Monetary Policy Rules with Model and Data Uncertainty [1]

Eric Ghysels*, Norman R. Swanson** and Myles Callan***

* Pennsylvania State University
and Centre Interuniversitaire de Recherche et Analyse des Organisations (CIRANO)

** Texas A&M Univeristy and *** Univeristy of Virginia

December 1999

Abstract

We examine the prevalence of data, specification, and parameter uncertainty in the formation of simple rules which mimic monetary policy-making decisions. Our approach is to build real-time datasets and simulate a real-time policy-setting environment in which we are able to assess the actual performance of rules, had they been followed in real time. This approach allows us not only to track the performance of alternative rules over time (hence facillitating a type of model selection among competing rules), but also allows us more generally to assess the importance of the data revision process in the formation of macroeconomic time series models. From the perspective of real time data, our results suggest that the use of data which are erroneous, in the sense that they were not available at the time decisions based on forecasts from the rules were used, can lead to the selection of quantitatively different models. From the perspective of policy rules, we find that: our version of "calibration" is better than naive estimation, although both are dominated by an approach to rule formation based on adaptive least squares learning using; rules based on seasonally unadjusted data are more reliable than those based on seasonally adjusted data; and rules based soly on preliminary data do not minimize mean square forecast error (MSE) risk. In particular, early releases of data can be noisy, and for this reason it is useful to also use data which have been revised when making decisions using policy rules.

JEL classification: C82, C53, C22.
Keywords: data revision process, prediction, adaptive and rational expectations, real-time data.

[1] The authors wish to thank Dean Croushore, Lars Hansen, Glenn Rudebusch and Thomas Sargent for stimulating conversations during the writing of the paper, and Brian Preslopsky for supplying us with target federal funds rate data. Also, comments made by seminar participants at the Federal Reserve Board, the Federal Reserve Bank of Kansas, the Federal Reserve Bank of St. Louis, Pennsylvania State University, Texas A&M University, the University of Kansas, the Univeristy of North Carolina, and the University of Pennsylvania are greatly appreciated. Corresponding Author: Norman R. Swanson (nswanson@econ.tamu.edu), Department of Economics, Texas A&M University, College Station, TX 77843-4228.

1 Introduction

In academic circles, model and data uncertainty are rarely discussed, even when the task at hand is the construction of real-time forecasts and forecast model selection. By model uncertainty, we mean that the specification and/or parameters of a model are no longer assumed to be fixed and known. While ignoring this type of uncertainty often leads to tractable models which are easily analyzed, it may also paint a picture of the world which is oversimplified. By data uncertainty, we mean that first released data are often noisy in the sense that incomplete and/or erroneous initial information has been used in their construction. Indeed, it may take many years of revisions before data are considered final.

From the perspective of monetary policy rules, which we here use as a vehicle to discuss model and data uncertainty, it is worth stressing that actual policy decisions are made in a real-time setting using preliminary and/or partially revised data. Thus, questions relating not only to which variables should be used, but also to which data releases should be used, make the process of policy-making much more complex than is typically assumed in abstract models of monetary policy. In this paper, we build real-time datasets and simulate a real-time policy-setting environment in which we are able to assess the actual performance of monetary policy rules, had they been followed in real time. This approach allows us not only to track the performance of alternative rules over time (hence facillitating a type of model selection among competing rules), but also allows us more generally to assess the importance of the data revision process in the formation of macroeconomic time series models.

