

Information in the Revision Process of Real-Time Datasets*

Valentina Corradi¹, Andres Fernandez², and Norman R. Swanson³

¹University of Warwick

²Rutgers University and Universidad de Los Andes

³Rutgers University and the Federal Reserve Bank of Philadelphia

April 2007
revised: June 2009

Abstract

Rationality of early release data is typically tested using linear regressions. Thus, failure to reject the null does not rule out the possibility of nonlinear dependence. This paper proposes two tests that have power against generic nonlinear alternatives. A Monte Carlo study shows that the suggested tests have good finite sample properties. Additionally, we carry out an empirical illustration using a real-time dataset for money, output, and prices. Overall, we find evidence against data rationality for output and prices, but not for money.

Keywords: bias; efficiency; generically comprehensive tests; rationality; preliminary, final, and real-time data.

JEL classification: C32, C53, E01, E37, E47.

* Valentina Corradi, Department of Economics, University of Warwick, Coventry, CV4 7AL, UK, v.corradi@warwick.ac.uk. Andres Fernandez, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, afernandez@fas-econ.rutgers.edu; and Department of Economics, Universidad de Los Andes, Cra. 1 No. 18A-10 Edificio C. Bogota, Colombia. Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. The paper was prepared for the conference entitled “Real-Time Data Analysis and Methods in Economics”, held at the Federal Reserve Bank of Philadelphia, and the authors are grateful to conference participants for many useful comments on early versions of this paper. We also wish to thank the editor, Serena Ng, two anonymous referees, Bruce Hansen, as well as seminar participants at the 2007 EC2 conference in Faro, the 2008 Royal Economic Society Conference, and the 2008 Summer Meeting of the Econometric Society for numerous useful suggestions. Corradi gratefully acknowledges ESRC grant RES-062-23-0311, and Swanson acknowledges financial support from a Rutgers University Research Council grant. The views expressed are solely our own and do not necessarily reflect the official positions or policies of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/.

1 Introduction

Testing for and assessing the rationality of early release economic data has long been a topic of considerable interest from both a theoretical and an empirical perspective. In many papers, two key questions have been addressed. The first question concerns the presence of a bias term in early releases of data. The second question concerns the predictability of later releases using information available at the time of the first release. Indeed, if subsequent data revisions are not predictable using information available at the time of the first release, then early releases are rational forecasts of final data and the revision error series only takes into account news, hence the “news hypothesis” versus the “noise hypothesis”. There is currently no consensus concerning these questions, as the empirical evidence differs across different macroeconomic series and time spans, as well as across different methodologies. For example, Mankiw, Runkle and Shapiro (1984), and Mankiw and Shapiro (1986) find that revisions in GNP are substantially news, but revisions in the money stock are better characterized as noise. Mork (1987) instead finds predictability in both GNP and the money stock. In a recent paper, Faust, Rogers and Wright (2005) find that for the U.S., later releases are only very marginally predictable, and so there is no strong evidence against the news hypothesis. On the other hand, for the U.K., Japan and Italy there is rather striking evidence in support of the noise hypothesis. Aruoba (2008) finds that for most U.S. macroeconomic series, revision errors have a positive bias, and are highly predictable using information available at the time of their first release. The issue of news versus noise is also important for forecast evaluation. For example, Clark and McCracken (2009) consider tests for comparing non-nested as well as nested forecasting models, when forecasts are produced using real-time data. They show that, under the news hypothesis, data revisions do not affect the limiting distributions of tests for predictive evaluation. On the other hand, the use of real-time data plays a crucial role whenever revisions are noisy, and effects differ depending whether one is comparing non-nested or nested models. Other recent related papers on the subject include Brodsky and Newbold (1994), Diebold and Rudebusch (1991), Kavajecz and Collins (1995), Keane and Runkle (1989,1990), Hamilton and Perez-Quiros (1996), Swanson, Ghysels, and Callan (1999), Croushore and Stark (2001, 2003), Swanson and van Dijk (2006), and the papers cited therein.

Most, if not all, of these papers have examined the issue of rationality using tests based on simple linear regressions, and such regressions are “in-sample”, in the sense that ex-ante prediction is not explicitly considered when examining data rationality. In this paper, we add to the literature by outlining two out-of-sample rationality tests which have power against generic non-linear alter-

natives. The tests we suggest are related to the conditional moment tests of Bierens (1982, 1990), de Jong (1996), and Corradi and Swanson (2002). Valid asymptotic critical values are obtained via a bootstrap procedure based on resampling functions of the entire history of the available information, along the lines of Goncalves and White (2004), and allowing for recursive estimation schemes, as in Corradi and Swanson (2007). In principle, one could augment linear regressions with non-linear terms, and still rely on simple regression-based tests. However, such tests would have power only against a fixed number of non-linear alternatives. Our tests, on the other hand, are consistent against generic alternatives, in the sense that they are able to detect any form of non-linearity.

In summary, we propose two tests. The first test can be used to determine whether preliminary data are rational, in the sense that subsequent data releases simply take “news” into account that were not available at the time of initial release. In cases where the null hypothesis fails to reject when applying the first test, we have evidence that first release data are those that should be used in out-of-sample prediction. On the other hand, if the null is rejected, we perform the second test, which is designed to determine whether non-rationality arises: (i) because of a bias in the preliminary estimate, in which case the preliminary data should be adjusted by including an estimate of the bias prior to use in out-of-sample prediction: or (ii) because available information has been used inefficiently when constructing first release data. Of course, as pointed out by Elliott, Komunjer and Timmermann (2005), if data agencies have asymmetric loss functions, then they may produce biased releases, even if available information is used efficiently.

In addition to discussing the asymptotic properties of the tests in the remaining sections of this paper, we also discuss the results of a small Monte Carlo experiment. In the experiment, we examine a variety of tests, including those discussed here, and we find that the proposed tests are reasonably sized and have adequate power. We also examine a real-time dataset for money, output, and prices, and find strong evidence against data rationality for output, in accord with recent findings by Aruoba (2008). Interestingly, for prices, rationality is not rejected when the standard linear test is implemented; but is rejected when our new test is used. For the money stock, the null of rationality fails to be rejected by all tests. The rest of the paper is organized as follows. Section 2 outlines some notation; and briefly explains the ideas behind testing for data rationality. Section 3 introduces our non-linear rationality tests; and establishes their asymptotic properties. Section 4 reports the results of a small Monte Carlo experiment. Section 5 discusses our empirical findings. Concluding remarks are given in Section 6, and all proofs are gathered in an appendix.

2 Setup

Let ${}_{t+k}X_t$ denote a variable (reported as an annualized growth rate) for which real-time data are available, where the subscript t denotes the time period to which the datum pertains, and the subscript $t+k$ denotes the time period during which the datum becomes available. In this setup, if we assume a one-month reporting lag, then first release or “preliminary” data are denoted by ${}_{t+1}X_t$. In addition, we denote fully revised or “final” data by ${}_fX_t$, where ${}_fX_t$ can be seen as the limit as $k \rightarrow \infty$ of ${}_{t+k}X_t$. Finally, data are grouped into so-called *vintages*, where the first vintage is preliminary data, the second vintage is 2^{nd} release data, and so on.

According to Muth’s (1961) definition of rational expectations, ${}_{t+1}X_t$, is a rational forecast of ${}_fX_t$ if and only if:

$${}_{t+1}X_t = E[{}_fX_t | \mathcal{F}_t^{t+1}], \quad (1)$$

where \mathcal{F}_t^{t+1} contains all information available at the time of release of ${}_{t+1}X_t$.

Standard rationality tests are based on the following regression model:

$${}_fX_t = \alpha + {}_{t+1}X_t\beta + {}_{t+1}W_t'\gamma + \varepsilon_{t+1}, \quad (2)$$

where ${}_{t+1}W_t$ is an $m \times 1$ vector of variables representing the conditioning information set available at time period $t+1$ and ε_{t+1} is an error term assumed to be uncorrelated with ${}_{t+1}X_t$ and ${}_{t+1}W_t$. The null hypothesis is $H_0 : \alpha = 0, \beta = 1, \text{ and } \gamma = 0$, and corresponds to the idea of testing for the rationality of ${}_{t+1}X_t$ for ${}_fX_t$, by finding out whether the conditioning information in ${}_{t+1}W_t$, available to the data issuing agency at the time of first release, has been efficiently used.

For money stock data, Mankiw, Runkle, and Shapiro (1984) find evidence against the null in (2), suggesting that preliminary money stock announcements are not rational. On the other hand, Kavajecz and Collins (1995) find that seasonally unadjusted money announcements are rational while adjusted ones are not. For GDP data, Mankiw and Shapiro (1986) find little evidence against the null hypothesis of rationality, while Mork (1987) and Rathjens and Robins (1995) find evidence of non-rationality, particularly in the form of prediction bias (i.e. $\alpha \neq 0$ in (2)). Keane and Runkle (1990) examine the rationality of survey price forecasts rather than preliminary (or real-time) data, using the novel approach of constructing panels of real-time survey predictions. This allows them to avoid aggregation bias, for example, and may be one of the reasons why they find evidence supporting rationality, even though previous studies focusing on price forecasts had found evidence to the contrary. Swanson and van Dijk (2006) consider the entire revision history

for each variable, and hence discuss the “timing” of data rationality by generalizing (2) as follows:

$${}_{t+k}X_t - {}_{t+1}X_t = \alpha + {}_{t+1}X_t\beta + {}_{t+1}W_t'\gamma + \varepsilon_{t+k}, \quad (3)$$

where $k = 1, 2, \dots$ defines the release (or vintage) of data. The objective is to assess whether there is information in the revision error between periods $t + k$ and $t + 1$ that could have been predicted when the initial estimate, ${}_{t+1}X_t$, was formed. They find that data rationality is most prevalent after 3 to 4 months, for unadjusted industrial production and producer prices.

