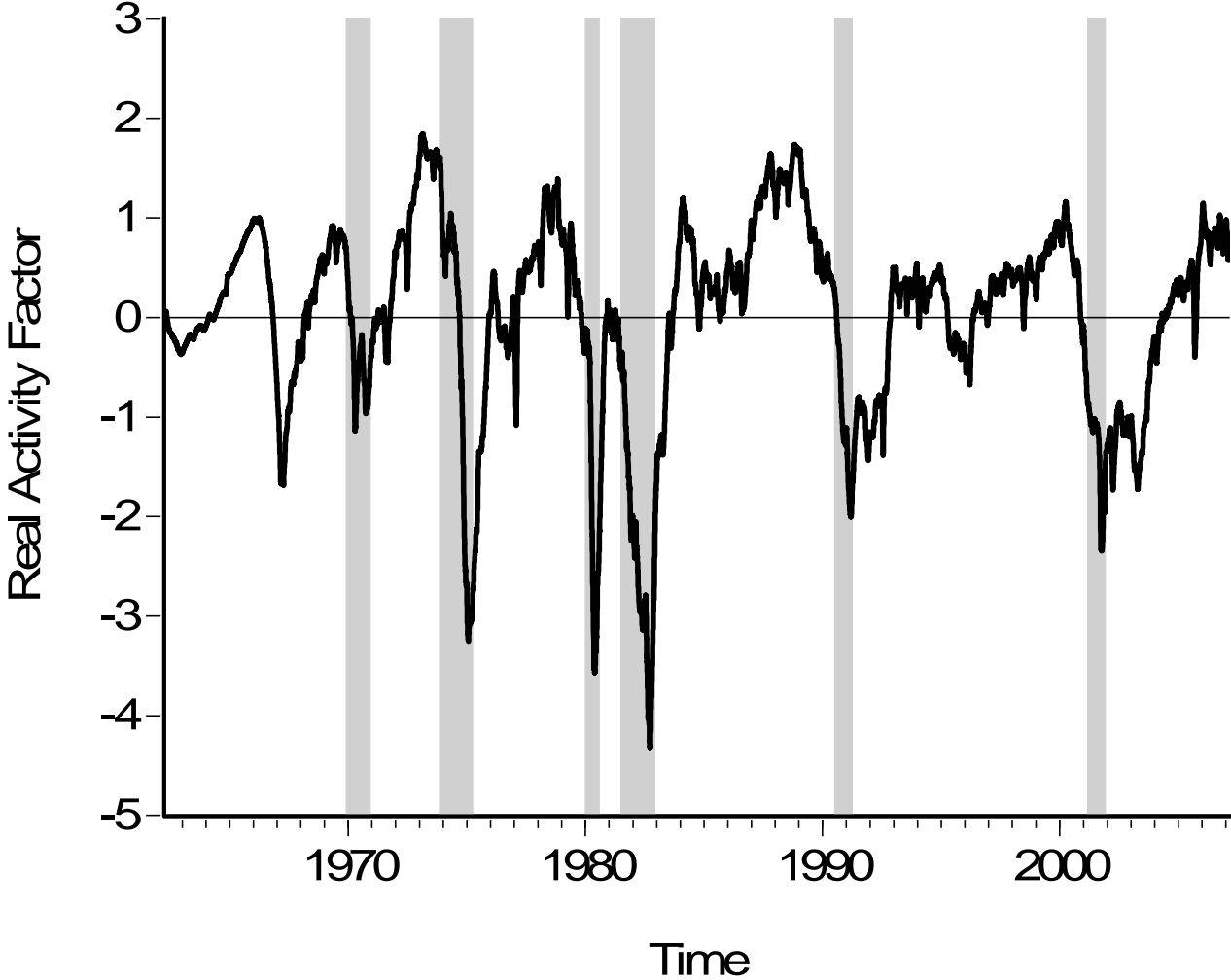


Figure 2
Smoothed Real Activity Factors: GE (Interval) and GEI (Point)