The class of rules which we consider are commonly referred to as Taylor's rules (see Taylor (1979, 1993a)), and are motivated by the apparent existence of trade-offs between inflation and output variability. Versions of these rules have been incorporated in and/or arise in a variety of different macroeconomic models. For example, Rotemberg and Woodford (1997, 1998) develop a rational expectations model with intertemporally optimizing agents in which various interest rate targeting rules arise as optimal responses of the monetary authority. This series of papers is not only important because monetary policy rules are shown to arise naturally when expected utility in a representative household is maximized, but also because the model allows for the computation of welfare measures for representative households under different monetary policy rule implementations. Based on their theoretical model, as well as on a thorough empirical evaluation, Rotemberg and Woodford (1998) find that low and stable inflation together with stable interest rates can be achieved when Taylor's rules of the type which we examine are augmented by including lagged federal funds rates. Many other extensions and variations of Taylor's rule have been proposed in recent years. For example, policy rules that focus on exchange rates or the money supply are alternatives to rules which focus on interest rates. Indeed, the recent literature on policy rules is large. A partial list of relevant papers includes: Bryant, Hooper and Mann (1993), Estrella and Mishkin (1998), Frankel and Chinn (1995), Fuhrer (1997), Fuhrer and Moore (1995), Hansen (1996), Henderson and McKibbin (1993), King and Wolman (1996), Levin, Wieland and Williams (1998), Matheny (1996), McCallum (1993,1997), and Taylor (1993b). We take our policy rules as given, and do not rationalize them with respect to any particular macroeconomic model. Thus, we do not attempt to offer new insights into the usefulness of policy rules per se (see e.g. Taylor (1993a,b), Sargent (1998a)). Moreover, unlike Hansen and Sargent (1998), we do not examine the deeper issue of the effect of model uncertainty on the *design* of policy rules, as we do not concern ourselves with the specification of a theoretical model. Rather, our approach is to emphasize two related but different issues, namely (i) model uncertainty viewed through the lens of parameter uncertainty and model specification, and (ii) the availability and timing of data with which to examine and implement rules.

Uncertainty in policy models is an issue which has recently received some attention in the literature, both from the perspective of model misspecification and from the perspective of learning. Examples of papers in this area include Anderson, Hansen and Sargent (1997), Chung (1990), Granger and Deutsch (1992), Hansen and Sargent (1998), Marcet and Nicolini (1997), and Sargent (1998). Related papers in the area of learning are: Bray (1982), Bray and Kreps (1987), Kuan and White (1994), Marcet and Sargent (1989a,b), and Woodford (1990), while a review of the learning literature is in Marimon (1997). Anderson et al. (1997) and Hansen and Sargent (1998) consider discrete and continuous time optimal policy control models where decision makers assume that their models are misspecified, and adopt robust strategies which are meant to hedge against certain types of model misspecification. In particular, closed form policy rules are formulated which are a function of a robustness parameter. The robustness parameter, in turn, resembles commonly used measures of risk aversion. In this approach, decision-makers no longer take for granted that their model is: (1) true, in the sense that their model coincides with the underlying data generating process, (2) known by the decision-maker, and (3) fixed or time invariant. In the same spirit as Anderson et al. (1997), although less formally, Granger and Deutsch (1992) examine the evaluation of policy decisions which arise when a particular economic variable is targeted. They propose numerous tests based on the comparison of actual target variable outcomes with forecast values arising from the implementation of competing policy models.