3 Consistent Out-of-Sample Tests for Rationality

3.1 The Framework

Consider using linear regression analysis in order to test for rationality, in which case failure to reject the null equates with an absence of linear correlation between the revision error and information available at the time of first data release. It follows that these tests do not necessarily detect non-linear dependence. Our objective is to provide a test for rationality which is consistent against generic non-linear alternatives. In other words, we propose a test that is able to detect any form of dependence between the revision error and information available at the time of the first data release. This is accomplished by constructing conditional moment tests which employ an infinite number of moment conditions (see, e.g., Bierens (1982, 1990), Bierens and Ploberger (1997), de Jong (1996), Stinchcombe and White (1998), Corradi and Swanson (2002)). To set notation, let ${}_{t+2}u_t^{t+1} = {}_{t+2}X_t - {}_{t+1}X_t$, and ${}_{t+1}W_t = ({}_{t+1}X_t, {}_{t+1}u_{t-1}^t)$, where ${}_{t+1}u_{t-1}^t = {}_{t+1}X_{t-1} - {}_tX_{t-1}$. Thus, ${}_{t+2}u_t^{t+1}$ and ${}_{t+1}u_{t-1}^t$ denote the errors between the 2nd and 1st vintages at times $t + 2$ and $t + 1$, respectively. Also, let $\mathcal{F}_t^{t+1} = \sigma({}_{s+1}W_s; 1 \leq s \leq t)$, so that \mathcal{F}_t^{t+1} contains information available at the time of first release, assuming a one month lag before the first datum becomes available. All of our results generalize immediately to the case where ${}_{t+1}W_t$ contains information other than ${}_{t+1}X_t$ and ${}_{t+1}u_{t-1}^t$. Furthermore, even though our discussion focuses primarily on ${}_{t+2}u_t^{t+1}$, our results generalize immediately to ${}_{t+k}u_t^{t+j}$ for $k \geq 2$, $j \geq 1$, and $k > j$.

In the sequel, consider testing the following hypotheses against their respective negations:

$$H_{0,1} : E({}_{t+2}u_t^{t+1} | \mathcal{F}_t^{t+1}) = 0 \text{ a.s. and}$$

$$H_{0,2} : E({}_{t+2}u_t^{t+1} | \mathcal{F}_t^{t+1}) = E({}_{t+2}u_t^{t+1}) \text{ a.s.}$$

Hypothesis $H_{0,1}$ is the null hypothesis that the first release is rational and unbiased, as the revision error in this case is a martingale difference sequence, adapted to the filtration generated by the entire history of past revision errors and past values of the variable to be predicted. This is consistent with the “news” version of rationality, according to which subsequent data revisions only take into

account news that was not available at the time of the first release. Thus, if we fail to reject $H_{0,1}$, it means that the first data release already incorporates all available information at the current time.

In principle, one might imagine forming a joint test for the null hypothesis that $E(t_{+k}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0$, for $k = 2, \dots, \bar{k}$, where $t_{+k}u_t^{t+1}$ denotes the revision error between the k^{th} and the 1^{st} releases. However, under the null of rationality, for $k > 2$, $t_{+k}u_t^{t+1}$ is perfectly correlated with $t_{+2}u_t^{t+1}$, as:

$$E(t_{+k}u_t^{t+1} | t_{+2}u_t^{t+1}) = E\left(\left(t_{+2}u_t^{t+1}\right)^2\right) + \sum_{j=3}^k E(t_{+j}u_t^{t+1} | t_{+2}u_t^{t+1}) = E\left(\left(t_{+2}u_t^{t+1}\right)^2\right),$$

which in turn follows because the revision error is uncorrelated, under the null hypothesis. In this sense, by considering additional revision errors, one gains no further information. Therefore, a test statistic for the joint null $E(t_{+k}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0$, for $k = 2, \dots, \bar{k}$ will be characterized by a degenerate limiting distribution. Broadly speaking, a joint test for $E(t_{+k}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0$, $k = 1, \dots, \bar{k}$ will converge to a \bar{k} -dimensional Gaussian process, with a covariance matrix of rank one. On the other hand, one can certainly use $t_{+k}u_t^{t+k-1}$, $k \geq 1$ in place of $t_{+2}u_t^{t+1}$ in $H_{0,1}$. Indeed, by sequentially testing $H_{0,1}$ using increasing values of k , one can estimate how many releases we need to achieve rationality. The only drawback is that, because of sequential test bias, there is a positive probability of overestimating the number of periods it takes before releases become efficient.

Hypothesis $H_{0,2}$ also forms the basis for a rationality test, because rationality implies that the revision error is independent of any function which is measurable in terms of information available at time $t + 1$. Nevertheless, the first release may be a biased estimator of the second release. In this sense, the first release would be unconditionally biased. Unconditional bias may arise due to the fact that the statistical reporting agency produces releases according to an asymmetric loss function (see, e.g., Elliott, Komunjer and Timmermann (2005)). For example, there may be a preference for pessimistic or “conservative” first releases, followed by more optimistic releases (see, e.g., Swanson and Van Dijk (2006)).

In summary, our first objective is to provide a test for $H_{0,1}$, which is consistent against all possible deviations from the null. Now, failure to reject $H_{0,1}$ would clearly suggest that one should use first release data for out-of-sample prediction. On the other hand, if $H_{0,1}$ is rejected, one remains with the problem of ascertaining the cause of the rejection. A logical next step would then be to construct a statistic for testing $H_{0,2}$ against its negation. If the null hypothesis fails to reject, then there is unconditional bias, but there is no issue of the inefficient use of available information. In this case, one can use the preliminary release plus the estimated mean of $t_{+2}u_t^{t+1}$, in prediction and policy applications. Needless to say, if $H_{0,2}$ is also rejected, then we have evidence that early

releases are both biased and inefficient forecasts of later releases. In this case, first releases are not appropriate for constructing out-of-sample forecasts of later releases.

3.2 Test Statistics and Assumptions

Bierens (Theorem 1, (1990)) shows that if $E(t_{+2}u_t^{t+1}|_{t+1}W_t) \neq 0$, then $E(t_{+2}u_t^{t+1} \exp_{t+1} W_t' \gamma) \neq 0$, for all $\gamma \in \Gamma$, except a subset of zero Lebesgue measure. Stinchcombe and White (1998) show that if $w(t_{+1}W_t', \gamma)$ is a generically comprehensive function, then whenever $\Pr(E(t_{+2}u_t^{t+1}|_{t+1}W_t) = 0) < 1$, $E(t_{+2}u_t^{t+1}w(t_{+1}W_t', \gamma)) \neq 0$, for all $\gamma \in \Gamma$, except a subset of zero Lebesgue measure. In addition to the Bierens exponential function, the class of generically comprehensive functions includes the logistic function, cumulative distribution functions, and other sigmoidal functions. Suppose that $t_{+1}W_t$ is a q -dimensional vector, then examples of $w(t_{+1}W_t', \gamma)$ include: $w(t_{+1}W_t', \gamma) = \exp(\sum_{i=1}^q \gamma_i \Phi(W_{i,t}))$ and $w(t_{+1}W_t', \gamma) = 1/(1 + \exp(c - \sum_{i=1}^q \gamma_i \Phi(W_{i,t})))$, with $c \neq 0$ and with Φ being a measurable one-to-one mapping from \Re^q to a bounded subset of \Re^q . In our context, we want to test whether the revision error is independent of the entire history. Thus, we need to ensure that if $\Pr(E(t_{+2}u_t^{t+1}|_{t+1}W_t, tW_{t-1}, \dots, 2W_1) = 0) < 1$, then $E(t_{+2}u_t^{t+1}w(\sum_{i=0}^{t-1} t_{+1-i}W_{t-i}'\gamma_i)) \neq 0$, for all $\gamma_i \in \Gamma$. In order to test $H_{0,1} : E(t_{+2}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0, a.s.$ versus $H_{A,1} : \Pr(E(t_{+2}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0) < 1$, we shall rely on the following statistic suggested by de Jong (1996):

$$M_{1,T} = \sup_{\gamma \in \Gamma} |m_{1,T}(\gamma)|, \quad (4)$$

where in our context: $m_{1,T}(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} t_{+2}u_t^{t+1}w(\sum_{j=0}^{t-1} \gamma_j' \Phi(t_{+1-j}W_{t-j}))$, and

$$\Gamma = \{\gamma_j : a_j \leq \gamma_j \leq b_j, j = 1, 2; |a_j|, |b_j| \leq B j^{-\kappa}, \kappa \geq 2\}. \quad (5)$$

As shown in Lemma 1 of de Jong, $(\Gamma, \|\gamma - \gamma'\|)$ is a compact metric space, with $\|\gamma - \gamma'\| = (\sum_{j=1}^{\infty} j^\kappa |\gamma_j - \gamma_j'|^2)^{1/2}$, where $|\cdot|$ denotes the Euclidean norm. In practice, one can allow for $\kappa = 2$ and choose $a_j = a(j+1)^{-2}$ and $b_j = b(j+1)^{-2}$, where a and b belong to some compact set in \Re^q . It is immediate to see that the weight attached to past observations decreases over time. Indeed, as stated in the assumptions below, the revision error is a short memory process, and therefore it is “independent” of its distant past, under both hypotheses.

If $H_{0,1}$ is not rejected, then one can conclude that the revision error is unpredictable, and thus the first release data already incorporate available information in an efficient way. On the other hand, if $H_{0,1}$ is rejected, then it is important to distinguish between the case of inefficiency and (unconditional) bias. Thus, whenever $H_{0,1}$ is rejected, it remains to consider $H_{0,2} : E(t_{+2}u_t^{t+1}|\mathcal{F}_t^{t+1}) = E(t_{+2}u_t^{t+1}), a.s.$ versus $H_{A,2} : \Pr(E(t_{+2}u_t^{t+1}|\mathcal{F}_t^{t+1}) = E(t_{+2}u_t^{t+1})) < 1$.

