FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall
Philadelphia, Pennsylvania 19106-1574
(215) 574-6428, www.phil.frb.org

Working Papers

Research Department

WORKING PAPER NO. 99-17

REGIME-SWITCHING IN EXPECTATIONS OVER THE BUSINESS CYCLE

Gwen Eudey
Board of Governors of the Federal Reserve System

Roberto Perli
University of Pennsylvania

Revised, November 1999

Regime-Switching in Expectations over the Business Cycle*

Gwen Eudey
Board of Governors of the Federal Reserve System

Roberto Perli
University of Pennsylvania

Revised, November 1999

Abstract

In this paper we argue that a plausible reason why output and other major US macroeconomic time series seem to follow a Markov switching process might be strictly related to expectations. We show that a time series of expectations of future output from the Survey of Professional Forecasters is the only one among the many we analyze that has switching properties compatible with those of output. Starting from this empirical evidence we then present a business cycle model with shocks to expectations (sunspots) that produces time series with the same properties as the US data.

*We wish to thank seminar participants at the Midwest Macroeconomics Meetings, the meetings of the Society of Economic Dynamics and Control, the Federal Reserve Bank of Cleveland, the Federal Reserve Bank of Philadelphia, the University of Iowa, Cornell University, IUPUI, Purdue University and the University of Essex for comments and suggestions. We are particularly grateful for comments from David DeJong, Richard Rogerson, and Randall Wright. Any remaining errors or omissions are our own.

1 Introduction

In this paper we build on a large literature that shows that major US macroeconomic variables follow cyclical Markov switching processes, and argue that a plausible reason why this is so has to do with expectations. We show that a time series of expectation about future output (from the Survey of Professional Forecasters) follows a switching process very similar to output's, whereas other obvious candidates as sources of this type of behavior (first and foremost the Solow residual, but also other monetary and fiscal policy variables) either cannot be classified as switching processes, or, if they can, their properties are very different from the properties of US output. This empirical evidence is consistent with a modeling strategy that emphasizes sunspots (or extrinsic uncertainty or animal spirits as they are alternatively known in the literature). We present a model where sunspots play a role that can replicate the main features of US output, consumption, labor and investment: not only the standard deviations relative to output and the correlation with output, as standard RBC models do, but also the switching properties. Our sunspots are calibrated based on the SPF time series for expectations.

As said above, a robust feature of the U.S. data seems to be that output, consumption, investment, and hours worked all follow a cyclical Markov switching process. Specifically, each of these series can be characterized as having two mean growth rates; a positive mean is associated with expansionary periods of the business cycle and a negative mean with recessions. It is possible to compute the probability of being in either one of the states: it is often very close to unity, indicating a good fit of the data to the two-state model[1]. Moreover, the probability of being in the negative growth state is always and only high during periods which the NBER has identified as recessions. Finally, these series share an additional time series property: they are more likely to remain in the high growth state than in the low growth state. In other words, expansions last longer than contractions.

We take the position that these features of the data are robust enough to be considered stylized facts of the U.S. business cycle. In addition to having interesting empirical implications that we do not address here,[2] this raises an

[1] Ang and Bekaert (1998) argue that if the data are not, with high probability, in one of the n states at all times, then the number of states may have been misspecified.

[2] For example, optimal forecasts of this type of variables should be made conditional on the belief about the current and lagged realization of the regime; forecasts about the state of the cycle (see Rotemberg and Woodford, 1996) which do not take this into account are not

important economic issue. A typical model, based on the stochastic optimal growth paradigm, will only be able to generate switching behavior over the business cycle if there is some exogenous driving process that behaves in that way. In the real business cycle framework, in particular, this disturbance would be the exogenous technology process, which can be measured by the Solow residual. We find, however, that the Solow residual does not follow a Markov process.[3] There must therefore be another driving force of the business cycle that has importance. We consider a broad set of policy variables (independently on whether or not they play a role in standard business cycle models): the federal funds rate (ffr), M1, the monetary base, domestic credit, government spending, and the government budget deficit. No set of these variables can be constructed which could be the source of this type of non-linearity in the data.

We then consider the possibility that regime-switching over the business cycle is the result of swings in expectations. Using data collected from the Survey of Professional Forecasters (SPF) about expectations for US output, we find that expectations can actually be described as a two-state Markov process. There are two distinct growth states in the survey data; expectations (forecasts) switch cyclically from positive to negative expected growth regimes; there is always a high probability of being in either state; and forecasts are more likely to remain rosy than gloomy. All this is consistent with the properties of US output.

Although there is no other viable source of exogenous regime-switching that we could find, it is still possible that expectations only follow the same process as does GDP because agents are rational. In other words, we cannot establish causality. Clearly, if we consider the probability that we are in the low growth state as a time series, then we could reject the hypothesis that expectations caused GDP movements if switches in GDP occurred (probabilistically) before switches in expectations. However, the opposite is true: regime-switching in the SPF Granger-causes that in GDP.[4] Given this fact, plus the timing and procyclicality of the SPF data, and the absence of any other plausible exogenous source of switching behavior, we consider the data consistent with a causal role for expectations in the non-linear (switching) component of the cycle, and we build a model accordingly.

---

optimal.

[3]Altug *et al* (1996) also analyzethe Solow residual, measured with and without variable capital use, and find no evidence of non-linearities of any kind.

[4]Matsusaka and Sbordone (1995), using the Michigan Survey *Index of Consumer Sentiment*, find that changes in that index Granger-cause changes in the log of real GNP. That is a somewhat different test since it does not account for regime-switching behavior in either variable, although both of those series follow Markov processes.

Our model is based upon the stochastic optimal growth model,and therefore belongs to the same category as RBC models; the main difference is that our model allows for multiple equilibria[5]. The model captures not only the switching feature of the data, but also traditional time-series properties of the data. Agents tend to think of the economy as being in either a boom or a recession; those expectations are self-fulfilling and the economy will switch between positive and negative growth rates according to the state of expectations. There are many ways of modeling switching expectations and sunspots in general; we choose to say that sunspots are overreactions to fundamental news. In particular, we make the expected regime dependent upon the value of the technology shock: if the latter is below a certain threshold, we assume that people become pessimistic (on average) about the state of the economy and think and act as if the economy were in a recession (and since expectations are self fulfilling, the economy actually enters a recession). If, on the other hand, the technology shock is observed to be above another given threshold, people are optimistic (on average) and act as if the economy were in a boom, thereby fueling an expansion. If the technology shock is observed to be between the two thresholds, agents do not change the state of their expectations.[6] So we can say that economic news affects the economy through two channels in our model: the direct effect on the resource constraint and therefore production inputs (the usual technology shock), and the indirect effect that this information has on perceptions about the state of the economy. Business cycles would exist in the model even if we were to shut down the role for expectations; *regime switching* during the cycle exists only in the case in which expectations are self-fulfilling. Thus expectations cause regime-switching to be a characteristic of the business cycle, but productivity shocks cause the business cycle itself.

We also see our data as being generally supportive of the multiple equilibria RBC framework. Self-fulfilling expectations are always a possible outcome in that framework, and yet there has been heretofore little empirical evidence demonstrating their importance over the business cycle. So far empirical tests

[5]Multiple equilibria are necessary for a model where expectations are the driving force. The first RBC model of this type was Farmer and Guo (1994); other models include Benhabib and Farmer (1996), Benhabib and Nishimura (1998) and Perli (1998a).