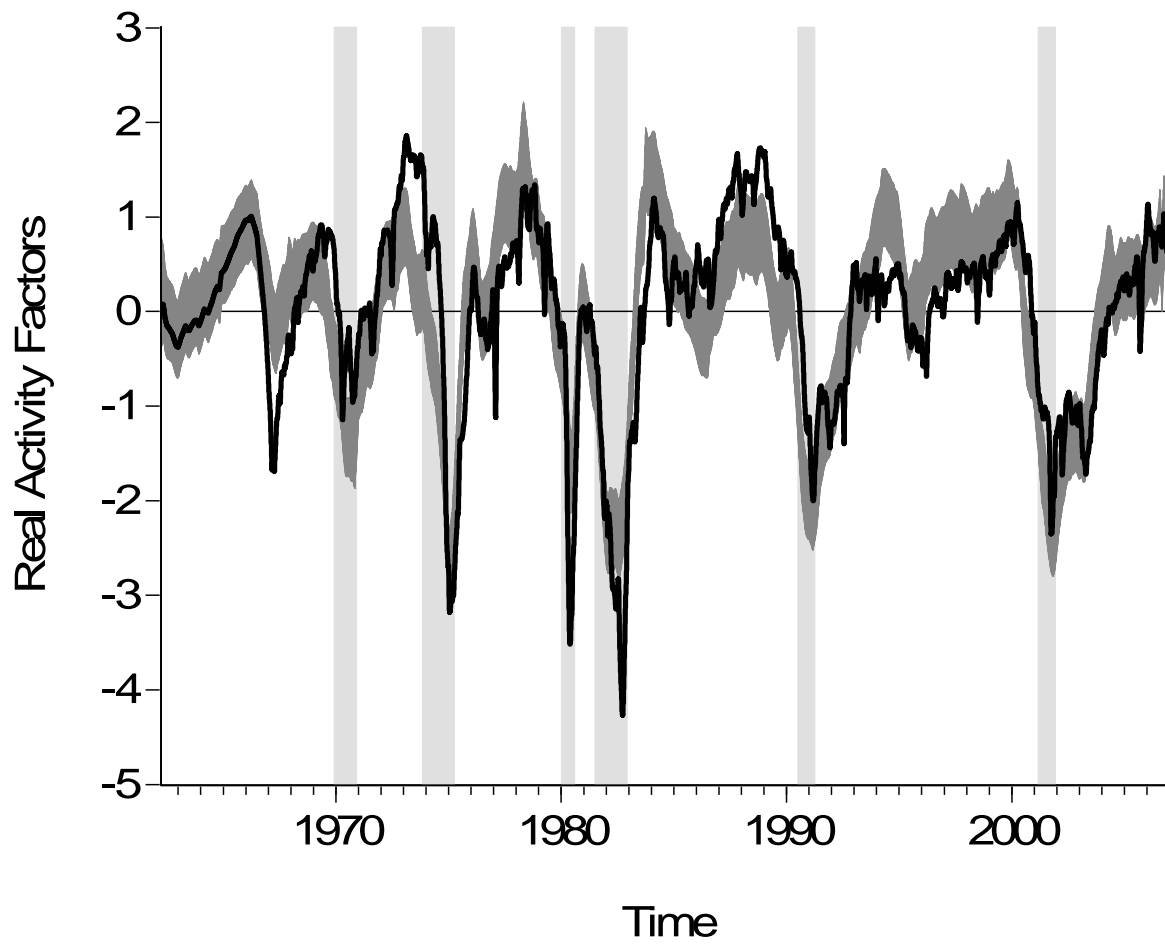


Figure 3
Smoothed Real Activity Factors Around NBER Recessions
GE (Interval) and GEI (Point)

