STRUCTURAL STABILITY AND MODELS OF THE BUSINESS CYCLE

BY

JAMES H. STOCK*

Summary

Ever since the pioneering work of Jan Tinbergen, econometric modelers have been aware of the danger that their models can be unstable over time and across policy environments. Work over the past fifteen years has produced a set of statistical procedures for identifying and modeling structural instability. This essay summarizes some of those procedures, uses them to discuss changes in the U.S. business cycle over the past four decades, and surveys some new research that tackles the widespread challenge of structural instability.

Key words: break tests, econometric models, structural breaks

1 INTRODUCTION

Jan Tinbergen's pioneering work on empirical macroeconomic models has shaped business cycle research ever since and thus has framed our current understanding of the business cycle. His models, first of the Dutch economy and then of the U.S. economy, have several key features that remain in many modern models. In Tinbergen's models, business cycles were treated as the outcome of shocks, or impulses, that propagate through the economy over time to produce complicated dynamic patterns. Although his individual equations were linear with simple, typically single-period lag structures, the resulting system could exhibit cyclical dynamics. The individual equations of the model were motivated by economic theory, and the model itself provided a framework for linking a large number of variables. Tinbergen's work, as described in the first volume of his report to the League of Nations, entailed many important details that continue to be part of modern macroeconometric methodology. Notably, his emphasis on testing business cycle theories led to evaluating models by their forecasting ability and to checking for the stability of their parameters over time. Finally, Tinbergen used his models both for positive and normative analysis, that is, both to evaluate economic theories and to provide a tool for the analysis of macroeconomic policy.

This essay looks at one aspect of modern business cycle analysis in the light of Tinbergen's early contributions: the current state of knowledge about the struc-

* Department of Economics, Littauer Center, Harvard University, Cambridge, MA 02138-3001, U.S.A.

De Economist **152**, 197–209, 2004.

tural stability of macroeconomic relations. In Tinbergen's time, methods for detecting structural changes entailed reestimating parameters using small subsets of the data and, absent measures of sampling uncertainty, drawing expert judgments about the results. Econometricians now have a large set of statistical tools for analyzing the stability of equations and systems of equations, and these tools have been applied to macroeconomic data sets. As I discuss below, the result has been an empirical finding of widespread instability in macroeconomic relations. This in turn poses substantial problems for practical forecasters and policy analysis.

2 METHODS FOR IDENTIFYING STRUCTURAL CHANGES

Over the past fifteen years, econometricians have produced a collection of methods for modeling structural breaks and time-varying parameters in economic relations. Two such methods are tests for structural breaks, notably the so-called Quandt (1960) likelihood ratio (QLR) test and analysis of stability based on pseudo out-of-sample forecasts. I show that, for the U.S., both methods suggest widespread evidence of structural instability among macroeconomic relations. I then turn to a particularly interesting change in the modern business cycle: the recent moderation in the cyclical volatility of aggregate output in many developed economies.

2.1 *The QLR Test for Structural Breaks*

The first step towards identifying a structural break in a macroeconomic time series is having a reliable test for a structural break, that is, a test that has controlled size under the null of no break and good power against the alternative of a break. One such test is the Quandt likelihood ratio test. To be concrete, consider the linear regression model with a single stationary regressor X_t,

$$Y_t = \beta_0 + \beta_{1t} X_t + u_t \tag{1}$$

where u_t is a serially uncorrelated error term, Y_t is the dependent variable, and β_0 and β_{1t} are regression coefficients.

The unusual feature of (1) is that the slope coefficient can change over time. In particular, suppose that the regression coefficient changes at some date τ:

$$\beta_{1t} = \begin{cases} \beta_1, t \leq \tau \\ \beta_1 + \delta, t > \tau \end{cases}. \tag{2}$$

If the break date τ is known, then the problem of testing the null hypothesis of no break (that is, $\delta = 0$) against the alternative of a nonzero break ($\delta \neq 0$) is equivalent to testing the hypothesis that the coefficient δ is zero in the augmented version of (1),

$$Y_t = \beta_0 + \beta_{1t} X_t + \delta Z_t(\tau) + u_t \,, \tag{3}$$

where $Z_t(\tau) = X_t$ if $t > \tau$ and $Z_t = 0$ otherwise. This test can be computed using a conventional t-statistic from estimating (3) by ordinary least squares; calling this $t(\tau)$, the hypothesis of no break is rejected at the 5% significance level if $|t(\tau)| > 1.96$.