[6]This modeling strategy is of course an oversimplification, since it is arguably difficult to observe a technology shock in the real world. Nonetheless, we want to give the sunspot the interpretation of a reaction (unnecessary from a fundamental point of view, but possible and consistent with rational expectations) to some fundamental variable. The way we do it is the simplest that we could think of to capture this idea. Alternatively, we could have simply said that the sunspot (expectations) switches exogenously, without any reason, and the results would not have changed.

of the implications of these models have been limited to a sort of "quantitative theory" exercise (see Kydland and Prescott (1996)): the authors fed some particular expectation shocks to their models and looked at various properties of the artificial time series the models generated (like standard deviations, correlations with output, etc.). In general those properties did not look very different from models driven by technology shocks only. Thus some have questioned the usefulness of the multiple-equilibria model as compared to a more parsimonious approach. That more parsimonious approach however is unable to explain the regime-switching behavior that we observe in the SPF, GDP, consumption, investment, and hours data.

The rest of the paper is organized as follows. In Section 2 we briefly discuss the implications of the Markov process for business cycle research and for identification of business cycle turning points. In Section 3 we describe the data and present our evidence that regime-switching is an important characteristic of the U.S. cycle. In Section 4 we present a multiple-equilibrium business cycle model of switching in expectations which matches features of the actual forecast and productivity data described in Section 3. Section 5 concludes the paper.

2 The Framework of Regime Switching

In this section we briefly review Hamilton's (1989) methodology to decide whether or not a given time series can be characterized as a Markov switching process; the details can be found in the original Hamilton's paper and in the other papers that we will cite. Hamilton (1989) and others have shown that U.S. output[7] can be characterized as following a particular type of ARIMA $(\rho, 1, \phi)$ process in which the mean, $\alpha_0 + \alpha_1 S_t$, changes stochastically and discretely with changes in the state S_t. Using Hamilton's notation[8], a series which contains a Markov component can be characterized in the following way

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 \cdot s_t + y_{t-1} + \rho_1 \cdot y_{t-1} + \rho_2 \cdot y_{t-2} + ... + \rho_p \cdot y_{t-p} + z_t \quad (1) \\ z_t &= \phi_1 \cdot z_{t-1} + \phi_2 \cdot z_{t-2} + ... + \phi_r \cdot z_{t-r} + \varepsilon_t \end{aligned}$$

where s_t is either zero or one, although its true value at any point in time can only be inferred probabilistically. In the case that it follows a first-order Markov

[7] Estimation of the Markov component is not sensitive to the definition of "output": Markov estimation for GNP (see Hamilton, 1989), GDP (used in this text), and the Index of Industrial Production (see Filardo, 1994), are qualitatively identical.

[8] See Hamilton (1989) for a technical description of the estimation of transition probabilities, state means, and historical turning points which we employ in what follows.

process, given the previous period's state, there is a fixed and known probability of remaining in that state or switching to the other state.

$$\begin{aligned} Prob[S_t &= 1 \mid S_{t-1} = 1] = p \\ Prob[S_t &= 0 \mid S_{t-1} = 1] = 1-p \\ Prob[S_t &= 0 \mid S_{t-1} = 0] = q \\ Prob[S_t &= 0 \mid S_{t-1} = 1] = 1-q \end{aligned} \tag{2}$$

In the next section we will collect quarterly data on several major macroeoc-nomic and policy variable for the US economy and use a modified program by Hamilton (based on his 1989 paper) to simultaneously estimate α_0 (the recessionary growth rate), $\alpha_0 + \alpha_1$(the expansionary state, in which $S_t = 1$), the transition probability between states (p and q), and the mathematical probability that we are in the recessionary state at any point in time in our sample.

In order for us to say that the Markov component of the series is a prominent feature of the business cycle, it must satisfy certain subjective criteria.[9] Firstly, we say that the switching process can be characterized as capturing the business cycle component of a series if the two means, α_0 and $\alpha_0 + \alpha_1$, are (statistically) significantly different and of opposite sign, so that one mean is associated with expansions and another with contractions. Secondly, the probability that we are in the negative growth state ($S_t = 0$) must always be very close to either zero or one, indicating a good fit of the two-state model to the data. Finally we will pay attention as to whether the periods with a high probability of being in a recession are compatible with the NBER recession dates.

Although NBER turning points are useful benchmarks for identifying business cycle contractions, however, we do not expect turning points identified by the NBER business cycle indicator, which are based on a *set* of data, to coincide *exactly* with those in any of our (univariate) time series. In fact, according to the particular model one has in mind for the economy, one might expect turning points in some variables to lead or lag those of the NBER dating. Thus, while the NBER helps us to identify recessionary periods, no effort is made in this paper to "score" or rank the individual series or lag specifications according to that yardstick.[10] We instead say that regime switching is not a quirky feature of the output data only, but rather a feature of all the U.S. economic variables that are typically studied in the business cycle literature (i.e., output, consumption,

[9]See for example Hamilton (1989), Filardo (1994), and Ang and Bekaert (1998).
[10]See Diebold and Rudebusch (1987).

investment and hours). Each of them might switch either at the same period as the official NBER turning points, or with a slight lead or lag, but the substantial feature of switching is what is of interest to us. In other words, although we lack a satisfactory statistical test of the hypothesis of regime switching for *each* of our variables at the exact NBER dates, we can still say that, when modelled as a two-state process, the U.S. business cycle data all behave in roughly the same way at roughly the same time and in a way consistent with our dating and understanding of the U.S. cycle.

Although we do not pursue them in this paper but rather plan to leave them for suture research, there are statistical implications of the fact that data follow a switching process that might be worth briefly mentioning here. To show the implications of this framework for standard methods of evaluating RBC models, rewrite **??** as an ARIMA process[11]:

$$s_t = (1-q) + \lambda s_{t-1} + \nu_t \tag{3}$$

$$\lambda = -1 + p + q \tag{4}$$

and where conditional on $S_{t-1} = 1$,

$$\begin{array}{ll} v_t = 1-p & \text{with probability } p \\ v_t = -p & \text{with probability } 1-p \end{array} \tag{5}$$

and conditional on S_{t-1}=0,

$$\begin{array}{ll} v_t = -(1-q) & \text{with probability } q \\ v_t = q & \text{with probability } 1-q \end{array} \tag{6}$$

Thus a Markov process can be modelled as an ARIMA with non-normally distributed errors.

One important implication of this particular type of ARIMA model is that if the series follows a regime-switching process then forecast errors may appear serially correlated *ex post* but have been rational given Equation 2.3. Thus serial correlation in forecast errors measured as the spread between the forecast and the actual value is not evidence against the assumption of rational expectations in the RBC model.[12]

A second implication of the Markov model for business cycle diagnostics is that the errors in forecasting regime-switching are not statistically independent

[11] This notation follows Hamilton (1989) exactly.

[12] Evans and Wachtel (1993) show that exactly this type of phenomena likely explains the apparent serial correlation in inflation forecast errors.

of lagged values of the series. Thus unconditional forecasts derived from a VAR are not optimal. Rotemberg and Woodford's (1996) method of model of evaluation which compares the forecasting power of the model to that of a VAR is therefore not appropriate if any of the series in the VAR follow a switching process. Similarly, the expected variance (or covariance) conditional on the belief about the current and lagged state of the world is not equal to the unconditional variance, which places asymptotic probabilities on the realization of either state.

