TECHNICAL WORKING PAPER SERIES

EVIDENCE ON STRUCTURAL INSTABILITY IN MACROECONOMIC TIME SERIES RELATIONS

James H. Stock
Mark W. Watson

Technical Working Paper No. 164

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1994

We are grateful to Neil Shephard for helpful discussions and participants at the 14th Annual International Symposium on Forecasting in Stockholm, Sweden for useful comments. This research was supported by National Science Foundation grant no. SES-91-22463. This paper is part of NBER's research program in Economic Fluctuations. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Technical Working Paper #164
September 1994

EVIDENCE ON STRUCTURAL INSTABILITY IN MACROECONOMIC TIME SERIES RELATIONS

ABSTRACT

An experiment is performed to assess the prevalence of instability in univariate and bivariate macroeconomic time series relations and to ascertain whether various adaptive forecasting techniques successfully handle any such instability. Formal tests for instability and out-of-sample forecasts from sixteen different models are computed using a sample of 76 representative U.S. monthly postwar macroeconomic time series, constituting 5700 bivariate forecasting relations. The tests indicate widespread instability in univariate and bivariate autoregressive models. However, adaptive forecasting models, in particular time varying parameter models, have limited success in exploiting this instability to improve upon fixed-parameter or recursive autoregressive forecasts.

James H. Stock
Kennedy School of Government
Harvard University
Cambridge, MA 02138
and NBER

Mark W. Watson
Department of Economics
Northwestern University
Evanston, IL 60208
and NBER

1. Introduction

Time series econometrics typically involves drawing inferences about the present or future using historical data. In some cases these inferences are about how the economy operates or how economic policy affects key variables. For example, much empirical work on monetary economics currently rests on inferences drawn from so-called structural vector autoregressions (VAR's); Bernanke and Blinder (1992) and Christiano, Eichenbaum and Evans (1993) provide two recent examples. In other cases these inferences are in the form of forecasts. Both applications typically require that the model at hand be stable (that the future be like the past) for such inferences to be valid. For example, giving advice to current policy makers based on a structural VAR requires, among other things, that the historically estimated model remains relevant today. Although studies occasionally include some analysis of stability, it is often limited in scope, perhaps consisting of reestimating the model on a single subsample. The importance of stability and the current lack of systematic evidence on it therefore leads us to ask, how generic is instability in multivariate time series relations?

To answer this, we undertake a two-part experiment. The first part assesses the prevalence of parameter instability in economic time series relations using a battery of recently developed tests for instability. This is done using a sample of 76 monthly time series for the postwar U.S. economy over the period 1959:1 - 1993:12 (420 observations), among which are 5700 distinct (although not independent) bivariate forecasting relations. These series are chosen to provide relations which are representative of those of interest to macroeconomists and macroeconomic forecasters. This sample is then used to compute empirical distributions of various tests for structural stability, including Nyblom's (1989) test for parameter stability, CUSUM tests, and break point tests such as the Quandt (1960) likelihood ratio statistic.

The second part of the experiment examines whether current state-of-the-art adaptive forecasting models capture the instability found by the stability tests and thereby improve upon more naive forecasts. This entails the empirical evaluation of different forecasting models which exhibit different degrees of adaptivity, ranging from no adaptivity (fixed parameter models) through moderate adaptivity (recursive least squares, rolling regression, and random walk coefficient time varying parameter (TVP) models with small coefficient evolution) to high adaptivity (TVP models with large coefficient evolution). Although work on regression models with stochastically time varying parameters (or "stochastic coefficients") dates to Cooley and Prescott (1973a, 1973b, 1976), Rosenberg (1972, 1973), and Sarris (1973), and although TVP models been applied to selected series, we know of no systematic evidence on whether these techniques might be widely useful in economic forecasting applications.[1] Eight univariate models are considered for each of the 76 series for a total of 608 univariate forecasting equations, and eight bivariate models are considered for each of the 5700 bivariate forecasting relations for a total of 45,600 bivariate forecasting equations. Models are compared using one-month-ahead mean square errors (MSE's). In the spirit of Makridakis et al. (1982) and Meese and Geweke (1984), who applied univariate forecasting techniques to large numbers of series, this part of this experiment yields a forecasting comparison suggesting which models typically do best in macroeconomic applications. The results also provide an opportunity to assess model robustness by identifying models which successfully guard against the most severe out-of-sample forecasting failures.

Looking ahead to the results, the tests indicate that instability is widespread. For example, one version of the Nyblom (1989) test rejects stability (at the 10% level) in more than 70% of the 5,700 bivariate relations. This instability is more prevalent in certain classes of series, such as measures of aggregate output, than in others. However, our results also suggest that forecasting models explicitly designed for time varying parameters (rolling regressions or random walk TVP models) often fail to perform as well as traditional fixed coefficient or recursive least squares models: in 57% of the 5700 pairs, fixed-coefficient or recursive least squares forecasts have the lowest out-of-sample MSE among the sixteen competing models, while in only 10% of the pairs do bivariate TVP models have the lowest out-of-sample MSE. When they are best, the gains associated with the TVP forecasting models typically are small.

In a small fraction of the cases, the TVP and recursive least squares models perform well while the fixed coefficient models perform quite poorly, and in this sense the TVP and recursive least squares models are more robust than the fixed coefficient models. Overall, however, the TVP models fail to exploit successfully the time variation found by the stability tests.

The outline of the paper is as follows. Section 2 describes the data set. The stability tests are described in section 3, and section 4 summarizes the empirical testing results. The forecasting models are described in section 5 and are evaluated empirically in section 6. Section 7 concludes.

2. The Data Set

Our objective in constructing the data set was to obtain a sample of economic time series for the U.S. which is representative of the relations of primary concern to macroeconomists and macroeconomic forecasters. While one could in principle draw series at random from a large macroeconomic database, a simple random sample would oversample certain classes of heavily represented series, such as industry-specific deflators, interest rates, or financial flows. Such a sample would be representative of the monthly data which are produced but not of the forecasting relations of interest to macroeconomists. Moreover, such a sample would omit important forecasting variables which are constructed from the primary data, such as interest rate spreads. In theory, stratification could eliminate the problem of oversampling certain classes of series which are produced in detail, but would not address the issue that many of the important forecasting series are constructed by researchers and thus are not contained in standard databases. Moreover, mechanical simple or stratified sampling would produce many series with definitional changes or other internal inconsistencies.

Our sample of series therefore was obtained by applying subjective judgment, using four criteria as guidelines:

1. The sample should include the main monthly economic aggregates and coincident indicators. This resulted in the inclusion of series such as industrial production, weekly hours, personal income and inventories.

2. The sample should include important leading economic indicators. This led us to include series such as monetary quantity aggregates, interest rates, interest rate spreads, stock prices, and consumer expectations.

3. The series should represent different broad classes of variables which can be expected to have quite different time series properties.

4. The series should have consistent historical definitions or, when the definitions are inconsistent (for example different base years for different segments of a real series) it should be possible to adjust the series with a simple additive or multiplicative splice.

These criteria were used to select 76 monthly U.S. economic time series. Most of the raw data were obtained from the CITIBASE data base, although many series were subsequently modified (for example by creating interest rate spreads). The series can be grouped into eight categories: output and sales; employment; new orders; inventories; prices; interest rates; money and credit; and other miscellaneous series including exchange rates, government spending and taxes, and miscellaneous leading indicators. The complete list of series and their mnemonics are given in the appendix.

The sample runs from 1959:1 to 1993:12. The starting date was chosen because this is the earliest date for which many of the series, in particular the monetary aggregates, are available. Four series (the series on government finance) start in 1967:6. The statistics in question were computed using the longest possible sample for which all relevant data were available.

Each series was screened to detect breaks and outliers due to changes in definitions or reporting practice. Most series were also transformed to induce approximate stationarity by taking either first differences or first differences of logarithms. For consistency, the stationarity transformation was in general applied to entire classes of series rather than on a case-by-case basis. For example, production, employment, prices, and money were all transformed using first differences of logarithms, and interest rates were transformed by first differencing. Some series which did not fit naturally into a broader category were analyzed on a case-by-case basis using visual inspection, *a-priori* reasoning, and unit root test statistics, and then transformed accordingly. The transformation for each series is listed in the appendix. It should be emphasized that many of the procedures are only slightly affected by the use of first differences vs. levels. In particular, the forecasting models in section 5 produce similar short-run forecasts using levels or first differences (this would not be the case if many of the series were cointegrated, but there are neither theoretical nor empirical reasons to suspect widespread cointegration among these series).

3. Description of Stability Tests

The empirical analysis uses variants of three classes of tests for parameter stability: tests for random (time-varying) coefficients; tests based on cumulative forecast errors (CUSUM tests); and tests based on sequential Wald tests for a single break. For completeness, we briefly summarize these tests here, although details are available in the original references. For additional discussion of these tests see the review by Stock (1994) and for additional references to stability tests more generally see Hackl and Westlund (1989, 1991).

The general model considered is,

(1) $$y_t = \mu_t + \alpha_t(L)y_{t-1} + \beta_t(L)x_{t-1} + \epsilon_t$$

where $\alpha_t(L)$ and $\beta_t(L)$ are p-th order lag polynomials which in general are time varying and where ϵ_t is serially uncorrelated with mean zero and variance σ^2. Let k denote the total number of regressors. Each test has as its null hypothesis that the parameters are constant, that is, $\mu_t = \mu$, $\alpha_t(L) = \alpha(L)$ and $\beta_t(L) = \beta(L)$. The derivation of the null distributions of the test statistics also assumes that the regressors are jointly second order stationary, along with additional technical conditions. When the discussion below refers to univariate tests, it is understood that the terms in x_{t-1} in (1) are omitted.

A. Tests for time-varying parameters

The first set of tests for randomly time-varying coefficients are Nyblom's (1989) locally most powerful tests against the alternative that the coefficients follow a random walk, where the random walk error is independent of ϵ_t. Nyblom (1989) derived his statistic against the alternative that all the coefficients are stochastic, and this requires some modification since we also test subsets of coefficients. Rewrite the regression (1) as $y_t = \theta_t' z_t + \epsilon_t$ in obvious notation. Suppose under the alternative that $q \leq k$ linear combinations of θ_t follows a random walk; that is, $R\theta_t = R\theta_{t-1} + \eta_t$, where R is a $q \times k$ matrix of constants which are either known or can be consistently estimated under the null, and η_t is i.i.d. and uncorrelated with ϵ_t. Then the modified Nyblom statistic is $L = T^{-2} \sum_{t=1}^{T} S_t' \hat{V}^{-1} S_t$, where $S_t = R \sum_{s=1}^{t} z_s e_s$, where $\{e_s\}$ are the residuals from OLS estimation of (1), and where $\hat{V} = (RT^{-1} \sum_{t=1}^{T} z_t z_t' R')\hat{\sigma}^2$, where $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} e_t^2$.

The test is evaluated for three different sets of coefficients:

(2a) L_{all}: test μ_t, $\alpha_t(L)$, $\beta_t(L)$;

(2b) $L_{\mu,\beta}$: test μ_t, $\beta_t(L)$;

(2c) $L_{\mu,\beta(1)}$: test μ_t, $\bar{\beta}_t$, where $\beta_t(L) = \bar{\beta}_t b(L)$ and b(1) is normalized so $b(1) = 1$.

For L_{all}, R is the $k \times k$ identity matrix; for $L_{\mu,\beta}$, R is the matrix of ones and zeros such that $Rz_t = (1, x_{t-1},\ldots,x_{t-p})'$; and for $L_{\mu,\beta(1)}$, R is the matrix such that $Rz_t = (1, \hat{\delta}(L)x_{t-1})$, where $\hat{\delta}(L) = \hat{\beta}(L)/\hat{\beta}(1)$ where $\hat{\beta}(L)$ is the OLS estimator of $\beta(L)$ under the null. The $L_{\mu,\beta(1)}$ statistic tests for stochastic evolution of the cumulative effect of x_t on the forecast of y_t.

Heteroskedasticity-robust variants of the Nyblom statistics were also computed by replacing $\hat{V}$ with $\tilde{V} = RT^{-1}\sum_{t=1}^{T} e_t^2 z_t z_t' R'$ (Hansen (1990)). The heteroskedasticity-robust versions of the tests in (2a), (2b) and (2c) are respectively denoted by L_{all}^r, $L_{\mu,\beta}^r$, and $L_{\mu,\beta(1)}^r$.

Also computed were two variants of the Breusch-Pagan (1979) LM test for random coefficients, for which the alternative hypothesis is that the coefficients are i.i.d. draws from a distribution with constant mean and finite variance. The two statistics are,

(3a) $BP_{all} = TR^2$ from the regression of e_t^2 onto $(1, y_{t-1}^2,\ldots,y_{t-p}^2, x_{t-1}^2,\ldots,x_{t-p}^2)$;

(3b) $BP_{\beta} = TR^2$ from the regression of e_t^2 onto $(1, x_{t-1}^2,\ldots,x_{t-p}^2)$;

(3c) $BP_{\beta(1)} = TR^2$ from the regression of e_t^2 onto $(1, (\hat{\delta}(L)x_{t-1})^2)$.

B. Tests based on cumulative forecast errors

One of the tests based on cumulative forecast errors is the maximal OLS CUSUM statistic proposed by Ploberger and Krämer (1992), which is similar to Brown, Durbin and Evans' (1975) CUSUM statistic except that the Ploberger-Krämer (1992) statistic is computed using OLS rather than recursive residuals. Let $\zeta_T(\delta) = T^{-1/2}\sum_{s=1}^{[T\delta]} e_s$, where $[\bullet]$ is the greatest lesser integer function. The Ploberger-Krämer (1992) maximal CUSUM statistic is,

(4) $$PK_{sup} = \sup_{\delta \in [0,1]} |\zeta_T(\delta)|.$$

A related statistic is the mean square of ζ_T:

(5) $$PK_{msq} = \int_0^1 \zeta_T(\delta)^2 d\delta.$$

The PK_{sup} and PK_{msq} statistics respectively have limiting representations as the supremum and the integral of the square of a one-dimensional Brownian bridge.

C. Tests based on sequential Wald statistics

The third set of tests statistics consists of functionals of the sequence of Wald test statistics, $F_T(\delta)$, which test the null hypothesis that the parameters are constant against the alternative that they have a single break at a fraction δ through the sample. The break date is treated as unknown *a-priori*, so that the tests involve computing the sequence $F_T(t/T)$ for $t=t_0,\ldots,t_1$, and then computing a functional of this sequence. Three such functionals are considered. The Quandt (1960) likelihood ratio (QLR) statistic, in Wald form, is given by

(6) $$QLR = \sup_{\delta\in(\delta_0,\delta_1)} F_T(\delta).$$

The mean Wald statistic (Hansen (1992), Andrews and Ploberger (1992)) is

(7) $$MW = \int_{\delta_0}^{\delta_1} F_T(\delta)d\delta.$$

The Andrews-Ploberger (1992) Wald statistic is the exponential average,

(8) $$APW = \ln\{ \int_{\delta_0}^{\delta_1} \exp(\tfrac{1}{2}F_T(\delta))d\delta\}.$$

These statistics have asymptotic representations as functionals of a k-dimensional Brownian bridge; see Andrews (1993) and Andrews and Ploberger (1992) for the details. The tests are implemented with 15% symmetric trimming ($\delta_0=1-\delta_1=.15$).

4. Stability Tests: Empirical Evidence

A. Univariate Tests

The values of the univariate stability test statistics, along with summary statistics on the fraction of rejections, are given for all 76 series in table 1. The final column contains the regression F statistic testing the hypothesis that the transformed series follows an AR(0). The first panel reports summary measures of rejections for each series. The second panel reports each of the individual test statistics. For all regressions, $p=6$.