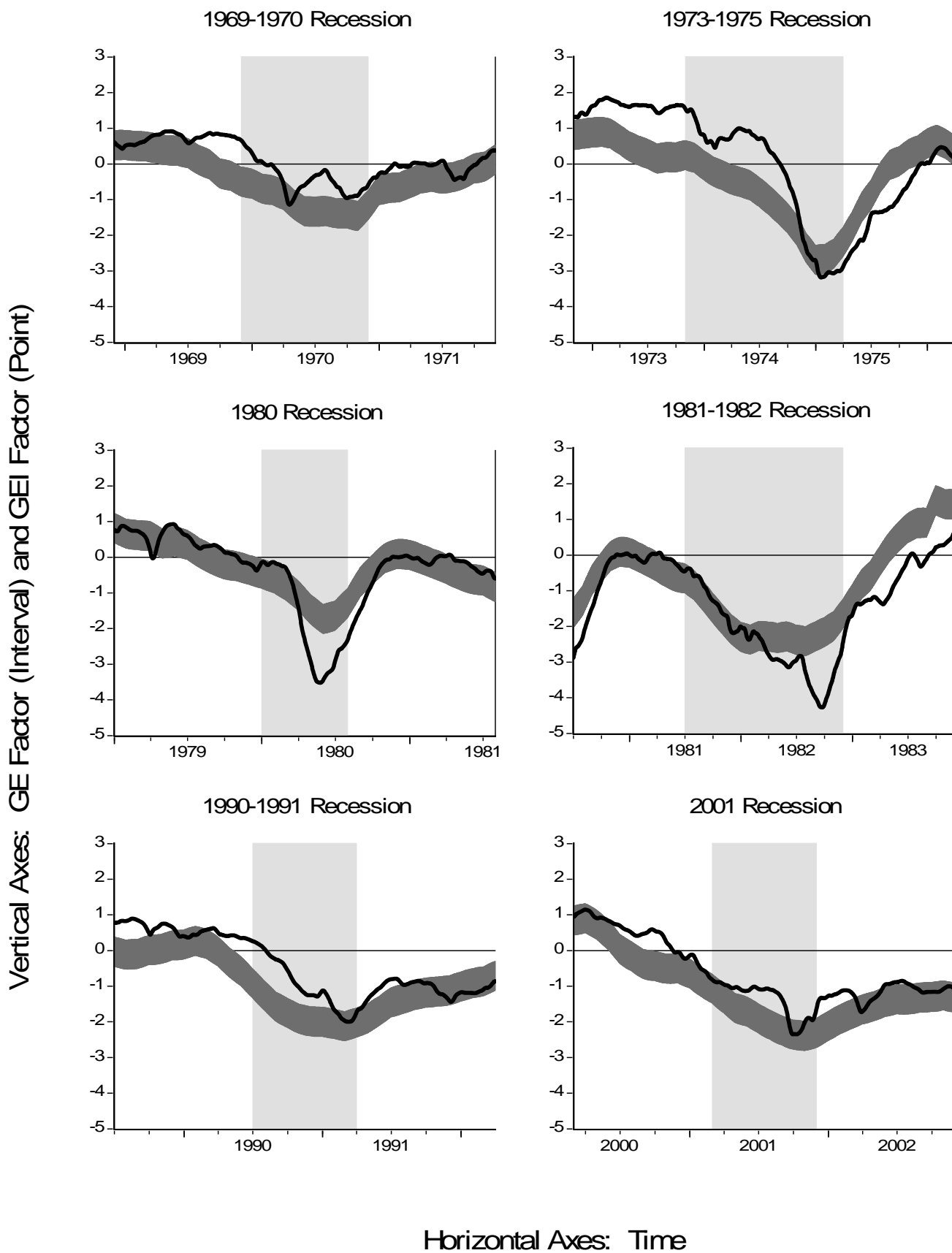


Figure 4
Filtered Real Activity Factors Around NBER Recessions
GE (Thin) and GEI (Thick)

