A Non-Linear Forecasting Model of GDP Growth

David N. DeJong* 
Department of Economics 
University of Pittsburgh 
Pittsburgh, PA 15217

Roman Liesenfeld 
Department of Economics 
Universität Tübingen 
72074 Tübingen, Germany

Jean-François Richard 
Department of Economics 
University of Pittsburgh 
Pittsburgh, PA 15217

October 14, 2002

Abstract

We develop a regime-switching model of GDP growth under which an observable indicator variable we refer to as a tension index is modeled as undergoing stochastic and unobservable regime changes. These changes then feed into the behavior of GDP growth. The tension index is constructed as the geometric sum of deviations of actual GDP growth from a corresponding “sustainable rate” (interpretable as the growth rate of potential GDP). We track its behavior using an autoregressive model that moves between expansionary and contractionary regimes. At the onset of an expansionary regime, the index begins to grow geometrically towards a newly established target level; as the index increases, the probability that a regime change will be triggered increases. We model this probability using a logistic specification that is increasing in the absolute value of the index. In the event of a regime change, the process becomes reversed, and the index begins to decline geometrically towards a new target level. Linking the behavior of the tension index to GDP growth enables us to capture floor and ceiling effects. Interpreted in this context, the longevity of the economic expansion of the 1990s is attributable to the moderate nature of output growth observed over this period, which generated correspondingly moderate levels of the tension index, and low probabilities of a reversal.

Keywords: regime switching; error correction; efficient importance sampling

*We thank Chris Briem and Kevin Sontheimer for providing comments and suggestions on a preliminary draft of the paper.
1 Introduction

It has long been recognized that success in modeling and forecasting GDP growth hinges critically on the characterization of non-linearities inherent in its behavior. The fact that the volatility of GDP growth fell dramatically in approximately the mid-1980s has made this characterization all the more challenging. Here, we are working to develop a univariate reduced-form model of GDP growth designed to capture both aspects of this behavior.

The model is in the regime-switching family, and thus has as antecedents the work of Hamilton (1989) and Tong (1990). Models in the tradition of Hamilton characterize movements between regimes as being governed by unobserved regime indicators. A recent example of work in this tradition is that of Kim, Morley and Piger (2002), who extend Hamilton’s Markov-switching model in order to capture the tendency of output growth to be particularly strong following recessions (i.e., to capture the post-recession “bounce-back” effect). Models in the tradition of Tong characterize movements between regimes as being governed by observed indicators; i.e., indicators triggered deterministically as functions of current and past GDP growth. For example, Beaudry and Koop (1993) supplement an autoregressive (AR) model of GDP growth with a “depth-of-recession” variable, which becomes non-zero when triggered by a negative-growth indicator. The inclusion of this variable enables them to capture the bounce-back effect. Pesaran and Potter (1997) extend this work by adding an additional “overheating” variable, which becomes non-zero when triggered by a rapid-growth indicator. This enables them to capture both a bounce-back effect and a “ceiling” effect, under which GDP growth is relatively low during incidences of “overheating”.

Our model is a hybrid, incorporating features of both traditions. Specifically, we track the behavior of an observed indicator variable we refer to as a “tension index”; the index is modeled as moving between expansionary and contractionary regimes when triggered by an unobservable regime indicator. We then model the behavior of GDP growth as a function of the tension index, and thus obtain a non-linear forecasting model that appears particularly adept at anticipating business-cycle turning points.

The tension index is illustrated in Figure 1, along with NBER-defined recessions represented as shaded bars. The index is constructed as the geometric sum of deviations of actual GDP growth from a corresponding “sustainable rate” (interpretable as the growth rate of potential GDP). The particular version of the index presented in Figure 1 was constructed using the geometric-decay factor 0.65 as the weight assigned to the deviation observed at time $t - j$ to the time-$t$ value of the index; alternative factors yield similar behavior. (For comparison, GDP growth is illustrated in Figure 2. This series is measured as logged differences in quarterly GDP, measured in chain-weighted 1996 prices, annualized by multiplying by 400, measured from 1947:1 through 2002:II.)

Note from Figure 1 that local peaks in the tension index tend to precede NBER-defined business-cycle peaks by several quarters, and local troughs in the index tend to coincide or slightly precede NBER-defined business-cycle troughs. Thus the index appears to have good potential for use in anticipating business-cycle turning points. Note also that the index tends to move fairly
systematically between its local peaks and troughs. In obtaining the results presented below, we have modeled the index as an AR process that moves between expansionary and contractionary regimes. At the onset of an expansionary regime, the index begins to grow geometrically towards a newly established target level. As the index increases, the probability that a regime change will be triggered increases. We model this probability using a logistic specification that is increasing in the absolute value of the index. In the event of a regime change, the process becomes reversed, and the index begins to decline geometrically towards a new target level. Interpreted in this context, the longevity of the economic expansion of the 1990s is attributable to the moderate nature of output growth observed over this period, which generated correspondingly moderate levels of the tension index, and low probabilities of a reversal.