Despite the recent interest in model uncertainty, it should be stressed that the bulk of the literature on monetary policy rules, and in macroeconomics in general, assumes that models are known and correctly specified. Within this context, rational expectations assumptions play a central role, and have led to numerous powerful policy prescriptions. The seminal papers by Phelps (1967), Friedman (1968) and Lucas (1972) laid much of the groundwork for these new results. Implications of this work were profound regarding the natural rate theory and the temporary character of trade-offs between inflation and unemployment, for example. The policy implications were far reaching in many other dimensions as well. For example, the Lucas critique suggested that the common practice of fitting econometric models, and simulating policy outcomes was in general incorrect. Despite the Lucas critique, however, the old approach of estimating new econometric models, and hence of formulating new policy prescriptions each time new data became available remained much in use. As Sargent (1998a) points out, this quandary can in many ways be equated with the distinction between policymaking under adaptive versus rational expectations. Sargent also stresses that early findings concerning the sub-optimality of adaptive expectations assumptions were based on the analysis of models which were assumed to be correctly specified (i.e. without model uncertainty). For example, when all models are viewed as approximations Sargent shows that simple "adaptive" forecasting techniques based on rolling regressions, where parameters are updated at each point in time, actually yield forecasts which are not inferior to those based on "optimal" rational expectations theories. This result is perhaps surprising, given that the rolling regression approach is certainly not optimal in a standard utility optimizing representative agent framework. It should perhaps also be noted that most papers which empirically examine policy rules abstain from parameter estimation and instead resort to calibration. As Sargent (1998b) points out, this approach implicitly assumes that the model from which the policy rule is derived is an approximation. Indeed, Rotemberg and Woodford (1998), for example, are clearly aware of this issue, as they empirically examine not only optimal policy rules which derive from their theoretical model (when appropriately calibrated), but also a variety of *other* related rules which are essentially alternative versions of Taylor's rule. This approach is consistent with the argument that since the models under investigation are approximations, the policy rules are also approximations. Furthermore, the "best" approximation might change as new information becomes available, not only in the sense of shifting parameters but also in the sense of changing specifications. Along these lines, one is left wondering what effects parameter estimation (or more generally parameter uncertainty) and model specification, have on the empirical evaluation of policy rules. We attempt to quantify these effects by jointly assessing their impact on the evaluation of policy rules when parameters are fixed (or calibrated), when parameters are estimated at a given point in time, when parameters are re-estimated as new information becomes available (a form of adaptive learning), and when the variables to include in the policy rule are chosen anew at each point in time (changing model specification).

With regard to data uncertainty, the importance of the timing and availability of the data which are used in the empirical evaluation of policy rules is crucial. In order to address this important issue, we use *real-time* datasets to replicate the information available to private agents and policy makers at any given point in time in the day-to-day process of policy setting. In this sense, we simulate a real-time policy setting environment. Our real-time data collection strategy ensures that "future information" due to the use of information which is temporally antecedent to the date under consideration is not (accidentally) incorporated into the dataset at the wrong point in time. This is particularly important for seasonally adjusted data, for example, as two sided filters are generally used in the construction of such data, and the re-estimation of the filters *after* date t, using data from $t + 1$ and $t + 2$, say, results in a revised seasonally adjusted figure for t which actually contains information which was available beyond period t. Before discussing the relative merits of using real-time datasets, however, it is worth pointing out that within the context of timing (or availability), economic data can easily be classified into three types: (1) *Preliminary Data*: These types of data consist of the first reported datum for each variable at each point in time. (2) *Partially Revised* or *Real-Time Data*: These types of data are much more difficult to collect than preliminary data, as they are made up of a full vector of observations at each point in time for each variable. (3) *Fully Revised* or *Final Data*: Final data are data which have been successively revised, and for which no further revisions will be made. This is the type of data that academics often have in mind when conducting economic time series studies, perhaps simply because it is data which is not subject to revision, and it is felt that if one could adequately forecast a *fully revised* figure, then there would be no need for further modeling. It is quite possible, however, that true *final data* will never be available for many economic series.[1] Interestingly, most datasets which are constructed by applied economists clearly consist of a mixture of preliminary data, partially revised data, and final revised data, but are clearly not real-time. This is an issue which has not been completely overlooked in the literature. For example, Orphanides (1997) recognizes the importance of real-time data when evaluating policy rules, and examines monetary policy rules using quarterly real-time data for the period 1987-1992. Our approach differs from Orphanides' in a number of respects. In particular, we use monthly data over a period of more than 20 years, examine parameter as well as model specification uncertainty, and also consider the effects of using seasonally adjusted versus unadjusted data. In related work, Ghysels (1987), Maravall and Pierce (1983, 1986), Pierce (1981), Sargent (1989), Trivellato and Rettore (1986) and Swanson, Ghysels and Callan (1997), examine revision process errors, while Croushore and Stark (1999), Diebold and Rudebusch (1991), Fair and Shiller (1990) and Swanson and White (1995,1997) point out that the comparison of econometric forecasts based on data from CITIBASE, for example, with forecasts made in real-time by professional forecasters (e.g. see Croushore (1993)) is invalid, strictly speaking, because real-time data are not used in the estimation of the econometric models.