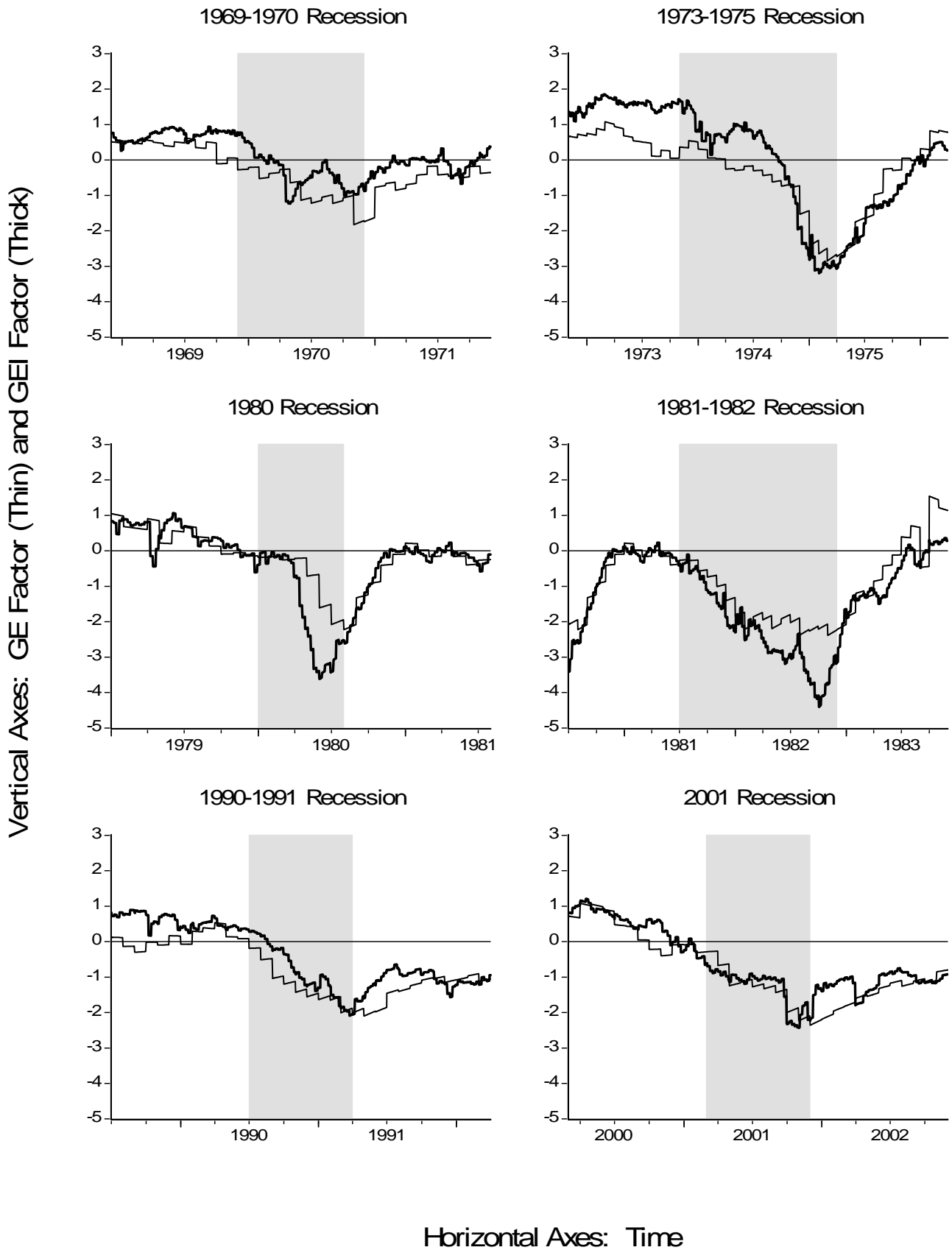


Figure 5
Simulated Real Activity Factors
Smoothed (Dashed) and True (Solid)