The answer to the question of whether there is evidence of widespread instability in these univariate autoregressions evidently depends on which stability test one uses. On the one hand, 50% of the series reject at the 5% level using the QLR statistic, and similar results obtain for the APW statistic. There are also many, if fewer, rejections using the MW statistic. These results provide evidence of one-time shifts in the parameters of the univariate autogressions. While the Breusch-Pagan (1979) test often rejects, this test also has power against heteroskedasticity so it is not clear whether this indicates heteroskedasticity or time variation in the parameters. The Nyblom test has a lower rejection rate (20% at the 5% level); when an adjustment is made for heteroskedasticity, the rejection rate drops to the level of the test. However, because the random walk time variation introduces heteroskedasticity, the heteroskedasticity-robust L^r statistics seems likely to have lower power than the L statistics, so the drop in significance from the L to L^r statistics need not be interpreted as evidence against time variation in the data but rather simply as evidence that the adjusted test has lower power. The rejection rates for the PK CUSUM statistics are low, suggesting that shifts in the intercept are not a major feature in these data.

The instability is more heavily concentrated in certain classes of series than others. For example, the QLR statistic rejects at the 5% level for all interest rate and inflation series. In contrast, other than the Breusch-Pagan test which could be detecting heteroskedasticity, none of the tests rejects for business failures, the government finance series, and several of the orders and inventories series.

B. Bivariate Tests

There are 5700 bivariate forecasting relations among our 76 series so rather than present all of the test statistics we present various graphical and tabular summaries. Summary rejection rates of the bivariate tests for parameter stability are presented in table 2. The final column reports the Granger causality Wald statistic testing the hypothesis that $\beta(L)=0$ in (1). For all regressions $p=6$, and all tests have level 10%.

The results for these 5700 bivariate relations are summarized in panel A of table 2. The main feature is the evidence of widespread instability in these relations, although this instability is only detected by a subset of the tests. The $L_{\mu,\beta}$ statistic rejects in over 70% of the cases, and its heteroskedasticity-robust variant rejects in almost 60% of the cases. The QLR and APW statistics also reject in approximately 60% of the cases. A large fraction (58%) of cases also have significant Granger causality statistics, a result which is perhaps surprising since no *a-priori* economic reasoning was used to select which variables should be used to forecast any particular dependent variable. As in the univariate results, the CUSUM-based tests have low rejection rates, which suggests that the instability does not arise from breaks or drift in the direction of the mean regressors. As the final two rows of panel A indicate, there is only slightly more instability among statistically significant predictive relationships (based on the Granger causality test) than among insignificant relationships.

These results can be used to examine stability in relations involving those variables which commonly appear in structural VAR modeling. Industrial production, real personal income, manufacturing employment, the CPI, the PPI, the 90-day Treasury bill rate, and the commercial paper-Treasury bill spread each reject stability in at least 93% of their 75 respective bivariate relations based on either the $L_{\mu,\beta}$ or QLR statistics, and M1 rejects in 77% of its bivariate relations based on the $L_{\mu,\beta}$ statistic, when these series are used as dependent variables (panel B). When these series appear as predictor variables (panel C), for each the QLR rejects in at least 59% of the 75 pairs. For five of the seven price series, the QLR statistic rejects stability in each of the 75 bivariate forecasting relations in which inflation is a dependent variable. When any of these five price series is instead used as a predictor, the QLR statistic again rejects in more than half the cases. If anything, it appears that instability in bivariate relations involving these key series is even more prevalent than on average across all 5700 relations.

These marginal distributions provide one window on the extent of instability in these 5700 relations. However, it is possible that some of this instability is in relations which would be of little interest from a forecasting perspective because they have low overall predictive content. Exploring this possibility requires examining the joint distribution of the instability and Granger-causality test statistics. This is done graphically in figures 1-4, which are scatterplots of selected stability test statistics against the Granger causality test statistic. These figures confirm that the stability and Granger causality test statistics are only weakly correlated. In a sense, each forecasting relation can be thought of as having a temporal average level of predictive content, and deviations from that predictive relation over time are largely uncorrelated with the average predictive content.

Figure 5 summarizes the estimated break dates ($[T\hat{\delta}]$, where $\hat{\delta}$ maximizes $F_T(\delta)$) for the bivariate relations for the which the corresponding QLR statistics are significant at the 5% level. Instability is concentrated around 1974-75, 1980-81, and at the endpoints (δ_0 and δ_1 in (6)).

Figure 6 is a scatterplot of the QLR vs. the APW statistics. Although these are rather different functionals of the sequential Wald statistics, respectively the maximum and an exponential average, the statistics are clearly highly correlated and give quite similar inferences in these data. Evidently little is lost by considering only one or the other of these statistics.

The large number of rejections when the Granger causality statistic is insignificant presents an intriguing opportunity. This is most easily seen for the QLR statistic. If there is in fact no predictive content in a bivariate relation in any subsample, then the fraction of QLR rejections will tend towards the level of the test. In contrast, a rejection by the QLR test but not by the Granger causality test suggests that the bivariate relation has predictive content in at least one continuous subsample, even though on average over the full sample it does not (ignoring type I errors). For example, oil prices (pw561) was a useful predictor for only 23% of the other series, but 61% of the corresponding QLR tests rejected. This raises the possibility that models which adapt to changing relations might find new, albeit transitory, forecasting relations to exploit.

5. Description of Forecasting Models

The second stage in this investigation is an examination of the performance of sixteen fixed and adaptive forecasting models. Eight of the forecasting models are univariate while eight are bivariate. Throughout, a (pseudo) in-sample estimation period is used for preliminary estimation of the parameters and a (pseudo) out-of-sample period is used for forecasting.

The eight univariate models consist of a fixed-parameter autoregression, two autoregressions estimated by rolling regression, one autoregression estimated by recursive least squares, and four random walk TVP models. The eight multivariate models are a fixed-parameter bivariate model, two bivariate models estimated by rolling regression, one model estimated by recursive least squares, and four bivariate models with random-walk time TVP. All models are of the form (1), with the coefficients fixed or time-varying as appropriate. The bivariate models will be referred to as vector autoregressions (VAR's), although because only one-step ahead forecasts are considered only the single equation (1) of the (y_t, x_t) VAR needs to be estimated.

The specification of the TVP models is conventional and assumes the coefficients follow a random walk. Let $\theta_t = (\mu_t, \alpha_{1t}, \ldots, \alpha_{pt}, \beta_{1t}, \ldots, \beta_{pt})$ (where $\{\beta_{it}\}$ are omitted in the univariate application). Then,

$$\theta_t = \theta_{t-1} + \eta_t,\ \eta_t \text{ i.i.d. } (0, \lambda^2\sigma^2 I_k) \qquad (9)$$

where I_k is the $k \times k$ identity matrix, so that λ^2 is the ratio of the variance of the parameter disturbance η_t to the variance of the regression error ϵ_t. The parameters of the TVP models are θ_0, σ^2, and λ. For each TVP model the value of θ_0 is set to its OLS estimate over the in-sample period, and σ^2 is estimated using the in-sample data. (The out-of-sample forecasts and their relative performances are insensitive to choice of initial conditions because of the long in-sample period.) One-step ahead forecasts are then produced using period-by-period updating with the Kalman filter. We consider four TVP models that differ only in their choice of λ. So that a single value of λ could be applied to series measured in different units, all series were first rescaled by dividing through by their in-sample standard deviation.

The eight univariate models are:

(10a) AR: AR(6); μ, $\alpha(L)$ estimated by OLS, then fixed at in-sample values for out-of-sample forecasts

(10b) RRA1: AR(6) estimated using rolling regression with 120 observations

(10c) RRA2: AR(6) estimated using rolling regression with 240 observations

(10d) RLSA: AR(6) estimated by recursive least squares

(10e) ATVP1: AR(6) estimated by TVP with λ = .005

(10f) ATVP2: AR(6) estimated by TVP with λ = .010

(10g) ATVP3: AR(6) estimated by TVP with λ = .020

(10h) ATVP4: AR(6) estimated by TVP with λ = .030.

The eight bivariate models are:

(11a) VAR: VAR(6); μ, $\alpha(L)$, $\beta(L)$ estimated by OLS, then fixed at in-sample values for out-of-sample forecasts

(11b) RRV1: VAR(6) estimated using rolling regression with 120 observations

(11c) RRV2: VAR(6) estimated using rolling regression with 240 observations

(11d) RLSV: VAR(6) estimated by recursive least squares

(11e) VTVP1: VAR(6) estimated by TVP with $\lambda = .005$

(11f) VTVP2: VAR(6) estimated by TVP with $\lambda = .010$

(11g) VTVP3: VAR(6) estimated by TVP with $\lambda = .020$

(11h) VTVP4: VAR(6) estimated by TVP with $\lambda = .030$.

Two related models are univariate and bivariate TVP models with λ estimated over the in-sample period. These models are not examined because this estimation is currently computationally impractical for the large number of forecasting relations under consideration.

For the forecast comparisions, the in-sample period ends in 1978:12. This cutoff date was chosen so that the models are tested in the turbulent economic conditions of the late 1970's and early 1980's. For series ending in 1993:12, 180 observations remain for the out-of-sample comparison.

6. Forecasting Model Comparison: Empirical Results

Comparing the various models using these data entails examining 608 univariate forecasting systems (76 variables, eight models each) and 45,600 bivariate forecasting systems (5700 bivariate forecasting relations, eight models each). All comparisons are made using out-of-sample one-month-ahead forecast MSE's, although in principle other loss functions could be used. The term "best model" will be used to refer to the model which minimizes this out-of-sample forecast MSE, relative to some comparison group. One objective of this comparison is to see which models do best most frequently. However, because of the instability found in section 4, another objective is to ascertain which if any of the models protect the forecaster from making extreme forecast errors resulting from parameter instability.

The question of which model performs best out-of-sample most frequently is examined in table 3. For each bivariate relation, MSE's from the eight bivariate and eight univariate models were computed; for the purposes of this tabulation, the model producing the lowest out-of-sample MSE among these sixteen was then deemed the "best" model for that (y_t, x_t) pair. Two sets of tabulations are presented. Panel B presents the fraction of times the column model is best among the 75 bivariate relations, broken down by forecasted variable. For example, for industrial production, in 16% of the 75 bivariate pairs, RRV2 produces the smallest out-of-sample MSE; in 61% of these pairs, ATVP1 outperforms not just the other seven univariate models but also the eight bivariate models. Panel C presents analogous results, broken down by forecasted variable, except that for each forecasted variable the comparison is only among the top ten of the 75 pairs, as measured by the Bayes Information Criterion (BIC) for the in-sample OLS estimation of (1) with fixed parameters. Thus, among forecasts of IP based on the ten variables with the lowest in-sample BIC's, in one case (10%) RRV2 has the lowest out-of-sample MSE, but in 7 cases ATVP1 outperforms the other univariate and the eight bivariate models. The first two rows of panel A respectively summarize the results of panels B and C, where the fractions are computed over all the forecasted variables. The final row of panel A presents results for bivariate relations with significant time variation, as measured by significance (at the 10% level) of the $L_{\mu,\beta}$ statistic evaluated for the in-sample period.

Several conclusions are evident from table 3. Overall, there is no clearly dominant model; no model performs best in more than 17% of the 5700 pairs. However, there is strong evidence that the adaptive models (the rolling, recursive and TVP models) outperform the two fixed-parameter models. Among the set of models with predictors based on the top 10 BIC's, 73% of the best-performing models are adaptive. However, the extreme TVP models (with $\lambda = .02$ and $\lambda = .03$) are rarely the top performers out of sample. Interestingly, comparing the final row of panel A with the first two rows indicates that the adaptive models perform similarly whether or not in-sample instability is detected. Consistent with the stability test evidence, the results in panels B and C show that different variables tend to be forecast best by different models. For example, exchange rates (exnwt2) have no rejections using the univariate stability tests but a moderately high fraction of rejections using the bivariate Nyblom tests, and in 88% of the 75 pairs the best forecasting model is the fixed-coefficients AR. In contrast, real personal income (gmyxp8) has widespread rejections by univariate and bivariate stability tests, and in 97% of the cases the best forecasting models are adaptive, primarily the univariate or bivariate recursive least squares models.

Table 4 summarizes pairwise comparisons of the sixteen models over all 5700 bivariate relations. The ATVP3 and ATVP4 models typically have MSE's worse than the other univariate models; the VTVP3 and VTVP4 models also typically perform worse than the other bivariate models. Among the univariate models, only RLSA and ATVP1 outperform the simple AR in more than 50% of the cases, and RLSA in turn outperforms ATVP1 in 63% of the cases. Among the bivariate models, RLSV outperforms all others at least 64% of the time. The RLSA and ATVP1 models typically outperform all bivariate models. While the univariate models often outperform the bivariate models, this is perhaps not surprising since *a-priori* reasoning would lead one to suspect that many of the 5700 pairs would have forecasting links which are weak at best.

Table 4 and the test results from section 4 provide additional evidence of instability. While 58% of the 5700 Granger causality statistics reject at the 10% level, only 38% of the recursive VAR forecasts (the best-performing bivariate forecast) outperform the recursive AR model out of sample. Presumably some of these Granger causality rejections are Type I errors, but with 75% power against the "true" VAR's and a 10% level test these 58% rejections correspond to 74% of the VAR's having nonzero coefficients on the predictor variable. The much lower fraction of pairs for which bivariate techniques improve performance out of sample thus is another indication of instability in the bivariate relations.

It is useful to go beyond these assessments of which model typically performs best to quantify the extent to which the various models reduce the possibility of extremely poor performance. Table 5 presents the empirical quantiles of the MSE's of the various models. To make results comparable across series, the MSE's are relative to the MSE for the recursive least squares AR forecast (RLSA). Panel A shows the distribution of these relative MSE's for the univariate forecasting models. The median values all exceed 1.00, consistent with the finding in Table 4 that RLSA has lower MSE than the other forecasts more than 50% of the time. The results also suggest that RLSA is the most robust univariate forecasting model, in the sense that its worst performance is significantly better than the worst performance of the other models. For example, the minimum relative MSE for the fixed coefficent AR model is .959 while its maximum relative MSE is 1.158. Thus, at its best, the AR forecast outperforms the RLSA forecast by 4.1%, while at its worst, the AR forecast underperforms the RLSA forecast by 13.6% ($=1-(1.158)^{-1}$). Panel B presents quantiles for the 5700 bivariate forecasts. Robustness at the α'th quantile can be determined by comparing the relative MSE at α to the inverse of the relative MSE at the ($1-\alpha$)'th quantile. For example, the relative MSE of VTVP1 at $\alpha=.001$ is .600, so that in .1% of the cases, VTVP1 outperforms RLSA by at least 40%. At $\alpha=.999$, the relative MSE of VTVP1 is 1.146, so that in .1% of the cases, RLSA outperforms VTVP1 by more than 12.7% ($=1-(1.146)^{-1}$). In this sense, in the .1% extremes, VTVP1 produces better forecasts than RLSA. The table indicates that at the .1% quantile, all of the bivariate models dominate RSLA, while at $\alpha=.5$% and 1%, only bivariate models with small time variation (RLSV and VTVP1) dominate RLSA. Similar results obtain at $\alpha=1$% for the best 10 BIC-selected models shown in panel C.

7. Conclusions

Some caveats are warranted. The stability tests are all evaluated using asymptotic critical values and some finite-sample size distortions are to be expected. However, Monte Carlo results for sequential Wald tests in Diebold and Chen (1992) suggest that these distortions are only moderate and could plausibly account for only a small fraction of the empirical rejections. Also, the quantitative results of the forecasting comparison depend on the choice of out-of-sample period.