Our findings can broadly be summarized as a set of prescriptions and diagnoses which are useful not only within the context of optimal monetary policy rule forecast models, but also within the context of the application of real-time data to macroeconomics in general. A partial list of our prescriptions and diagnoses is as follows. Vintage matters. For example, it is clear that using *only* "final" data does not yield optimal forecasting models. Thus, prediction precision and hence monetary authority credibility is affected by the vintage (or release) of data used. Adaptive least squares learning yields improved results. In particular, while "calibration" is better than naive estimation, both are dominated by an approach to model formation based on adaptive least squares learning. Dynamic information sets are useful. Put another way, policy rules based on distributed lag polynomials of target variables outperform simpler rules. In addition, the correct application of real-time information leads to policy rule precision which is comparable to that achieved by the use of ex-post data. Thus, the use of the standard (ex-post) sorts of datasets routinely applied in empirical economics not only invalidates any claim that later empirical findings are representative of

[1]This is because benchmark and definitional changes are ongoing and may continue into the indefinite future, for example. Moreover, seasonal adjustment filters involve two-sided filters which in principle have infinite leads and lags.

the real-time flow of events in the economy, but also yields no noteable performance enhancement. Seasonally unadjusted data are better. This may be surprising as it is often argued that seasonal adjustment filters extract the "relevant" component of the data (see e.g. Ghysels (1994)). However, even cursory examination of our policy simulation results reveals that rules based on seasonally unadjusted data are more reliable than when seasonally adjusted data are used. Moreover, unadjusted data are directly available and avoid filtering problems which are exaggerated in real-time datasets. Our last main finding is the following. Patience pays off. Forecast models based soly on preliminary data do not minimize mean square forecast error (MSE) risk. For example, using data which have been revised for 9 months leads to around a 50% decrease in MSE. This suggests that preliminary data should be used with care by both policy-makers and private agents, and such data should perhaps be "down-weighted" when used to revise and update models and coefficient estimates. This result does not suggest, however, that one need throw out preliminary data, and indeed Amato and Swanson (1999) show that the "vintages" of data which yield lowest MSE risk are dependent on the "vintage" of actual data which one uses when forming forecast errors by subtracting actual realizations from forecasts.

The rest of the paper is organized as follows. In Section 2 we broadly discuss monetary policy rules. Section 3 contains details of the datasets which we have constructed. Empirical considerations are discussed in Section 4, while our findings are presented in Section 5. Conclusions are gathered in the final section.

2 A Brief Background of Monetary Policy Rules

John Taylor, in his seminal 1979 paper, introduces nominal rigidities into a rational expectations framework, and derives a model of the macroeconomy in which monetary policy irrelevance does not hold. Optimal monetary policy in this setting exploits a second-order Phillips curve (i.e. a long-run trade-off between inflation and output volatility), implying that business cycle fluctuations can be reduced by increasing the variability of inflation through accommodating monetary policies. In addition, the most important policy instrument in the U.S. since the 1960's has arguably been the Federal Funds rate.[2] Thus, it is not surprising that recent research in optimal monetary policy

[2]See Bernanke and Blinder (1992) and Bernanke and Mihov (1998) for related discussion and for a detailed account of recent U.S. monetary policy operations.

has focused primarily on the use of short-term interest rates as policy instruments. For example, Bryant, Hooper and Mann (1993) report on a series of policy rule simulations, in which short-term interest rates are adjusted in response to deviations (from predetermined targets) in (a) the exchange rate, (b) the money supply, (c) nominal output, and (d) a combination of inflation (or the price level) and real output. Their findings suggest that rules which target inflation and output are the most successful in terms of reducing and stabilizing output and price variability. Taylor (1993a) draws on this finding and suggests a simple interest rate policy rule:

$$R_t = 1 + 1.5\pi_t + 0.5y_t, \tag{1}$$

where, R_t is the federal funds rate, π_t is the rate of inflation, and y_t is the output gap (i.e. the percentage deviation of output from its long-run trend).[3] With this rule, the monetary authority raises the Federal Funds rate if either inflation rises above a target rate (which is assumed to be 2) or if real output rises above its long-term trend, with equal weights applied in either case. Using quarterly data, Taylor demonstrates that this rule successfully mimics U.S. monetary policy for the period 1987 to 1992. Since our analysis is based on monthly data we replicate his graphical evidence (see Figure 1) using real-time industrial production and the Consumer Price Index series. Even though Taylor's original analysis is not based on real-time data and uses quarterly data, our findings based on various versions of Taylor's rule generally agree with his observations.