Now, note that $m_{1,T}(\gamma)$ does not contain estimated parameters, so that there is no difference between in-sample and out-of-sample tests, when testing $H_{0,1}$ versus $H_{A,1}$. This is no longer true when testing $H_{0,2}$ versus $H_{A,2}$, as implementation of the test requires the computation of the deviation from zero of the revision error. In this case, we thus propose splitting the sample T , such that $T = R + P$, where only the last P observations are used for testing rationality. The mean is estimated recursively as:

$$\hat{\mu}_t = \frac{1}{t} \sum_{j=1}^{t-2} j+2 u_j^{j+1}, \text{ for } t = R + 1, \dots, R + P.$$

It follows that the statistic of interest is: $M_{2,P} = \sup_{\gamma \in \Gamma} |m_{2,P}(\gamma)|$, where

$$m_{2,P}(\gamma) = \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (t+2 u_t^{t+1} - \hat{\mu}_t) w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j W_{t-j}) \right).$$

In the sequel, we require the following assumptions. We state the assumptions for the case in which ${}_{t+1}W_t \in \mathcal{R}^q$, so that if we set ${}_{t+1}W_t = ({}_{t+1}X_{t,t+1} u_t^{t+1})$, $q = 2$.

Assumption A1: (i) The weights γ_j are defined as in (5). (ii) ${}_{t+1-j}W_{t-j}$ is strictly stationary and strong mixing sequence with size $-2r/(r-2)$, $r > 2$, (iii) $E \left(({}_{t+2}u_t^{t+1})^{2r} \right) < \infty$, and $E \left(w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j W_{t-j}) \right)^{2r} \right) < \infty$, uniformly in γ , for $r > 2$, and $E \left(\left({}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j W_{t-j}) \right) \right)^2 \right) > 0$, (iv) $\gamma'_j \Phi(t+1-j W_{t-j}) \in \mathcal{U} = (\underline{u}, \bar{u})$, with $-\infty < \underline{u} < \bar{u} < \infty$, (v) $w(\cdot, \gamma)$ is a generically comprehensive function, such that $\sup_{\gamma \in \Gamma} \|w(\cdot, \gamma)\|_{s,2,\mathcal{U}} < \infty$, for $s > (q+1)/2$, where for each $\gamma \in \Gamma$, $\|w(\cdot, \gamma)\|_{s,2,\mathcal{U}} = \left(\sum_{|\alpha| \leq s} \int_{\mathcal{U}} |D^\alpha w(x, \gamma)|^2 dx \right)^{1/2}$, with $D^\alpha w(x, \gamma) = \frac{\partial^{|\alpha|} w(x, \gamma)}{\partial x_1 \times \dots \times \partial x_q}$ and $|\alpha| = \sum_{l=1}^q \alpha_l$.

Remark 1: We can immediately see that **A1(iv)** is satisfied if we choose Φ to be the atan function, as is customary in the consistent test literature (see, e.g., Bierens (1982, 1990)).

Remark 2: The norm $\|w(\cdot)\|_{s,2,\mathcal{U}}$ defined in **A1(v)** is known as the Sobolev norm, and it ensures that the function has derivatives up to order s which are square integrable. This condition is satisfied for all s , by most of the test functions commonly used in the consistent test literature, such as the exponential, the logistic, and cumulative distribution functions.

Our test statistics, $M_{1,T}$ and $M_{2,P}$, are based on a continuum of moment conditions, thus ensuring power in all directions. As we shall see in the next sections, the “price to pay” is that the limiting distributions are suprema over Gaussian processes, and thus there are no “ready to use” critical values. Hence, the need for bootstrap critical values. One alternative to our tests is given by the neural network test of Lee, White and Granger (1993). They suggest selecting

p , p finite, random draws for γ from a uniform distribution, for example, and then constructing $w(W'_{t-1}\gamma_1), \dots, w(W'_{t-p}\gamma_p)$ and considering the first $p^* < p$ principal components thereof. Then, one can proceed by regressing ${}_{t+2}u_t^{t+1}$ on the p^* principal components and constructing TR^2 , using the R^2 from the latter regression. Under the null, and under the additional assumption of conditional homoskedasticity, $TR^2 \xrightarrow{d} \chi_{p^*}^2$.

3.3 Asymptotics

We now state the limiting behavior of the two statistics suggested above.

Theorem 1: *Let Assumption A1 hold. Then: (i) Under $H_{0,1}$, $M_{1,T} \xrightarrow{d} \sup_{\gamma} |m_1(\gamma)|$, where $m_1(\gamma)$ is a zero mean Gaussian process with covariance kernel given by:*

$$C(\gamma_1, \gamma_2) = E \left(({}_{t+2}u_t^{t+1})^2 w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi(t+1-j) W_{t-j} \right) w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi(t+1-j) W_{t-j} \right) \right).$$

(ii) Under $H_{A,1}$, there exist an $\varepsilon > 0$, such that $\Pr \left(\frac{1}{\sqrt{T}} M_{1,T} > \varepsilon \right) \rightarrow 1$.

From the statement in part (i) of Theorem 1, it is immediate to see that the covariance kernel does not contain cross terms that capture the correlation between ${}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right)$ and ${}_{s+2}u_s^{s+1} w \left(\sum_{j=0}^{s-1} \gamma'_j \Phi(s+1-j) W_{s-j} \right)$, for all $s \neq t$. This is because the revision error is a martingale difference sequence under the null hypothesis. Whenever $w(\cdot)$ is an exponential function, the statement in Theorem 1 corresponds to that in Theorem 4 of de Jong (1996). Given A1, the proof of the theorem above follows from the empirical process central limit theorem of Andrews (1991), for heterogeneous near epoch dependent (NED) arrays. Hansen (1996b) provides an empirical process central limit theorem for the case in which A1(iv) fails to hold and $\gamma'_j \Phi(t+1-j) W_{t-j}$ is unbounded, while de Jong (1996) requires $\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j}$ to be v -stable on $t+1-j) W_{t-j}$ instead of near epoch dependent, where v -stability is a slightly weaker concept than NED; indeed his proof is much more involved, though he confines his attention to the exponential function. For a discussion of the relationship between NED and v -stability see Pötscher and Prucha (1997, Ch. 6.2).

As mentioned above, if we reject $H_{0,1}$, then we want to be able to distinguish between the case of unconditional biasedness and the case of inefficient use of information. The theorem below establishes the asymptotic properties of $M_{2,P}$.

Theorem 2: *Let Assumptions A1 hold. Then, as $R, P \rightarrow \infty$, $P/R \rightarrow \pi$, $0 \leq \pi < \infty$,*

(i) Under $H_{0,2}$, $M_{2,P} \xrightarrow{d} \sup_{\gamma} |m_2(\gamma)|$, where $m_2(\gamma)$ is a zero mean Gaussian process with covariance kernel given by:

$$\begin{aligned}
& C(\gamma_1, \gamma_2) \\
&= E \left(\left(({}_{t+2}u_t^{t+1}) - \mu_u \right)^2 w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi(t+1-j) W_{t-j} \right) w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi(t+1-j) W_{t-j} \right) \right) \\
&+ 2\Pi\sigma_u^2 E \left(w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi(t+1-j) W_{t-j} \right) \right) E \left(w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi(t+1-j) W_{t-j} \right) \right) \\
&- \Pi E \left(\left(({}_{t+2}u_t^{t+1}) - \mu_u \right)^2 w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi(t+1-j) W_{t-j} \right) \right) E \left(w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi(t+1-j) W_{t-j} \right) \right) \\
&- \Pi E \left(\left(({}_{t+2}u_t^{t+1}) - \mu_u \right)^2 w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi(t+1-j) W_{t-j} \right) \right) E \left(w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi(t+1-j) W_{t-j} \right) \right), \quad (6)
\end{aligned}$$

where $\mu_u = E({}_{t+2}u_t^{t+1})$, and $\sigma_u^2 = \text{Var}({}_{t+2}u_t^{t+1})$, and $\Pi = 1 - \pi^{-1} \ln(1 + \pi)$.

(ii) Under $H_{A,2}$, there exist an $\varepsilon > 0$, such that $\Pr\left(\frac{1}{\sqrt{P}}M_{2,P} > \varepsilon\right) \rightarrow 1$.

The difference between the limiting distributions in Theorems 1 and 2 is apparent in the last three lines of (6), which reflect the contribution of the recursively estimated sample mean. Note that if $\pi = 0$, then $\Pi = 0$ and thus the last three terms on the right-hand side of (6) are zero.

The limiting distributions in both theorems depend on the nuisance parameter $\gamma \in \Gamma$, and thus standard critical values are not available. In this case, asymptotically valid critical values for $M_{1,T}$ can be obtained via the conditional p-value approach of Hansen (1996a) and de Jong (1996), and asymptotically valid critical values for $M_{2,P}$ can be obtained using the generalization of the conditional p-value approach due to Corradi and Swanson (2002), which takes into proper account the contribution of parameter estimation error. A drawback of this approach is that it requires direct estimation of the contribution of parameter estimation error to the covariance kernel. On the other hand, the construction of bootstrap critical values does not require direct estimation of parameter estimation components. Hence, in the next section we suggest an easy-way-to-implement bootstrap procedure, and we establish first order validity thereof.