With these caveats in mind, these results suggest some general observations about the two time series applications laid out in the introduction, structural VAR modeling and forecasting. One finding relevant to VAR modeling is that relations involving key macroeconomic variables such as industrial production, personal income, employment, prices, and interest rates are, if anything, more likely to be unstable than the average of these 5700 bivariate relations. While most VAR modeling involves more than two variables, a stable multivariate VAR implies a stable VAR for any subset of variables, so these tests can be seen as testing an implication of the hypothesis of stability of larger models. While it is possible for some (but not all) equations in a multivariate VAR to be stable despite instability of the bivariate VAR formed from two of the variables, impulse responses involve all equations in the VAR so instability in the bivaraite relations implies instability in at least some multivariate impulse responses. One practical lesson which this emphasizes is the importance of performing systematic stability analysis as part of a structural VAR modeling exercise.

Given the widespread evidence of structural instability, the comparison of the forecasting models yielded some surprising results. While the fixed-parameter models occasionally worked very poorly, models with only small degrees of adaptivity performed well. In particular the univariate and bivariate recursive least squares models typically were either best or nearly best. In 38% of the cases, the recursive VAR outperformed the recursive AR out of sample; in 99% of the 5700 bivariate relations, the recursive VAR model produced an out-of-sample MSE which was at most 7.8% higher than the recursive AR model. Moreover, in 99% of the bivariate relations, the recursive VAR model produced an out-of-sample MSE which was at most 8.7% higher than the best-performing model for each bivariate relation (including the univariate models) and in 50% of the cases its MSE was within 1.4% of the best model. Of course, if the parameters are in fact constant, then the recursive estimator will be efficient relative to the fixed parameter models. One striking result is that the more adaptive models such as rolling regression or the TVP models typically failed to improve upon recursive least squares, and indeed did not even guard against extreme failures as well as did recursive least squares. This negative finding suggests that the class of models considered here, which are the models most widely studied in adaptive forecasting, are largely unsuccessful in modeling and exploiting the instability we found in typical macroeconomic applications.

Appendix: Definitions of Series

The entries for each series are the series mnemonic, the transformation code, and the definition of the series. For series obtained from CITIBASE, the CITIBASE mnemonic has been used. The transformation codes are: 0=first difference, 1=log first difference, 2=level.

A. Output and Sales
ip 1 index of industrial production
ipxmca 2 capacity util rate: manufacturing,total(% of capacity,sa)(frb)
gmpy 1 personal income: total (bil$,saar)
gmyxp8 1 personal income (real) less transfers
rtql 1 retail trade: total (mil.87$)(s.a.)
gmcq 1 personal consumption expenditure:total (bill.1987$)
ipcd 1 industrial production: durable consumer gds (1987=100,sa)
ced87m 1 personal consumption expenditures:durable goods,87$
xci 1 stock-watson index of coincident indicators
mt82 1 manuf. and trade sales

B. Employment
lpmhuadj 1 total employee hours (adjusted)
lphrm 2 avg. weekly hrs. of production wkrs.: manufacturing (sa)
lhell 1 index of help-wanted adv.
lhnaps 1 persons at work: part time eas-slack wk,nonag (thous,sa)
luinc 2 avg wkly initial claims,state unemploy.ins.,exc p.rico(thous;sa)
lhu5 1 unemploy.by duration: persons unempl.less than 5 wks (thous.,sa)
lhur 0 unemployment rate: all workers, 16 years & over (%,sa)
lhelx 2 employment: ratio; help-wanted ads:no. unemployed clf

C. New Orders
hsbp 2 housing authorized: index of new priv housing units (1967=100;sa)
mdu82 1 mfg unfilled orders: durable goods industries, 82$
mpcon8 1 contracts & orders for plant & equipment in 82$(bil$,sa)! 2
mocm82 1 mfg new orders: consumer goods & material,82$(bil$,sa)! 2
mdo82 1 mfg new orders: durable goods industries,82$(bil$,sa)! 2
ivpac 2 vendor performance: % of co's reporting slower deliveries(%,nsa)
pmi 2 purchasing managers' index (sa)
pmno 2 napm new orders index (percent)

D. Inventories
invmt87 1 manufacturing & trade inventories:total,87$(bil$,sa)
invrd 1 inventories, retail (sa)
invwd 1 inventories, wholesale (sa)
ivm1d8 1 mfg inventories: materials & supplies, all mfg indus 87$(sa)
ivm2d8 1 mfg inventories: work in process, all mfg indus 87$(sa)
ivm3d8 1 mfg inventories: finished goods, all mfg industries 87$(sa)
ivmtd 1 manufacturing & trade inventories:total
ivm1d 1 mfg inventories: materials & supplies, all mfg indus (mil$,sa)

ivm2d 1 mfg inventories: work in process, all mfg indus (mil$,sa)
ivm3d 1 mfg inventories: finished goods, all mfg industries (mil$,sa)
invrd8 1 inventories, retail 87$ (sa)
invwd8 1 inventories, wholesale 87$ (sa)

E. Prices
gmdc 1 pce,impl pr defl:pce (1987=100)
punew 1 cpi-u: all items (82-84=100,sa)
pw 1 producer price index: all commodities (82=100,nsa)
pw561 1 producer price index: crude petroleum (82=100,nsa)
pw561r 1 pw561/punew
jocci 1 dept. of commerce commodity price index
joccir 1 jocci/punew

F. Interest Rates
fyff 0 interest rate: federal funds (effective) (% per annum,nsa)
fygm3 0 interest rate: u.s.treasury bills,sec mkt,3-mo.(% per ann,nsa)
fygm6 0 interest rate: u.s.treasury bills,sec mkt,6-mo.(% per ann,nsa)
fygt1 0 interest rate: u.s.treasury const maturities,1-yr.(% per ann,nsa)
fybaac 0 bond yield: moody's baa corporate (% per annum)
fygt10 0 interest rate: u.s.treasury const maturities,10-yr.(% per ann,nsa)
cp6_gm6 2 yield on 6 month commercial paper - fygm6
g10_g1 2 fyt10 - fygt1
g10_ff 2 fygt10 - fyff
baa_g10 2 fybaac - fygt10

G. Money and Credit
fcbcuc 2 change in bus and consumer credit outstand.(percent,saar)(bcd111)
fcbcucy 2 fcbcuc-annual percentage growth in GMPY
delinqcr 0 delinq. rate, total install. credit
cci30m 0 consumer instal.loans: delinquency rate,30 days & over, (%,sa)
fm1d82 1 money stock: m-1 in 1982$ (bil$,sa)(bcd 105)
fm2d82 1 money stock: m-2 in 1982$(bil$,sa)(bcd 106)
fmbase 1 monetary base, adj for reserve req chgs(frb of st.louis)(bil$,sa)
fm1 1 money stock: m1(curr,trav.cks,dem dep,other ck'able dep)(bil$,sa)
fm2 1 money stock:m2(m1+o'nite rps,euro$,g/p&b/d mmmfs&sav&sm time dep(bil$,
fm3 1 money stock: m3(m2+lg time dep,term rp's&inst only mmmfs)(bil$,sa)
fmbaser 1 monetary base: fmbase/punew

H. Other Variables
exnwt2 1 Trade weighted average nominal exhange rate
fspcom 1 s&p's common stock price index: composite (1941-43=10)
fspcomr 1 fspcom/punew
fail 1 business failures: current liabilities (mil$,nsa)
failr 1 fail/punew
gfosa 1 federal government outlays seasonally adjusted
gfrsa 1 federal government receipts seasonally adjusted
gfor 1 Real federal government outlays, gfosa/punew
gfrr 1 Real federal government receipts, gfrsa/punew
hhsntn 2 u. of mich. index of consumer expectations(bcd-83)

Footnotes

1. Applications of adaptive forecasting include Baudin, Nadeau, and Westlund (1984), Guyton, Zhang and Foutz (1986), Engle, Brown and Stern (1988), Sessions and Chatterjee (1989), Schneider (1991), Young, Ng, Lane and Parker (1991), Zellner, Hong and Min (1991), Edlund and Søgaard (1993), Min and Zellner (1993), and the time-varying VAR's developed in Doan, Litterman and Sims (1984), Highfield (1986), and Sims (1982, 1993). Surveys of TVP models are provided by Chow (1984), Nichols and Pagan (1985), Engle and Watson (1987), and Harvey (1989).

References

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

Andrews, D.W.K., I. Lee and W. Ploberger (1992), "Optimal Changepoint Tests for Normal Linear Regression," Cowles Foundation Discussion Paper #1016.

Baudin, A., S. Nadeau, and A. Westlund (1984), "Estimation and Prediction Under Structural Instability: The Case of the U.S. Pulp and Paper Market," *Journal of Forecasting*, 3, 63-78.

Bernanke, B.S. and A.S. Blinder (1992), "The Federal Funds Rates and the Channels of Monetary Transmission," *American Economic Review*, 82, 901-921.

Breusch, T.S. and A.R. Pagan (1979), "A Simple Test for Heteroskedasticity and Random Coefficient Variation," *Econometrica*, Vo. 47, pp. 1287-1294.

Brown, R.L., Durbin, J. and Evans, J.M. (1975), "Techniques for Testing the Constancy of Regression Relationships over Time with Comments," *Journal of the Royal Statistical Society, Series B*, 37, 149-192.

Chow, G.C. (1984), "Random and Changing Coefficient Models," in Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics*, North Holland, Amsterdam, Chapter 21.

Christiano, L., M. Eichenbaum, and C. Evans (1993), "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," manuscript, Department of Economics, Northwestern University.

Cooley, T.F. and E.C. Prescott (1973a), "An Adaptive Regression Model," *International Economic Review*, 14, pp. 364-371.

Cooley, T.F. and E.C. Prescott (1973b), "Tests of an Adaptive Regression Model," *Review of Economics and Statistics*, 55, pp. 248-256.

Cooley, T.F. and E.C. Prescott (1976), "Estimation in the Presence of Stochastic Parameter Variation," *Econometrica*, 44, pp. 167-184.

Diebold, F.X. and C. Chen (1992), "Testing Structural Stability with Endogenous Break Point: A Size Comparison of Analytic and Bootstrap Procedures," manuscript, Department of Economics, University of Pennsylvania.

Doan, T., R. Litterman and C. Sims (1984), "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, Vol. 3, 1-100.

Edlund, P.-O. and H.T. Søgaard (1993), "Fixed versus Time-Varying Transfer Functions for Modelling Business Cycles," *Journal of Forecasting*, 12, 345-364.

Engle, R.F. and M.W. Watson (1987), "The Kalman Filter Model: Applications to Forecasting and Rational Expectations Models," in *Advances in Econometrics: Fifth World Congress of the Econometric Society*, T. Bewley (ed.), Cambridge University Press, Cambridge, U.K.

Engle, R.F., S. Brown and G. Stern (1988), "A Comparison of Adaptive Structural Forecasting Methods for Electricity Sales," *Journal of Forecasting*, Vol. 7, 149-172.

Guyton, D.A., N-F Zhang, and R.V. Foutz (1986), "A Random Parameter Process for Modeling and Forecasting Time Series," *Journal of Time Series Analysis*, 7, 105-115.

Hackl, P. and A. Westlund (1989), "Statistical Analysis of 'Structural Change': An Annotated Bibliography," *Empirical Economics*, 143, 167-192.

Hackl, P. and A.H. Westlund (eds) (1991), *Economic Structural Change: Analysis and Forecasting*, Springer-Verlag, Berlin.

Hansen, B.E. (1990), "Lagrange Multiplier Tests of Parameter Instability in Non-linear Models," manuscript, Department of Economics, University of Rochester.

Hansen, B.E. (1992), "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics*, 10, 321-336.

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

Highfield, R.A. (1986), "Forecasting with Bayesian State Space Models," Unpublished Ph. D. thesis, University of Chicago, Graduate School of Business.

Makridakis, S., A. Anderson, R. Carbonne, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, and R. Winkler (1982), "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," *Journal of Forecasting*, 1, 111-153.

Meese, R. and J. Geweke (1984), "A Comparison of Autoregressive Univariate Forecasting Procedures for Macroeconomic Time Series," *Journal of Business and Economic Statistics*, 2, 191-200.

Min, C. and A. Zellner (1993), "Bayesian and non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates," *Journal of Econometrics*, 56, pp. 89-118.

Nichols, D.F. and A.R. Pagan, (1985), "Varying Coefficient Regression," in E.J. Hannan, P.R. Krishnaiah and M.M. Rao, eds., *Handbook of Statistics*, Vol. 5, North Holland, Amsterdam, pp. 413-449.

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223-230.

Ploberger, W. and W. Krämer (1992), "The CUSUM test with OLS Residuals," *Econometrica*, 60, 271-286.

Quandt, R.E. (1960), "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association*, 55, 324-330.

Rosenberg, B. (1972), "The Estimation of Stationary Stochastic Regression Parameters Re-Examined," *Journal of the American Statistical Association*, 67, pp. 650-654.

Rosenberg, B. (1973), "The Analysis of a Cross-Section of Time Series by Stochastically Convergent Parameter Regression," *Annals of Economic and Social Measurement*, 2, 461-484.

Sarris, A.H. (1973), "A Bayesian Approach to Estimation of Time Varying Regression Coefficients," *Annals of Economic and Social Measurement*, 2, 501-523.

Schneider, W. (1991), "Stability Analysis Using Kalman Filtering, Scoring, EM, and an Adaptive EM Method," in Hackl, P. and A.H. Westlund (eds), *Economic Structural Change: Analysis and Forecasting*, Springer-Verlag, Berlin.

Sessions, D.N. and S. Chatterjee (1989), "The Combining of Forecasts Using Recursive Techniques with Non-stationary Weights," *Journal of Forecasting*, 8, 239-251.

Sims, C.A. (1982), "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity*, 1, 107-164.

Sims, C.A. (1993), "A Nine-Variable Probabilistic Macroeconomic Forecasting Model," in J.H. Stock and M.W. Watson (eds.), *Business Cycles, Indicators and Forecasting*, University of Chicago Press for the NBER, 179-204.

Stock, J.H. (1994), "Unit Roots, Structural Breaks, and Trends," in R.F. Engle and D. McFadden (eds), *The Handbook of Econometrics, Vol. IV*. Amsterdam: Elsevier.

Young, P.C., C.N. Ng, K. Lane, and D. Parker (1991), "Recursive Forecasting, Smoothing and Seasonal Adjustment of Non-stationary Environmental Data," *Journal of Forecasting*, 10, 57-89.

Zellner, A., C. Hong and C. Min (1991), "Forecasting Turning Points in International Output Growth Rates Using Bayesian Exponentially Weighted Autoregression, Time-Varying Parameter, and Pooling Techniques," *Journal of Econometrics*, 49, pp. 275-304.