In addition to Taylor's (1993a,b) evidence, other recent research which further explores rational expectations models with sticky prices suggests that simple policy rules successfully mimic the dynamic properties of the economy (e.g. Fuhrer and Moore (1995), Fuhrer (1997) and Rotemberg and Woodford (1997,1998)). Fuhrer estimates the "optimal policy frontier" which dictates the optimal trade-off between deviations of inflation around a target and output around its potential. This trade-off rises rapidly when the standard deviation of either inflation or output falls below 2%, which suggests that a balanced policy is preferable. Also, he shows that Taylor's (1979) model lies close to this optimal frontier. Rotemberg and Woodford (1998) suggest an alternate optimal policy, one that responds positively to both the lagged funds rate itself, with a parameter larger than one, and inflation. Rotemberg and Woodford (1998, pp. 52) conclude that: "Probably our

[3]It should be stressed that Taylor does not advocate this particular rule, and notes that "... simple, algebraic formulations of such rules cannot and should not be mechanically followed by policymakers" (Taylor (1993a, pp. 213)). However, we view (1) as his "best" specification.

most important conclusion is that a simple interest-rate feedback rule of the kind proposed by Taylor (1993) can achieve outcomes nearly as good as are achievable in principle by any policy, assuming that the commitment of the monetary authority to the rule can be made sufficiently credible"

A novel view of the history of U.S. monetary policy is proposed by Sargent (1998), in which a new form of adaptive expectations is introduced. The methodology is applied to the hypothesized "regime shift" that followed Paul Volker's election as chairman the Board of Governors of the Federal Reserve. Sargent suggests that the shift in policy associated with Volker's chairmanship may not have come from a sudden adherence to a rational expectations philosophy but rather gradually by a learning process that was adaptive, driven by the accumulation of information on the success and/or failure of past policies. In particular, Sargent (1998, pp. 141) notes that: "The regime shifts occur, not from a change in the government's econometric or policy-making procedures, but from a change in the beliefs created by its econometric procedure"

As mentioned above, our approach is to assess the actual performance of a variety of rules, had they been followed in real-time. The rules which we examine can be summarized as follows, where m denotes the vintage of data used (m is a fixed integer between 2 and 25) and t indexes the horizon over which the policy rule is implemented:

Rule 1: $R_t = 1 + 1.5\pi_{t-1}(t-m) + 0.5y_{t-1}(t-m)$,

Rule 2: $R_t = \hat{a}_0 + \hat{b}_1\pi_{t-1}(t-m) + \hat{c}_1 y_{t-1}(t-m)$,

Rule 3: $R_t = \hat{a}_0 + \sum_{j=0}^{p_1} \hat{b}_j\pi_{t-1}(t-j-m) + \sum_{j=0}^{p_2} \hat{c}_j y_{t-1}(t-j-m)$,

Rule 4: $R_t = \hat{a}_0 + \sum_{j=0}^{p_1(t-1)} \hat{b}_j\pi_{t-1}(t-j-m) + \sum_{j=0}^{p_2(t-1)} \hat{c}_j y_{t-1}(t-j-m)$,

Rule 5: $R_t = \hat{a}_0 + \sum_{j=1}^{p_1}\sum_{k=j-1}^{p_2} \hat{b}_{jk}\pi_{t-j}(t-k-m) + \sum_{j=1}^{p_3}\sum_{k=j-1}^{p_4} \hat{c}_{jk} y_{t-j}(t-k-m)$.