3.4 Bootstrap Critical Values

First order validity of the block bootstrap in the context of recursive estimation is established in Corradi and Swanson (2007), for the case in which the test function, w , depends only on a finite number of lags. Intuitively, in the current context, if we resample the data, say ${}_{t+2}u_t^{t+1}$ and ${}_{t+1}W_t$, then the statistic computed using the resampled observations cannot mimic the behavior of the

original statistic, simply because the correct temporal ordering is no longer preserved. A scenario analogous to this one arises in the context of the conditional variance of a GARCH model, which is a near epoch dependent map on all of the history of the process. Goncalves and White (2004), in order to obtain QMLE estimation of GARCH parameters, suggest resampling (blocks of) the likelihood. More recently, Corradi and Iglesias (2008) show higher order refinement for QMLE GARCH estimation based on a similar resampling methodology. In the sequel, we use the same idea and jointly resample $({}_{t+2}u_t^{t+1}, w_t(\gamma))$, where $w_t(\gamma) = w\left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j)W_{t-j}\right)$. Under the null, ${}_{t+2}u_t^{t+1}w_t(\gamma)$ is a martingale difference sequence. Therefore, first order asymptotic validity of the bootstrap statistic can be achieved by simply resampling blocks of length one, as in the *iid* case. On the other hand, in order to achieve higher order refinement, one has to use the block bootstrap with a block size increasing with the sample, even in the case of martingale difference sequences. This is because for refinements it no longer suffices to merely mimic the first two sample moments when showing asymptotic validity, and the martingale difference assumption does not help for higher moments (see, e.g., Andrews (2002)). However, our statistics are not pivotal, because of the presence of the nuisance parameters, γ , that are unidentified under the null, and thus we cannot obtain higher order bootstrap refinements. For this reason, when considering $M_{1,T}$ it suffices to make $T-2$ independent draws from $({}_3u_1^2, w_1(\gamma), \dots, {}_T u_{T-2}^{T-1}, w_{T-2}(\gamma))$. Let $I_i, i = 1, \dots, T-2$ be $T-2$ independent discrete uniform variates on $1, \dots, T-2$; it follows that the resampled series $({}_3u_1^{*2}, w_1^*(\gamma), \dots, {}_T u_{T-2}^{*T-1}, w_{T-2}^*(\gamma))$ can be written as $({}_{I_1+2}u_{I_1}^{I_1+1}, w_{I_1}(\gamma), \dots, {}_{I_T} u_{I_{T-2}}^{I_{T-1}}, w_{I_{T-2}}(\gamma))$. Thus, for all $i, w_{I_i}(\gamma) = \left(\sum_{j=0}^{I_i-1} \gamma'_j \Phi(I_i+1-j)W_{I_i-j}\right)$ which is equal to $w\left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j)W_{t-j}\right)$, for $t = 1, \dots, T-2$, with equal probability $1/(T-2)$. Note that for each bootstrap replication, we use the same set of resampled values across all $\gamma \in \Gamma$. Hereafter, E^* and Var^* denote the mean and the variance with respect to the law governing the bootstrap samples, P^* ; furthermore $O_{p^*}(1)$ and $o_{p^*}(1)$ denote terms that are asymptotically bounded and approach zero in probability, respectively, according to P^* .

The bootstrap analog of $M_{1,T}$, say $M_{1,T}^*$, is then defined to be:

$$M_{1,T}^* = \sup_{\gamma \in \Gamma} |m_{1,T}^*(\gamma)|,$$

where

$$m_{1,T}^*(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} ({}_{t+2}u_t^{*t+1}w_t^*(\gamma) - {}_{t+2}u_t^{t+1}w_t(\gamma)).$$

We next turn to constructing the bootstrap analog of $M_{2,P}$, say $M_{2,P}^*$. First, construct the

bootstrap analog of $\hat{\mu}_t$, say $\hat{\mu}_t^*$, which is defined as:

$$\hat{\mu}_t^* = \arg \min_{\mu} \left(\frac{1}{t} \sum_{j=0}^t \left(j+2u_j^{*j+1} - \mu \right)^2 + 2\mu \frac{1}{T} \sum_{i=1}^{T-2} \left(i+2u_i^{i+1} - \hat{\mu}_t \right) \right). \quad (7)$$

That is,

$$\hat{\mu}_t^* = \frac{1}{t} \sum_{j=0}^t \left(j+2u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} i+2u_i^{i+1} \right) + \hat{\mu}_t \text{ for } t = R, \dots, R+P-2. \quad (8)$$

From (8), we can immediately see that $\hat{\mu}_t^*$ is not the exact bootstrap counterpart of $\hat{\mu}_t$. This is due to the recentering term in (7), which is necessary to ensure that $E^*(\hat{\mu}_t^*) = \hat{\mu}_t$.

Next, construct:

$$M_{2,P}^* = \sup_{\gamma \in \Gamma} |m_{2,P}^*(\gamma)|,$$

where

$$m_{2,P}^*(\gamma) = \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left((t+2u_t^{*t+1} - \hat{\mu}_t^*) w_t^*(\gamma) - \frac{1}{T} \sum_{i=1}^{T-2} (i+2u_i^{i+1} - \hat{\mu}_t) w_i(\gamma) \right), \quad (9)$$

and where $t+2u_t^{*t+1}$ and $w_t^*(\gamma)$ are the same resampled series used in the construction of $M_{1,T}^*$ (except that we are no longer using w_1^*, \dots, w_{R-1}^*). Note that we need to recenter the above bootstrap statistic around the full sample mean, as observations have been resampled from the full sample.

Theorem 3: *Let Assumption A1 hold. Then:*

(i)

$$P \left[\omega : \sup_{x \in \mathbb{R}} \left| P^* \left(\sup_{\gamma \in \Gamma} |m_{1,T}^*(\gamma)| \leq x \right) - P \left(\sup_{\gamma \in \Gamma} |m_{1,T}^\mu(\gamma)| \leq x \right) \right| > \varepsilon \right] \rightarrow 0,$$

where $m_{1,T}^\mu(\gamma) = \sqrt{T} (m_{1,T}(\gamma) - E(m_{1,T}(\gamma)))$.

(ii) As $R, P \rightarrow \infty$, $P/R \rightarrow \pi$, $0 \leq \pi < \infty$, then:

$$P \left[\omega : \sup_{x \in \mathbb{R}} \left| P^* \left(\sup_{\gamma \in \Gamma} |m_{2,P}^*(\gamma)| \leq x \right) - P \left(\sup_{\gamma \in \Gamma} |m_{2,P}^\mu(\gamma)| \leq x \right) \right| > \varepsilon \right] \rightarrow 0,$$

where $m_{2,P}^\mu(\gamma) = \sqrt{P} (m_{2,P}(\gamma) - E(m_{2,P}(\gamma)))$.

Note that if $\pi = 0$, then the contribution of parameter estimation error is asymptotically negligible and one can simply replace $\hat{\mu}_t^*$ with $\hat{\mu}_t$. The above results suggest proceeding in the following manner. For any bootstrap replication, compute the bootstrap statistic, $M_{1,T}^*$ ($M_{2,P}^*$). Perform B bootstrap replications (B large) and compute the quantiles of the empirical distribution of the B bootstrap statistics. Reject $H_{0,1}$ ($H_{0,2}$) if $M_{1,T}$ ($M_{2,P}$) is greater than the $(1 - \alpha)^{th}$ -percentile of the corresponding bootstrap distribution. Otherwise, do not reject. Now, for all samples except a set with probability measure approaching zero, $M_{1,T}$ ($M_{2,P}$) has the same limiting

distribution as the corresponding bootstrapped statistic, ensuring asymptotic size equal to α . Under the alternative, $M_{1,T}$ ($M_{2,P}$) diverges to (plus) infinity, while the corresponding bootstrap statistic has a well-defined limiting distribution, ensuring unit asymptotic power.

4 Monte Carlo Findings

We consider the following data generating process for first and second releases:

$$\begin{aligned} {}_{t+1}X_t &= \alpha + \theta {}_tX_{t-1} + \varepsilon_{t+1}^D; \quad \varepsilon_{t+1}^D \sim N(0, \sigma_D^2), \\ {}_{t+2}X_t &= \beta_1 {}_{t+1}X_t + \beta_2 g(Z_t) + \varepsilon_{t+2}^R; \quad \varepsilon_{t+2}^R \sim N(0, \sigma_R^2), \end{aligned}$$

where Z_t is assumed to follow an AR(1) process. Namely, $Z_t = \psi Z_{t-1} + \varepsilon_t^Z$; $\varepsilon_t^Z \sim N(0, \sigma_Z^2)$. Given this structure, we consider five different cases.

Case I: Rationality of first release data $\{\beta_1 = 1; \beta_2 = 0\}$.

Case II: First revision error linearly correlated with available information $\{\beta_1 = 1; \beta_2 \neq 0; g(x) = x^K, K = 1\}$.

Case III: Linear lack of rationality $\{\beta_2 = 0; \beta_1 > 1\}$.

Case IV: First revision error dependent on a quadratic function of available information $\{\beta_1 = 1; \beta_2 \neq 0; g(x) = x^K, K = 2\}$.

Case V: First revision error dependent on a more complex non-linear function of available information $\{\beta_1 = 1; \beta_2 \neq 0; g(x) = \exp(\tan^{-1}(x))\}$.

The values for the parameters governing the process of first releases, $\{\alpha, \theta, \sigma_D\}$ are calibrated to be $\{0.003, 0.6, 0.007\}$, so as to match the sample mean and variance of output, prices and money, respectively, over the period considered in the empirical illustration. In addition, we set $\sigma_R = 0.0001$ and $\sigma_Z = 0.001$. Also, we consider different parameterizations for β_1, β_2, θ and ϕ , as indicated in Table 1. We consider multiple sample sizes, including $T = 200$, $T = 300$, $T = 500$, and $T = 1000$, carry out 1000 Monte Carlo replications, and for each of them perform 500 bootstrap replications. In this version of the paper, however, we report results only for the case where $T = 500$. For our cases where $T < 500$, power results, particularly for Case II, were appreciably worse for the CFS test, although the CFS test continues to yield appreciably higher power for Case IV than all of the other tests, as might be expected. Complete results are available upon request from the authors.

The properties of three different rationality tests are examined. The first one based on testing $H_0 : \gamma_1 = \gamma_2 = \gamma_3 = 0$, in the linear regression below (see Swanson and van Dijk (2006)):

$${}_{t+2}u_t^{t-1} = \gamma_1 + \gamma_2 {}_{t+1}X_t + \gamma_3 Z_t + \varepsilon_t, \quad (10)$$

where ${}_{t+2}u_t^{t-1} = {}_{t+2}X_t - {}_{t+1}X_t$. The second one is based on testing $H_0 : E({}_{t+2}u_t^{t+1} \cdot Z_t) = 0$ versus $H_0 : E({}_{t+2}u_t^{t+1} \cdot Z_t) \neq 0$, and is implemented via the statistic suggested by Chao, Corradi and Swanson (CCS: 2001). Namely:

$$m_T = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} \frac{{}_{t+2}u_t^{t+1} \cdot Z_t}{\sqrt{\widehat{var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} \cdot Z_t \right)}}, \quad (11)$$

whenever we use normal critical values or

$$mu_T = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} \cdot Z_t, \quad (12)$$

when we use bootstrap critical values. Finally, the third test is $M_{1,T}$ as defined in Section 3.2, and hereafter labelled the CFS test. The findings for the three tests are reported in Table 1, for the $T=500$.