Table 1
Univariate Tests for Stability

A. Summary: Percent rejections over all series

Test Size	L_{all}	L^r_{all}	PK_{sup}	PK_{msq}	BP_{all}	QLR	MW	APW	F
	Test								
10%	28.9	10.5	21.1	11.8	72.4	53.9	34.2	52.6	98.7
5%	19.7	5.3	10.5	6.6	65.8	50.0	25.0	42.1	97.4
1%	9.2	0.0	0.0	0.0	55.3	30.3	14.5	27.6	96.1

B. Results for Individual Series

Series	Sample	L_{all}	L^r_{all}	PK_{sup}	PK_{msq}	BP_{all}	QLR	MW	APW	F
A. Output and Sales										
ip	59:2 93:12	1.72*	1.08	0.83	0.19	18.61***	21.94**	8.74	6.70	14.22***
ipxmca	59:1 93:12	1.88*	1.17	0.81	0.13	27.73***	18.87*	9.30	7.05*	2450.44***
gmpy	59:2 93:12	3.58***	1.14	1.23*	0.45**	145.24***	58.81***	20.98***	25.51***	3.04***
gmyxp8	59:2 93:12	5.48***	1.54	1.02	0.25	133.62***	70.15***	30.78***	32.03***	3.85***
rtql	59:2 93:12	0.88	0.72	0.60	0.04	43.37***	18.66	6.70	5.20	4.88***
gmcq	59:2 93:12	1.83*	1.58	1.23*	0.37*	47.51***	24.47**	12.17**	8.50**	2.83**
iped	59:2 93:12	1.50	1.08	0.66	0.09	87.19***	23.61**	9.77	7.90**	1.78
ced87m	59:2 93:10	3.02***	1.73*	0.54	0.05	55.30***	46.80***	20.44***	18.47***	5.37***
xci	59:3 93:12	1.68	1.20	0.87	0.21	23.47***	18.33	9.63	5.90	20.08***
mt82	59:2 93:12	1.32	1.17	0.68	0.07	10.65*	13.93	8.00	4.95	2.08*
B. Employment										
lpnbuadj	59:2 93:12	1.48	1.28	0.58	0.07	15.84**	17.53	9.20	5.92	6.27***
lphrm	59:1 93:12	0.83	1.33	1.30*	0.24	27.85***	20.92*	5.61	5.10	317.38***
lhell	59:2 93:12	1.85*	1.70*	0.44	0.03	13.72**	30.11***	11.30*	11.52***	18.43***
lhmsps	59:2 93:12	1.26	1.08	0.56	0.05	32.22***	11.96	8.29	4.60	3.07***
luinc	59:1 93:12	1.41	1.17	1.04	0.24	56.08***	22.81**	10.07	7.77*	1220.68***
lhu5	59:2 93:12	0.89	0.96	0.90	0.16	10.89*	12.36	5.64	3.68	14.78***
lhur	59:2 93:12	1.25	1.10	0.78	0.04	18.23***	15.85	8.22	5.13	9.68***
lhelx	59:1 93:12	1.12	0.78	0.89	0.17	82.00***	25.58**	8.36	8.20**	8275.90***
C. New Orders										
hsbp	59:1 93:12	1.45	1.13	0.82	0.06	21.81***	21.95**	10.98*	7.57*	1005.56***
mdu82	59:2 93:12	2.23**	1.44	0.79	0.17	3.54	24.38**	14.08**	9.83**	55.09***
mpcon8	59:2 93: 9	0.56	0.42	0.58	0.05	24.72***	9.66	4.14	2.59	17.85***
mocm82	59:2 93: 9	1.03	0.86	0.69	0.08	11.31*	15.50	7.13	4.99	3.15***
mdo82	59:2 93: 9	1.47	1.39	0.69	0.10	3.81	28.64***	9.95	9.50**	8.28***
ivpac	59:1 93:12	1.31	1.11	0.89	0.12	15.17**	22.27**	8.84	7.73*	690.03***
pmi	59:1 93:12	0.98	0.90	0.87	0.16	7.72	11.48	5.99	3.74	504.28***
pmno	59:1 93:12	1.01	0.85	0.65	0.08	4.78	12.32	6.25	4.15	225.84***
D. Inventories										
invmt87	59:2 93: 9	1.12	0.98	0.92	0.24	4.90	12.60	6.26	4.09	20.88***
invrd	59:1 93: 9	1.33	1.24	0.87	0.24	9.04	20.34*	7.62	5.87	4.47***
invwd	59:1 93: 9	1.82*	1.32	1.12	0.23	25.33***	22.12**	11.41*	8.28**	12.63***
ivm1d8	59:2 93: 9	3.48***	2.20**	1.13	0.23	38.93***	38.58***	19.54***	14.39***	19.83***
ivm2d8	59:2 93: 9	0.97	1.01	0.97	0.27	6.00	13.66	5.41	4.35	25.43***
ivm3d8	59:2 93: 9	1.55	1.52	1.33**	0.59**	4.35	19.44	9.89	6.44	7.67***

Table 1, continued

Series	Sample	L_{all}	L^{c}_{all}	PK_{sup}	PK_{msq}	BP_{all}	QLR	MW	APW	F
ivmtd	59:1 93: 9	0.88	0.91	1.03	0.23	15.26**	12.44	6.20	3.66	71.58***
ivm1d	59:1 93: 9	3.70***	2.08**	0.84	0.20	60.32***	36.31***	21.71***	14.87***	55.00***
ivm2d	59:1 93: 9	0.99	0.99	1.01	0.21	7.18	13.28	5.08	3.84	51.62***
ivm3d	59:1 93: 9	1.45	1.37	1.48**	0.37*	10.11	19.43	9.47	6.30	24.58***
invrd8	59:2 93: 9	1.38	1.22	0.68	0.06	14.87**	15.40	8.32	5.69	4.92***
invwd8	59:2 93: 9	1.77*	1.23	1.00	0.16	23.71***	19.54	10.37*	7.50*	3.30***
E. Prices										
gmdc	59:2 93:12	1.40	1.05	0.77	0.18	4.28	45.15***	14.54***	18.49***	42.31***
pwnew	59:2 93:12	2.44***	1.42	0.73	0.16	18.06***	44.09***	20.66***	18.10***	73.50***
pw	59:2 93:12	2.35**	0.66	0.95	0.23	113.95***	93.39***	19.78***	41.02***	17.85***
pw561	59:2 93:12	2.22**	0.71	1.27*	0.28	54.89***	56.30***	13.08**	24.13***	11.06***
pw561r	59:2 93:12	1.91**	0.64	1.15	0.23	47.17***	49.08***	10.61*	20.37***	10.24***
jocci	59:2 93:11	1.09	0.92	0.81	0.08	22.54***	25.88**	7.23	9.22**	21.14***
joccir	59:2 93:11	0.98	0.82	0.82	0.07	18.85***	22.50**	6.45	7.61*	19.25***
F. Interest Rates										
fyff	59:2 93:12	1.15	0.49	1.18	0.14	32.68***	39.82***	11.53**	15.13***	15.05***
fygm3	59:2 93:12	1.15	0.54	1.41**	0.21	71.15***	34.74***	9.70	12.45***	19.84***
fygm6	59:2 93:12	1.32	0.57	1.36**	0.21	73.86***	32.45***	10.50*	11.49***	18.75***
fygt1	59:2 93:12	1.05	0.50	1.36**	0.20	87.81***	30.78***	9.03	10.53***	23.86***
fybaac	59:2 93:12	0.76	0.56	1.42**	0.27	74.80***	27.13***	6.25	8.80**	27.23***
fygt10	59:2 93:12	0.83	0.50	1.41**	0.24	48.68***	25.84**	8.90	8.18**	17.81***
cp6_gm6	59:1 93:12	0.65	0.39	0.78	0.07	145.60***	34.04***	5.25	11.38***	218.52***
g10_g1	59:1 93:12	1.18	1.16	1.27*	0.28	26.07***	22.48**	9.37	7.23*	821.55***
g10_ff	59:1 93:12	1.45	0.45	0.65	0.08	114.79***	49.81***	11.56**	18.43***	218.13***
baa_g10	59:1 93:12	1.51	1.59	1.14	0.37*	13.04**	26.78***	12.54**	9.64**	938.84***
G. Money and Credit										
fcbcuc	59:1 92:11	0.95	0.80	1.05	0.25	6.21	15.11	5.70	4.65	149.80***
fcbcucy	59:1 92:11	1.81	1.65	1.10	0.52**	29.15***	17.79	8.69	6.14	15.71***
delinqcr	59:2 93: 6	1.32	1.45	0.88	0.18	6.89	12.49	8.97	4.74	5.33***
cci30m	59:2 93: 9	1.19	1.03	0.99	0.17	5.27	19.09	7.52	6.13	6.47***
fm1d82	59:2 93: 9	1.93**	1.65	1.23*	0.27	24.52***	28.96***	12.79**	9.46**	23.05***
fm2d82	59:2 93: 9	2.08**	2.00**	0.97	0.24	1.59	28.43**	14.98***	10.08***	50.49***
fmbase	59:2 93:12	1.79*	1.87*	1.31*	0.58**	6.11	19.19	10.60*	7.18*	15.89***
fm1	59:2 93:12	1.42	1.33	1.16	0.82**	38.42***	15.38	8.51	4.94	21.47***
fm2	59:2 93:12	1.06	1.07	1.32**	0.37*	3.25	18.13	6.91	5.68	73.20***
fm3	59:2 93:12	2.44***	1.80*	1.11	0.33	3.98	24.97**	15.56***	10.17***	141.76***
fmbaser	59:2 93:12	1.91**	1.53	1.30*	0.23	3.83	29.84***	13.12**	12.54***	19.72***
H. Other Variables										
exnwt2	59:2 93:12	0.90	0.67	0.99	0.11	15.07**	16.86	6.63	5.46	9.33***
fspcomr	59:2 93:12	1.14	1.04	0.99	0.18	18.60***	14.72	7.42	5.45	7.49***
fspcom	59:2 93:12	1.15	1.07	0.96	0.18	15.32**	14.99	7.36	5.47	6.93***
fail	59:2 93:12	1.25	1.06	0.63	0.06	11.06*	16.56	7.92	5.91	36.23***
failr	59:2 93:12	1.25	1.05	0.62	0.05	11.53*	16.56	7.91	5.93	36.49***
gfosa	67:8 93:10	1.01	0.85	0.91	0.22	43.70***	11.07	6.43	3.78	4.29***
gfrsa	67:8 93:10	1.00	1.12	0.89	0.18	4.16	10.07	5.85	3.34	11.94***
gfor	67:8 93:10	0.81	0.68	0.49	0.04	33.18***	10.21	5.07	3.27	4.88***
gfrr	67:8 93:10	1.04	1.20	0.55	0.06	5.30	10.72	5.82	3.31	11.46***
hhsntn	59:1 93:12	1.95**	2.03**	0.84	0.15	41.74***	47.34***	20.23***	21.00***	1077.94***

Notes: Tests are significant at the: *10%, **5%, and ***1% levels. All tests were performed for AR(6) models including a constant term. See the appendix for series definitions and the text for descriptions of the tests.

Table 2
Bivariate Tests for Stability
Percent of Tests Significant at 10% Level

A. Summary of All Regressions

	Test Statistic														
	L_{all}	$L_{\mu,\beta}$	$L_{\mu,\beta(1)}$	L^{r}_{all}	$L^{r}_{\mu,\beta}$	$L^{r}_{\mu,\beta(1)}$	PK_{sup}	PK_{meq}	BP_{all}	BP_{β}	$BP_{\beta(1)}$	QLR	MW	APW	GC
Combined	23.3	70.2	18.8	10.7	59.6	18.8	18.8	15.3	66.3	22.6	23.8	60.2	37.9	59.3	58.4
GC significant	25.5	74.1	21.0	13.0	62.9	21.0	19.7	16.3	66.6	27.6	30.3	61.8	42.0	60.5	100.0
GC insignif.	20.2	64.8	15.6	7.4	55.0	15.6	17.6	13.8	65.9	15.7	14.7	57.9	32.2	57.6	0.0

B. Percent rejections, listed by variable being forecasted

	Test Statistic														
Series	L_{all}	$L_{\mu,\beta}$	$L_{\mu,\beta(1)}$	L^{r}_{all}	$L^{r}_{\mu,\beta}$	$L^{r}_{\mu,\beta(1)}$	PK_{sup}	PK_{meq}	BP_{all}	BP_{β}	$BP_{\beta(1)}$	QLR	MW	APW	GC
A. Output and Sales															
ip	49.3	93.3	17.3	5.3	54.7	17.3	13.3	25.3	98.7	22.7	25.3	68.0	52.0	62.7	66.7
ipxmca	50.7	94.7	21.3	8.0	81.3	21.3	1.3	5.3	100.0	16.0	18.7	60.0	48.0	62.7	89.3
gmpy	100.0	100.0	22.7	14.7	54.7	22.7	62.7	93.3	100.0	18.7	12.0	100.0	100.0	100.0	29.3
gmyxp8	100.0	100.0	25.3	22.7	97.3	25.3	16.0	40.0	100.0	18.7	16.0	100.0	100.0	100.0	58.0
rtql	0.0	20.0	13.3	0.0	37.3	13.3	4.0	1.3	100.0	12.0	14.7	30.7	9.3	30.7	60.0
gmcq	46.7	68.0	24.0	28.0	89.3	24.0	77.3	81.3	100.0	10.7	9.3	81.3	70.7	77.3	57.3
ipcd	17.3	80.0	17.3	14.7	82.7	17.3	13.3	16.0	100.0	24.0	32.0	86.7	52.0	88.7	78.7
ced87m	93.3	100.0	17.3	20.3	98.7	17.3	2.7	4.0	100.0	14.7	16.0	100.0	100.0	100.0	72.0
xci	37.3	90.7	12.0	5.3	88.0	12.0	18.7	28.7	100.0	32.0	32.0	52.0	49.3	54.7	64.0
mt82	8.0	88.7	9.3	2.7	68.0	9.3	5.3	5.3	41.3	21.3	33.3	12.0	10.7	14.7	70.7
B. Employment															
lpmhuadj	24.0	97.3	18.7	22.7	94.7	18.7	1.3	2.7	44.0	21.3	22.7	32.0	38.7	34.7	73.3
lphrm	13.3	32.0	22.7	29.3	60.0	22.7	56.0	34.7	98.7	17.3	34.7	38.7	16.0	32.0	78.7
lhel	53.3	98.7	10.7	42.7	97.3	10.7	2.7	2.7	45.3	20.0	26.7	98.7	72.0	87.3	74.7
lhmaps	16.0	80.0	13.3	10.7	84.0	13.3	0.0	1.3	100.0	28.0	30.7	42.7	45.3	46.7	77.3
luinc	1.3	80.0	14.7	1.3	46.7	14.7	9.3	10.7	100.0	34.7	45.3	58.7	30.7	58.7	74.7
lhu5	2.7	62.7	17.3	1.3	65.3	17.3	13.3	16.0	28.0	18.7	22.7	40.0	6.7	33.3	73.3
lhur	10.7	78.0	10.7	5.3	62.7	10.7	8.0	5.3	94.7	37.3	52.0	34.7	28.0	37.3	72.0
lhelx	4.0	70.7	18.7	8.7	52.0	18.7	0.0	1.3	100.0	13.3	13.3	92.0	24.0	84.0	56.0
C. New Orders															
hsbp	5.3	88.0	12.0	4.0	86.7	12.0	5.3	2.7	94.7	16.0	22.7	45.3	24.0	37.3	48.0
mdu82	65.3	90.7	6.7	10.7	85.3	6.7	0.0	0.0	13.3	24.0	22.7	64.0	86.7	84.0	40.0
mpcon8	0.0	13.3	10.7	0.0	14.7	10.7	4.0	2.7	100.0	8.0	22.7	1.3	1.3	1.3	65.3
mocm82	0.0	24.0	18.3	0.0	24.0	13.3	10.7	12.0	33.3	17.3	22.7	8.0	9.3	9.3	78.7
mdo82	20.0	54.7	14.7	14.7	88.0	14.7	6.7	9.3	13.3	22.7	36.0	90.7	41.3	78.7	77.3
ivpac	13.3	80.0	24.0	9.3	72.0	24.0	5.3	0.0	82.7	17.3	10.7	77.3	28.0	74.7	42.7
pmi	1.3	49.3	12.0	4.0	30.7	12.0	17.3	16.0	34.7	33.3	25.3	21.3	6.7	20.0	77.3
pmno	8.0	65.3	16.0	4.0	53.3	16.0	16.0	20.0	14.7	25.3	22.7	25.3	10.7	29.3	84.0
D. Inventories															
invmt87	12.0	81.3	8.0	4.0	64.0	8.0	2.7	6.7	17.3	22.7	24.0	37.3	16.0	33.3	88.0
invrd	21.3	69.3	21.3	10.7	73.3	21.3	24.0	30.7	12.0	5.3	4.0	66.7	28.0	64.0	66.7
invwd	20.0	97.3	13.3	2.7	57.3	13.3	17.3	6.7	100.0	28.0	32.0	65.3	44.0	62.7	58.7
ivm1d8	94.7	97.3	8.0	38.7	97.3	8.0	22.7	8.7	100.0	22.7	14.7	97.3	94.7	94.7	69.3
ivm2d8	0.0	48.7	8.7	0.0	64.0	8.7	8.0	6.7	6.7	6.7	4.0	12.0	8.0	12.0	52.0
ivm3d8	17.3	98.0	16.0	14.7	97.3	16.0	60.0	68.0	1.3	8.0	10.7	56.0	42.7	45.3	61.3