We estimate the model by maximizing its corresponding likelihood function. Evaluation of this function poses two challenges. First, since target levels that prevail in each regime are treated as unobservable random variables, likelihood evaluations must be obtained by integrating over the distribution of these variables. This is accomplished using the efficient importance sampling (EIS) procedure of Richard and Zhang (1998). This procedure produces as a by-product smoothed estimates of the target levels; inferences regarding the end-of-sample target level facilitate forecasts of the future trajectory of the index.

Second, likelihood evaluation cannot be accomplished without first conditioning on the complete set of regime-change dates realized over the sample period. This requires the use of an iterative procedure to insure the selection of a coherent set of dates. The procedure we employ works as follows. We begin by specifying an initial sequence of regime-change dates, with which we obtain conditional ML estimates. We then use the estimated model to assess the validity of each of the chosen dates. For a given date, we do this by calculating the probability (according to the estimated model) that the date in fact featured a break, relative to a sequence of alternative possibilities. These include the possibility that the break occurred at an alternative date in a given neighborhood of the chosen date, and that no break occurred during the time period in question. If any of the chosen dates in the original sequence is not validated in this process, the dates are then re-aligned to coincide with those assigned maximum probabilities in this first stage, and the process is repeated. The procedure ends when the pre-chosen dates represent the most likely scenario according to the estimated model. This iterative procedure is akin to the EM-type algorithms typically employed in estimated Markov-switching models (as outlined, e.g., in Hamilton, 1994).

Once model estimates are obtained, forecasts of the tension index, GDP growth, regime-change probabilities, regime-change dates, and target levels are obtained via simulation. Also, we use output from the model to forecast the occurrence of NBER turning points. This is done by fitting the historical relationship observed between contemporaneous values of the tension index, smoothed target levels, and regime-change probabilities with NBER-defined expansions and recessions using a probit model. Feeding forecasted values of these variables into the estimated probit model yields forecasted probabilities of the occurrence (or continuation) of an NBER-defined recession.
2 The Model

The model we use to describe and forecast GDP growth consists of two components. The first is the observed tension index illustrated in Figure 1 which is designed to reflect the state of the economy between expansionary and contractionary regimes. The second is a non-linear autoregressive model for this index which exploits predictable patterns in its behavior. The corresponding specification of the GDP growth arises from the stochastic model of the tension index.

We begin by characterizing the tension index. As noted above, this index is constructed as a geometric sum of deviations of the actual GDP growth from a corresponding sustainable rate, which might be interpreted as the growth rate of the potential GDP. Let $g_t$, $t : 1 \rightarrow T$ denote the growth rate of real GDP observed at period $t$ and $g_t^*$ the corresponding sustainable growth rate. According to standard theories of economic growth, the sustainable rate is determined by factors like technical progress, accumulation of capital and population growth. While this rate is not directly observable, it can be estimated. Based upon its economic interpretation one would expect that $g_t^*$ evolves smoothly over time, if at all. In the results presented here we estimated $g_t^*$ by fitting a third-order polynomial in time to the series of $g_t^1$.

The tension index denoted by $y_t$ is given by the following filter of the deviations $(g_t - g_t^*)$:

$$y_t = \sum_{i=0}^{\infty} \delta (g_{t-i} - g_{t-i}^*).$$

(1)

Since it cumulates weighted deviations of the actual from sustainable growth rate, $y_t$ can be interpreted as an error integral. During an expansionary (contractionary) phase which is typically characterized by $g_t > g_t^*$ ($g_t < g_t^*$) the tension index grows (declines), and the longer the expansionary (contractionary) phase is ongoing the greater (smaller) the tension index becomes.

The parameter $\delta$ controls the persistence of the influence of past deviations on current $y_t$. Possible sources of persistence in positive deviations include, for example, inventory depletions and labor-market rigidities. The series of $y_t$ in Figure 1 was produced by setting $\delta$ equal to 0.65. This is the value resulting from optimizing a criterion function specified in footnote 2 below, but the empirical results which are presented below are fairly robust to a wide range of alternative values between 0 and 1. An alternative representation of the tension index is given by

$$y_t = \delta y_{t-1} + (g_t - g_t^*).$$

(2)

Thus the tension index is a first-order autoregressive process where the innovations are the deviation of the actual from the sustainable growth rate.

Having characterized the tension index, we now discuss the stochastic non-linear model designed to capture its distributional and dynamic behavior. The key features of this model are, first, that the index alternates between two regimes: one in which the index increases on average, and one in which the index decreases on average; and second, that the regime changes

---

1We obtained similar results using the mean of $g_t$ as a constant value for $g_t^*$ and a $g_t^*$-series based upon the Hodrick-Prescott filter which is more volatile than the third-order polynomial time trend.
are stochastic, in which case the probability of a change depends upon the absolute value of the tension index.