3 Switching in the US Data

In this section we apply the methodology described above to a number of US time series. We divide our analysis between the variables most commonly associated with the business cycle, i.e., oupput, consumption, investment and hours (which we label "endogenous" variables); the variables that could be interpreted as being the exogenous reason why the endogenous variables exhibit switching, such as the Solow residual and monetary and fiscal policy variables (the "exogenous variables"); and the variable that we think might play the most important role in the explanation of the switching behavior of the endogenous variables, namely the expectations series from the Survey of Professional Forecasters.

3.1 Endogenous Variables

We use real, seasonally adjusted, quarterly data on consumption, investment[13], output (GDP)[14], and hours worked[15] over the period 1960:1-1997:4, collected from the Survey of Current Business.

The Markov model we choose to estimate requires stationarity in the data. It is standard in the real business cycle literature to apply the HP filter to the log data. It seems inappropriate, however, to fit a non-linear Markov model to data that has had a non-linear trend removed. It is also contrary to our purpose, which is to explain non-linearity in the data, rather than filter it out. For none of the series described above could we reject the hypothesis of a unit

[13] Our measure is gross private nonresidential investment.

[14] Estimation of the Markov model was not sensitive to the choice of GNP, GDP, or the Index of Industrial Production as a measure for output, as said above.

[15] Hours worked is measured as manufacturing hours time total employment (relative to the labor force). We also estimated the Markov model using household hours worked as a labor input measure but with no qualitative difference in the results. Since the latter series was not available to us for the full sample, however, we chose to present the data described in the text.

root, and so the Markov model was estimated, as in **??**, under the hypothesis of regime-switching in the growth rate of each of the series.

We considered ARIMA$(\rho, 1, 0)$ systems with up to six AR terms for each variable. The lag length which maximized the likelihood ratio varied for each series. For output and hours worked, selection of lag length had no qualitative effect on the estimation of the state means or transition probabilities. For both consumption and investment, however, only that AR specification presented in Table 1 resulted in convergence to a local maximum. Since the two-state model appears to fit these data very tightly (there is always a near-unit probability of being in either state), we interpret this as a sample-size problem, or in other words as indicating there are things going on in these two series not captured by the AR specification.

.

variable	**# of lags**	**state 1 mean**	**state 0 mean**	**p**	**q**
Output	4	.9485	-.8468	.9503	.4527
		(0.1146)	(0.4041)		
Investment	4	1.5436	-2.7206	.9583	.5505
		(0.3734)	(0.9968)		
Consumption	1	1.0189	.07180	.9570	.8544
		(.0741)	(.1837)		
Hours Worked	2	.4133	-.7691	.9107	.6031
		(.1084)	(.4161)		

Table 1: Results of Markov Estimation of Endogenous Variables

Table 1 presents the estimation of the state means ($\alpha_0 + \alpha_1$and α_0) and probability of remaining in the expansionary (p) or contractionary (q) state. We cannot reject the presence of two distinct means of opposite sign for any of the series. For all there is greater probability of remaining in an expansion than a contraction. This implies that booms last longer than recessions, which is a broadly held view of the cycle.

Figure 1 shows the probability of being in the contractionary growth state for consumption, non-residential investment, output, and hours. For Real GDP growth, we include the NBER business cycle turning points for the period. The probability of being in the contractionary state rises sharply only during recessions. This feature is particularly prominent for output, consumption, and hours worked.[16] For each of the series, there is always a high probability of being in either state, indicating a good fit of the data to the two-state system.

To summarize, we are able to identify regime-switching in all of these series. There are common features to each of the series: the two-state model appears to be a good fit of each series and discrete changes in the U.S. data occur not only at business cycle frequencies but only at or very close to those periods thought to be business cycle turning points in the US economy. There are also differences between these series which are brought out by the estimation. For each of the series, the expected variance conditional on being in a contraction is smaller than that in a boom, but the degree of asymmetry varies across the series. For example, the conditional covariance between output and consumption does not fall as much during contractions as does that between output and investment. Although we do not pursue these issues in this paper, there are interesting implications for forecasting that derive from these data.

3.2 Exogenous Variables

A necessary but not sufficient criteria for the standard business cycle model to exhibit switching behavior is that there be switching in at least one source of exogenous disturbance. Thinking from a modeling perspective (especially a stochastic optimal growth perspective), output and the other endogenous variables are not going to exhibit switching characteristics unless there is something that forces them to behave in that way.[17] The most obvious canditate as a switching-inducing variable is the Solow residual, or the technology shock that plays such a prominent role in the real business cycle literature. We also look at several other possible candidates, essentially most of the monetary and fiscal policy variables. If they are the reason why the endgenous variables switch, all or at least some of these variables should show regime switches to contractionary states at each of the turning points identified in Figure 1, but should not predict turning points at any other periods.

We consider the following candidates: the Solow residual[18], the federal gov-

[16]Although not reported here, U.S. imports also follow a regime-switching process associated with the business cycle. This is consistent with our findings for the consumption and investment data.

[17]Or there could be *deterministic* switching if the model implied deterministic, periodic cycles. This, however, typically requires parameterizations that are not compatible with rigorous calibrations.

[18]We allowed the weight on capital input to take the values .7, .6 or zero. We also considered two approximations to variable capactity utilization: one with industrial energy use as a proxy for utilization, and one using Burnside and Eichenbaum's (1998) capital series in which there is

ernment budget deficit, the monetary base, M1, domestic credit[19], the federal funds rate (ffr) [20], and the spread between the ffr and the t-bills. As before, these are real[21], seasonally adjusted, quarterly data over 1960-1997 from the Survey of Current Business. Again, we cannot reject the presence of a unit root for any of these data in our sample and therefore proceed to estimate the Markov model in growth rates as in Section 3.1.

Table 2 indicates the degree to which each of these series meet our criteria for regime changes as a characteristic of the business cycle: convergence to a local maximum in the two-state estimation, two statistically distinct means, high probability of being in either state at all times (fit), and timing of regime changes corresponding to the timing of U.S. business cycles. Neither the real ffr, domestic credit, nor the Solow residual (see earlier footnote) converged to a local maximum with two distinct means, indicating that neither of these series could be the source of such movements in consumption, investment, hours worked, or output. The fed funds / t-bill spread, which might plausibly be stripped of inflation, reveals two states corresponding to decade-long monetary regimes (loose, then tight) not at business cycle frequencies. While the budget deficit could be characterized as a switching process, the switches were not associated with the U.S., cycle, but with several longer-term regime changes over the period.[22]

depreciation through use. None of these measures, loosely labelled "the Solow residual" in this paper, demonstrated regime-switching. This is consistent with Altug, Ashley, and Patterson's (1996) result that there is no type of non-linearity in any of these aggregate productivity measures.

[19]We use the currency-depost ratio as a proxy of the amount of resources allocated to the banking system for credit allocation.

[20]In real terms, this is the *ex post* federal funds rate, or realized real rate of return. Our finding that this is locally non-stationary does not imply that this series can rise or fall without bound. See, for example, Ang and Beckaert (1998) for a discussion of other non-stationary models of the real interest rate.

[21]U.S. inflation follows a switching process that is not associated with the business cycle. Consequently each of the available nominal counterparts to these series shared that characteristic.