Table 2, continued

Series	Test Statistic														
	L_{all}	$L_{\mu,\beta}$	$L_{\mu,\beta(1)}$	L^r_{all}	$L^r_{\mu,\beta}$	$L^r_{\mu,\beta(1)}$	PK_{sup}	PK_{msq}	BP_{all}	BP_{β}	$BP_{\beta(1)}$	QLR	MW	APW	GC
ivmtd	13.3	49.3	16.0	0.0	28.0	16.0	8.0	10.7	84.0	36.0	41.3	33.3	18.7	33.3	72.0
ivm1d	94.7	97.3	14.7	45.3	97.3	14.7	8.0	8.0	100.0	20.0	14.7	96.0	97.3	94.7	66.7
ivm2d	2.7	40.0	5.3	0.0	56.0	5.3	2.7	4.0	4.0	5.3	1.3	13.3	5.3	10.7	60.0
ivm3d	10.7	76.0	18.7	9.3	73.3	18.7	58.7	25.3	16.0	10.7	13.3	46.7	30.7	42.7	78.7
invrd8	12.0	88.0	10.7	5.3	77.3	10.7	2.7	1.3	46.7	8.0	8.0	40.0	24.0	45.3	69.3
invwd8	10.7	84.7	10.7	2.7	86.7	10.7	1.3	1.3	98.7	14.7	17.3	49.3	29.3	52.0	46.7
E. Prices															
gmdc	10.7	52.0	9.3	5.3	41.3	9.3	1.3	0.0	6.7	8.0	5.3	100.0	92.0	100.0	32.0
pwnew	86.7	100.0	24.0	14.7	82.7	24.0	10.7	1.3	69.3	18.7	22.7	100.0	100.0	100.0	68.0
pw	62.7	100.0	20.0	0.0	20.0	20.0	12.0	8.0	100.0	25.3	29.3	100.0	100.0	100.0	36.0
pw561	50.7	100.0	52.0	8.0	42.7	52.0	42.7	5.3	100.0	12.0	17.3	100.0	64.0	100.0	32.0
pw561r	38.7	98.7	50.7	6.7	38.0	50.7	4.0	4.0	100.0	12.0	17.3	100.0	34.7	100.0	26.7
jocci	4.0	50.7	16.0	1.3	53.3	16.0	0.0	0.0	97.3	13.3	16.0	93.3	17.3	92.0	41.3
joccir	4.0	52.0	14.7	0.0	56.0	14.7	4.0	1.3	82.7	17.3	14.7	76.0	12.0	66.7	58.7
F. Interest Rates															
fyff	10.7	76.0	42.7	2.7	36.0	42.7	32.0	5.3	100.0	41.3	48.0	100.0	53.3	100.0	85.3
fygm3	10.7	77.3	46.7	2.7	38.7	46.7	77.3	5.3	100.0	49.3	48.0	98.7	41.3	97.3	70.7
fygm6	12.0	89.3	46.7	2.7	36.0	46.7	72.0	5.3	100.0	42.7	45.3	94.7	45.3	94.7	86.7
fygt1	9.3	68.0	42.7	1.3	29.3	42.7	73.3	4.0	100.0	41.3	42.7	100.0	22.7	94.7	60.0
fybaac	0.0	32.0	30.7	2.7	30.7	30.7	77.3	12.0	100.0	41.3	37.3	93.3	8.0	89.3	46.7
fygt10	0.0	41.3	41.3	0.0	24.0	41.3	81.3	6.7	100.0	32.0	36.0	89.3	10.7	82.7	48.0
cp6_gm6	1.3	28.0	22.7	1.3	10.7	22.7	4.0	1.3	100.0	49.3	46.7	100.0	10.7	97.3	40.0
g10_g1	0.0	41.3	26.7	1.3	16.0	26.7	42.7	6.7	100.0	38.7	41.3	68.0	22.7	62.7	62.7
g10_ff	9.3	66.7	26.7	0.0	10.7	26.7	0.0	0.0	100.0	38.7	38.7	100.0	40.0	100.0	80.0
bas_g10	4.0	89.3	33.3	10.7	82.7	33.3	13.3	45.3	49.3	37.3	42.7	78.7	49.3	77.3	62.7
G. Money and Credit															
fcbcuc	5.3	56.0	10.7	5.3	44.0	10.7	21.3	20.0	13.3	16.0	25.3	33.3	12.0	30.7	82.7
fcbcucy	18.7	100.0	10.7	26.7	100.0	10.7	8.0	82.7	100.0	16.0	14.7	34.7	21.3	34.7	80.0
delinqcr	18.7	84.0	10.7	12.0	78.7	10.7	2.7	0.0	5.3	8.0	9.3	32.0	33.3	30.7	60.0
cci30m	6.7	70.7	22.7	10.7	81.3	22.7	5.3	4.0	2.7	9.3	10.7	64.0	14.7	80.0	69.3
fm1d82	49.3	92.0	29.3	29.3	80.0	29.3	33.3	8.0	100.0	34.7	18.7	92.0	80.0	93.3	78.7
fm2d82	66.7	100.0	30.7	62.7	100.0	30.7	13.3	26.7	14.7	29.3	26.7	93.3	93.3	93.3	70.7
fmbase	28.0	92.0	12.0	18.7	86.7	12.0	54.7	81.3	21.3	28.0	33.3	26.7	44.0	37.3	54.7
fm1	16.0	77.3	24.0	14.7	58.7	24.0	24.0	78.7	100.0	32.0	25.3	42.7	25.3	36.0	56.0
fm2	9.3	54.7	25.3	16.0	62.7	25.3	72.0	64.0	14.7	18.7	13.3	45.3	26.7	42.7	69.3
fm3	89.3	100.0	14.7	37.3	100.0	14.7	4.0	44.0	17.3	22.7	14.7	84.0	100.0	93.3	42.7
fmbaser	44.0	57.3	13.3	20.0	54.7	13.3	45.3	13.3	16.0	25.3	24.0	89.3	73.3	96.0	88.0
H. Other Variables															
exnwt2	2.7	45.3	20.0	4.0	33.3	20.0	2.7	0.0	64.0	28.0	24.0	34.7	10.7	33.3	21.3
fspcomr	0.0	56.0	2.7	1.3	58.7	2.7	4.0	0.0	82.7	21.3	20.0	16.0	8.0	16.0	48.0
fspcom	0.0	52.0	2.7	1.3	48.0	2.7	1.3	0.0	56.0	22.7	14.7	16.0	5.3	16.0	46.7
fail	2.7	68.0	9.3	4.0	58.7	9.3	0.0	0.0	42.7	26.7	20.0	38.7	4.0	37.3	16.0
failr	2.7	68.0	10.7	4.0	57.3	10.7	0.0	0.0	45.3	29.3	20.0	38.7	4.0	38.7	14.7
gfosa	2.7	20.0	17.3	2.7	33.3	17.3	0.0	1.3	100.0	8.0	8.0	12.0	4.0	9.3	46.7
gfrsa	0.0	24.0	4.0	0.0	61.3	4.0	2.7	2.7	26.7	30.7	36.0	2.7	0.0	4.0	45.3
gfor	1.3	13.3	10.7	1.3	21.3	10.7	0.0	0.0	100.0	5.3	10.7	9.3	2.7	8.0	42.7
gfrr	0.0	36.0	2.7	0.0	89.3	2.7	0.0	0.0	25.3	29.3	34.7	1.3	0.0	4.0	48.0
hhsntn	41.3	100.0	30.7	49.3	74.7	30.7	2.7	9.3	100.0	28.0	29.3	100.0	97.3	100.0	54.7

Table 2, continued

C. Percent rejections, listed by variable used as predictor

Series	L_{all}	$L_{\mu,\beta}$	$L_{\mu,\beta(1)}$	L^F_{all}	$L^F_{\mu,\beta}$	$L^F_{\mu,\beta(1)}$	PK_{sup}	PK_{msq}	BP_{all}	BP_{β}	$BP_{\beta(1)}$	QLR	MW	APW	QC
	Test Statistic														
<u>A. Output and Sales</u>															
ip	33.3	89.3	34.7	16.0	78.7	34.7	16.0	12.0	70.7	22.7	17.3	72.0	49.3	69.3	76.0
ipxmca	40.0	85.3	18.7	25.3	81.3	18.7	16.0	21.3	73.3	49.3	44.0	68.0	56.0	69.3	82.7
gmpy	28.0	80.0	14.7	17.3	85.3	14.7	9.3	12.0	58.7	4.0	10.7	57.3	38.7	56.0	61.3
gmyxp8	38.7	72.0	13.3	28.0	69.3	13.3	16.0	8.0	52.0	2.7	1.3	64.0	49.3	66.7	52.0
rtql	14.7	68.0	20.0	4.0	53.3	20.0	20.0	10.7	82.7	12.0	25.3	58.7	25.3	60.0	66.7
gmcq	22.7	72.0	17.3	6.7	81.3	17.3	20.0	13.3	65.3	16.0	21.3	56.0	32.0	56.0	72.0
ipcd	22.7	73.3	30.7	8.0	82.7	30.7	14.7	8.0	78.7	25.3	30.7	61.3	30.7	56.0	45.3
ced87m	20.0	69.3	20.0	5.3	54.7	20.0	17.3	9.3	82.7	13.3	24.0	53.3	30.7	53.3	49.3
xci	40.0	90.7	29.3	16.0	82.7	29.3	20.0	12.0	76.0	25.3	13.3	73.3	48.0	68.0	88.0
mt82	37.3	88.0	36.0	16.0	70.7	36.0	16.0	13.3	66.7	26.7	34.7	68.0	46.7	69.3	81.3
<u>B. Employment</u>															
lpmhuadj	29.3	80.0	14.7	16.0	65.3	14.7	22.7	16.0	84.0	4.0	10.7	68.0	42.7	66.7	72.0
lphrm	21.3	73.3	26.7	12.0	62.7	26.7	29.3	29.3	70.7	40.0	41.3	73.3	46.7	73.3	62.7
lhel	32.0	78.7	24.0	16.0	65.3	24.0	22.7	18.7	62.7	21.3	38.7	62.7	42.7	65.3	92.0
lhnaps	29.3	78.7	45.3	13.3	74.7	45.3	17.3	14.7	61.3	10.7	30.7	68.0	42.7	88.0	86.7
luinc	20.0	73.3	13.3	6.7	60.0	13.3	6.7	8.0	70.7	54.7	42.7	69.3	40.0	70.7	88.0
lhu5	24.0	82.7	34.7	12.0	77.3	34.7	21.3	10.7	85.3	5.3	18.7	58.7	38.7	58.7	49.3
lhur	24.0	82.7	24.0	14.7	72.0	24.0	32.0	24.0	69.3	16.0	29.3	66.7	42.7	65.3	81.3
lhelx	33.3	80.0	53.3	22.7	82.7	53.3	16.0	24.0	68.0	33.3	20.0	80.0	65.3	80.0	85.3
<u>C. New Orders</u>															
hsbp	13.3	60.0	9.3	5.3	44.0	9.3	25.3	13.3	86.7	33.3	25.3	56.0	30.7	56.0	85.3
mdu82	21.3	84.0	29.3	10.7	74.7	29.3	18.7	16.0	82.7	13.3	13.3	49.3	33.3	48.0	57.3
mpcon8	17.3	68.0	10.7	0.0	56.0	10.7	20.0	16.0	60.0	1.3	13.3	56.0	26.7	49.3	53.3
mocm82	30.7	77.3	29.3	17.3	70.7	29.3	13.3	10.7	85.3	16.0	29.3	48.0	36.0	46.7	77.3
mdo82	26.7	78.7	28.0	10.7	85.3	28.0	13.3	10.7	69.3	18.7	21.3	60.0	37.3	53.3	80.0
ivpac	12.0	70.7	13.3	4.0	54.7	13.3	17.3	12.0	65.3	21.3	20.0	61.3	34.7	58.7	77.3
pmi	30.7	84.0	13.3	14.7	64.0	13.3	9.3	13.3	70.7	42.7	29.3	64.0	40.0	62.7	85.3
pmno	36.0	77.3	16.0	16.0	65.3	16.0	18.7	16.0	72.0	41.3	41.3	68.0	41.3	66.7	81.3
<u>D. Inventories</u>															
invmt87	22.7	74.7	14.7	8.0	58.7	14.7	16.0	13.3	65.3	5.3	9.3	54.7	28.0	53.3	32.0
invrd	26.7	68.0	10.7	12.0	62.7	10.7	26.7	10.7	86.7	10.7	10.7	57.3	28.0	54.7	45.3
invwd	22.7	89.3	5.3	12.0	54.7	5.3	6.7	8.0	86.7	26.7	30.7	53.3	30.7	54.7	38.7
ivm1d8	26.7	70.7	16.0	10.7	62.7	16.0	14.7	17.3	70.7	28.0	30.7	66.7	32.0	65.3	42.7
ivm2d8	20.0	73.3	16.0	12.0	65.3	16.0	8.0	14.7	68.0	8.0	6.7	52.0	21.3	50.7	38.7
ivm3d8	28.0	78.7	13.3	12.0	73.3	13.3	34.7	37.3	73.3	25.3	21.3	68.0	45.3	72.0	48.0
ivmtd	26.7	74.7	5.3	17.3	82.7	5.3	14.7	17.3	64.0	25.3	28.0	62.7	36.0	58.7	68.0
ivm1d	30.7	74.7	14.7	14.7	69.3	14.7	8.0	14.7	73.3	37.3	38.0	62.7	40.0	61.3	66.7
ivm2d	22.7	74.7	14.7	16.0	66.7	14.7	10.7	18.7	68.0	22.7	25.3	49.3	24.0	45.3	85.3
ivm3d	29.3	85.3	9.3	21.3	70.7	9.3	29.3	24.0	70.7	37.3	32.0	66.7	48.0	66.7	56.0
invrd8	22.7	72.0	17.3	10.7	65.3	17.3	24.0	13.3	61.3	9.3	12.0	60.0	24.0	60.0	38.7
invwd8	18.7	66.7	17.3	9.3	53.3	17.3	13.3	8.0	68.0	20.0	13.3	50.7	28.0	48.0	12.0