In particular, the tension index is modeled as a Gaussian first-order autoregressive process given by

$$y_t = \mu_{jt} + \rho y_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d.} \mathcal{N}(0, \sigma^2),$$

where the index $j_t$ denotes regime $j$, $(j : 1 \rightarrow J)$ prevailing in period $t$ and $\mu_{jt}$ represents the stochastic latent target level in regime $j$. The following considerations are relevant to the selection of a stochastic representation for the latent target level: Since $\mu_{jt}$ determines the average tendency of the movements of $y_t$, it should be positive during periods of economic expansion and negative during periods of contraction. Furthermore, Figure 1 reveals that relatively sharp (moderate) increases in the tension index tend to give way to relatively sharp (moderate) decreases. Such a form of persistence in the behavior of the index suggests that we model the absolute value of the target level as an autoregressive process. Based on these considerations, the following specification for the latent target level is used:

$$\mu_{jt} = s_t \exp\{\lambda_j\}, \quad \text{with} \quad \lambda_j = \kappa_0 + \kappa_1\lambda_{j-1} + \eta_j, \quad \eta_j \sim \text{i.i.d.} \mathcal{N}(0, \sigma^2),$$

where $s_t$ is a variable indicating the sign of the target level, that is

$$s_t = \begin{cases} 
1, & \text{if } t \text{ is in an expansionary regime} \\
-1, & \text{if } t \text{ is in a contractionary regime}. 
\end{cases}$$

Furthermore, it is assumed that the innovations $\epsilon_t$ and $\eta_j$ are uncorrelated. Note that this specification implies that the target level of the tension index is decomposed into two components - the direction represented by the observable process $s_t$ and the absolute value given by the latent autoregressive lognormal process $\exp\{\lambda_j\}$.

In order to model the probability of a transition from the state of an expansionary regime in period $t$ ($s_t = 1$) into the state of a contractionary regime in period $t + 1$ ($s_{t+1} = -1$), and vice versa, we use the following logistic model:

$$\pi_t = P(s_{t+1} = -s_t | s_t = s_t, y_t) = \frac{1}{1 + \exp\{b_0 - b_1 s_t y_t\}}.$$  

This specification reflects the idea that as the absolute value of the index $s_t y_t$ increases during an expansionary or a contractionary phase, the probability that a regime change will be triggered increases².

Since we are interested in characterizing and forecasting the GPD growth $y_t$, it is useful to present its specification implied by the stochastic model of the tension index (3) together with

²Note that under this specification, the average value of the index which is necessary to trigger a regime change is given by $b_0 / b_1$ and the corresponding variance is proportional to $1/b_1^2$. We chose $\delta = 0.65$ in Equation (1) by searching for the value of $\delta$ that yielded the highest corresponding maximum likelihood estimate of $b_0$ and hence the highest value of the $t$-statistic associated with a regime change. As noted, this turned out to not be particularly important, as the results presented below are robust to a wide range of alternative values of $\delta$.  

4
its definition (1), that is

\[
g_t - g_t^* = \mu_{j_t} + (\rho - \delta) y_{t-1} + \epsilon_t \quad (7)
\]

\[
= \mu_{j_t} + (\rho - \delta) \sum_{i=1}^{\infty} \delta^i(g_{t-i} - g_{t-i}^*) + \epsilon_t .
\]

Hence, the GDP growth measured in deviation from the sustainable rate follows an infinite-order autoregressive process with a stochastic drift. Note that this specification is similar to the Markov-switching models proposed by Hamilton (1989,1994). The major difference in our case is that regime changes are triggered by the behavior of an observable index according to Equation (6).

3 Estimation

3.1 ML-EIS

In order to estimate the model (3)-(6), the persistence parameter in the definition of the tension index \( \delta \) and the complete set of regime-change dates realized over the sample period must be known. In this subsection we describe the technique employed to estimate the model for a given value of \( \delta \) (used to construct the series of \( y_t \)) and a set of regime-change dates. As noted in footnote 2, \( \delta \) was chosen in order to maximize the estimated value of \( b_0 \) in Equation (6). In the following subsection, we discuss the iterative procedure we employ for the selection of the regime-change dates.

According to Equation (4), the target level of the tension index \( \mu_{j_t} \) is driven by the latent autoregressive process \( \lambda_{j_t} \). Hence, the evaluation of the likelihood function associated with the observable variables requires that the \( \lambda_{j_t} \)'s be integrated out from the corresponding joint density. Since the \( \lambda_{j_t} \)'s enter the model nonlinearly this integration cannot be accomplished by standard integration techniques. To overcome this problem we apply a Monte Carlo (MC) integration technique based upon the Efficient Importance Sampling (EIS) procedure developed by Richard and Zhang (1998).