From Table 1, we can immediately see that all three tests have similar rejection rates under the null, with an empirical size ranging from 8% to 17% against a nominal size of 10%. Interestingly, in Cases II and III (i.e., linear correlation between the revision error and available information), the linear and CCS tests do not outperform the CFS test, even though the first two tests have power against a fixed linear alternative, while the third one has power against all possible deviations from the null. In case IV, characterized by a quadratic relationship between the revision error and the available information, the CFS substantially outperforms the linear and CCS tests. This is a case where there is a relatively small deviation from linearity. Finally, in Case IV, which is characterized by a very large deviation from linearity, both the linear and CFS tests have rejection rates equal to one, and strongly outperform the CCS test.

5 Empirical Results

5.1 Data

In an illustrative empirical analysis, we use the following series: real GDP (seasonally adjusted), GDP chain-weighted price index (seasonally adjusted), and the money stock (measured as M1, seasonally adjusted). All series have a quarterly frequency and were obtained from the Federal Reserve Bank of Philadelphia’s real-time data set for macroeconomists (RTDSM), which can be accessed online at <http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/>. The series were obtained from the “By-Variable” files of the “Core Variables/Quarterly Observations/Quarterly Vintages” dataset. The data we use are discussed in detail in Croushore

and Stark (2001). The first vintage in our sample is 1965:Q4, for which the first calendar observation is 1959:Q3. This means that the first observation in our dataset is the observation that was available to researchers in the fourth quarter of 1965, corresponding to the calendar date for the third quarter of 1959. The datasets continue through 2006:Q4, with calendar date corresponding to 2006:Q3, allowing us to keep track of the exact data that were available at each vintage for every possible calendar dated observation up to one quarter before the vintage date. This makes it possible to trace the entire series of revisions for each observation across vintages. We use log-differences throughout our analysis. Figure 1 contains plots of the rate of growth of real GDP, and its first and second revision errors. We can immediately see upon inspection of the plots that while the first revision error is roughly normally distributed, the second revision error is instead rather concentrated around zero. This is one of the reasons why much of our analysis focuses only on the impact of first and second revision errors - later revision errors appear to offer little useful information, other than signalling the presence of definitional and related methodological changes.

5.2 Rationality Tests

In order to test $H_{0,1} : E(t_{+2}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0$, a.s. versus $H_{A,1} : \Pr(E(t_{+2}u_t^{t+1}|\mathcal{F}_t^{t+1}) = 0) < 1$, we construct the statistic $M_{1,T}$, as defined in (4), setting ${}_{t+1}\mathcal{F}_t^{t+1} = {}_{t+1-j}W_{t-j}$, $j = 0, 1, \dots, t-1$ with (i) ${}_{t+1-j}W_{t-j} = {}_{t+1-j}X_{t-j}$; and (ii) ${}_{t+1-j}W_{t-j} = [{}_{t+1-j}X_{t-j}, {}_{t+1-j}u_{t-1-j}^{t-j}]'$. Following Corradi and Swanson (2002) we set w to be the exponential function, and Φ the inverse tangent function. We also set $\gamma_j \equiv \gamma \cdot (j+1)^{-2}$, where γ is defined over a fine grid, $\gamma \in [0, 3]$ for case (i); and

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \in [0, 3] \times [0, 3],$$

for case (ii). To sum up, the test statistic, under case (i), is computed as the supremum of

$$m_{1,T}(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} \exp \left(\sum_{j=0}^{t-1} (\gamma_1(j+1)^{-2} \tan^{-1}(X_{t-j})) \right),$$

and under case (ii), as the supremum of

$$m_{1,T}(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} \exp \left(\sum_{j=0}^{t-1} \left(\gamma_1(j+1)^{-2} \tan^{-1}({}_{t+1-j}X_{t-j}) + \gamma_2(j+1)^{-2} \tan^{-1}({}_{t+1-j}u_{t-1-j}^{t-j}) \right) \right).$$

We also construct linear rationality tests based on the regression in (10), where setting $Z_t = {}_{t+1}u_{t-1}^{t-2}$ corresponds to ‘case’ (ii) above, and where omitting Z_t from the test regression corresponds to ‘case’ (i) above; as well as the test as defined in (12). As our findings are qualitatively the same for cases (i) and (ii) above, we only report results for case (i); and these are gathered in Table 2. It is

immediately apparent, upon inspection of the results, that for output, the null of rationality is rejected by all tests at a 5% level, with the only exception being that the null is rejected only at a 10% level, when using the linear test. These findings correspond to those of Aruoba (2008). For prices, the null is rejected (at both the 5% and 10% levels) based on the CCS and CFS test, but not based on the linear test. Finally, for money, the null fails to reject based on all tests. We leave to future research the issue of whether bias adjustment may be sufficient to address the irrationality in these datasets.

6 Concluding Remarks

We outline two new tests for data rationality, both of which are designed to assess rationality from an ex-ante forecasting perspective. The distinctive feature of our tests is that, contrary to existing tests, they have power against generic non-linear alternatives. The price paid for such generality is that the limiting distribution is not nuisance parameter free. Hence, “ready to use” critical values are not available. Valid asymptotic critical values are instead obtained via a bootstrap procedure based on resampling functions of the entire history of available information, along the lines of Goncalves and White (2004), and allowing for recursive estimation schemes, as in Corradi and Swanson (2007). An illustrative empirical implementation of the tests yields strong evidence against data rationality, in the sense that early releases of U.S. output, price, and money do not take into account all of the available information at the time of data release. Moreover, for money, our non-linear test suggests rejection of the rationality null, whereas a standard linear regression based test and the so-called linear CCS test do not.

Many issues in the burgeoning real-time literature remain unresolved. For example, from an empirical perspective it remains to extend the analysis that we carry out to later vintages (only the first two vintages are examined in our empirical analysis). From a theoretical perspective, it remains to extend the predictive density accuracy tests of Corradi and Swanson (2006a,b) to the case of real-time data, thus furthering the pointwise predictive accuracy test results reported in Clark and McCracken (2009).

7 References

- Andrews, D.W.K. (1991), An Empirical Process Central Limit Theorem for Dependent Non-Identically Distributed Random Variables, *Journal of Multivariate Analysis*, **38**, 187-203.
- Andrews, D. W. K. (2002), Higher-Order Improvements of a Computationally Attractive k -Step Bootstrap for Extremum Estimators, *Econometrica* **70**, 119-162.
- Aruoba, S.B. (2008), Data Revisions are not Well-Behaved, *Journal of Money, Credit and Banking*, **40**, 319-340.
- Bierens, H.B. (1982), Consistent Model Specification Tests, *Journal of Econometrics*, **20**, 105-134.
- Bierens, H.B. (1990), A Conditional Moment Test of Functional Form, *Econometrica* , **58**, 1443-1458
- Bierens, H.J. and W. Ploberger (1997), Asymptotic Theory of Integrated Conditional Moment Tests, *Econometrica*, **65**, 1129-1152.
- Brodsky, N. and P. Newbold (1994), Late Forecasts and Early Revisions of United States GNP, *International Journal of Forecasting* **10**, 455-460.
- Chao, J., V. Corradi and N.R. Swanson (2001), An out-of-sample Test for Granger Causality, *Macroeconomic Dynamics* **5**, 598-620.
- Clark, T. E. and M.W. McCracken (2009), Tests of Equal Predictive Ability with Real-Time Data, this issue.
- Corradi, V. and E.M. Iglesias (2008), Bootstrap Refinements for QML Estimators of the GARCH(1,1) Parameters, *Journal of Econometrics*, **144**, 500-510.
- Corradi, V. and N.R. Swanson (2002), A Consistent Test for non-linear out-of-sample Predictive Accuracy, *Journal of Econometrics* **110**, 353-381.
- Corradi, V. and N.R. Swanson (2006a), Predictive Density Evaluation, in: **Handbook of Economic Forecasting**, eds. Clive W.J. Granger, Graham Elliot and Allan Timmermann, Elsevier, Amsterdam, pp. 197-284.
- Corradi, V. and N.R. Swanson (2006b), Predictive Density and Conditional Confidence Interval Accuracy Tests, *Journal of Econometrics* **135**, 187-228.
- Corradi, V. and N.R. Swanson (2007), Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes, *International Economic Review*, **48**, 67-109.
- Croushore, D. and T. Stark (2001), A Real-Time Dataset for Macroeconomists, *Journal of Econometrics* **105**, 111-130.