In practice, τ is typically unknown so the test in the preceding paragraph cannot be implemented. However, the t-statistic can be computed for all possible values of τ in some range. If the largest of the absolute values of these t-statistics exceeds some critical value, then the hypothesis of no break can be rejected. The difficulty with this method is that the critical value is not 1.96, because you are providing yourself with multiple opportunities to reject the null hypothesis. The distribution has, however, been worked out (Andrews (1993)), and if the values of τ considered are those in the central 70% of the sample then the relevant 5% critical value is 2.95.

To extend this method to multiple regressors, one needs to compute the F-statistics testing the hypothesis that the additional regressors are zero for each value of τ, then compute the maximum of these. This is commonly called the QLR (or sup-Wald) test statistic for a structural break; for a table of critical values, see Stock and Watson ((2003), Table 12.5). Andrews and Ploberger (1994) developed other related tests for structural stability.

2.2 *Pseudo Out-of-Sample Forecast Comparisons*

An alternative approach to assessing structural stability builds on Tinbergen's intuition that one way to test a model is to look at its forecasting performance: a change in its forecasting performance over the sample suggests changes in the model coefficients. Although true forecasting occurs in real time, historical forecasting experience can be simulated by first estimating the model using data through an earlier date (for example, 1990), then making a forecast one or more periods ahead (for example, for 1991), then moving forward one period and repeating the exercise. This produces a sequence of forecasts and forecast errors that simulate the forecasts that would have been made using the model in real time. If the resulting pseudo out-of-sample performance of the model changes, relative to a benchmark forecasting model, then this can be taken as evidence of a shift in the parameters of the model. When this comparison is made using nonnested models, formal statistical comparisons can be made using the statistics proposed in West (1996).

3 EVIDENCE OF INSTABILITY IN TIME SERIES FORECASTING MODELS

What happens when these methods are applied to relations that are used for macroeconomic forecasting? Are the coefficients of econometric models stable, or do they appear to change over time?

One way to address these questions is to examine Tinbergen's 'final equation' for econometric models, which (absent exogenous variables) is the univariate autoregressive representation of the time series entering the model. Table 1 presents the results of applying QLR tests to quarterly U.S. data from 1959 – 2002 on major macroeconomic variables. The first column reports the *p*-value of the test of stability of all the autoregressive parameters. If the test rejects at the 5% significance level, the second column reports the estimated break date, where the break date is estimated by least squares estimation of the model augmented by the interaction variables (as in (3)), where τ is estimated along with the regression coefficients by least squares. The results in Table 1 point towards substantial instability in these autoregressions; although the hypothesis of stability is not rejected for GDP or the goods production series, it is rejected for other production measures, employment, inflation, and interest rates.

A similar conclusion of widespread instability is reached when pseudo out-of-sample forecasting methods are used. In Stock and Watson (2003b), we review the ability of asset prices to forecast real output growth and inflation over the past four decades in seven developed economies. Using pseudo out-of-sample forecasts, we find that good forecasting performance of an asset price in one period or in one country does not imply that the same asset price will be a good predictor in another time period or another country.

The results in Table 2 are indicative of the findings in the broader study of Stock and Watson (2003b). Table 2 reports the root mean squared forecast error for pseudo out-of-sample forecasts, relative to a benchmark model, where the benchmark is a univariate autoregression with the number of lags chosen using the Bayes information criterion (BIC); a value of 1.0 means that the forecast errors of the candidate model is the same as that of the benchmark autoregression. For the series shown in Table 2, all the candidate models outperformed the autoregressive benchmark during 1971–1984, but most produced relatively poor forecasts during the later period of 1985–1999. In some cases, such as forecasts based on the term spread, both the early gain and the later deterioration are very large. Forecasts based on housing prices, however, exhibited better performance in the second period than in the first. The general pattern in Table 2 of changes in relative forecast performance is typical for other predictors, other forecast horizons, and countries other than the U.S.