Let \( d_t \) denote a dummy variable indicating whether or not the regime prevailing in period \( t \) is an expansionary one, i.e. \( d_t = (s_t + 1)/2 \), and \( e_t \) the corresponding dummy variable indicating whether or not period \( t \) is a regime-change period, i.e. \( e_t = d_{t+1}(1 - d_t) + d_t(1 - d_{t+1}) \). Furthermore, let \( f(X, \Lambda, \theta) \) represent the joint density of \( X = \{y_t, e_t\}_{t=1}^T \) and \( \Lambda = \{\lambda_j\}_{j=1}^J \) indexed by the unknown parameter vector \( \theta \). The likelihood associated with the observable variables \( X \) is given by the \( J \)-dimensional integral

\[
L(\theta) = \int f(X, \Lambda, \theta)d\Lambda . \quad (8)
\]

Let \( t(j) \) denote the regime-change period from regime \( j \) to regime \( j + 1 \), with \( t(0) \equiv 1 \) and \( t(J) \equiv T \). The integrand in Equation (8) can be factorized as follows convonably with the
specification of the model and its built-in conditional independence assumption

\[ L(\theta) = \int \prod_{j=1}^{J} h_j(\lambda_j, \theta)p_j(\lambda_j|\lambda_{j-1}, \theta)d\Lambda, \]  

(9)

with

\[ h_j(\lambda_j, \theta) = \prod_{t=t(j-1)+1}^{t(j)} \left\{ \frac{1}{\sigma} \phi \left( \frac{y_t - s_t \exp(\lambda_j) - p_j y_{t-1}}{\sigma} \right) \right\}, \]  

(10)

\[ p_j(\lambda_j|\lambda_{j-1}, \theta) = \frac{1}{\sigma_\eta} \phi \left( \frac{\lambda_j - \kappa_0 - \kappa_1 \lambda_{j-1}}{\sigma_\eta} \right), \]  

(11)

where \( \phi(\cdot) \) denotes the standard normal density. The function \( h_j(\lambda_j, \theta) \) represents the conditional joint density of \( y_{t(j-1)+1}, ..., y_{t(j)} \) and \( e_{t(j-1)+1}, ..., e_{t(j)} \) given \( \lambda_j \) resulting from Equations (3) and (6), whereas the function \( p_j(\lambda_j|\lambda_{j-1}, \theta) \) is the conditional density of \( \lambda_j \) given \( \lambda_{j-1} \) resulting from Equation (4). Following Equation (9), the likelihood can be interpreted as the expected value of \( \prod_j h_j \) on the distribution \( \prod_j p_j \). Hence, a natural MC estimate of \( L(\theta) \) for a given value of \( \theta \) based upon this factorization is given by

\[ \hat{L}_N(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left[ \prod_{j=1}^{J} \hat{h}_j(\lambda_j, \theta) \right], \]  

(12)

where \( \hat{\lambda}_j(\theta) \) denotes a trajectory drawn from the sequence of \( p_j \) densities. This MC-estimate based upon a sequence of sampling densities directly obtained from the statistical specification of the model is highly inefficient, since it completely ignores critical information of the latent process conveyed by the observation of the \( y_t \)’s. Accordingly, the MC sampling variance of \( \hat{L}_N(\theta) \) increases dramatically with the dimension of the integration problem \( J \). This implies that for all practical purposes a prohibitively large MC sample size \( N \) would be required in order to achieve a reasonable degree of accuracy for the estimates of \( L(\theta) \). See, for example, Danielsson and Richard (1993).

To overcome this efficiency problem, we employ the EIS procedure of Richard and Zhang (1998). EIS searches for a sequence of auxiliary samplers \( \{m_j(\lambda_j|\lambda_{j-1}, a_j)\}_{j=1}^{J} \), that exploits the sample information of the \( \lambda_j \)’s conveyed by the \( y_t \)’s. Typically, this sequence of auxiliary samplers are straightforward parametric extensions of the sequence of natural samplers \( \{p_j(\lambda_j|\lambda_{j-1}, \theta)\}_{j=1}^{J} \), which are, according to Equation (11), Gaussian samplers. For a given value of the auxiliary parameters \( A = \{a_j\}_{j=1}^{J} \), the integral (9) can be rewritten as

\[ L(\theta) = \int \prod_{j=1}^{J} h_j(\lambda_j, \theta)p_j(\lambda_j|\lambda_{j-1}, \theta)m_j(\lambda_j|\lambda_{j-1}, a_j)d\Lambda \]  

(13)

and the corresponding importance sampling MC estimate of the likelihood is given by

\[ \hat{L}_N(\theta, A) = \frac{1}{N} \sum_{i=1}^{N} \left[ \prod_{j=1}^{J} \frac{\hat{h}_j(\lambda_j(\theta), \theta)p_j(\lambda_j(\theta)|\lambda_{j-1}(\theta))\hat{\lambda}_j(\theta)(a_j, \theta)}{m_j(\lambda_j(a_j)|\lambda_{j-1}(a_j), a_j)} \right], \]  

(14)

\(^3\)For ease of notation it is assumed that the initial condition \( \lambda_0 \) is a known constant, but alternative assumptions on it can easily be accommodated.
where \( \{ \hat{\lambda}^{(i)}_j(a_j) \}_{j=1}^J \) denotes a trajectory drawn from the sequence of auxiliary importance samplers \( m_j \).