- Croushore, D. and T. Stark (2003), A Real-Time Dataset for Macroeconomists: Does Data Vintage Matter?, *Review of Economics and Statistics* **85**, 605–617.
- de Jong, R.M. (1996), The Bierens Test Under Data Dependence, *Journal of Econometrics*, **72**, 1-32.
- Diebold, F.X. and G.D. Rudebusch (1991), Forecasting Output with the Composite Leading Index: A Real-Time Analysis, *Journal of the American Statistical Association* **86**, 603–610.
- Elliott, G., I. Komunjer and A. Timmermann (2005), Estimation and Testing of Forecast Rationality Under Flexible Loss, *Review of Economic Studies*, **72**, 1107-1125.
- Faust, J., J.H. Rogers and J.H. Wright (2005), News and Noise in G-7 GDP Announcements, *Journal of Money, Credit and Banking*, **37**, 403-417.
- Gallant, A.R., and H. White (1988), *A Unified Theory of Estimation and Inference for non-linear Dynamic Models*, Blackwell, Oxford.
- Goncalves, S. and R.M. de Jong (2003), Consistency of the Stationary Bootstrap under Weak Moment Conditions, *Economic Letters*, **81**, 273-278.
- Gonçalves, S. and H. White (2002), The Bootstrap for the Mean of Dependent Heterogeneous Arrays, *Econometric Theory*, **18**, 1367-1384.
- Gonçalves, S. and H. White (2004), Maximum Likelihood and the Bootstrap for non-linear Dynamic Models, *Journal of Econometrics* **119**, 199-219.
- Hamilton, J.D. and G. Perez-Quiros (1996), What Do the Leading Indicators Lead?, *Journal of Business* **69**, 27–49.
- Hansen, B.E. (1996a), Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis, *Econometrica*, **64**, 413-430.
- Hansen, B.E. (1996b), Stochastic Equicontinuity for Unbounded Dependent Heterogeneous Arrays, *Econometric Theory*, **12**, 347-359.
- Kavajecz, K.A. and S. Collins (1995), Rationality of Preliminary Money Stock Estimates, *Review of Economics and Statistics* **77**, 32–41.
- Keane, M.P. and D.E. Runkle (1989), Are Economic Forecasts Rational?, Federal Reserve Bank of Minneapolis *Quarterly Review* **13** (Spring), 26–33.
- Keane, M.P. and D.E. Runkle (1990), Testing the Rationality of Price Forecasts: New Evidence from Panel Data, *American Economic Review* **80**, 714–735.
- Künsch, H.R., (1989), The Jackknife and the Bootstrap for General Stationary Observations. *Annals of Statistics* **17**, 1217–1241.

Lee, T.-H., H. White and C.W.J. Granger (1993), Testing for Neglected non-linearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests, *Journal of Econometrics*, **56**, 269-290.

Mankiw, N.G., D.E. Runkle and M.D. Shapiro (1984), Are Preliminary Announcements of the Money Stock Rational Forecasts?, *Journal of Monetary Economics* **14**, 15–27.

Mankiw, N.G. and M.D. Shapiro (1986), News or Noise: an Analysis of GNP Revisions, *Survey of Current Business* **66**, 20–25.

Mork, K.A. (1987), Ain't Behavin': Forecast Errors and Measurement Errors in Early GNP Estimates, *Journal of Business & Economic Statistics* **5**, 165–175.

Muth, J. F. (1961), Rational Expectations and the Theory of Price Movements, *Econometrica*, **29**, 315–335.

Pötscher, B.M., and I.R. Prucha (1997), *Dynamic non-linear Econometric Models*, Springer, Berlin.

Rathjens, P., and Robins, R. P. (1995), Do Government Agencies Use Public Data? The Case of GNP, *Review of Economics and Statistics*, **77**, 170–172.

Stinchcombe, M.B. and H. White (1998), Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative, *Econometric Theory*, **14**, 3, 295-325.

Swanson, N.R., E. Ghysels and M. Callan (1999), A Multivariate Time Series Analysis of the Data Revision Process for Industrial Production and the Composite Leading Indicator, in R.F. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger*, Oxford: Oxford University Press, pp. 45-75.

Swanson, N.R. and D. van Dijk (2006), Are Statistical Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry, *Journal of Business & Economic Statistics* **24**, 24–42.

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

Appendix

Proof of Theorem 1:

(i) Given A1(i), letting $\mathcal{F}_s = \sigma (_{i+1}W_i, i \leq s - 1)$, $s < t$, and letting $\|z_t\|_2^{1/2} = (E(z_t^2))^{1/2}$ for any, possibly vector-valued, random variable z_t , we have that:

$$\begin{aligned} & \sup_{s \leq t \leq T} \left\| \sum_{j=1}^{t-1} j^{-2} \Phi (_{t+1-j}W_{t-j}) - E \left(\sum_{j=0}^{t-1} j^{-2} \Phi (_{t+1-j}W_{t-j}) | \mathcal{F}_s \right) \right\|_2^{1/2} \\ & \leq \sum_{j=s}^{T-1} j^{-2} \sup_t \|\Phi (_{t+1}W_t)\|_2^{1/2} \simeq O (s^{-1-\varepsilon}), \quad \varepsilon > 0, \end{aligned}$$

and thus, from the definition of NED mapping (see e.g. Gallant and White (1988, Ch. 4)), it follows that $\sum_{j=0}^{t-1} j^{-2} \Phi (_{t+1-j}W_{t-j})$ is L_2 -NED of size -1 on $\Phi (_{t+1-j}W_{t-j})$.

Given A1(iv)-(v), for each $\gamma \in \Gamma$, $w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi (_{t+1-j}W_{t-j}) \right)$ is Lipschitz of order 1 on \mathcal{U} . Thus, given A1(ii), from Theorem 4.2 in Gallant and White (1988), it follows that $w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi (_{t+1-j}W_{t-j}) \right)$ is NED of size -1 on the strong mixing base $\Phi (_{t+1-j}W_{t-j})$ of size $-2r/(r-2)$. Now, under the null hypothesis, $_{t+2}u_t^{t+1}$ is a martingale difference sequence. Given A1(ii)-(iii), it follows from Lemma 2 in Andrews (1991) that for any $\gamma_1, \gamma_2 \in \Gamma$, $Cov (m_T(\gamma_1), m_T(\gamma_2))$ exists. Straightforward calculation shows that:

$$\begin{aligned} & Cov (m_T(\gamma_1), m_T(\gamma_2)) \\ & = E \left((_{t+2}u_t^{t+1})^2 w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi (_{t+1-j}W_{t-j}) \right) w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi (_{t+1-j}W_{t-j}) \right) \right), \end{aligned}$$

as for $t > s$,

$$\begin{aligned} & E \left(_{t+2}u_t^{t+1} \ _{s+2}u_s^{s+1} w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi (_{t+1-j}W_{t-j}) \right) w \left(\sum_{j=0}^{s-1} \gamma'_{2,j} \Phi (_{s+1-j}W_{s-j}) \right) \right) \\ & = E \left(_{s+2}u_s^{s+1} w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi (_{t+1-j}W_{t-j}) \right) w \left(\sum_{j=0}^{s-1} \gamma'_{2,j} \Phi (_{s+1-j}W_{s-j}) \right) E (_{t+2}u_t^{t+1} | \mathcal{F}_t^{t+1}) \right) \\ & = 0. \end{aligned}$$

Thus, Assumption C in Andrews (1991) is satisfied, and the statement follows from his Theorem 4, recalling his remark 4 and the fact that given A1(i), Γ is a bounded set in \mathcal{R}^2 .

(ii)

$$m_{1,T}(\gamma)$$

$$\begin{aligned}
&= \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} \left({}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right) - E \left({}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right) \right) \right) \\
&+ \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} E \left({}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right) \right).
\end{aligned}$$

The first term on the RHS above is $O_P(1)$ as it converges in distribution by part (i). As $w(\cdot, \gamma)$ is generically comprehensive, $E \left({}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right) \right) \neq 0$, for all $\gamma \in \Gamma$. The statement then follows.

Proof of Theorem 2: (i) For any given γ , let $\mu_u = E \left({}_{t+2}u_t^{t+1} \right)$, $\psi = \sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j}$, $\mu(\gamma) = E(w(\psi))$, and

$$\begin{aligned}
m_{2,P}(\gamma) &= \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} ({}_{t+2}u_t^{t+1} - \mu_u) w(\psi) - \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(\frac{1}{t} \sum_{i=0}^t ({}_{i+2}u_i^{i+1} - \mu_u) \right) w(\psi) \\
&= \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} ({}_{t+2}u_t^{t+1} - \mu_u) w(\psi) - \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(\frac{1}{t} \sum_{i=0}^t ({}_{i+2}u_i^{i+1} - \mu_u) \right) \mu(\gamma) + o_p(1), \quad (13)
\end{aligned}$$

where the $o_p(1)$ term holds uniformly in γ , because of the uniform law of large numbers for NED processes on a strong mixing base (see, e.g., Gallant and White, (1988, Ch. 3)). Note also that the first term on the RHS of (13), $(m_{2,P}(\gamma_1), m_{2,P}(\gamma_2))$, converges in distribution for each for any $\gamma_1, \gamma_2 \in \Gamma$ and is stochastic equicontinuous on Γ , by Theorem 4 in Andrews (1991). Finally, consider the second term on the RHS of (13). It can be treated by a similar argument to that used in the proof of Theorem 1 in Corradi and Swanson (2002).

(ii) As in Part (ii) of Theorem 1.

Proof of Theorem 3: (i) First, note that, as I_i are identically and independently distributed, $w_t^*(\gamma)$ is *iid*, conditional on the sample. Now,

$$E^* \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{*t+1} w_t^*(\gamma) \right) = \sqrt{T} E^* ({}_{t+2}u_t^{*t+1} w_t^*(\gamma)) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} w_t(\gamma), \quad (13)$$

and thus for all γ , $E^* \left(m_{1,T}^*(\gamma) \right) = m_{1,T}(\gamma)$. Also,

$$\begin{aligned}
Var^* \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{*t+1} w_t^*(\gamma) \right) &= E^* \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} \left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) - \frac{1}{T} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} w_t(\gamma) \right) \right)^2 \\
&= \frac{1}{T} \sum_{t=1}^{T-2} \left({}_{t+2}u_t^{t+1} w_t(\gamma) - \frac{1}{T} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} w_t(\gamma) \right)^2 = E \left(({}_{t+2}u_t^{t+1})^2 w \left(\sum_{j=0}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right) \right) + o_p(1).
\end{aligned} \quad (14)$$

We need to show that, pointwise in γ , $\sqrt{T} \left(\text{Var}^* \left(m_{1,T}^*(\gamma) \right) \right)^{-1/2} \left(m_{1,T}^*(\gamma) - E^* \left(m_{1,T}^*(\gamma) \right) \right)$ is asymptotically standard normal, conditional on the sample, and for all samples except for a subset with probability measure approaching zero. This follows from Theorem 3.5 in Künsch (1989), once we have shown that the Linderberg condition in his proof (p. 1238) holds, pointwise in γ . Hence, we need to show that:

$$E^* \left(T^{-1} \left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) - E^* \left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) \right) \right)^2 1_{\{T^{-1/2} | {}_{t+2}u_t^{*t+1} w_t^*(\gamma) - E^* \left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) \right) | > \varepsilon\}} \right) = o_p(T^{-1}). \quad (15)$$

Now,

$$\begin{aligned} & P^* \left(T^{-1/2} \left| {}_{t-2}u_t^{*t+1} w_t^*(\gamma) - E^* \left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) \right) \right| > \varepsilon \right) \\ & \leq \frac{1}{\varepsilon^4 T^2} E^* \left(\left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) - E^* \left({}_{t+2}u_t^{*t+1} w_t^*(\gamma) \right) \right)^4 \right) \\ & \xrightarrow{p} 0, \end{aligned}$$

given (13) and (14), and thus (15) follows. Note that, for every $\gamma \in \Gamma$, even if $w_t(\gamma)$ is NED of size -1 , ${}_{t+2}u_t^{*t+1} w_t(\gamma)$ is a martingale difference sequence under the null. Thus, we can rely on the *iid* bootstrap and the verification of the Linderberg condition in (15) is straightforward. On the other hand, Goncalves and White (2002 and 2004) and Goncalves and de Jong (2003) consider the case of the block bootstrap for extremum estimators when the score is near epoch dependent on a heterogeneous mixing sequence; and has a long-run variance which is $O(1)$ but does not possess a well-defined probability limit. In this case, they require the score to be NED of size $-(2 + \delta)$, $\delta > 0$, and show that a Lyapunov (stronger) condition is indeed satisfied.