[22]Note that this does not at all imply that productivity or fiscal policy shocks have no affect on the business cycle, only that they do not cause that particular characteristic of the cycle.

variable	convergence	distinct states	fit	timing
Solow residual	no	-	–	–
domestic credit	no	-	–	–
real ffr	**yes**	no	–	–
budget deficit	**yes**	**yes**	no	no
real ffr / t-bill spread	**yes**	**yes**	**yes**	no
real M-base	**yes**	**yes**	**yes**	no
real M1	**yes**	**yes**	**yes**	**?**

Table 2: Markov Estimation of Exogenous Variables

Two money-supply variables were considered: real monetary base, and real M1. Both measures of the monetary base followed regime-switching processes with distinct means and greater probability of remaining in the expansionary than in the contractionary state. For neither, however, were changes in regime regularly and only associated with changes in the state of the U.S. business cycle. Both series, if they were the source of switching behavior in endogenous variables, would have caused regime changes that did not occur in the endogenous data. Real M1 is the most successful series by our criteria: it follows a two-state process with one positive and one negative growth state, high probability of being in either state at all times, and switches with high probability to the low-growth state in only two periods, both associated with U.S. recessions. The latter is shown in Figure 2, which plots the probability that real M1 is in the negative growth state. M1 alone cannot explain regime switching as a stylized fact of the U.S. cycle, but our data are consistent with the description of those two regime changes as being triggered by "monetary recessions", whatever the original cause of the recession itself.

3.3 Expectations Data

In this section we consider the possibility that expectations are the driving force behind regime-switching over the cycle. As shown in multiple-equilibria models (for example Farmer-Guo (1994), Perli (1998), and Azariadis-Smith (1998) among many others), expectations can be self-fulfilling and therefore have real economic effects. There is also considerable intuitive appeal to the idea that people think of the economy as discretely switching between two states, characterized as boom or bust, bull or bear. For example, the wording in the Michigan Survey of consumer sentiment of the two questions that refer to the macroeconomy reflect that appeal:

"Now turning to business conditions in the country as a whole–do you think that during the **next twelve months** we'll have **good** times financially, or **bad** times, or what?"

"Looking ahead, which would you say is more likely–that in the country as a whole we'll have continuous good times during the **next five years** or so, or that we will have periods of widespread **un**employment or depression, or what?"[23]

As indicated above, the Michigan Survey of consumer expectations gives a forecast about the direction of economic changes in the economy, but does not give a point estimate. We use instead data from the Survey of Professional Forecasters (SPF). The SPF asks for seasonally adjusted estimates of both GDP and prices for the current and following four quarters. From these data we derived a measure of expected real GDP growth over the coming year.[24] The SPF data are reported as both mean and median responses, but there was no interesting difference between the Markov estimates of the two series. We present results for the median response only. The SPF data, aside from a few missing observations in the middle of the sample, were available quarterly beginning in 1968 though 1997.

We are looking for several things in this data. If expectations are rational, not only should expectations follow a switching process but the means and transitional probabilities should be the same as for GDP. The timing of those regime changes should also be synchronous with those of GDP. In addition, if expectations are rational (rather than adaptive), and even more so if there is to be a causal role for expectations over the business cycle, regime changes in GDP should not lead those in the forecast.

[23] Bold appears in actual survey.

[24] The survey referred to GNP until 1992. Presumably this in unimportant, as the timing and persistence of business cycles identified with GDP are nearly identical to those using GNP.

US GDP		state 1 mean	state 0 mean
		.9068	-.9735
	s.e.	(.1109)	(.4274)
	p(q)	.9569	.4079

Survey		state 1 mean	state 0 mean
		.8060	-.6240
	s.e.	(.3736)	(.3736)
	p(q)	.9468	.4620

Table 3: Markov estimation over 1968-1997 sample

Table 3 presents a comparison of the Markov estimation results for the SPF and GDP over the common sample. The mean estimates for each state do not differ significantly across the two series, and the estimated transitional probabilities are nearly identical. Figure 3 shows the historical probability that the forecasters expected to enter a contraction over the next year. Two things are clear in that figure: the two-state model is a good fit to the data, and the changes in regime are highly correlated with the business cycle.

Matsusaka and Sbordone (1995) present empirical evidence that expectations (measured using the Michigan index) Granger-cause GNP growth. While that is supportive of the 'sunspot' class of RBC model, asymmetric behavior implied by the Markov model draws that linear specification into question: the data in Tables 1 and 3 suggest smaller covariances in expansions than in recessions. We therefore construct a slightly different causality test: we hypothesize that the probability that SPF expectations are pessimistic does not Granger-cause the probability that GDP is in a contractionary state, and find that we can reject that hypothesis with a high degree of confidence. These results are presented in Table 4. Clearly, Granger-causality of expectations does not necessarily imply a causal role: expectations may respond more quickly to news than does actual output because of time-to-build or other real rigidities. The findings are therefore consistent with, but not proof of, self-fulfilling expectations.

lags	one year ahead forecast
2	.0001
4	.0008
6	.002

Table 4: Probability that SPF does not Granger-cause GDP

4 A model of Regime-Switching over the Business Cycle

There are at least two plausible channels through which the business cycle might exhibit regime-switching. One is that there is a trigger point below or above which the economy jumps to a different steady state equilibrium. A second approach, which we adopt, is that a stationary exogenous variable follows a switching process coincident with the cycle. Our SPF data follow such a process. We use this data to calibrate a business cycle model with self-fulfilling expectations and a single steady-state growth path.

Ours is a two-sector model based on Perli (1998a); on of the sectors has the interpretation of a non-market sector. We could have used a simple one-sector model (such as the seminal Farmer and Guo (1994)), but we chose not to because to have expectations (sunspots) that matter in that model one needs high increasing returns to scale, in particular higher than it is now considered plausible relative to US data estimations. Note however that we also simulated a version of the Farmer and Guo (1994) model, and all the results below held true for that model also.

As said, in our model there are two sectors: one market (m) and one non-market (n). In the market sector, a good Y_t is produced in each period t, which can be either consumed, C_{m_t}, or invested. In the home sector, the agents produce a non-tradeable good C_{n_t} that can only be consumed. Both goods are produced using capital, K_{m_t} and K_{n_t}, and labor, H_{m_t} and H_{n_t}. There is a representative agent with utility defined over C_{m_t}, C_{n_t}, H_{m_t}, and H_{n_t}; assuming that the total time available to the agent is one, leisure is given by $1 - H_{m_t} - H_{n_t}$. A CES aggregator creates a composite consumption good.

$$C_t = [aC^{\varepsilon}_{m_t} + (1-a)C^{\varepsilon}_{n_t}]^{1/\varepsilon}$$

Depending on the value of ε, the elasticity of substitution between C_{m_t} and C_{n_t}, given by $1/(1-\varepsilon)$, can be anything between zero and infinity. The greater the substitutability, the larger the resource transfer as relative productivity or expected productivity changes between the two sectors, thereby lowering the increasing returns necessary to have there be multiple equilibria in the model. H_{m_t} and H_{n_t} are perfect substitutes.[25] Capital is assumed to move freely between sectors; again, the greater the degree of substitutability the greater are resource transfers in response to new information.

[25] This is compatible with some of the estimates in Rupert, Rogerson, and Wright (1995).