TABLE 1 – QLR TESTS FOR CHANGES IN AUTOREGRESSIVE PARAMETERS: QUARTERLY U.S. DATA, 1959 - 2002

Series	*p*-value	break date
GDP (growth rate)	0.98	.
Production of goods (total) (growth rate)	0.92	.
Production of nondurable goods (growth rate)	0.09	.
Production of durable goods (growth rate)	0.77	.
Production of services (growth rate)	0.00	1968:3
Production of structures (growth rate)	0.02	1991:3
Nonagricultural employment (growth rate)	0.03	1981:2
Price inflation (GDP deflator)	0.00	1973:2
90-day T-bill rate (first difference)	0.00	1981:1
10-year T-bond rate (first difference)	0.02	1981:1

Notes: The first column shows the *p*-value for the QLR statistic testing the stability of all the autoregressive coefficients in an AR(4). The second column shows the least squares estimate of the break date when the QLR statistic is significant at the 5% level. Source: Stock and Watson (2002a, Table 3)

TABLE 2 – RELATIVE MEAN SQUARE FORECAST ERRORS FOR PSEUDO OUT-OF-SAMPLE FORECASTS OF US REAL GDP GROWTH, 4 QUARTERS AHEAD (UNIVARIATE AUTOREGRESSION = 1.0): QUARTERLY DATA, 1959 - 1999

Predictor	1971 – 1984	1985 – 1999
Federal Funds rate (level)	0.78	1.42
90-day Treasury bill rate (level)	0.85	1.06
Term spread (long bond rate minus Federal Funds rate)	0.48	2.51
Real stock return	0.90	1.27
Percentage changes in housing prices	1.06	0.93
Percentage change in real M2	0.57	1.41

Notes: Each entry is the mean squared pseudo out-of-sample forecast errors (MSFE) for a candidate model with lags of GDP growth and lags of the additional predictor, relative to a the corresponding MSFE for a univariate autoregression; all lags are chosen by BIC. An entry less than 1.0 means that the candidate model outperformed the univariate benchmark during this period. Source: Stock and Watson (2003b, Table 3)

4 CAN WE PRODUCE RELIABLE FORECASTS WHEN FORECASTING MODELS ARE UNSTABLE?

The obvious initial answer to this question is no, one cannot produce reliable forecasts if the underlying forecasting model is unstable. But recent research suggests that this answer might be too pessimistic. At least for US macroeconomic time series for the past four decades, much of the instability in the predictive relations appears to be idiosyncratic. For example, the results in Table 2 suggest that a good strategy would have been to forecast using the term spread through 1984, then to switch to forecasts based on the growth in housing prices. Unfortunately, this recommendation is based on retrospective information, so an interesting research question is how such a strategy could be made operational.

Research on real-time forecasting in the presence of model instability has focused on three general approaches. The first approach involves estimating an explicit model of the time variation of the parameters, typically by treating the time varying parameters as unobserved components and using Kalman filtering methods (or its generalizations to nonlinear models), to estimate the hyperparameters of the model. This in turn permits inferences about the parameters of the forecasting relation and, hence, permits making adaptive real-time forecasts (see Harvey (1989) for a textbook treatment of linear models with unobserved components).

A second approach to forecasting in the presence of model instability entails developing real-time methods for switching between candidate forecasting models. The simplest of these are ad-hoc rules such as, 'use the candidate model that has worked best for the past two years.' The problem with such a simple rule is that it is easily fooled by sampling variability to put weight on a model that is lucky rather than good, and not surprisingly such simple rules fair poorly in forecasting comparisons (e.g. Stock and Watson (2003c)). A more sophisticated approach along these lines is to estimate a Markov switching model that switches among a variety of candidate models, an approach sometimes referred to as a hidden Markov model. Hidden Markov models have not been explored much in economic forecasting as a tool for switching among models with different candidate predictors, although they have widespread application other fields, including pattern recognition (speech recognition, protein classification, etc.).

A third approach to forecasting with unstable models is to use time-invariant forecasting models that involve a large number of predictors. The key idea in this approach is that, if the instability is sufficiently idiosyncratic across predictors, then using many predictors might nonetheless produce reliable forecasts. At least for the US over the past four decades (the case that has received the most research attention), the instability in predictive relations does appear to be idiosyncratic, and empirical results based on this approach are promising. One specific strategy is to use large-scale (many predictor) dynamic factor models to derive a small number of common factors, then to use these few common factors as predictors; despite idiosyncratic instability in the factor loadings, the common factors can be estimated precisely. Recent research by Forni et al. (2000, 2001) and by Stock and Watson (1999, 2002b) suggests that large dynamic factor models can provide a coherent framework for economic forecasting and can produce forecasts that improve upon those from smaller time series models. Another promising strategy in the same general vein is to use forecast combining methods to produce combination forecast that, in effect, average out the instability in the constituent individual forecasting models. Recent empirical work (Stock and Watson (2003c)) suggests that combination forecasts are competitive with those produced by large dynamic factor methods. Research in these areas is ongoing.