In order to minimize the MC sampling variance of \( \hat{L}_N(\theta, A) \) for a given MC sample size \( N \), EIS searches values of the \( a_j \)'s which provide a good match between numerator and denominator in Equation (14). Since this represents a very high-dimensional minimization problem, it has to be broken down into manageable subproblems. In particular, the factorized form of the Equations (13) and (14) suggests constructing a sequence of subproblems for each regime \( j \) separately. Let, therefore, \( k(\lambda_j, \lambda_{j-1}, a_j) \) be a functional approximation for the density \( h_j(\lambda_j, \theta)p_j(\lambda_j|\lambda_{j-1}, \theta) \) with the properties, that it is a density kernel for the auxiliary sampler \( m_j(\lambda_j|\lambda_{j-1}, a_j) \) and analytically integrable w.r.t \( \lambda_j \). The auxiliary sampler is then given by

\[
m_j(\lambda_j|\lambda_{j-1}, a_j) = \frac{1}{\chi(\lambda_{j-1}, a_j)} k(\lambda_j, \lambda_{j-1}, a_j)
\]

with

\[
\chi(\lambda_{j-1}, a_j) = \int k(\lambda_j, \lambda_{j-1}, a_j) d\lambda_j.
\]

Since \( m_j \) is a Gaussian samplers for \( \lambda_j|\lambda_{j-1} \) the density kernel \( k_j \) is an exponential function in \( \lambda_j \) and \( \lambda_j^2 \) and the integration constant \( \chi \) is a function in \( \lambda_{j-1} \) only. Note that a good match between the density \( h_j(\lambda_j, \theta)p_j(\lambda_j|\lambda_{j-1}, \theta) \) alone and the kernel \( k(\lambda_j, \lambda_{j-1}, a_j) \) would leave \( \chi(\lambda_{j-1}, a_j) \) unaccounted for. Since, however \( \chi(\lambda_{j-1}, a_j) \) does not depend on \( \lambda_j \) it can be transferred back into regime \( j-1 \) minimization subproblem. Based upon the factorization of the auxiliary sampler (15), EIS requires solving a simple back-recursive sequence of low-dimensional least squares problems of the form\(^4\):

\[
\hat{a}_j(\theta) = \arg \min_{a_j} \sum_{i=1}^N \left\{ \ln \left[ h_j(\hat{\lambda}^{(i)}_j(\theta), \theta) p_j(\hat{\lambda}^{(i)}_j(\theta)|\hat{\lambda}^{(i)}_{j-1}(\theta), \theta) \chi(\hat{\lambda}^{(i)}_j(\theta), \hat{a}_{j+1}(\theta)) \right] - c_j - \ln k(\hat{\lambda}^{(i)}_j(\theta), \hat{\lambda}^{(i)}_{j-1}(\theta), a_j) \right\}^2,
\]

for \( j : J \to 1 \), with \( \chi(\lambda_j, a_{j+q}) \equiv 1 \). As in Equation (12), \( \{ \hat{\lambda}^{(i)}_j(\theta) \}_{i=1}^J \) denotes a trajectory drawn from the sequence of \( p_j \) densities and the \( c_j \)'s are unknown constants to be estimated jointly with the \( a_j \)'s. Since the density kernels \( k \) are exponential functions in the latent variables and their squares, the EIS least squares problems (17) are linear in the auxiliary parameters \( a_j \)'s to be estimated. While the natural samplers \( p_j \) are highly inefficient for likelihood MC estimation, they generally suffice to produce vastly improved importance samplers; a second and occasionally third iteration of the EIS algorithm where the \( p_j \)-samplers are replaced by the previous stage importance samplers suffice to produce maximally efficient importance samplers. Finally, the EIS estimate of the likelihood function for a given value of \( \theta \) is obtained by substituting \( \{ \hat{a}_j(\theta) \}_{j=1}^J \) for \( \{ a_j \}_{j=1}^J \) in Equation (14).

ML-EIS estimates of \( \theta \) are obtained by maximizing \( \hat{L}_N(\theta, A) \) with respect to \( \theta \), using an iterative numerical optimizer. For such an optimizer to converge, it is important to use Common

---

\(^4\)For details of the implementation of EIS for models with a Gaussian latent process as we are considering here, see, for example, Liesenfeld and Richard (2002).
Random Numbers (CRNs). The use of CRNs means, that the $N$ trajectories \( \{ \lambda_j(t) \}_{j=1}^N \) used to evaluate the likelihood for each value of $\theta$ during the optimization are obtained by transformation of a fixed set of $J \cdot N$ canonical random numbers, here standardized normals\(^5\).

### 3.2 Selecting Regime-Change Dates

As noted, likelihood evaluation cannot be accomplished without first conditioning on a set of regime-change dates realized over the sample period. This requires the use of an iterative procedure to insure the selection of a coherent set of dates. The procedure we employ works as follows. We begin by specifying an initial sequence of regime-change dates, with which we obtain conditional ML estimates. We then use the estimated model to assess the validity of each of the chosen dates. For a given date, we do this by calculating the probability (according to the estimated model) that the date in fact featured a break, relative to a sequence of alternative possibilities. These include the possibility that the break occurred at an alternative date in a given neighborhood of the chosen date, and that no break occurred during the time period in question.