In Rules 1-5, R_t is a short-term interest rate instrument (in our case either the effective or target Federal Funds rate), π_t is the rate of inflation, and y_t is the output gap. Notice that for Rules 1-2, only one vintage of data is used, and this data is always released at time $t-1$ and is thus available at time t. As the vintage, m, varies between 2 and 25, twenty four different versions of these rules exist, leading to twenty-four real-time policy simulations for each rule. Rules 3 and 4 are the same as Rules 1 and 2, except that a sequence of different vintages of the target variables which are all released at time $t-1$ are incorporated in the policy rule. Recall that the different vintages correspond to updates of historical data. Rule 5 not only fixes a vintage "starting point" (as do Rules 1 to 4), but also allows previous releases of data for the same calendar time observation to be used by policy setters. This rule, thus, allows for every dimension of our real-time datasets to be used in the construction of policy rules. Note that rules 3-5 nest what might be called "standard" linear rational expectations type forecasting models for π and y, as lag dynamics and parameters are estimated based on various learning criteria which are discussed below.

Extensions to Rules 2-5 which include lagged values of the short-term interest rate instrument are given as follows:

Rule 6: $R_t = \hat{a}_0 + \hat{b}_1\pi_{t-1}(t-m) + \hat{c}_1 y_{t-1}(t-m) + \hat{d}R_{t-1}$,

Rule 7: $R_t = \hat{a}_0 + \sum_{j=0}^{p_1}\hat{b}_j\pi_{t-1}(t-j-m) + \sum_{j=0}^{p_2}\hat{c}_j y_{t-1}(t-j-m) + \hat{d}R_{t-1}$,

Rule 8: $R_t = \hat{a}_0 + \sum_{j=0}^{p_1(t-1)}\hat{b}_j\pi_{t-1}(t-j-m) + \sum_{j=0}^{p_2(t-1)}\hat{c}_j y_{t-1}(t-j-m) + \hat{d}R_{t-1}$,

Rule 9: $R_t = \hat{a}_0 + \sum_{j=1}^{p_1}\sum_{k=j-1}^{p_2}\hat{b}_{jk}\pi_{t-j}(t-k-m) + \sum_{j=1}^{p_3}\sum_{k=j-1}^{p_4}\hat{c}_{jk} y_{t-j}(t-k-m) + \hat{d}R_{t-1}$.

The above rules can all be interpreted as adaptations of (1) to a real-time setting. In Rule 1, the response coefficients are fixed. Hence, there is no parameter estimation by the agent (or the monteray authority, if they use our schemes to parametrize their policy rules). One interpretation of Rule 1 is that it is an optimal solution to some calibrated macroeconomic model. Notice that for any given simulation (across t), Rule 1 involves one vintage of data (i.e. m is fixed). Thus, by varying m and comparing simulation results we are able to assess the relevance of different vintages (or releases) of economic data for model construction. One aspect of this feature of Rule 1 is that by examining the performance of this rule across vintages, we can quantify the benefits associated with waiting for more precise updates of the relevant target variables. Note also that we use monthly data whereas Taylor (1993) uses quarterly data. Thus, Rule 1 is not the same as the Taylor (1993) rule. Rather, the appropriate analog to Taylor's rule in our context is Rule 2, as long as parameters are estimated only once at the beginning of the ex-ante simulation (corresponding to our No Window case discussed below).

As Taylor (1993a) points out, it is not clear whether the response coefficients in Rule 1 are optimal. In addition, from the perspective of an agent forming policy forecast models it is reasonable to allow some flexibility in the model specification process. For these reasons, Rules 2 through 9 are based on *estimated* response coefficients. (In the sequel we use OLS in all of our estimations.) In our analysis, we consider three different coefficient estimation schemes. The schemes are based on the amount of data used in, and the frequency of, response coefficient estimation. In the first scheme, all parameters are estimated once at the beginning of each real-time policy simulation, and remain fixed thereafter. This scheme is referred to as the No Window case. In the second scheme (the Fixed Window case) we use fixed rolling 50 and 100 month real-time data windows to re-estimate response parameters before each new policy decision is made. The third scheme, our Increasing Window case, is the same as the $Fixed\ Window$ case except that we use an increasing real-time data window to estimate the coefficients, beginning with a window width of 50 months. It should be stressed that these schemes are practically feasible from the perspectives of both real-time adaptive forecast model construction and policymaking, as they only entail using information available in real-time.