5 RECENT CHANGES IN THE BUSINESS CYCLE

The discussion so far has focused on changes in the coefficients of reduced form models, which implies that some important coefficients of the original structural model have been changing. Recently, there has been considerable interest in a different apparent change in the postwar business cycle: a marked decline in the volatility of overall economic activity in a number of developed economies, that is, an overall moderation in the business cycle.

This moderation, originally noted for the U.S. by Kim and Nelson (1999) and McConnell and Perez-Quiros (2000), can be seen by examining the evolution of the standard deviation of annual GDP growth. As is summarized in Table 3, US GDP growth was relatively quiescent in the 1960s, was more volatility in the 1970s and 1980s, and during the 1990s was even less volatile than during the 1960s. Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) independently concluded that there was a break in the volatility of GDP in the mid 1980s, a conclusion that they reached using different methods (Kim and Nelson (1999) used a Bayesian stochastic volatility model and a nonGaussian smoother, whereas McConnell and Perez-Quiros (2000) used classical break tests like the QLR). Although there has been some debate about whether this reduction in volatility is actually a sharp break or, as suggested by Blanchard and Simon (2001), part of a longer trend to increasingly moderate cycles, there is little debate about whether there actually was a decline in volatility. Moreover, as shown by van Dijk et al. (2002), this moderation is evident in other developed economies as well.

This moderation of the business cycle, and its different nature in different countries, can be seen in Figure 1. This figure plots the four-quarter percentage growth rate of GDP in the US, Germany, and Italy, along with an estimate of the instantaneous standard deviation of quarterly GDP growth. The instantaneous standard deviation was estimated using the stochastic volatility smoother described in Stock and Watson (2002a). In each country, annual fluctuations in GDP growth now are much smaller than they were thirty years ago. The nature of this reduction of volatility is different in each country, however. In the US, the reduction appears to be sharp, occurring in the mid-1980s; in Germany the reduction has been ongoing over the past four decades; and in Italy, the reduction seems to have occurred in approximately 1980, although the estimate of the instantaneous standard deviation does not suggest a sharp break at which to date the moderation.

Although there is little disagreement about whether this moderation occurred, there is no consensus as to its source. As discussed in Stock and Watson (2002a,

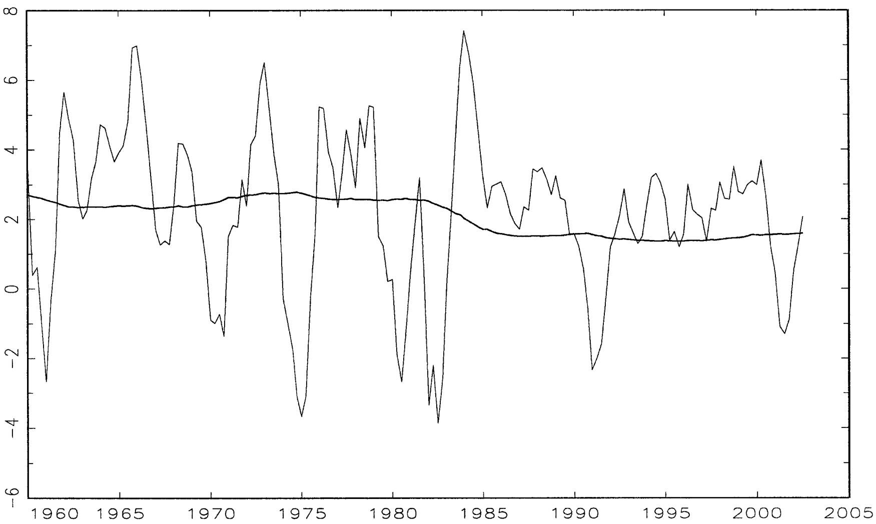
TABLE 3 – SUMMARY STATISTICS FOR FOUR-QUARTER GROWTH IN US REAL GDP, 1960 – 2001

Sample Period	Standard deviation (%)
1960 - 1969	2.0
1970 - 1979	2.7
1980 - 1989	2.6
1990 – 2001	1.5