Table 2, continued

Series	L_{all}	$L_{\mu,\beta}$	$L_{\mu,\beta(1)}$	L^r_{all}	$L^r_{\mu,\beta}$	$L^r_{\mu,\beta(1)}$	PK_{sup}	PK_{msq}	BP_{all}	BP_{β}	$BP_{\beta(1)}$	QLR	MW	APW	GC
	Test Statistic														
E. Prices															
gmdc	28.0	78.7	12.0	22.7	74.7	12.0	29.3	25.3	74.7	46.7	42.7	50.7	36.0	50.7	50.7
punew	20.0	69.3	10.7	13.3	60.0	10.7	26.7	26.7	70.7	46.7	37.3	62.7	40.0	61.3	70.7
pw	18.7	52.0	0.0	8.0	38.7	0.0	18.7	21.3	64.0	24.0	28.0	61.3	33.3	60.0	64.0
pw561	9.3	45.3	17.3	1.3	28.0	17.3	17.3	10.7	62.7	5.3	9.3	61.3	41.3	61.3	22.7
pw561r	9.3	45.3	17.3	0.0	29.3	17.3	18.7	10.7	62.7	5.3	9.3	52.0	40.0	52.0	20.0
jocci	24.0	66.7	21.3	9.3	60.0	21.3	17.3	14.7	61.3	17.3	18.7	54.7	34.7	56.0	73.3
joccir	22.7	66.7	18.7	5.3	57.3	18.7	16.0	13.3	72.0	32.0	38.7	54.7	34.7	54.7	80.0
F. Interest Rates															
fyff	16.0	41.3	4.0	1.3	20.0	4.0	13.3	8.0	54.7	17.3	29.3	69.3	37.3	66.7	70.7
fygm3	14.7	42.7	1.3	1.3	21.3	1.3	10.7	8.0	61.3	25.3	30.7	77.3	44.0	74.7	58.7
fygm6	17.3	53.3	1.3	2.7	26.7	1.3	13.3	8.0	60.0	24.0	38.7	76.0	40.0	76.0	68.0
fygt1	16.0	50.7	2.7	1.3	29.3	2.7	12.0	8.0	65.3	25.3	34.7	73.3	41.3	77.3	77.3
fybaac	13.3	56.0	14.7	2.7	38.7	14.7	24.0	21.3	70.7	29.3	30.7	65.3	45.3	64.0	76.0
fygt10	13.3	54.7	6.7	2.7	33.3	6.7	16.0	12.0	76.0	29.3	33.3	69.3	42.7	61.3	78.7
cp6_gm6	25.3	70.7	17.3	8.0	44.0	17.3	22.7	17.3	66.7	38.7	36.0	68.0	46.7	70.7	76.0
g10_g1	30.7	73.3	34.7	13.3	56.0	34.7	25.3	36.0	65.3	26.7	24.0	68.0	57.3	72.0	89.3
g10_ff	12.0	52.0	8.0	1.3	25.3	8.0	18.7	17.3	69.3	42.7	37.3	56.0	28.0	53.3	68.0
baa_g10	22.7	72.0	8.0	8.0	60.0	8.0	17.3	20.0	81.3	61.3	34.7	73.3	53.3	74.7	81.3
G. Money and Credit															
fcbcuc	21.3	86.7	21.3	12.0	50.7	21.3	6.7	6.7	64.0	18.7	17.3	58.7	30.7	54.7	60.0
fcbcucy	25.3	65.3	26.7	13.3	65.3	26.7	9.3	5.3	64.0	9.3	9.3	62.7	32.0	57.3	21.3
delinqcr	16.0	61.3	22.7	6.7	50.7	22.7	17.3	12.0	58.7	13.3	21.3	46.7	25.3	48.0	26.7
cci30m	13.3	66.7	22.7	1.3	53.3	22.7	20.0	13.3	58.7	16.0	25.3	54.7	37.3	56.0	36.0
fm1d82	36.0	81.3	46.7	28.0	76.0	46.7	30.7	24.0	70.7	33.3	36.0	66.7	50.7	65.3	65.3
fm2d82	22.7	70.7	9.3	9.3	57.3	9.3	24.0	20.0	64.0	2.7	17.3	49.3	30.7	50.7	65.3
fmbase	26.7	80.0	33.3	16.0	72.0	33.3	29.3	26.7	72.0	21.3	13.3	56.0	46.7	60.0	28.0
fm1	26.7	78.7	29.3	12.0	69.3	29.3	38.7	32.0	68.0	25.3	32.0	61.3	53.3	61.3	53.3
fm2	18.7	72.0	1.3	12.0	69.3	1.3	9.3	10.7	81.3	1.3	13.3	58.7	45.3	60.0	68.0
fm3	30.7	82.7	16.0	17.3	76.0	16.0	5.3	5.3	88.0	22.7	21.3	68.0	54.7	68.0	42.7
fmbaser	26.7	77.3	37.3	18.7	72.0	37.3	28.0	24.0	69.3	32.0	26.7	57.3	41.3	60.0	56.0
H. Other Variables															
exnwt2	20.0	72.0	24.0	2.7	58.7	24.0	24.0	10.7	66.7	25.3	20.0	58.7	38.7	57.3	29.3
fspcomr	28.0	76.0	18.7	9.3	65.3	18.7	24.0	10.7	62.7	22.7	24.0	50.7	37.3	49.3	66.7
fspcom	26.7	76.0	17.3	9.3	64.0	17.3	22.7	13.3	64.0	20.0	25.3	50.7	38.7	50.7	65.3
fail	18.7	62.7	18.7	4.0	62.7	18.7	18.7	12.0	65.3	14.7	9.3	49.3	28.0	44.0	20.0
failr	18.7	64.0	18.7	2.7	61.3	18.7	18.7	12.0	65.3	14.7	9.3	49.3	26.7	45.3	18.7
gfosa	18.7	70.7	20.0	9.3	68.0	20.0	24.0	16.0	60.0	26.7	13.3	42.7	33.3	45.3	45.3
gfrsa	16.0	42.7	14.7	9.3	53.3	14.7	20.0	12.0	60.0	4.0	8.0	46.7	21.3	40.0	18.7
gfor	20.0	73.3	29.3	10.7	69.3	29.3	25.3	17.3	57.3	18.7	12.0	44.0	32.0	46.7	41.3
gfrr	16.0	45.3	14.7	10.7	50.7	14.7	21.3	14.7	60.0	4.0	9.3	44.0	21.3	41.3	16.0
hhsntn	14.7	82.7	13.3	5.3	42.7	13.3	21.3	12.0	72.0	50.7	29.3	58.7	29.3	56.0	49.3

Notes: All statistics are based on regression (1) with 6 lags. See the appendix for series definitions and the text for descriptions of the tests.

Table 3
Best Out-of-Sample Forecasting Models:
Percentage of cases in which model was best out-of-sample

A. Summary

	AR	RRA1	RRA2	RLSA	ATVP1	ATVP2	ATVP3	ATVP4	VAR	RRV1	RRV2	RLSV	VTVP1	VTVP2	VTVP3	VTVP4
All Models	17	4	8	15	8	4	4	0	13	1	5	12	8	3	1	0
10 BIC Sel.	15	2	8	14	10	3	3	0	12	2	8	12	8	8	1	0
Stab. rej.	13	5	6	14	10	4	5	0	10	1	6	12	8	4	1	0

B. Best models among all bivariate pairs, by variable being forecasted

	AR	RRA1	RRA2	RLSA	ATVP1	ATVP2	ATVP3	ATVP4	VAR	RRV1	RRV2	RLSV	VTVP1	VTVP2	VTVP3	VTVP4
A. Output and Sales																
ip	0	0	0	0	81	0	0	0	1	0	16	7	5	0	0	0
ipmmsa	0	0	0	81	0	0	0	0	1	1	3	15	11	7	1	0
gmpy	76	0	0	0	0	0	0	0	20	1	0	3	0	0	0	0
gmyxp8	0	0	0	47	0	0	0	0	3	4	1	44	1	0	0	0
rtql	71	0	0	0	0	0	0	0	20	0	0	8	1	0	0	0
gmeq	0	0	73	0	0	0	0	0	0	0	21	1	1	3	0	0
iped	0	0	0	0	48	0	0	0	3	0	5	17	16	0	0	0
cod87m	0	0	56	0	0	0	0	0	0	0	43	0	1	0	0	0
xei	0	0	0	88	0	0	0	0	0	0	0	12	7	4	0	0
mt82	0	0	0	85	0	0	0	0	11	0	0	20	4	0	0	0
B. Employment																
lpmbuadj	0	81	0	0	0	0	0	0	0	13	5	1	13	4	0	0
lphrm	38	0	0	0	0	0	0	0	21	0	7	13	13	4	0	0
lhel	0	40	0	0	0	0	0	0	0	3	0	16	37	4	0	0
lhmspe	0	0	0	0	51	0	0	0	0	3	1	7	12	24	3	0
luine	0	0	51	0	0	0	0	0	12	0	18	17	0	1	0	0
lhu5	52	0	0	0	0	0	0	0	12	0	1	8	20	5	0	0
lhur	0	0	0	58	0	0	0	0	4	1	0	16	18	1	0	0
lhelx	0	0	0	68	0	0	0	0	0	0	1	16	11	3	0	0
C. New Orders																
habp	0	0	67	0	0	0	0	0	0	0	5	20	5	1	1	0
mdu82	0	0	78	0	0	0	0	0	0	0	20	0	0	1	0	0
mpcon8	60	0	0	0	0	0	0	0	20	0	4	12	3	1	0	0
mocm82	51	0	0	0	0	0	0	0	28	0	0	18	5	0	0	0
mdo82	0	0	0	0	48	0	0	0	4	7	0	25	15	1	0	0
ivpac	0	0	0	87	0	0	0	0	1	1	1	5	4	0	0	0
pmi	0	0	0	65	0	0	0	0	8	0	0	21	3	1	0	0
pmno	55	0	0	0	0	0	0	0	17	0	4	18	4	1	0	0
D. Inventories																
invmt87	0	0	0	45	0	0	0	0	3	1	11	23	13	1	0	0
invrd	0	0	0	47	0	0	0	0	20	0	1	23	8	1	0	0
invmd	0	0	0	72	0	0	0	0	3	0	0	24	1	0	0	0
ivm1d8	0	63	0	0	0	0	0	0	0	4	8	3	0	15	7	0
ivm2d8	58	0	0	0	0	0	0	0	17	0	3	16	8	0	0	0
ivm3d8	0	0	0	0	41	0	0	0	0	3	13	12	25	5	0	0

Table 3, continued

	AR	RRA1	RRA2	RLSA	ATVP1	ATVP2	ATVP3	ATVP4	VAR	RRV1	RRV2	RLSV	VTVP1	VTVP2	VTVP3	VTVP4
ivm1d	0	68	0	0	0	0	0	0	0	11	12	3	0	3	4	0
ivm2d	0	0	0	44	0	0	0	0	27	0	8	18	4	0	0	0
ivm3d	0	0	0	0	32	0	0	0	0	5	27	25	7	4	0	0
invrd8	0	0	0	48	0	0	0	0	18	0	16	13	7	0	0	0
invwd8	0	0	0	0	75	0	0	0	3	0	0	8	13	1	0	0
E. Prices																
gmdc	0	0	0	0	0	73	0	0	18	0	0	1	3	4	1	1
pwnew	0	0	0	0	0	0	85	0	0	0	3	0	0	5	0	0
pw	0	0	0	0	0	0	83	0	0	0	0	0	1	0	5	0
pw381	25	0	0	0	0	0	0	0	75	0	0	0	0	0	0	0
pw361r	36	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0
jocci	65	0	0	0	0	0	0	0	15	0	0	16	4	0	0	0
joccir	0	0	0	64	0	0	0	0	11	0	0	23	3	0	0	0
F. Interest Rates																
fyff	33	0	0	0	0	0	0	0	57	0	3	7	0	0	0	0
fygm3	49	0	0	0	0	0	0	0	48	0	0	1	0	1	0	0
fygm6	43	0	0	0	0	0	0	0	38	0	1	16	0	1	0	0
fygt1	47	0	0	0	0	0	0	0	38	0	1	12	1	0	0	0
fybaac	58	0	0	0	0	0	0	0	41	0	0	0	0	0	0	0
fygt10	51	0	0	0	0	0	0	0	47	0	0	3	0	0	0	0
ep6_gm6	91	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0
g10_g1	63	0	0	0	0	0	0	0	37	0	0	0	0	0	0	0
g10_ff	0	0	44	0	0	0	0	0	36	0	8	9	3	0	0	0
bas_g10	0	65	0	0	0	0	0	0	17	1	12	3	0	1	0	0
G. Money and Credit																
fcbcuc	0	0	0	55	0	0	0	0	12	0	0	11	23	0	0	0
fcbcucy	0	0	0	0	0	56	0	0	8	0	0	3	20	11	0	1
delinqcr	0	0	0	52	0	0	0	0	0	1	0	44	3	0	0	0
cci30m	33	0	0	0	0	0	0	0	37	0	0	19	11	0	0	0
fm1d82	0	0	0	0	40	0	0	0	1	0	5	28	21	4	0	0
fm2d82	0	0	0	0	0	63	0	0	0	1	1	3	8	17	7	0
fmbase	0	0	0	0	53	0	0	0	0	0	5	23	18	0	0	0
fm1	61	0	0	0	0	0	0	0	21	0	0	11	7	0	0	0
fm2	0	0	0	0	0	48	0	0	18	0	0	1	5	28	0	0
fm3	0	0	0	0	0	0	93	0	0	0	0	0	1	3	3	0
fmbaser	0	0	0	0	0	57	0	0	0	0	3	5	13	21	0	0
H. Other Variables																
exnwt2	88	0	0	0	0	0	0	0	11	0	1	0	0	0	0	0
fspcomr	0	0	0	0	78	0	0	0	0	0	3	15	4	0	0	0
fspcom	0	0	0	0	84	0	0	0	0	0	1	13	1	0	0	0
fail	0	0	0	0	73	0	0	0	15	0	0	8	4	0	0	0
failr	75	0	0	0	0	0	0	0	18	0	0	7	3	0	0	0
gfosa	0	0	0	68	0	0	0	0	0	3	0	28	0	0	0	0
gfrsa	0	0	87	0	0	0	0	0	0	0	7	7	0	0	0	0
gfor	0	0	0	78	0	0	0	0	0	1	1	21	0	0	0	0
gfrr	0	0	84	0	0	0	0	0	0	0	11	4	1	0	0	0
hhsntn	0	0	55	0	0	0	0	0	0	7	19	20	0	0	0	0