To explain how this is done, let $B_0$ be a $J$-vector containing the initial sequence of chosen dates, the $j$th element of which is $t(j)$. Also, let $B_{0,-j}$ denote a corresponding $(J - 1)$-vector constructed by excluding the $j$th element of $B_0$. Next, let $\theta_0$ denote the ML estimate of $\theta$ obtained using $B_0$. Finally, let $t_1 = (t(j - 1) + 1)$, and $t_2 = (t(j + 1) - 1)$, so that $[t_1, t_2]$ represents the complete set of dates between the $(j-1)$st and $(j+1)$st break dates. Then the probability that the $j$th break occurred at any point $t$ over the range $[t_1, t_2]$ is given by

$$ P(t(j) = t | B_{0,-j}) = \frac{L(\theta_0 | t(j) = t)}{L(\theta_0 | t(j) = 0) + \sum_{s=t_1}^{t_2} L(\theta_0 | t(j) = s)} \quad (18) $$

where $L(\theta_0 | t(j) = t)$ denotes the value of the likelihood function evaluated using $\theta_0$, conditional on the augmentation of $B_{0,-j}$ with the additional breakpoint $t(j) = t$, and $L(\theta_0 | t(j) = 0)$ denotes the value of the likelihood function obtained without augmenting $B_{0,-j}$ with an additional break date. To assess the validity of the $j$th break date specified under $B_0$, we evaluated Equation (18) for each date $t$ in the range $[t_1, t_2]$, as well as for $t = 0$. If $t = t(j)$ was assigned the highest probability, it was validated; otherwise $t(j)$ was re-aligned to coincide with the date that was assigned the highest probability. Once this process was completed for each of the $J$ elements of $B_0$, we specified a second vector $B_1$, obtained a second set of parameter estimates $\theta_1$, and repeated the process. The procedure ends when the dates chosen in a given round represent the most likely scenario according to the parameter estimates obtained during that round. In our experience, this process typically converged within three rounds.

\(^5\)See, for example Devroye (1986).
4 Results

4.1 Model Estimates

Parameter estimates are reported in Table 1, and the final break dates used to obtain them are reported in Table 2. Figure 3 illustrates the break dates, along with the tension index and NBER recessions. Regarding the parameter estimates, the autoregressive coefficient for the tension index $p$ is estimated at 0.79. This indicates that the half-life of a given deviation of the index from the unconditional mean prevailing in a given regime is roughly three quarters. Thus in response to a regime change, the movement of the index towards its new target level is fairly rapid. The autoregressive parameter estimated for the latent drift process is 0.49, indicating that knowledge of the drift prevailing under the current regime is helpful in forecasting the drift expected in the following regime. Finally, the estimated logit specification (6) implies that as the absolute value of the tension index rises from 10 to 14 to 17, transition probabilities from one regime to the next rise from 0.25 to 0.75 to 0.95.

These parameter estimates were obtained using the 16 break dates listed in Table 2 and illustrated in Figure 3. Of the 16 dates, six are assigned a probability of at least 90%, 10 receive probabilities of at least 70%, and 13 receive probabilities of at least 50%. With the exception of the final break - chosen at date 1999:IV - the probability in favor of a break at or in the neighborhood of the breaks we actually chose is no less than 87%.

Regarding the break chosen at 1999:IV, it receives a relatively low probability assignment from the model (7%) because it coincides with a relatively moderate value of the index (approximately 5.8, compared with an average absolute value of 12.6 at break dates). In this sense, this break is unusual in comparison with its counterparts: it does not appear to have resulted from an excessive build-up of tension. However, it is not possible to account for the ensuing business-cycle downturn without assigning a preceding break, and 1999:IV is the point at which a break is most likely to have occurred. Moreover, the tension index fell during 6 of the subsequent 7 quarters following this specified break, so even though an excessive build-up of tension does not seem to be a likely culprit in this case, the notion that a break occurred nevertheless seems plausible.

We view the failure of the current specification of the model to account for this break as a serious problem that will require us to adapt significant modifications; we are currently working towards this end. The fundamental source of difficulty lies in the well-known fact that post-war business-cycle activity became significantly dampened in approximately the mid-1980s. For example, returning to Figure 2, note that the volatility of output growth fell significantly following the recession of 1982: the standard deviation of annualized growth in quarterly output was approximately 4.7 through 1982, and 2.2 thereafter. This change has clearly translated into a change in the behavior of the tension index, which has evolved relatively smoothly and towards less extreme latent target levels following the recession of 1982. One change we plan to adopt is to model the innovations to the tension index in (3) as following a GARCH process, which seems to capture the heteroskedasticity evident in the data. We are also experimenting with modifications to the model's characterization of the trajectories followed by the tension index.
towards newly established target levels.