Notes: Summary statistics are shown for $100\times\ln(GDP_t / GDP_{t-4})$, where GDP_t is the quarterly value of real GDP. Source: Stock and Watson (2002a, Table 1)

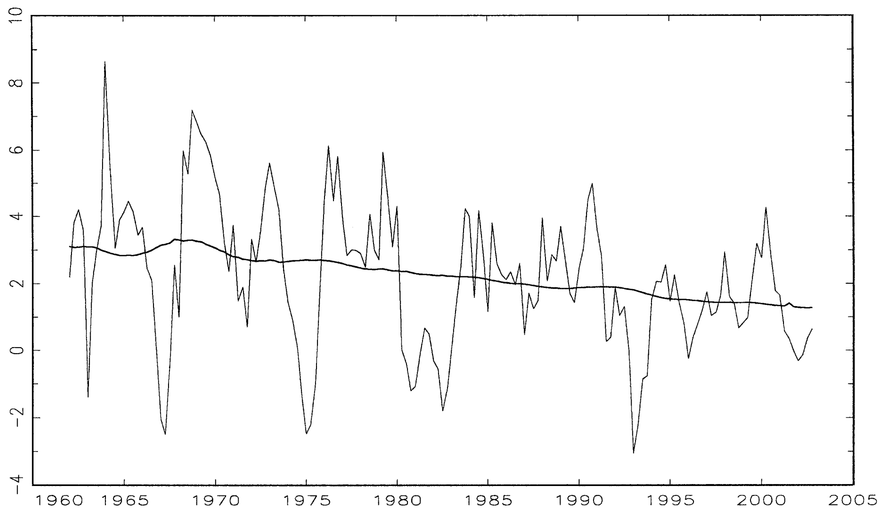
2003d), a variety of explanations have been proposed for this reduction in volatility. These explanations include changes in the structure of modern economies, such as reduced credit constraints among consumers or changes in production technologies; changes in policy regimes, in particular a switch towards more steadfast monetary policies that result in reduced sensitivity of the economy to economic shocks; and simply a reduction in the magnitude of recent economic shocks.

In Tinbergen's framework, some of these explanations entail changes in the magnitude of the impulses to the macroeconomy, while others involve changes in

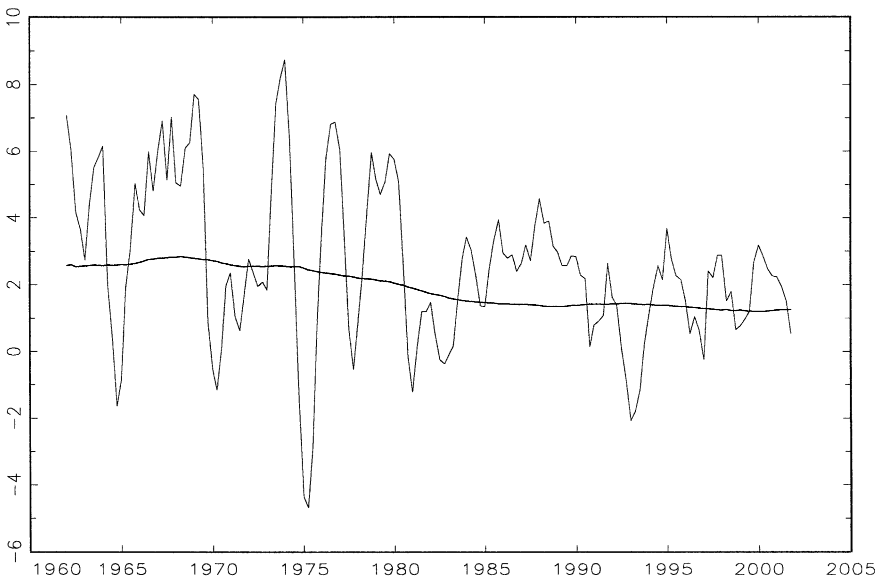


(A) US

Figure 1 – Four-quarter percentage growth of real GDP and a smoothed estimate of its time-varying standard deviation (A) US (B) Germany (C) Italy



(B) Germany



(C) Italy

Figure 1 – Continued

the propagation mechanism. It is important to understand which of these proposed explanations is actually right. If the source of the moderation is improved policy or changes in the structure of the economy, then we might expect this moderation to persist. On the other hand, if this moderation is merely reflects a decade of good luck – that is, small macroeconomic shocks – then macroeconomic managers and forecasters should be ready for a return to the more turbulent business cycles of the past as soon as this period of good luck ends.