Utility is derived according to:

$$u(C_{m_t}, C_{n_t}, H_{m_t}, H_{n_t}) = \log[aC_{m_t}^{\varepsilon} + (1-a)C_{n_t}^{\varepsilon}]^{1/\varepsilon} + A\frac{(1 - H_{m_t} - H_{n_t})^{1-\gamma}}{1-\gamma}.$$

The production technology is Cobb-Douglas in both sectors. In the market sector Y_t is produced using K_{m_t} and H_{m_t} according to $Y_t = B_t \theta_t K_{m_t}^b H_{m_t}^{1-b}$. Here θ_t is an exogenous technology parameter, while the term B_t is assumed to include the external effects to production arising from the aggregate levels of K_{m_t} and H_{m_t} which are not taken into account by the single agents; in particular: $B_t = \bar{K}_{m_t}^{b_K} \bar{H}_{m_t}^{b_H}$. In the home sector, C_{n_t} is produced according to $C_{n_t} = K_{n_t}^s H_{n_t}^{1-s}$; no externalities or exogenous technology parameters are assumed to exist in home production.

Since what is produced at home is immediately consumed, total capital can increase only because of investments made in the market sector. We assume a common depreciation rate to capital in each sector. The law of motion of total capital is given by:

$$K_{t+1} = Y_t - C_{m_t} + (1-\delta)K_t = B_t K_{m_t}^b H_{m_t}^{1-b} - C_{m_t} + (1-\delta)K_t \qquad (7)$$

Total capital is the only endogenous state variable of the model.

The representative agent chooses how much to consume, how much to work, and the allocation of capital in the two sectors, so that the infinite discounted sum of utility is maximized, subject to the relations among the variables that were described above, and subject to the law of motion of capital. Formally,

$$\max_{C_{i_t}, H_{i_t}, K_{i_t}} \mathrm{E}_0 \sum_{t=0}^{\infty} \rho^t \left\{ \log[aC_{m_t}^{\varepsilon} + (1-a)C_{n_t}^{\varepsilon}]^{1/\varepsilon} + A\frac{(1 - H_{m_t} - H_{n_t})^{1-\gamma}}{1-\gamma} \right\}$$

subject to:

$$\begin{aligned} K_{t+1} &= B_t \theta_t K_{m_t}^b H_{m_t}^{1-b} - C_{m_t} + (1-\delta)K_t \\ C_{n_t} &= K_{n_t}^s H_{n_t}^{1-s} \\ K_t &= K_{m_t} + K_{n_t} \\ \theta_{t+1} &= \theta_t^z u_t \\ K_0 &= \bar{K}_0 > 0 \text{ given} \end{aligned}$$

As described above, u_t is an i.i.d. random variable.

Expectations and productivity are are described in Section 4.1. The rest of the model we calibrate exactly as in Perli (1998a). Based on findings in

Benhabib, Rogerson and Wright (1991) and Hill (1985), we assume that 17% of time is spent in market work, and 15% in non-market work. Greenwood and Hercowitz (1991) measure the ratio between consumer durables plus residential capital and business plus non–residential capital to be 1.13, which is the value we use for K_n/K_m. Furthermore, we set b, the private share of capital in the production function of the market good, equal to 0.3, the discount factor ρ equal to 0.9898, and the depreciation rate δ equal to 0.025 to simulate quarterly data.

Estimates of ε and γ vary, with implied elasticities of substitution ranging from 1 to 4.[26] We choose a fairly conservative estimate of 1.5 (ε equal to .33)[27] Following Hanson (1985) and Rogerson (1988), we set the value of the intertemporal elasticity of substitution of leisure equal to zero[28].

These elasticities are key because they put a lower bound on the extent of increasing returns necessary to generate multiple equilibria (and therefore self-fulfilling expectations) Here we use $b_K = 0.1$ and $b_H = 0.1$, which are well within the region that implies multiple equilibria.[29] The rest of the parameter values are derivative of this set.[30]

We assume that expectations follow a Markov switching process. In particular, we assume that agents may become "optimistic" or "pessimistic", based upon the technology innovation that they observe. Assume technology is a random walk with disturbance u_t. If u_t is observed to be high enough, say higher than a certain value u_2, we assume that agents are more likely to be "optimistic", whereas if u_t is low enough, say lower than u_1, we assume that agents are more likely to be "pessimistic". Any values of u_t in between leave the agents in the same state as they were before. Since we assume that the innovation u_t is log–normally distributed, we also know the probabilities that agents will be in a certain state next period given the current state; in other words, we know the transition probabilities between the two states. If we call S_t the state of agents' expectations, we have the structure described in section 2:

[26] See, for example, McGrattan, Rogerson, and Wright (1993), or Perli (1998a).

[27] Perli (1998a) shows that expectation–driven cycles are possible also with ε within a very wide range of values.

[28] Multiple equilibria are possible also with higher values of γ,for example $\gamma = 1$, as is again shown in Perli (1998a).

[29] These are not the minimum values of the externalities required for multiple equilibria. We have to choose slightly higher values, since parameters that are at the border of the indeterminacy region generate very high frequency fluctuations that are counterfactual (for more on this point see Perli (1998b)).

[30] See Perli (1998a) for the details.

$$
\begin{aligned}
Prob[S_t &= 1 \mid S_{t-1} = 1] = p \\
Prob[S_t &= 0 \mid S_{t-1} = 1] = 1 - p \\
Prob[S_t &= 0 \mid S_{t-1} = 0] = q \\
Prob[S_t &= 0 \mid S_{t-1} = 1] = 1 - q
\end{aligned} \tag{8}
$$

where 1 is the optimistic state and 0 is the pessimistic state.

The actual value of the sunspot is drawn from a probability distribution which is different in the two states. In particular we assume that when the agents are pessimistic the sunspot is drawn from a distribution which has mean zero, but is skewed towards negative values. When the agents are optimistic, instead, the sunspot is drawn from a distribution which has mean zero, but is skewed towards positive values.

This structure is necessary to ensure that in each period the conditional mean of the sunspot is zero, while at the same time capturing the idea that pessimistic (optimistic) agents are more inclined towards negative (positive) expectations. Moreover, this also captures the fact that, whereas agents' expectations might be coordinated about the state (optimistic or pessimistic), they might not be in agreement about the exact amount of optimism or pessimism. In other words, the actual sunspot that is observed here could be interpreted as the result of the aggregation of many different individual views of the world; indeed, it is possible, but not very likely, that we could observe a positive sunspot when the agents are pessimistic and *vice-versa*. Thus, observing expectations, we can only probabilistically infer the expectations regime. This corresponds to our understanding of the actual (SPF) expectations data as well.

We choose F distributions for the optimistic and pessimistic case which are the mirror image of each other. Both are mean zero. We calibrate the relationship between u_1 and u_2 to yield the same transition probabilities that we observe for output in the U.S. data.

Note that we assume that all movements in expectations are the result of changes in productivity.[31] In other words, the "animal spirits" component of economic forecasts is in the way in which linear movements in fundamentals

[31] This is not directly testable in the data; although it is be possible to estimate the importance of productivity movements in predicting turning points in expectations, these tests are difficult to interpret for several reasons. Firstly, the fact that productivity is procyclical means that it may help explain turning points even if there is no causal relationship. Secondly, such a framework would be one in which there were time-varying transitional probabilities, which would not be consistent with our stylized model.

cause swings in mood. Animal spirits do not cause business cycles in our model, but do cause the regime-switching (and associated) features of the cycle.