We conclude the discussion of the specification of break points with a note regarding the possibility that the economy may have experienced a more recent break that has ushered in a new "increase" regime. This possibility is assessed in the last row of Table 1. Two factors suggest that this is not particularly likely. First, while the tension index attains a local minimum of -8.07 in 2001:II, this would once again constitute a small break-point value by historical standards (the probability assigned by the model to a break at this point is only 12%). Second, the tension index fell in 2002:II, following only two quarters of growth. Thus according to the model, the most likely scenario is that the economy is continuing to proceed under a "decrease" regime. The forecasts presented below are based on the presumption that this is indeed the case. We conclude this section by describing how we use information from the model to forecast NBER turning points. This is done using a probit model to characterize the probability of the occurrence (or continuation) of an NBER-defined recession in any given period. We estimate the model by linking historical observations of recession/expansion outcomes with the behavior of three explanatory variables: the tension index, target levels estimated using the tension-index model, and regime-change probabilities estimated using the tension-index model. We then generate forecasts of these explanatory variables using the tension-index model, and calculate corresponding forecasts of recession probabilities by plugging the forecasted explanatory variables into the fitted probability model.

Figure 4 illustrates the ability of the explanatory variables we employ to account for recession/expansion behavior by reporting probabilities of recession occurrences inferred from their historical behavior (bottom panel). Figure 4 also illustrates actual recession occurrences (upper panel), and probabilities of recession occurrences inferred from the behavior of output growth (middle panel); this is done merely to establish a modest benchmark for comparison. Two aspects of Figure 4 are particularly noteworthy. First, relative to the probabilities inferred from the behavior of output growth, the probabilities inferred from the behavior of the explanatory variables we employ are markedly lower during non-recessionary periods in general. Second, for the last two recessions, probabilities associated with the tension index are markedly higher than those associated with output growth alone. The message from Figure 4 is that the tension-index model is reasonably effective in accounting for the occurrence of recessions/recoveries.

4.2 Forecasts

Forecasts are presented graphically in Figures 5-8. Figure 5 illustrates probabilities that the tension index will undergo a break for each quarter over the next two years (top panel), along with cumulative probabilities of a break on or before each quarter (bottom panel). Moving clockwise and beginning in the upper-left diagram, Figure 6 reports forecasts of the tension index, its target level, the probability of being in an NBER-defined recession, and the probability of undergoing a break at a one-year horizon. Figure 7 replicates Figure 6 at a two-year horizon. Finally, Figure 8 reports growth-rate forecasts at one- and two-quarter and one- and two-year horizons.

The forecasts presented in Figures 6-8 are in the form of distributions. Ranges of the variables
are indicated on the horizontal axes, and the corresponding heights of the distributions are depicted on the vertical axes. Comparisons of the height of a given distribution at alternative locations along the horizontal axis indicate the relative likelihood that the variable in question will take on the corresponding value indicated on the horizontal axis. There are three sources of dispersion in the forecasts: the shocks that interrupt the smooth transition of the tension index towards its current target level; the fact that regime changes are random rather than deterministic functions of the size of the tension index; and the fact that new target levels resulting from regime changes are random rather than deterministic functions of the target level that prevailed under the previous regime.

In interpreting our results, it is important to recall the economic scenario we have imposed in generating them: the tension index last underwent a break in 1999:IV, and is continuing along a "decrease" trajectory. The current value of the tension index is -2.9, and our estimate of its current target level \((1 - \rho)\mu_J\) is approximately -6. Recall that according to our estimates, gaps between the tension index and its current target level are cut in half within three quarters. If this pattern holds true over the next four quarters, the index will be approximately -4.5 at that time, thus we will have experienced an extended period of economic stagnation, with growth averaging approximately 1.15%. Note from Figure 6, however, that there is considerable uncertainty associated with this forecast: the modal value of the distribution associated with the index is approximately -6, but the distribution is quite diffuse. In fact, this distribution assigns 19% of its weight to positive values of the index, and the corresponding weight assigned at the eight-quarter horizon (Figure 7) is 28%.

As Figure 5 indicates, the probability that the index will undergo a break at some point in the coming year is 25%, and 50% over the next two years (bottom panel). Also, the probability of a break in any given quarter is no greater than 7% (top panel). Thus while the model foresees an upcoming period of stagnation, the forecasted decline is not so great that it carries with it a substantial threat to generate a quick reversal due to an excessive build-up of tension.

If a break does occur, of course, the average trajectory of the tension index will become reversed, and it will begin moving towards a positive target level. Although it is difficult to see in Figures 6 and 7, the forecasted value of the new target level lies in the range of 2 to 16. Note from the figures that the spike at the current value of approximately -6 dominates the forecasted distribution of the target level.

The fact that the economy seems to be continuing along a "decrease" regime raises the possibility that we may be heading towards a "double-dip" recession, like that experienced in the early 1980s. The 1-quarter-ahead forecast generated by the model (not depicted) assigns a 6% probability to the recurrence of a recession at that time; recurrence probabilities at the one- and two-year horizons are 15% and 10% (illustrated by the distributional spikes at unity in the lower-right diagrams in Figures 6 and 7). These relatively modest probabilities reflect the fact that, although the model deems the economy to be continuing along a "decrease" regime, the predicted decline is fairly moderate by historical standards.