One prominent candidate explanation for why volatility has reduced is that US monetary policy changed from being overly accommodative (which led to instability in inflation and output) to being actively anti-inflationary. Evaluating this hypothesis entails answering a counterfactual question: what would volatility have been in the 1990s, had US monetary policy of the 1990s instead been that of the 1970s? A key motivation for Tinbergen's work was to develop models that could be used to provide quantitative answers counterfactual questions that entail hypothetical alternative policies. Accordingly, to provide some quantitative insight into the question of the source of the volatility reduction, in Stock and Watson (2003d) we performed some counterfactual simulations using four different econometric models. The four models ranged from a structural vector autoregression to Smets and Wouter's (2003) nine-equation forward-looking dynamic stochastic general equilibrium model. In each model, the change in US monetary policy circa 1980 was modeled as a change in the coefficients of a dynamic Taylor (1993) rule from one in which the response of the short-term interest rate to a change in the inflation rate was less than one-for-one, to a rule in which the response was greater than unity. The simulations indicated that the change in monetary policy accounted for at most a small fraction of the observed reduction in volatility; the estimates range from 26% (that is, the change in monetary policy, *ceteris paribus*, accounts for 26% of the reduction in volatility of four-quarter GDP growth) to -10% (so that the change in monetary policy is estimated to have *increased* GDP volatility slightly). These estimates suggest that the change in monetary policy was not the source of the reduction in the volatility in the business cycle. Instead, the source of the volatility reduction appears to have been either changes in the private economy or a reduction of the size of the impulses. Of course, these conclusions are based on the analysis of four specific econometric models, and regrettably there has been more divergence than convergence among economists since Tinbergen's time about which econometric model provides the most reliable representation of the macroeconomy.

6 NEW STRUCTURAL MODELS: A SOLUTION TO THE PROBLEM OF INSTABILITY?

So far I have focused on the instability of correlations, predictive relations, and standard deviations, and have argued that there is mounting evidence that such instability is widely present in measures of movement and comovement of macroeconomic time series. As the foregoing discussion of the moderation of the business cycle makes clear, such instability in measures of comovement could arise even if the true underlying structural relations governing macroeconomic dynamics are constant, as long as there are external changes such as changes in policy or in the types of shocks impinging on the economy. Said differently, the Lucas critique implies that reduced-form and final-form equations will be unstable even if the so-called deep parameters governing agents' decisions are constant, as long as there are changes in the external economic environment.

This observation suggests that a possible solution to the problem of instability is to develop models based on 'deep parameters' that are invariant to changes in the external environment, and to use those models to evaluate economic policies; indeed, this was a central goal of Tinbergen's research program. Achieving this goal, however, has proven elusive, for it requires estimation of dynamic causal effects from macroeconomic data. We now understand that many of the relations appearing in older large simultaneous equations macroeconomic models are not really structural, but rather are themselves reduced-form equations that do not in general lead to consistent estimates of dynamic causal effects.

These deficiencies have led to a new generation of structural models, dynamic stochastic general equilibrium (DSGE) models, which follows in Tinbergen's tradition, even though the particulars of DSGE models have evolved far from his initial equations. Broadly speaking, DSGE models differ from earlier generations by postulating a decision *problem* – the method by which agents make their decision (for example, a rational expectations model in which agents maximize the expected discounted utility of consumption), rather than by writing down the decision *rule* directly (for example, a consumption function). These decision problems are solved jointly to yield an equilibrium that is stochastic, because it depends on shocks impinging on the economy (such as technological innovations), and is dynamic, because the levels of the endogenous variables evolve over time. Aside from its intellectual elegance, the main practical advantage of the DSGE approach is that it is more plausible that the method of making a decision (utility maximization) is invariant to policy changes than are the parameters of a derived decision rule (a rule relating consumption to income).

Will DSGE models finally solve the problem of parameter instability and produce stable models for policy analysis and forecasting? It is too early to say, in part because these models still incorporate many simplifications and in part because it has only been recently that these models have been able to provide even approximate fits to macroeconomic time series data (see Smets and Wouters (2003) for an important recent success in this regard). The challenge confronting DSGE models are considerable, however: the lesson of the Lucas critique is that the model parameters must truly be 'deep parameters' if they are to be policy invariant, that is, they must reflect the decision problem actually used by agents to reach their decisions. Yet the decision problems implemented in DSGE models entail utility and profit maximization by representative agents, frameworks inconsistent with mounting evidence of heterogeneity at the individual level and of economic behavior that departs from rational expectations and utility maximization. Whether such compromises, made for numerical tractability, lead to parameter instability in practice is an empirical question that currently remains unanswered.