Although there are two sectors in our model economy, only the market sector data has a direct counterpart to the actual consumption, hours worked, investment, and output data in Section 3. All of the statistics presented for the model in this section refer only to that market sector data. Table 5 shows the implied variances from the model and the U.S. data after detrending with the HP filter. There is of course an inconsistency in assuming the unit root for model estimation and then using the HP filter, rather than first-differencing. We present the HP-filtered data for ease of comparison of the model output with more familiar business cycle models. As is clear from the table, the model with Markov-switching in expectations has nearly identical implications as does the model with normally distributed expectations shocks (as in Farmer-Guo (1994) or Perli (1998a)), or as most other RBC models. The cause is clear: the shocks in either case are constrained to be mean-zero and therefore have no long-run implications for the variance (the unconditional variance unaffected). Moreover, the model output under the assumption of no expectations transmission has the same implications for variance. This is a well-known feature of the sunspot model: that the empirical implications with and without sunspots are difficult to distinguish, lending little support for or against the theory.

both shocks	**output**	**consumption**	**investment**	**hours**
s.d.(x) / s.d.(output)	1	.34	3.7	.83
corr(x,output)	1	.75	.98	.98

tech shocks only	**output**	**consumption**	**investment**	**hours**
s.d.(x) / s.d.(output)	1	.23	3.78	.85
corr(x,output)	1	.79	.99	.98

actual data	**output**	**consumption**	**investment**	**hours**
s.d.(x) / s.d.(output)	1	.86	7.78	1.6
corr(x,output)	1	.77	.84	.88

Table 5: Standard Regression Output
Regime-switching model and actual data

The thrust of the paper of course is that the model output should also exhibit regime-switching over the business cycle. Since the model is calibrated to the data, only the model in which there is a role for expectations can explain that feature of the data.

both shocks	**convergence**	**distinct states**	**fit**	**timing**
output	**yes**	**yes**	**yes**	**yes**
consumption	**yes**	**yes**	**yes**	**yes**
investment	**yes**	**yes**	**yes**	**yes**
hours	**yes**	**yes**	**yes**	**yes**

tech shocks only	**convergence**	**distinct states**	**fit**	**timing**
output	no	-	-	-
consumption	no	-	-	-
investment	no	-	-	-
hours	no	-	-	-

Table 6: Markov Estimation of Model Output

Table 6 evaluates the model according to the criteria established in Sections 2 and 3. Figure 4 shows the probability of being in the contractionary state for each of the endogenous series; there is always a high probability of being in either state, as in the actual data, and each of the series switches regimes only at business cycle turning points. The historical probability of being in a contraction is least convincing for the model Investment series (this is also true in the actual data), although there is only a high probability of being in a contraction when that probability is also high in the other series. Nonetheless it is clear that even under our stylized specification, the probability of entering a contraction does not rise in perfect harmony in each of the series, so that there are leading as well as lagging indicators over the cycle.

5 Conclusions

We believe this paper has three interesting contributions. We provide broad evidence that regime-switching is a characteristic of the U.S. business cycle that complements evidence already presented by other authors, chiefly Hamilton (1989). Secondly, we show that this is not a radical assertion, in the sense that it is quite consistent with established business cycle theory so long as a role is provided for expectations shocks and the model is calibrated to match regime-switching in the expectations survey data. Finally, by looking at the data in this way, we are able to distinguish between the empirical implications of the model with self-fulfilling expectations and that with productivity shocks alone, with the evidence more supportive of the former. Thus our findings could be interpreted as being supportive of both the sunspot- and Markov (Hamilton, 1989) models of the U.S. business cycle.

6 References

Altug, S., R.A. Ashley and D.M. Patterson (1996): "Are Technology Shocks Nonlinear?" Virginia Tech Economics Department Working Paper E95-55.

Ang, A. and G. Bekaert (1998): "Regime Switches in Interest Rates." NBER Working Paper n. 6508.

Azariadis, C. and B. Smith (1998): "Financial Intermediation and Regime Switching in Business Cycles." *American Economic Review*, vol.88, pp.516-536.

Benhabib, J. and R.E.A. Farmer (1994): "Indeterminacy and Increasing Returns." *Journal of Economic Theory*, 63, pp. 19–41.

Benhabib, J. and K. Nishimura (1998): "Indeterminacy Under Constant Returns in Multisector Economies." *Journal of Economic Theory*, forthcoming.

Benhabib J., R. Rogerson, and R. Wright (1991): "Homework in Macroeconomics: Household Production and Aggregate Fluctuations." *Journal of Political Economy,* vol. 99, pp. 1166-1187.

Burnside,C. and M.Eichenbaum (1998): "Factor-Hoarding and the Propagation of Business Cycle Shocks." *American Economic Review,* vol 86, pp.1154-1174.

Diebold, F.X. and G.D. Rudebusch (1989): "Scoring the Leading Indicators." *Journal of Business*, vol. 62 n.3, pp. 369–391.

Evans, M. and P. Wachtel (1993): "Inflation Regimes and the Sources of Inflation Uncertainty." *Journal of Money, Credit and Banking*, vol. 25 n. 3, pp. 475–520.

Farmer, R.E.A. and J.T. Guo (1994): "Real Business Cycles and the Animal Spirit Hypothesis." *Journal of Economic Theory* 63, pp. 42–72.

Filardo, A. (1994): "Business-Cycle Phases and Their Transitional Dynamics." *Journal of Business and Economic Statistics,* vol 12, pp.299-308.

Greenwood J. and Z. Hercowitz (1991): “The Allocation of Capital and Time over the Business Cycle.” *Journal of Political Economy,* vol. 99, pp. 1188-1214.

Greenwood J., R. Rogerson, and R. Wright (1995). “Household Production in Real Business Cycle Theory,” in T. Cooley, ed., *Frontiers of Business Cycle Research* (Princeton University Press, Princeton).

Hamilton, J.D. (1989(): “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle.” *Econometrica*, vol. 57 n. 2, pp. 357–384.

Hansen, G. (1985): “Indivisible Labor and the Business Cycle.” *Journal of Monetary Economics*, vol. 16, pp. 309–327.

Hill M.S. (1985): “Patterns of Time Use,” in F.T. Juster and F.P. Stafford, eds, *Time, Goods, and Well–Being* (University of Michigan Press, Ann Arbor).

Kydland, F.E. and E.C. Prescott (1996): “The Computational Experiment: An Econometric Tool.” *Journal of Economic Perspectives*, vol. 10 n. 1, pp. 69–86.

Matsusaka, J.G. and A.M. Sbordone (1995): ”Consumer Confidence and Economic Fluctuations.” *Economic Inquiry*, vol. 33, pp. 296–318.

McGrattan E., R. Rogerson, and R. Wright (1993): “Household Production and Taxation in the Stochastic Growth Model.” Manuscript, University of Pennsylvania.

Perli, R. (1998a): “Indeterminacy, Home Production, and the Business Cycle: a Calibrated Analysis.” *Journal of Monetary Economics*, vol. 41, pp. 105–125.

Perli R. (1998b): “Increasing Returns, Home Production, and Persistence of Business Cycles.” *Journal of Economic Dynamics and Control*, vol 22, pp. 519–543.

Rogerson, R. (1988): “Indivisible Labor, Lotteries, and Equilibrium.” *Journal of Monetary Economics*, vol 21, pp.2-16.

Rotemberg J. and M. Woodford (1996): "Real–Business–Cycle Models and the Forecastable Movements in Output, Hours, and Consumption." *American Economic Review*, vol. 86, pp. 71–89.

Rupert P., R. Rogerson, and R. Wright (1995): "Estimating Substitution Elasticities in Household Production Models." *Economic Theory*, forthcoming.