The distributions in Figure 8 are accompanied by two solid bars, one drawn at 0, the other at the estimated "sustainable" growth rate (approximately 3%). Note that the modes of these
distributions drift rightward as the forecast horizon increases, but fail to reach the sustainable rate even after two years. Again, this reflects the relatively extended period of economic stagnation foreseen by the model.

5 Summary

In the post-war U.S., we have observed a tendency for the economy to follow extended periods of output growth in excess (or below) rates that can be sustained indefinitely. We have explored an interpretation of this behavior under which these episodes generate tension in the economy that ultimately leads to a reversal in the average trajectory of output growth. Our goal has been to exploit regularities in this pattern of behavior in order to anticipate transitions between acceleration and deceleration regimes. While promising, the results we have obtained to date clearly suggest that extensions of the model are necessary if we are to successfully characterize the dampened pattern of business-cycle activity we have experienced since the mid-1980s. We intend to pursue this by building additional flexibility into the model, so that we can account for both tranquil and volatile growth episodes within the context of a coherent statistical framework.
References


Table 1. ML-EIS Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\rho$</th>
<th>$\sigma_e$</th>
<th>$\kappa_0$</th>
<th>$\kappa_1$</th>
<th>$\sigma_\eta$</th>
<th>$b_0$</th>
<th>$b_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates</td>
<td>0.794</td>
<td>3.323</td>
<td>0.321</td>
<td>0.492</td>
<td>0.518</td>
<td>6.638</td>
<td>0.554</td>
</tr>
<tr>
<td>Asy. s.e.</td>
<td>0.042</td>
<td>0.042</td>
<td>0.039</td>
<td>0.0400</td>
<td>0.039</td>
<td>1.041</td>
<td>0.039</td>
</tr>
</tbody>
</table>

Table 2. Break Dates and Estimated Probabilities

<table>
<thead>
<tr>
<th>Break Date</th>
<th>Estimated Probability</th>
<th>Alternative Date #1</th>
<th>Probability of Alternative Date #1</th>
<th>Alternative Date #2</th>
<th>Probability of Alternative Date #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1949 : IV</td>
<td>0.93</td>
<td>1949 : II</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1950 : III</td>
<td>0.99</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1954 : I</td>
<td>0.64</td>
<td>1953 : IV</td>
<td>0.19</td>
<td>1954 : II</td>
<td>0.16</td>
</tr>
<tr>
<td>1955 : II</td>
<td>0.48</td>
<td>1955 : III</td>
<td>0.27</td>
<td>1955 : I</td>
<td>0.24</td>
</tr>
<tr>
<td>1958 : I</td>
<td>0.99</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1959 : II</td>
<td>0.99</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1960 : IV</td>
<td>0.82</td>
<td>1961 : I</td>
<td>0.18</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1966 : I</td>
<td>0.72</td>
<td>1965 : IV</td>
<td>0.24</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1970 : IV</td>
<td>0.90</td>
<td>1970 : II</td>
<td>0.04</td>
<td>1970 : I</td>
<td>0.03</td>
</tr>
<tr>
<td>1973 : I</td>
<td>0.94</td>
<td>1973 : II</td>
<td>0.05</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1975 : I</td>
<td>0.82</td>
<td>1974 : IV</td>
<td>0.14</td>
<td>1974 : III</td>
<td>0.03</td>
</tr>
<tr>
<td>1978 : II</td>
<td>0.86</td>
<td>1978 : III</td>
<td>0.07</td>
<td>1978 : IV</td>
<td>0.04</td>
</tr>
<tr>
<td>1982 : I</td>
<td>0.46</td>
<td>1980 : II</td>
<td>0.32</td>
<td>1982 : III</td>
<td>0.10</td>
</tr>
<tr>
<td>1984 : I</td>
<td>0.54</td>
<td>1984 : II</td>
<td>0.23</td>
<td>1983 : IV</td>
<td>0.11</td>
</tr>
<tr>
<td>1991 : I</td>
<td>0.53</td>
<td>1990 : IV</td>
<td>0.25</td>
<td>1991 : II</td>
<td>0.09</td>
</tr>
<tr>
<td>1999 : IV</td>
<td>0.07</td>
<td>2000 : II</td>
<td>0.04</td>
<td>2000 : I</td>
<td>0.03</td>
</tr>
<tr>
<td>2001 : III</td>
<td>0.12</td>
<td>2001 : II</td>
<td>0.05</td>
<td>2001 : IV</td>
<td>0.03</td>
</tr>
</tbody>
</table>

*NOTE:* -- denotes that alternative received probability of less than 0.01.
Figure 1: Tension Index and NBER Dates
Figure 2: Actual and “Sustainable” Growth in Real GDP (annual %)
Figure 3: Tension Index, NBER Dates, and Break Dates
Figure 4: In-Sample Predictions of NBER Dates
Figure 5: Forecasted Probabilities of Breaks
Figure 6: Forecasts, 1-Year Horizon
Figure 7: Forecasts, 2-Year Horizon
Figure 8: Growth-Rate Forecasts