7 CONCLUSION

The findings of Tables 1 – 3 are typical of broader evidence of instability in macroeconomic relations; for additional references and results, see Stock and Watson (1996) and Hansen (2001). As the results in Table 2 make clear, this instability poses a substantial practical problem for real-time forecasters and for other users of macroeconomic models. Developing methods for detecting this instability in real time, for modeling this instability, and for making forecasts that are robust to this instability continues to be an important line of research that follows in the path of Tinbergen's early checks for coefficient stability. Recent work suggests that new methods for time series analysis, especially ones that draw on large numbers of time series, can prove useful for forecasting even in the presence of parameter instability. How much closer we are to achieving Tinbergen's vision of a stable quantitative structural macroeconomic model for policy design is a matter of debate, however, and in my view achieving that part of Tinbergen's research program remains a formidable challenge.

REFERENCES

Andrews, D.W.K. (1993), 'Tests for Parameter Instability and Structural Change with Unknown Change Point,' *Econometrica*, 61, pp. 821-856.

Andrews, D.W.K. and W. Ploberger (1994), 'Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative,' *Econometrica*, 62, pp. 1383-1414.

Blanchard, O. and J. Simon (2001), 'The Long and Large Decline in U.S. Output Volatility,' *Brookings Papers on Economic Activity*, 2001:1, pp. 135-164.

van Dijk, D., D.R. Osborn, and M. Sensier (2002), Changes in Variability of the Business Cycle in the G7 Countries, manuscript, Erasmus University Rotterdam.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), 'The Generalized Factor Model: Identification and Estimation,' *Review of Economics and Statistics*, 82, pp. 540-554.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2001), Do Financial Variables Help Forecasting Inflation and Real Activity in the EURO Area?, CEPR Working Paper

Hansen, B.E. (2001), 'The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity,' *Journal of Economic Perspectives*, 15(4), pp. 117-128.

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

Kim, C.-J. and C.R. Nelson (1999), 'Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle,' *The Review of Economics and Statistics*, 81, pp. 608-616.

McConnell, M.M and G. Perez-Quiros (2000), 'Output Fluctuations in the United States: What has Changed Since the Early 1980's,' *American Economic Review*, 90(5), pp. 1464-1476.

Smets, Frank, and Raf Wouters (2003), Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach, manuscript.

Stock, J.H. and M.W. Watson (1996), 'Evidence on Structural Instability in Macroeconomic Time Series Relations,' *Journal of Business and Economic Statistics*, 14, pp. 11-29.

Stock, J.H. and M.W. Watson (1999), 'Forecasting Inflation,' *Journal of Monetary Economics*, 44, pp. 293-335.

Stock, J.H. and M.W. Watson (2002a), 'Has the Business Cycle Changed and Why?,' forthcoming, *NBER Macroeconomics Annual* 2002, Cambridge, MIT Press.

Stock, J.H. and M.W. Watson (2002b), 'Macroeconomic Forecasting Using Diffusion Indexes,' *Journal of Business and Economic Statistics*, 20(2), pp. 147-162.

Stock, J.H. and M.W. Watson (2003a), *Introduction to Econometrics*, Boston, Addison Wesley Longman.

Stock, J.H. and M.W. Watson (2003b), 'Forecasting Output and Inflation: The Role of Asset Prices,' *Journal of Economic Literature*, 41, pp. 788-829.

Stock, J.H. and M.W. Watson (2003c), Combination Forecasts of Output Growth in a Seven-Country Data Set, manuscript.

Stock, J.H. and M.W. Watson (2003d), 'Has the Business Cycle Changed? Evidence and Explanations,' forthcoming in *Monetary Policy and Uncertainty*, Federal Reserve Bank of Kansas City.

Taylor, John B. (1993), 'Discretion versus Policy Rules in Practice,' *Carnegie-Rochester Conference Series on Public Policy*, 39, pp. 195-214.

West, K.D. (1996), 'Asymptotic Inference about Predictive Ability,' *Econometrica*, 64, pp. 1067-1084.