The first rule which uses estimated coefficients is Rule 2 (or Rule 6 with lagged R_t). This rule has the same policy response structure as Rule 1, but with estimated coefficients. By comparing Rule 1 and Rule 2 we are able to assess the relative merits of using estimated rather than calibrated response coefficients. As Rotemberg and Woodford (1997,1998) show, however, optimal policy may involve distributed lag polynomials of target variables and policy instruments. Thus, we also consider various Rotemberg and Woodford type rules. First, Rule 3 (or Rule 7 with lagged R_t) defines R_t as a function of p_1 vintages of π and p_2 vintages of y, where p_1 and p_2 are selected using the Schwarz Information Criterion (SIC), and the maximum number of lags is 24.[4] Notice although the use of the word *lags* refers to calendar time lags, by fixing the release date of our data at $t-1$, we ensure that the most up to date revisions of all lags of our target variables are used. In Rule 3, p_1 and p_2 do not change as new information becomes available, but rather are based on a real-time "startup" sample of observations. Rule 4 (or Rule 8 with lagged R_t) is the same as Rule 3, except that p_1 and p_2 are selected anew every time a forecast or policy decision is made. A comparison of Rule 3 and 4 thus allows us to assess the impact of parameter uncertainty (which occurs in Rules 3 and 4) and model specification uncertainty (which occurs only in Rule 4). Note

[4] Recall that when referring to "vintages" of data, we are referring to data available at some calendar date which have been revised. So that the m^{th} vintage of data available at time $t-1$ is the revised datum for period $t-m$.

that the type of model specification uncertainty which we consider is limited to lag order selection.[5] Moreover, comparing Rule 3 or 4 with Rule 2 highlights the impact of including distributed lags dynamics in our policy rules, and comparing any of Rules 2-4 with Rule 1 allows us to assess the impact of parameter and/or specification uncertainty on policy rules. In summary, the different ways in which the various rules can be compared allows us to disentangle, at least to some extent, the different effects that parameter, specification and data uncertainty have on Taylor's rule.

Thus far, we have only discussed policy rules based on data which are available at time $t-1$. In particular, Rules 1 through 4 assume that optimal policy is based on a single release of data. For example, when the subscript on our target variables is $t-1$, only data available in period $t-1$ are used in policy formation. While these rules clearly entail real-time policy setting, they may be naive in one sense. If private agents (or policy-makers) believe that different releases of data for the same calendar period contain different information, then they may want to formulate models based not only on different vintages of data released in $t-1$, but also on different vintages of data released in $t-2$, say. For example, if the first and second release observation for January of 1990 are formed using information sets which are non-nested, both variables may be useful. Although this sort of scenario might seem surprising, Swanson, Ghysels and Callan (1997) find some evidence that it is indeed true. In order to allow for this eventuality, we also consider Rule 5 (or Rule 9 with lagged R_t), which mixes vintages and releases of data, thereby allowing private agents or policy setters to explicitly include models of the revision process in their models.

3 Real-Time Data

In the sequel we use four monthly U.S. time series. Two of the variables (the target and effective federal funds rates) are not subject to revision. The other two variables (industrial production and the consumer price index) are subject to revision, and real-time datasets for each of them (both seasonally unadjusted and adjusted) have been constructed. In particular, industrial production (IP) data for the period 1950:04 to 1998:03 have been gathered, while consumer price index (CPI) data for the period 1978:01 to 1998:03 have been gathered. It is worth noting that the data matrix for IP contains more than 170,000 nontrivial entries while the data matrix for the CPI contains more than 29,000 nontrivial entries. As discussed above, the large number of observations in our