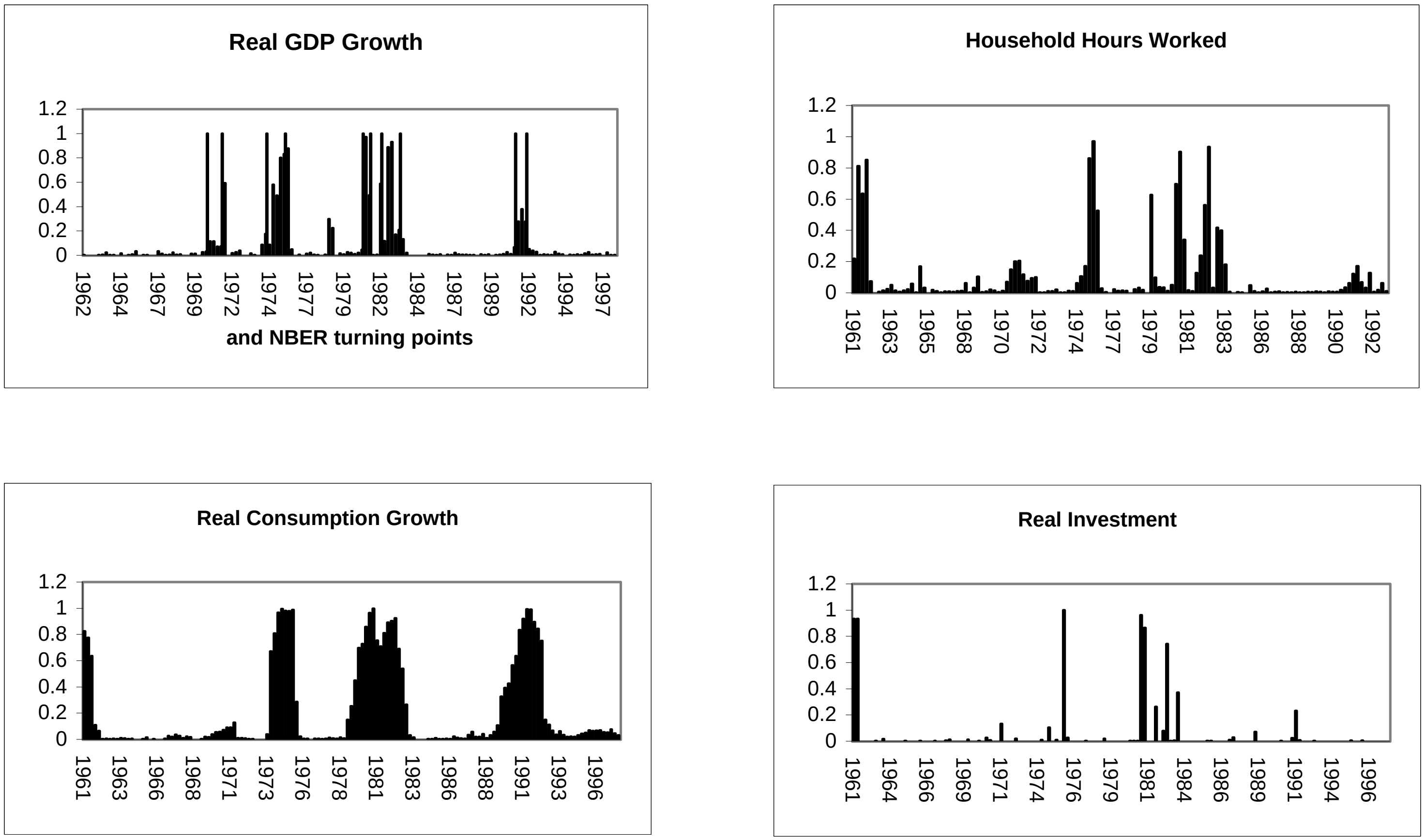
Real GDP Growth
1.2
1
0.8
0.6
0.4
0.2
0
1962 1964 1967 1969 1972 1974 1977 1979 1982 1984 1987 1989 1992 1994 1997
and NBER turning points
Household Hours Worked
1.2
1
0.8
0.6
0.4
0.2
0
1961 1963 1965 1968 1970 1972 1974 1977 1979 1981 1983 1986 1988 1990 1992
Real Consumption Growth
1.2
1
0.8
0.6
0.4
0.2
0
1961 1963 1966 1968 1971 1973 1976 1978 1981 1983 1986 1988 1991 1993 1996
Real Investment
1.2
1
0.8
0.6
0.4
0.2
0
1961 1964 1966 1969 1971 1974 1976 1979 1981 1984 1986 1989 1991 1994 1996

Figure 1

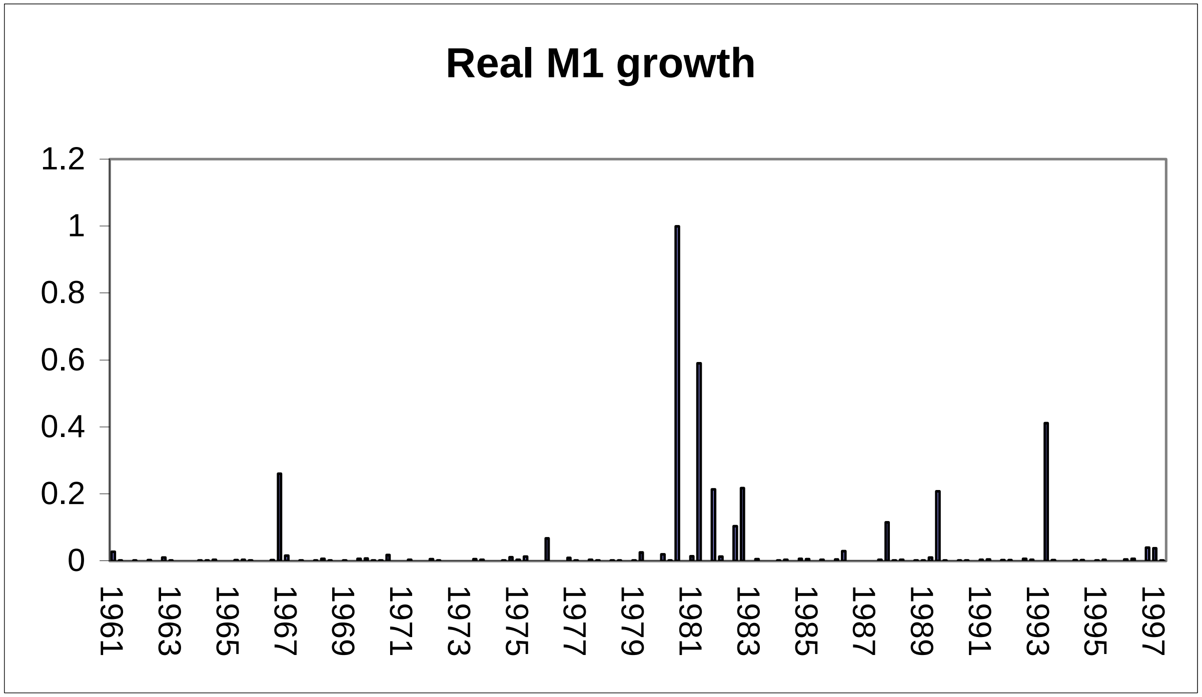
Real M1 growth
1.2
1
0.8
0.6
0.4
0.2
0
1961
1963
1965
1967
1969
1971
1973
1975
1977
1979
1981
1983
1985
1987
1989
1991
1993
1995
1997