Joint convergence of $\left(\sqrt{T} \left(m_{1,T}^*(\gamma_1) - E^* \left(m_{1,T}^*(\gamma_1) \right) \right), \sqrt{T} \left(m_{1,T}^*(\gamma_2) - E^* \left(m_{1,T}^*(\gamma_2) \right) \right) \right)$ follows immediately as a consequence of the Wold device. Given A1(iv)-(v), $m_{1,T}^*(\gamma) - E^* \left(m_{1,T}^*(\gamma) \right)$ has a series expansion with smooth (first order Lipschitz) coefficients, as described in equations (3.3) and (3.4) in Andrews (1991). Thus, stochastic equicontinuity on Γ follows from Theorem 1 in Andrews (1991). The statement in part (i), then follows by noting that for each $\gamma_1, \gamma_2 \in \Gamma$,

$$\begin{aligned} & \text{Cov}^* \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{*t+1} w_t^*(\gamma_1), \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{*t+1} w_t^*(\gamma_2) \right) \\ & = E \left(\left({}_{t+2}u_t^{*t+1} \right)^2 w \left(\sum_{j=0}^{t-1} \gamma'_{1,j} \Phi(t+1-j) W_{t-j} \right) w \left(\sum_{j=0}^{t-1} \gamma'_{2,j} \Phi(t+1-j) W_{t-j} \right) \right) + o_p(1). \end{aligned}$$

(ii) Let $w_t(\gamma) = w \left(\sum_{j=1}^{t-1} \gamma'_j \Phi(t+1-j) W_{t-j} \right)$, $\mu(\gamma) = E(w_t(\gamma))$, and $\mu_u = E({}_{t+2}u_t^{*t+1})$.

$$m_{2,P}^*(\gamma) = \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(\left({}_{t+2}u_t^{*t+1} - \hat{\mu}_t^* \right) w_t^*(\gamma) - \frac{1}{T} \sum_{i=1}^{T-2} \left({}_{i+2}u_i^{i+1} - \hat{\mu}_i \right) w_i(\gamma) \right)$$

$$\begin{aligned}
&= \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(({}_{t+2}u_t^{*t+1} - \hat{\mu}_t) w_t^*(\gamma) - \frac{1}{T} \sum_{i=1}^{T-2} ({}_{i+2}u_i^{i+1} - \hat{\mu}_t) w_i(\gamma) \right) \\
&\quad - \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) w_t^*(\gamma) \\
&= \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(({}_{t+2}u_t^{*t+1} - \hat{\mu}_t) w_t^*(\gamma) - \frac{1}{T} \sum_{i=1}^{T-2} ({}_{i+2}u_i^{i+1} - \hat{\mu}_t) w_i(\gamma) \right) \\
&\quad - \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) \mu(\gamma) - \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) (w_t^*(\gamma) - \mu(\gamma)) \\
&= I_P(\gamma) - II_P(\gamma) - III_P(\gamma)
\end{aligned}$$

We need to show that, conditional on the sample, and for all samples except a subset with probability measure approaching zero, $I_P(\gamma)$ has the same limiting distribution as in the statement in (i), $II_P(\gamma)$ has the same limiting distribution as $\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(\frac{1}{t} \sum_{i=1}^t ({}_{i+2}u_i^{i+1} - \mu_u) \right) \mu(\gamma)$, and that $III_P(\gamma)$ is $o_p^*(1)$. Now,

$$E^* \left(\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (({}_{t+2}u_t^{*t+1} - \hat{\mu}_t) w_t^*(\gamma)) \right) = \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \frac{1}{T} \sum_{i=1}^{T-2} ({}_{i+2}u_i^{i+1} - \hat{\mu}_t) w_i(\gamma),$$

and

$$\begin{aligned}
&Var^* \left(\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} ({}_{t+2}u_t^{*t+1} - \hat{\mu}_t) w_t^*(\gamma) \right) \\
&= \frac{1}{P} \sum_{t=R}^{T-2} \frac{1}{T} \sum_{i=1}^{T-2} \left(({}_{i+2}u_i^{i+1} - \hat{\mu}_t) w_i(\gamma) - \frac{1}{T} \sum_{i=1}^{T-2} ({}_{i+2}u_i^{i+1} w_t(\gamma) - \hat{\mu}_t) w_i(\gamma) \right)^2.
\end{aligned}$$

Thus, $I_P(\gamma)$ can be treated along the same lines as in Part (i).

Recalling the definition of $\hat{\mu}_t^*$, given in (8),

$$\begin{aligned}
- II_P(\gamma) &= \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) \mu(\gamma) \\
&= \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(\frac{1}{t} \sum_{j=1}^t \left({}_{j+2}u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} {}_{i+2}u_i^{i+1} \right) \right) \mu(\gamma) \\
&= \mu(\gamma) \frac{a_{R,0}}{\sqrt{P}} \left(\sum_{j=1}^R \left({}_{j+2}u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} {}_{i+2}u_i^{i+1} \right) \right) \\
&\quad + \frac{1}{\sqrt{P}} \sum_{i=1}^{P-2} a_{R,i} \left({}_{R+2+i}u_{R+i}^{*R+1+i} - \frac{1}{T} \sum_{i=1}^{T-2} {}_{i+2}u_i^{i+1} \right),
\end{aligned}$$

where for $i = 0, 1, \dots, P$, $a_{R,i} = (R+i)^{-1} + \dots + (R+P-1)^{-1}$. From Lemma A5 and Lemma A6

in West (1996), it follows that:

$$\begin{aligned}
& Var^* \left(\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) \mu(\gamma) \right) \\
&= 2\Pi\mu(\gamma)^2 \left(E^* \left(j+2u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} i+2u_i^{i+1} \right)^2 \right) \\
&= 2\Pi\mu(\gamma)^2 \sigma_u^2 + o_p(1),
\end{aligned}$$

with $\Pi = 1 - \pi^{-1} \ln(1 + \pi)$. So, for all $\gamma_1, \gamma_2 \in \Gamma$,

$$Cov^* \left(\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) \mu(\gamma_1), \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \hat{\mu}_t) \mu(\gamma_2) \right) = 2\Pi\mu(\gamma_1)\mu(\gamma_2)\sigma_u^2.$$

Thus, $II_P(\gamma)$ properly captures the contribution of parameter estimation error. Finally, recalling (8),

$$III_P(\gamma) = \frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} \left(\frac{1}{t} \sum_{j=1}^t \left(j+2u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} i+2u_i^{i+1} \right) \right) (w_t^*(\gamma) - \mu(\gamma)).$$

As, $w_j^*(\gamma)$ and $j+2u_j^{*j+1}$ are *iid* conditional on the sample, for all $t \neq j$,

$$E^* \left(\left(j+2u_j^{*j+1} - E^* \left(j+2u_j^{*j+1} \right) \right) (w_t^*(\gamma) - \mu(\gamma)) \right) = 0,$$

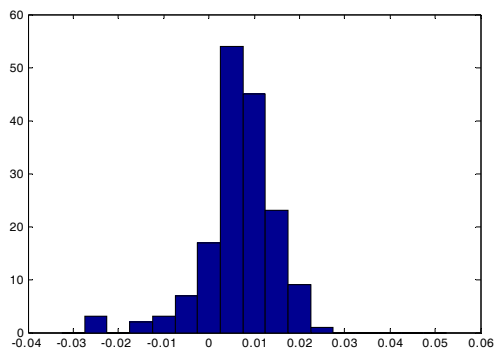
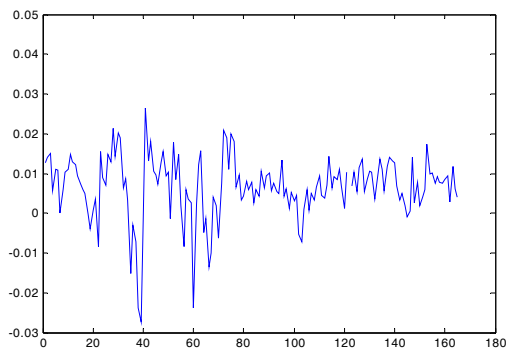
and thus

$$\begin{aligned}
& Var^* \left(\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} (\hat{\mu}_t^* - \mu_u) (w_t^*(\gamma) - \mu(\gamma)) \right) \\
&= \frac{1}{P} \sum_{t=R}^{T-2} E^* \left(\left(\frac{1}{t} \sum_{j=1}^t \left(j+2u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} i+2u_i^{i+1} \right) \right)^2 (w_t^*(\gamma) - \mu(\gamma))^2 \right) \\
&\leq \sup_{t \geq R} \left(E^* \left(\frac{1}{t} \sum_{j=1}^t \left(j+2u_j^{*j+1} - \frac{1}{T} \sum_{i=1}^{T-2} i+2u_i^{i+1} \right) \right)^4 \right)^{1/2} \left(E^* (w_t^*(\gamma) - \mu(\gamma))^4 \right)^{1/2} \\
&= o_p(1),
\end{aligned}$$