Table 3, continued

C. Best models among those bivariate pairs with the 10-lowest in-sample BIC, by variable being forecasted

	AR	RRA1	RRA2	RLSA	ATVP1	ATVP2	ATVP3	ATVP4	VAR	RRV1	RRV2	RLSV	VTVP1	VTVP2	VTVP3	VTVP4
A. Output and Sales																
ip	0	0	0	0	70	0	0	0	0	0	10	0	0	20	0	0
ipxmca	0	0	0	70	0	0	0	0	0	0	0	0	30	0	0	0
gmpyq	60	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0
gmyxp8	0	0	0	20	0	0	0	0	20	0	0	60	0	0	0	0
rtql	40	0	0	0	0	0	0	0	20	0	0	40	0	0	0	0
gmcq	0	0	20	0	0	0	0	0	0	0	70	0	0	10	0	0
ipcd	0	0	0	0	80	0	0	0	0	0	10	10	0	0	0	0
cod87m	0	0	10	0	0	0	0	0	0	0	90	0	0	0	0	0
xci	0	0	0	60	0	0	0	0	0	0	0	10	20	10	0	0
mt82	0	0	0	60	0	0	0	0	20	0	0	20	0	0	0	0
B. Employment																
lpnbuadj	0	20	0	0	0	0	0	0	0	30	0	0	30	0	0	0
lphrm	0	0	0	0	0	0	0	0	0	0	10	20	40	30	0	0
lhel	0	30	0	0	0	0	0	0	0	0	0	10	50	10	0	0
lhnape	0	0	0	0	40	0	0	0	0	0	0	0	10	30	20	0
luinc	0	0	50	0	0	0	0	0	10	0	30	0	0	10	0	0
lhu5	40	0	0	0	0	0	0	0	0	0	0	10	30	20	0	0
lhur	0	0	0	20	0	0	0	0	0	10	0	50	20	0	0	0
lhelx	0	0	0	60	0	0	0	0	0	0	10	10	20	0	0	0
C. New Orders																
hsbp	0	0	80	0	0	0	0	0	0	0	0	10	0	10	0	0
mdu82	0	0	70	0	0	0	0	0	0	0	30	0	0	0	0	0
mpcon8	30	0	0	0	0	0	0	0	30	0	10	30	0	0	0	0
mocm82	40	0	0	0	0	0	0	0	40	0	0	20	0	0	0	0
mdo82	0	0	0	0	30	0	0	0	0	20	0	30	20	0	0	0
ivpac	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
pmi	0	0	0	50	0	0	0	0	10	0	0	30	0	10	0	0
pmno	40	0	0	0	0	0	0	0	0	0	10	30	10	10	0	0
D. Inventories																
invmt87	0	0	0	20	0	0	0	0	0	10	20	20	20	10	0	0
invrd	0	0	0	50	0	0	0	0	0	0	0	20	20	10	0	0
invwd	0	0	0	80	0	0	0	0	0	0	0	20	0	0	0	0
ivm1d8	0	20	0	0	0	0	0	0	0	10	20	0	0	40	10	0
ivm2d8	90	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
ivm3d8	0	0	0	0	70	0	0	0	0	10	0	0	10	10	0	0
ivmtd	0	0	0	20	0	0	0	0	40	0	10	20	10	0	0	0
ivm1d	0	20	0	0	0	0	0	0	0	20	30	0	0	20	10	0
ivm2d	0	0	0	80	0	0	0	0	0	0	0	10	10	0	0	0
ivm3d	0	0	0	0	0	0	0	0	0	10	60	20	0	10	0	0
invrd8	0	0	0	20	0	0	0	0	0	0	50	20	10	0	0	0
invwd8	0	0	0	0	90	0	0	0	0	0	0	0	10	0	0	0
E. Prices																
gmdc	0	0	0	0	0	90	0	0	0	0	0	0	0	10	0	0
punew	0	0	0	0	0	0	90	0	0	0	0	0	0	0	10	0
pw	0	0	0	0	0	0	80	0	0	0	0	0	10	0	10	0
pw561	20	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0
pw561r	30	0	0	0	0	0	0	0	70	0	0	0	0	0	0	0
jocci	90	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
joccir	0	0	0	70	0	0	0	0	0	0	0	20	10	0	0	0

Table 3, continued

	AR	RRA1	RRA2	RLSA	ATVP1	ATVP2	ATVP3	ATVP4	VAR	RRV1	RRV2	RLSV	VTVP1	VTVP2	VTVP3	VTVP4
F. Interest Rates																
fyff	0	0	0	0	0	0	0	0	90	0	0	10	0	0	0	0
fygm3	30	0	0	0	0	0	0	0	60	0	0	10	0	0	0	0
fygm6	30	0	0	0	0	0	0	0	50	0	0	20	0	0	0	0
fygt1	50	0	0	0	0	0	0	0	30	0	0	10	10	0	0	0
fybaac	50	0	0	0	0	0	0	0	50	0	0	0	0	0	0	0
fygt10	40	0	0	0	0	0	0	0	50	0	0	10	0	0	0	0
cp6_gm6	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
g10_g1	80	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
g10_ff	0	0	0	0	0	0	0	0	40	0	50	0	10	0	0	0
baa_g10	0	50	0	0	0	0	0	0	40	0	10	0	6	0	0	0
G. Money and Credit																
fcbcuc	0	0	0	30	0	0	0	0	20	0	0	20	30	0	0	0
fcbcucy	0	0	0	0	0	0	0	0	0	0	0	10	50	40	0	0
delinqcr	0	0	0	80	0	0	0	0	0	0	0	20	0	0	0	0
cci30m	20	0	0	0	0	0	0	0	40	0	0	20	20	0	0	0
fm1d82	0	0	0	0	40	0	0	0	0	0	0	50	10	0	0	0
fm2d82	0	0	0	0	0	50	0	0	0	0	0	0	0	40	10	0
fmbase	0	0	0	0	60	0	0	0	0	0	0	30	10	0	0	0
fm1	50	0	0	0	0	0	0	0	40	0	0	10	0	0	0	0
fm2	0	0	0	0	0	40	0	0	0	0	0	0	20	40	0	0
fm3	0	0	0	0	0	0	90	0	0	0	0	0	0	10	0	0
fmbaser	0	0	0	0	0	20	0	0	0	0	20	30	10	20	0	0
H. Other Variables																
exnwt2	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
fspcomr	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
fspcom	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
fail	0	0	0	0	90	0	0	0	0	0	0	10	0	0	0	0
failr	90	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
gfosa	0	0	0	70	0	0	0	0	0	0	0	30	0	0	0	0
gfrsa	0	0	70	0	0	0	0	0	0	0	0	30	0	0	0	0
gfor	0	0	0	80	0	0	0	0	0	0	0	20	0	0	0	0
gfrr	0	0	70	0	0	0	0	0	0	0	10	10	10	0	0	0
hhsntn	0	0	80	0	0	0	0	0	0	0	20	20	0	0	0	0

Notes: See the appendix for series descriptions. All models included six lags plus a constant. The first row of panel A shows results for all 76x76 models. Entries in the second row are the corresponding fraction, except that the set of bivariate relations is restricted from 75 to 10 for each forecasted variable, where the 10 predictors are chosen to be those with the lowest in-sample BIC for the forecasted variable at hand. Entries in the third row are for the set of bivariate relations restricted to be those for which the $L_{u,g}$ statistic is significant at the 10% level, when calculated through 1978:12. Panel B shows detailed results for all model for each variable, and panel C shows detailed results for the 10 best fitting (in-sample) models for each variable. The in-sample period was from the later of 1959:1 or the first data for which data are available, to 1978:12, and the out-of-sample period is from 1979:1 through the earlier of the final date for the series or 1993:12.

Table 4
Comparison of out-of-sample forecasts
among all 5700 bivariate bivariate combinations

Percentage of times that row forecast MSE is less than column forecast MSE

Model	AR	RRA1	RRA2	RLSA	ATVP1	ATVP2	ATVP3	ATVP4	VAR	RRV1	RRV2	RLSV	VTVP1	VTVP2	VTVP3	VTVP4
AR	--	66	55	36	45	53	79	87	65	79	60	50	54	63	79	88
RRA1	34	--	26	12	20	30	72	89	53	80	48	36	41	58	82	91
RRA2	45	74	--	20	34	62	84	93	61	85	65	51	57	70	85	93
RLSA	64	88	80	--	63	82	89	93	69	88	72	62	68	76	88	94
ATVP1	55	80	66	37	--	86	93	93	66	86	67	56	63	77	89	95
ATVP2	47	70	38	18	14	--	93	97	58	78	55	45	48	67	88	94
ATVP3	21	28	16	11	7	7	--	100	44	60	33	27	27	36	73	90
ATVP4	13	11	7	7	7	3	0	--	34	42	20	16	15	19	48	77
VAR	35	47	38	31	34	41	56	66	--	65	42	26	33	45	68	80
RRV1	21	20	15	12	14	22	40	58	35	--	9	7	6	20	61	87
RRV2	40	52	35	28	33	45	67	80	58	91	--	22	37	63	88	96
RLSV	50	64	49	38	44	55	73	84	74	93	78	--	64	78	91	96
VTVP1	46	59	43	32	37	52	73	85	67	94	63	36	--	86	95	98
VTVP2	37	42	30	24	23	33	64	81	55	80	37	22	14	--	98	99
VTVP3	21	18	15	12	11	12	27	54	32	39	11	9	5	2	--	100
VTVP4	12	8	7	6	5	6	10	23	20	13	4	4	2	1	0	--

Notes: See the notes to table 3.

Table 5

Selected Quantiles of Distributions of Mean Square Forecast Errors Relative to MSE for the AR Recursive Least Squares (RLSA) Forecast

A. Univariate Forecasts

		Percentile					
Model	Min	0.050	0.250	0.500	0.750	0.950	Max
AR	0.959	0.972	0.988	1.003	1.016	1.032	1.158
RLA1	0.967	0.987	1.006	1.017	1.022	1.049	1.076
RLA2	0.977	0.991	1.001	1.004	1.010	1.023	1.076
ATVP1	0.988	0.992	1.000	1.002	1.004	1.025	1.035
ATVP2	0.976	0.987	1.002	1.007	1.014	1.089	1.107
ATVP3	0.969	0.984	1.014	1.024	1.039	1.233	1.269
ATVP4	0.973	0.989	1.029	1.040	1.063	1.368	1.413

B. Bivariate Forecasts (All)

		Percentile											
Model	Min	0.001	0.005	0.010	0.050	0.250	0.500	0.750	0.950	0.990	0.995	0.999	Max
VAR	0.512	0.607	0.914	0.931	0.960	0.994	1.018	1.048	1.145	1.276	1.338	1.570	2.278
RLV1	0.493	0.612	0.920	0.934	0.979	1.016	1.037	1.057	1.093	1.131	1.145	1.200	1.392
RLV2	0.491	0.602	0.906	0.922	0.964	0.998	1.013	1.029	1.059	1.093	1.105	1.159	1.268
RLSV	0.486	0.598	0.903	0.919	0.961	0.992	1.006	1.019	1.047	1.078	1.094	1.152	1.263
VTVP1	0.491	0.600	0.903	0.921	0.961	0.995	1.009	1.024	1.050	1.076	1.098	1.146	1.274
VTVP2	0.499	0.602	0.905	0.923	0.964	1.001	1.020	1.038	1.094	1.123	1.133	1.183	1.320
VTVP3	0.517	0.615	0.917	0.935	0.978	1.020	1.048	1.074	1.220	1.280	1.289	1.329	1.454
VTVP4	0.534	0.639	0.934	0.951	0.996	1.043	1.077	1.114	1.331	1.428	1.446	1.485	1.574

C. Best 10 bivariate models as selected using in-sample BIC

		Percentile									
Model	Min	0.005	0.010	0.050	0.250	0.500	0.750	0.950	0.990	0.995	Max
VAR	0.512	0.886	0.901	0.940	0.980	1.017	1.068	1.195	1.519	1.947	2.278
RLV1	0.493	0.712	0.889	0.936	1.003	1.031	1.052	1.101	1.175	1.240	1.392
RLV2	0.491	0.688	0.875	0.928	0.981	1.009	1.031	1.071	1.157	1.193	1.268
RLSV	0.486	0.689	0.870	0.922	0.977	1.002	1.024	1.061	1.146	1.194	1.263
VTVP1	0.491	0.701	0.873	0.922	0.978	1.008	1.027	1.059	1.130	1.149	1.274
VTVP2	0.499	0.724	0.884	0.926	0.983	1.018	1.042	1.096	1.157	1.201	1.320
VTVP3	0.517	0.772	0.904	0.939	1.001	1.043	1.077	1.222	1.306	1.368	1.454
VTVP4	0.534	0.809	0.926	0.958	1.022	1.070	1.112	1.331	1.437	1.517	1.574

Note: Mean square forecast errors are relative to mean square error of the recursive least squares AR forecast. See the note to Table 3 for a description of the models.

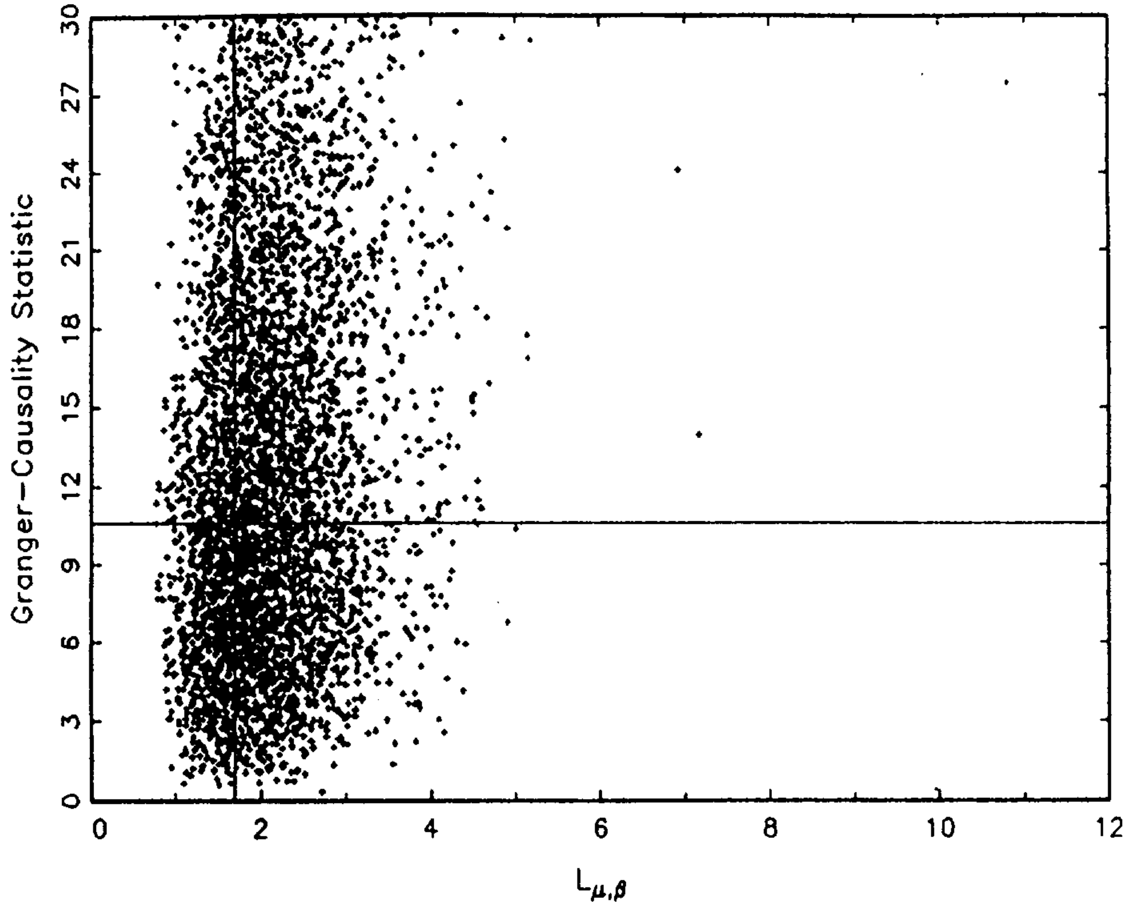

Figure 1

Scatterplot of Nyblom (1989) $L_{\mu,\beta}$ statistic vs. Granger causality F-statistic