Figure 2

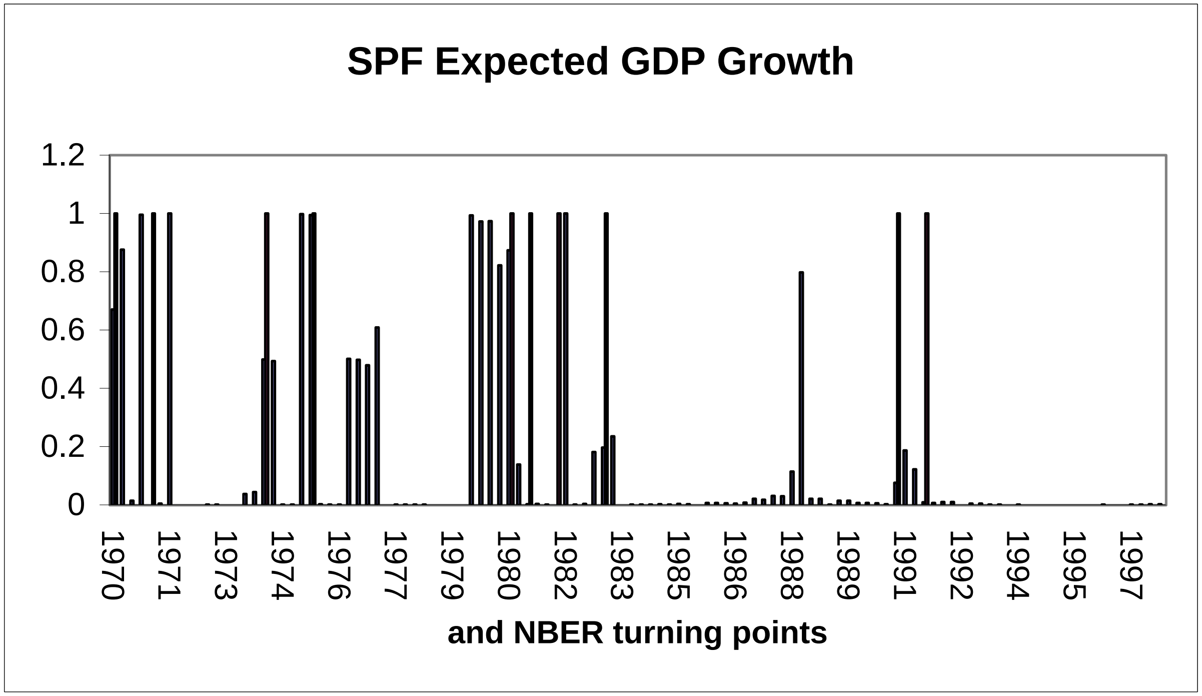
SPF Expected GDP Growth
1.2
1
0.8
0.6
0.4
0.2
0
1970
1971
1973
1974
1976
1977
1979
1980
1982
1983
1985
1986
1988
1989
1991
1992
1994
1995
1997
and NBER turning points

Figure 3

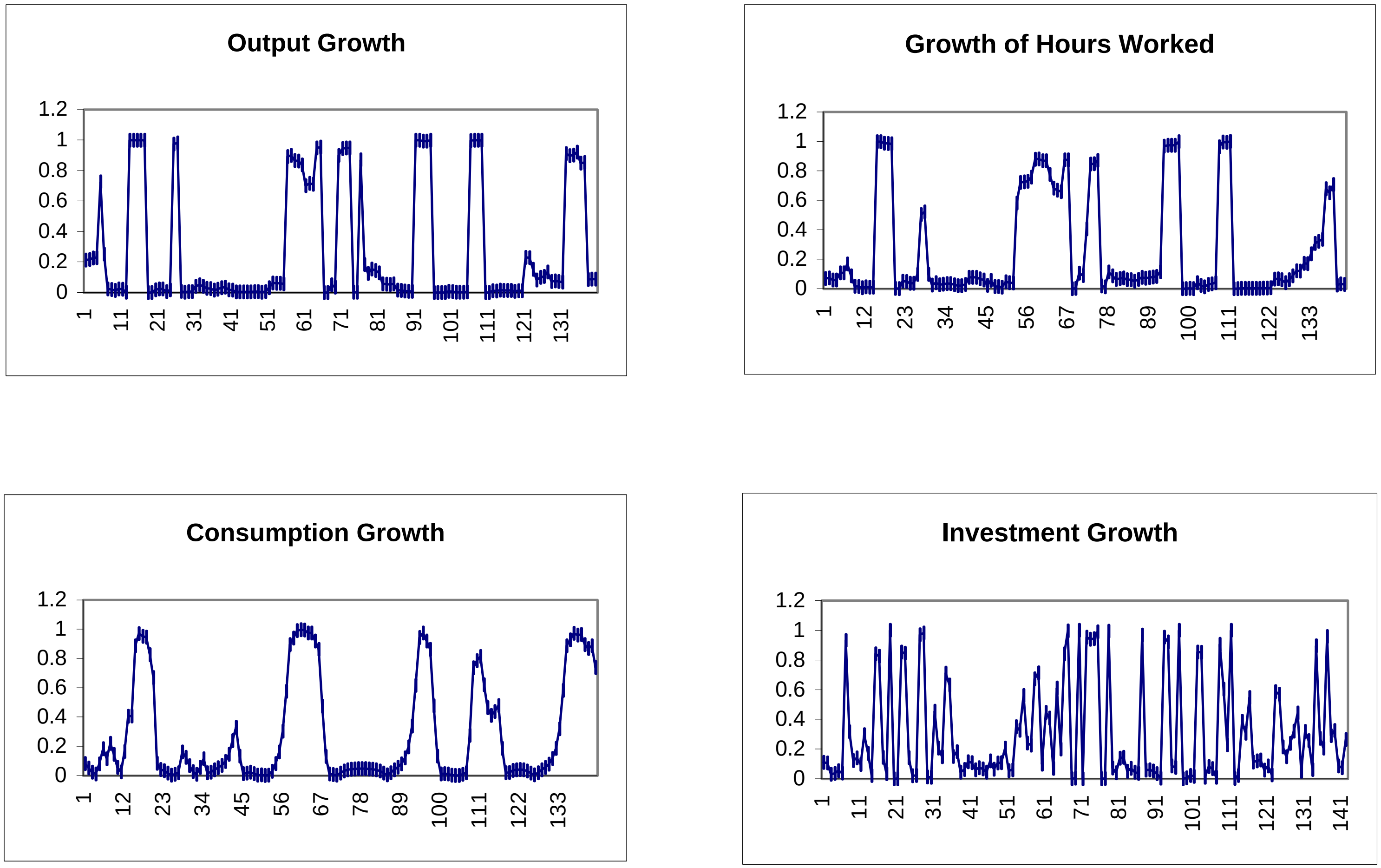
Output Growth
1.2
1
0.8
0.6
0.4
0.2
0
1 11 21 31 41 51 61 71 81 91 101 111 121 131
Growth of Hours Worked
1.2
1
0.8
0.6
0.4
0.2
0
1 12 23 34 45 56 67 78 89 100 111 122 133
Consumption Growth
1.2
1
0.8
0.6
0.4
0.2
0
1 12 23 34 45 56 67 78 89 100 111 122 133
Investment Growth
1.2
1
0.8
0.6
0.4
0.2
0
1 11 21 31 41 51 61 71 81 91 101 111 121 131 141

Figure 4