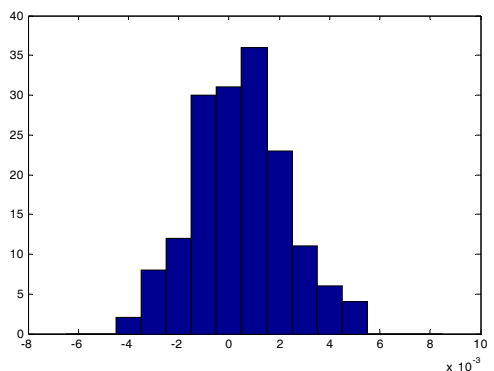
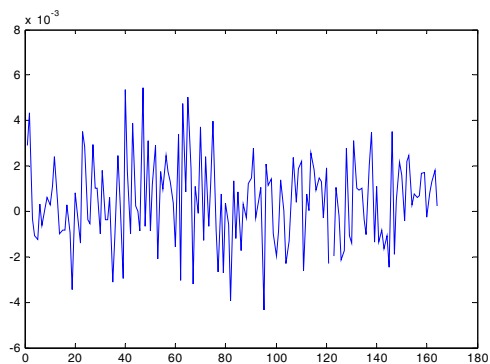
which ensures that $\frac{1}{\sqrt{P}} \sum_{t=R}^{T-2} ((\hat{\mu}_t^* - \hat{\mu}_t) (w_t^*(\gamma) - \mu(\gamma))) = o_{p^*}(1)$, where the $o_{p^*}(1)$ term does not depend on γ . Hence, $III_P(\gamma) = o_{p^*}(1)$, uniformly in γ .

Figure 1: Output Growth Rates, First, and Second Revision Errors — 1965:4 - 2006:4

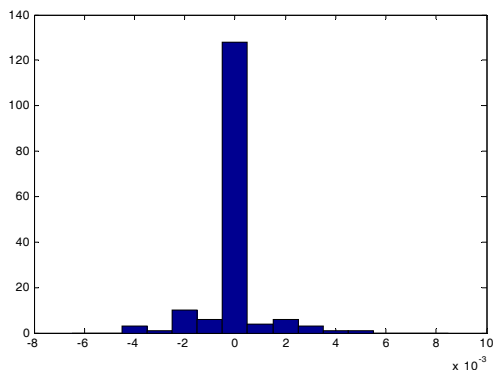
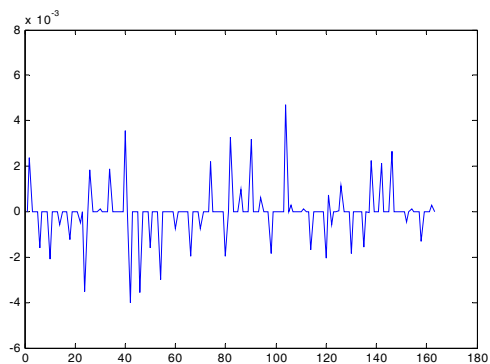
Growth Rates



First Revision Errors



Second Revision Errors



Note: First revision errors are defined as follows: ${}_{t+2}u_t^{t+1} = {}_{t+2}X_t - {}_{t+1}X_t$, where ${}_{t+1}X_t$ is the annualized growth rate of output pertaining to calendar date t , and available at time $t + 1$. Similarly, second revision errors are defined as ${}_{t+3}u_t^{t+2} = {}_{t+3}X_t - {}_{t+2}X_t$. See Sections 2 and 3 for further details.

Table 1: Monte Carlo Results - Empirical Size and Power of Various Tests - T=500^(*)

		<i>Linear Regression Test</i>	<i>CCS Test</i>		<i>CFS Test</i>
			<i>Normality</i>	<i>Bootstrap</i>	
<i>Case I</i>	$\theta = 0.6$	0.114	0.110	0.084	0.172
	$\theta = 0.9$	0.114	0.110	0.086	0.168
<i>Case II</i>	$\beta_2 = 0.3, \psi = 0.6$	1.000	1.000	1.000	1.000
	$\beta_2 = 0.3, \psi = 0.9$	1.000	1.000	1.000	1.000
	$\beta_2 = 0.6, \psi = 0.6$	1.000	1.000	1.000	1.000
	$\beta_2 = 0.6, \psi = 0.9$	1.000	1.000	1.000	1.000
	$\beta_2 = 0.9, \psi = 0.6$	1.000	1.000	1.000	1.000
	$\beta_2 = 0.9, \psi = 0.9$	1.000	1.000	1.000	1.000
	$\beta_2 = 2.0, \psi = 0.6$	1.000	1.000	1.000	1.000
<i>Case III</i>	$\beta_1 = 1.01$	1.000	1.000	1.000	1.000
	$\beta_1 = 1.02$	1.000	1.000	1.000	1.000
	$\beta_1 = 1.05$	1.000	1.000	1.000	1.000
	$\beta_1 = 1.10$	1.000	1.000	1.000	1.000
<i>Case IV</i>	$\beta_2 = 0.3, \psi = 0.6$	0.106	0.086	0.090	0.208
	$\beta_2 = 0.3, \psi = 0.9$	0.102	0.092	0.092	0.258
	$\beta_2 = 0.6, \psi = 0.6$	0.108	0.092	0.070	0.224
	$\beta_2 = 0.6, \psi = 0.9$	0.118	0.114	0.122	0.354
	$\beta_2 = 0.9, \psi = 0.6$	0.108	0.098	0.086	0.244
	$\beta_2 = 0.9, \psi = 0.9$	0.122	0.148	0.140	0.490
	$\beta_2 = 2.0, \psi = 0.6$	0.120	0.116	0.090	0.348
<i>Case V</i>	$\beta_2 = 2.0, \psi = 0.9$	0.262	0.326	0.212	0.844
	$\beta_2 = 0.3, \psi = 0.6$	1.000	0.434	0.266	1.000
	$\beta_2 = 0.3, \psi = 0.9$	1.000	0.718	0.382	1.000
	$\beta_2 = 0.6, \psi = 0.6$	1.000	0.434	0.260	1.000
	$\beta_2 = 0.6, \psi = 0.9$	1.000	0.718	0.396	1.000
	$\beta_2 = 0.9, \psi = 0.6$	1.000	0.434	0.254	1.000
	$\beta_2 = 0.9, \psi = 0.9$	1.000	0.718	0.376	1.000
	$\beta_2 = 2.0, \psi = 0.6$	1.000	0.434	0.252	1.000
$\beta_2 = 2.0, \psi = 0.9$	1.000	0.718	0.386	1.000	

^(*)In this table, test rejection frequencies are reported for samples generated using $T = 500$ observations. "Linear Regression Test" is a test of the null of rationality in first releases based on the following regression model: ${}_{t+2}u_t^{t-1} = \gamma_1 + \gamma_2 {}_{t+1}X_t + \gamma_3 Z_t + \varepsilon_t$, where ${}_{t+2}u_t^{t+1} = {}_{t+2}X_t - {}_{t+1}X_t$. The first column reports the percentage of rejections of the null hypothesis, $H_0 : \gamma_1 = \gamma_2 = \gamma_3 = 0$, at a 10% significance level. Following Chao, Corradi and Swanson (2001), CCS is the test based on the statistic $m_{1,T} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} \cdot Z_t$. The second column of numerical entries in the table reports the percentage of rejections of the null hypothesis, $H_0 : E({}_{t+2}u_t^{t+1} \cdot Z_t) = 0$, at a 10% significance level, using critical values gotten from the normal distribution, while the third column reports rejection frequencies based upon the use of bootstrapped critical values. Finally, the fourth column of numerical entries is the percentage of rejections of the null hypothesis of rationality based upon implementation of the $M_{1,T}$ test (called here the CFS test) described in the Section 3.

Table 2: Rationality Test Results^(*)

Output	Linear Regression Test		
	F-Test Statistic	5% CV	10% CV
	2.893	3.053	2.336
	CCS Test & Bootstrap Critical Values		
	Statistic	5% CV	10% CV
	0.000078	0.000044	0.000030
	CFS Test Statistic & Bootstrap Critical Values		
	Statistic	5%	10%
	0.023944	0.011842	0.009845
Prices	Linear Regression Test		
	F-Test Statistic	5% CV	10% CV
	1.851	3.053	2.336
	CCS Test & Bootstrap Critical Values		
	Statistic	5% CV	10% CV
	0.000046	0.000030	0.000026
	CFS Test Statistic & Bootstrap Critical Values		
	Statistic	5% CV	10% CV
	0.056355	0.043220	0.034450
Money	Linear Regression Test		
	F-Test Statistic	5% CV	10% CV
	0.52381	3.053	2.336
	CCS Test & Bootstrap Critical Values		
	Statistic	5% CV	10% CV
	0.000021	0.000042	0.000032
	CFS Test Statistic & Bootstrap Critical Values		
	Statistic	5% CV	10% CV
	0.002549	0.015691	0.011907

^(*) For the linear regression test, case (i) is reported on, where we fit the regression ${}_{t+2}u_t^{t+1} = \gamma_1 + \gamma_2 {}_{t+1}X_t + \varepsilon_t$ and we report the F-statistic associated with the null hypothesis that $\gamma_1 = \gamma_2 = 0$. Case (ii), which includes the lagged revision error as an additional regressor in the test regression, yields qualitatively similar results. The CCS test reported on in this table is the Chao, Corradi and Swanson (2001) test with associated bootstrap critical values (see Corradi and Swanson (2007)). The CFS test reported on in this table is $M_{1,T}$, where the null hypothesis of data rationality is defined as $H_{0,1}$ (see Section 2 for details).