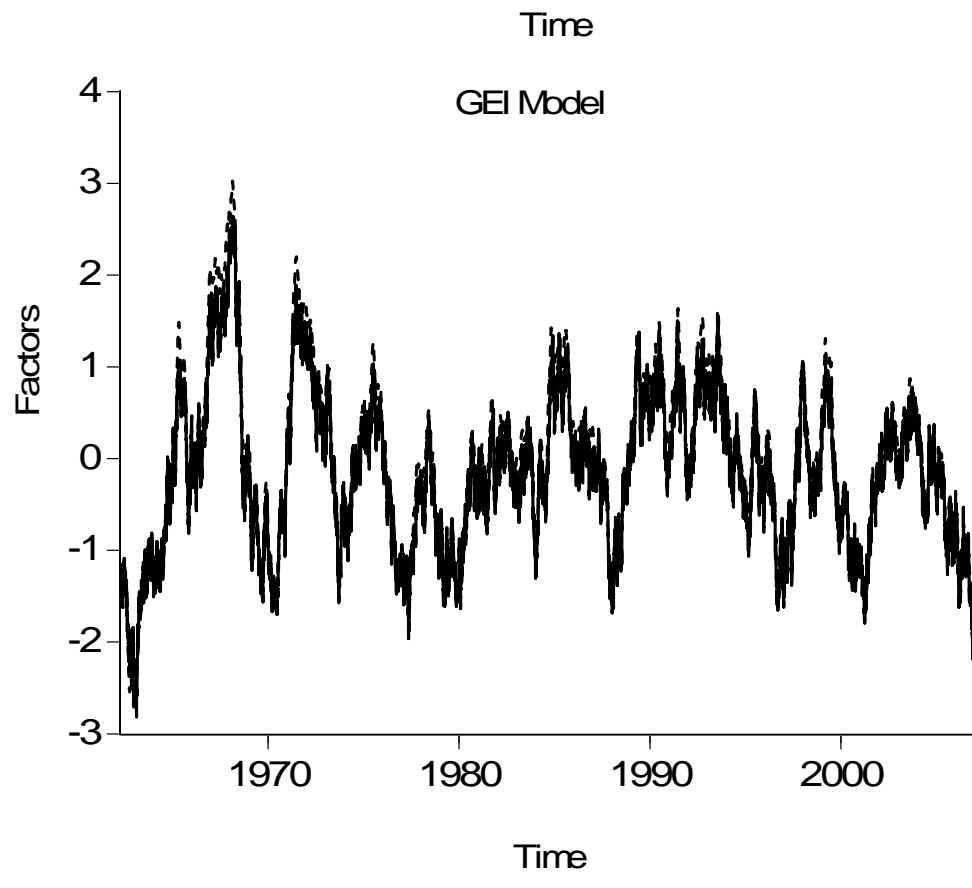
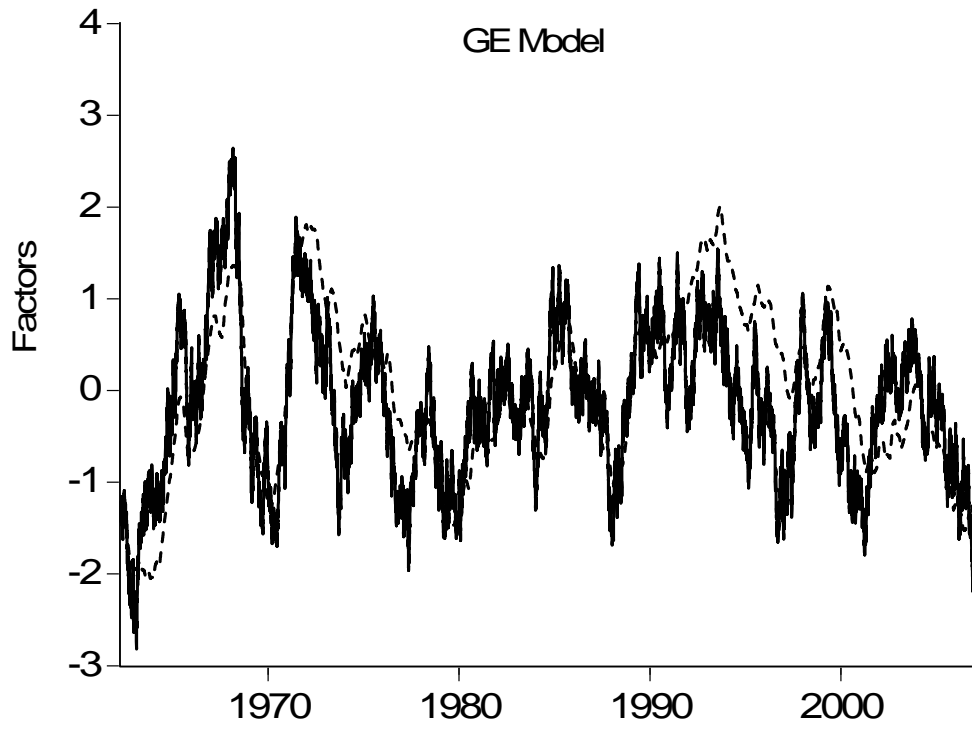


Figure 6
Tentacle Plot: Six Real Activity Paths Assessed in 2008