Solid lines denote 10% critical values

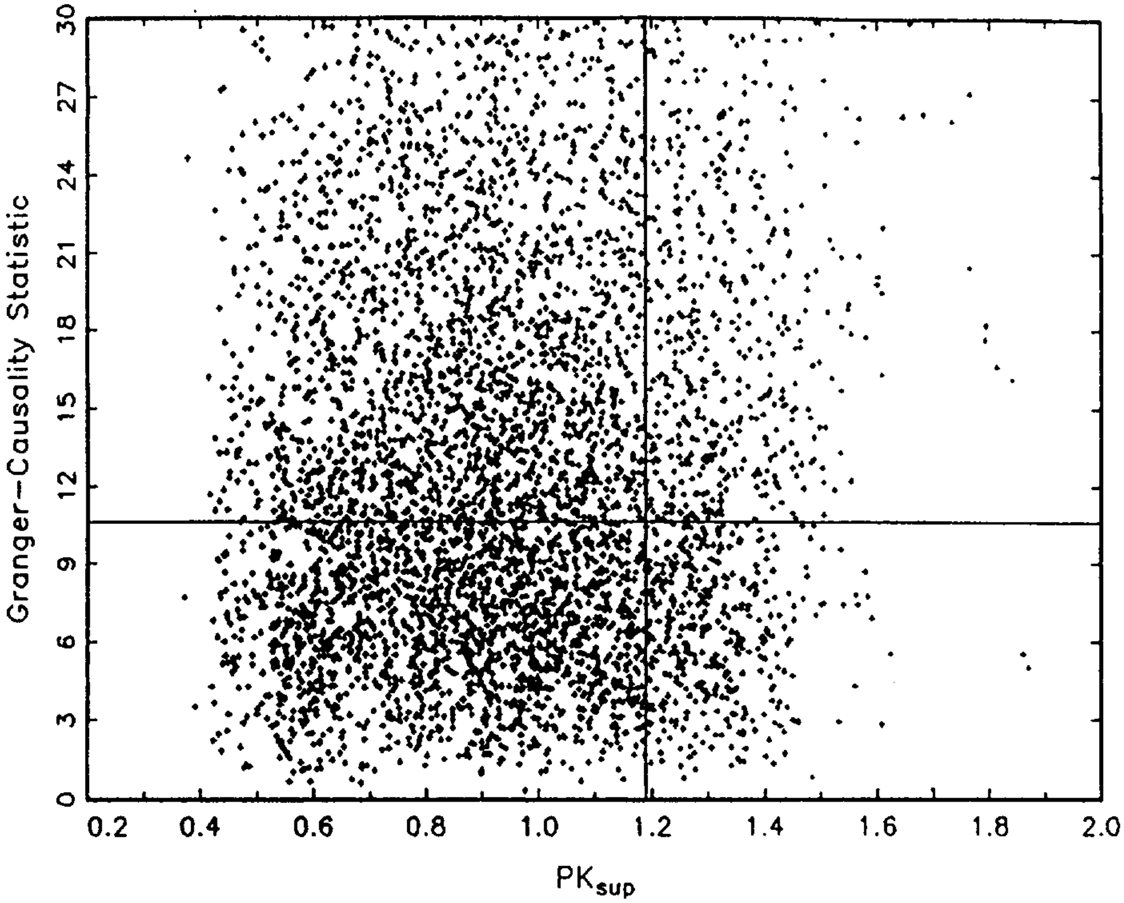

Figure 2

Scatterplot of Ploberger-Krämer (1992) PK_{sup} statistic vs. Granger causality F-statistic

Solid lines denote 10% critical values

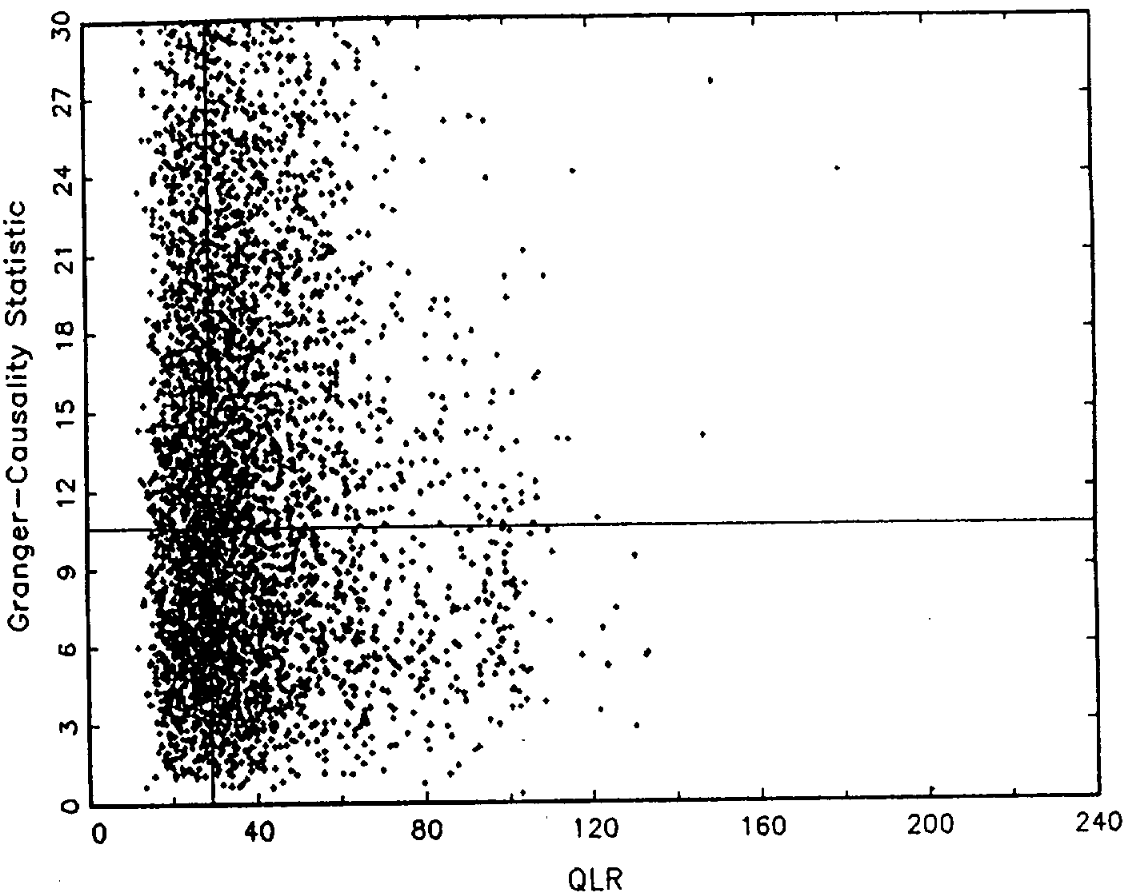

Figure 3

Scatterplot of Quandt (1960) QLR statistic vs. Granger causality F-statistic

Solid lines denote 10% critical values

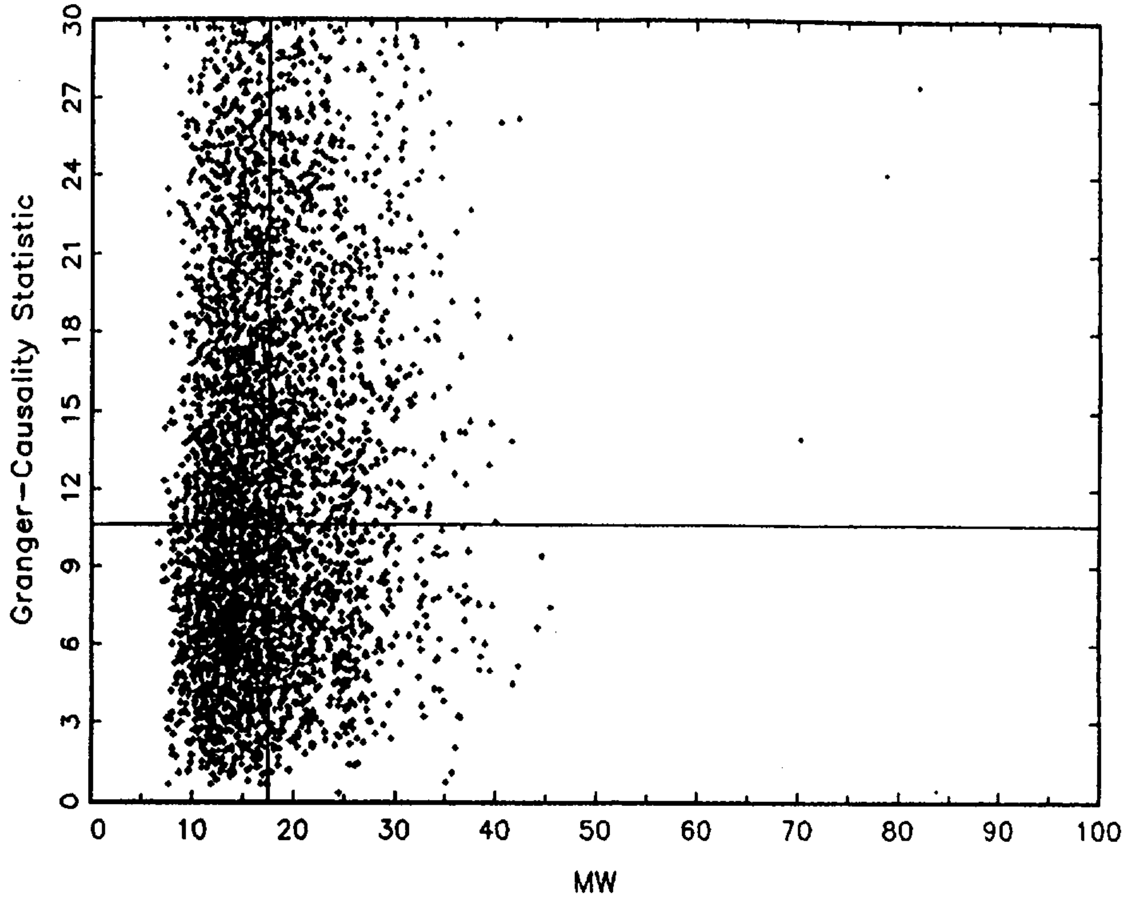

Figure 4

Scatterplot of MW statistic vs. Granger causality F-statistic

Solid lines denote 10% critical values

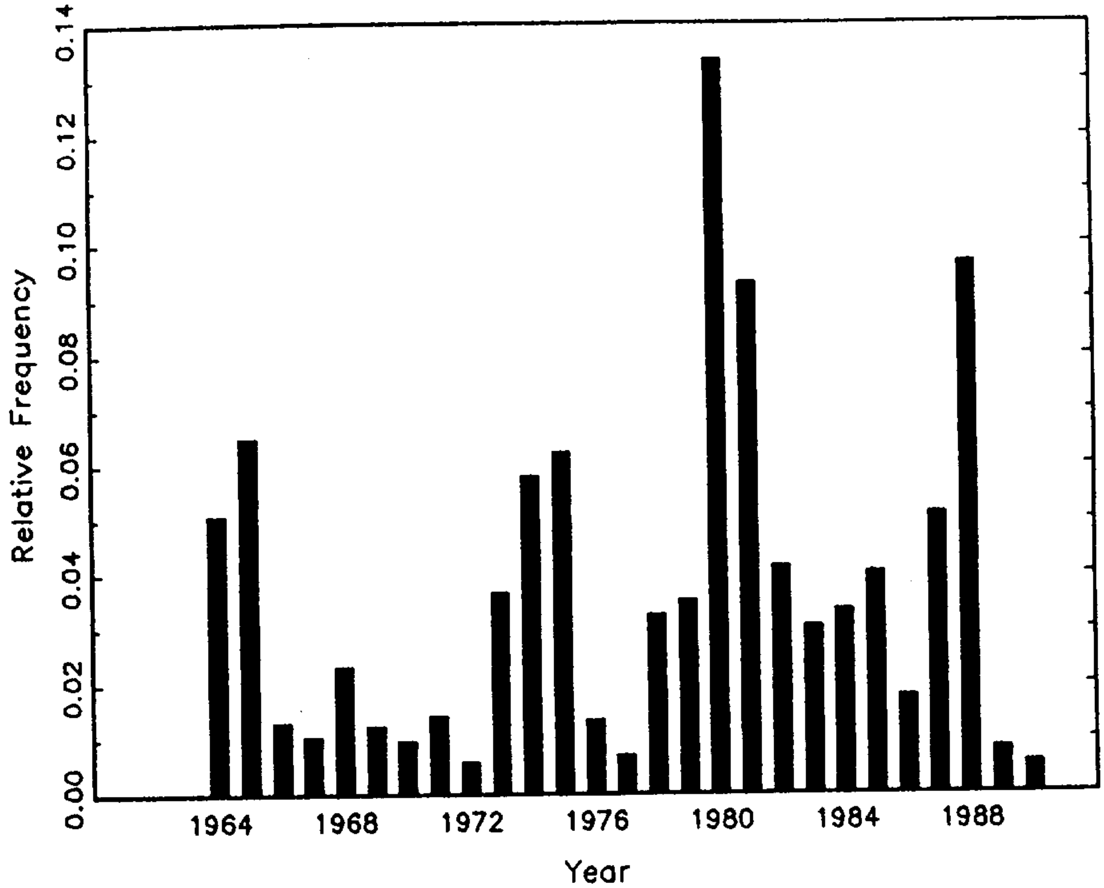

Figure 5

Histogram of break-dates from QLR statistic

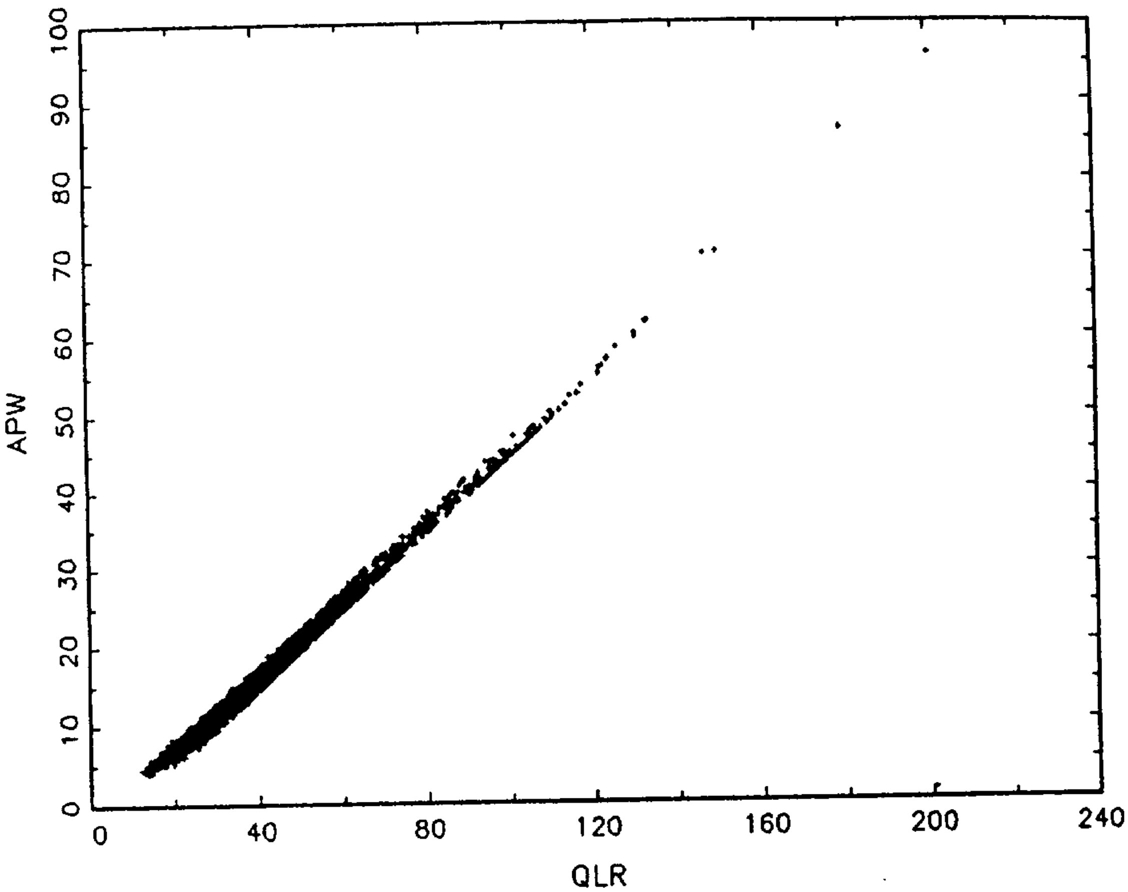

Figure 6

Scatterplot of Andrews-Ploberger (1992) APW statistic vs. QLR statistic