[5]For example, although nonlinear rules might be relevant, they are not investigated.

real-time datasets is due to the nature of the data collection - at each point in time an entirely new sequence of data is collected, going back to the beginning of the sample period. In order to further illustrate these features of our data, we reproduce the matrix structure of a generic real-time dataset in Table 1. The entries in the table are denoted $X_t(\tau)$, where the subscript t refers to the release date of the data pertaining to period τ, the date which is in parentheses. Therefore, the diagonal elements in the matrix correspond to the first released or preliminary data. For example, the first entry in Table 1 shows the May 1950 release of the April IP or CPI figure. Keeping τ fixed the first row in the table shows the series of revisions from $t =$ May 1950 up until the end of our dataset March 1998.

Both IP and CPI data are released on or around the 15th of each month, and a typical months' release of data for these variables is comprised of a first, or preliminary release for the previous month, and 1 to 5 months of revisions to data previously released. In addition, more comprehensive benchmark and base year revisions occur from time to time for each of these variables. Turning first to the IP data, the following details are worth noting. Seasonally adjusted and unadjusted IP figures are compiled by the Federal Reserve Board. The primary source for seasonally adjusted IP data is the Federal Reserve Bulletin. For unadjusted IP (before October 1995) the main source is the Bureau of Economic Analysis' publication Survey of Current Business. Additional data for these series were obtained from Federal Reserve monthly statistical releases. Federal Reserve releases for IP are called G.12.3 before May 1990, and G.17 thereafter. Recent releases, and a partial real-time dataset from 1972 onwards for seasonally adjusted data, can also be obtained from the Board of Governor's of the Federal Reserve's website (http://www.bog.frb.fed.us/releases). Also, for three of the major (benchmark) revisions to IP, the Federal Reserve Board released separate publications - (1) Industrial Production 1957-59 Base, (2) Industrial Production 1971 and (3) 1976 Revision.

Aside from typical monthly revisions to recently released data, there have various major updates to IP. Numerous updates are "benchmark" updates for which at least 10 years of data are revised. In the January 1997 benchmark update, it was announced that the primary feature of the benchmark update was to reformulate indexes based on weights that are updated annually rather than every five years, as had been previously done. In general, updates involve updating seasonal adjustment weights (for seasonally adjusted data) and incorporating more complete information on important individual series, while benchmark updates additionally revise series definitions. Another type of update is the base year change. Examples of dates for updates of these types are: July 1971, July

1985, March 1990 and January 1997. Each of these base year updates coincides with a benchmark update, and in fact there are only four recent benchmark updates which do not correspond to base year updates (October 1967, July 1976, January 1994, and January 1998). Recent updates which are not benchmark updates occurred during June 1964, August 1965, October 1966, June 1972, June 1973, August 1977, July 1979, August 1980, August 1981, August 1986, September 1987, April 1993 and November 1994, for example. For IP, there are 3 missing entries due to two major revisions, they are November and December 1953 and March 1985. We replaced each missing observation with the first available data for that period (which in each case is the second release).

Our real-time dataset for seasonally adjusted CPI was constructed from Federal Reserve Bank of St. Louis publications. The main source of this data is National Economic Trends. However, recent releases of the CPI can be obtained from http://www.stls.frb.org/fred/dataindx.html. In general, benchmark revisions to the CPI occur every 12 months, at which time revisions to the data for the preceding 12 months are reported. There is one base year revision - January 1988 (see National Economic Trends, pp. iii - August 1988 for details.)

For unadjusted CPI data, the series is compiled by the Department of Labor. The sources for the data are the Survey of Current Business, and the Department of Labor Publications entitled Monthly Labor Review and Consumer Price Index Detailed Report. There have been 3 recent base year updates in January of 1971, 1988, and 1995. These base year updates coincide with benchmark updates. Current releases, and detailed information regarding this series are available at the Bureau of Labor Statistics' CPI website (http://stats.bls.gov/cpifact8.htm).

The remaining variables, the target federal funds rate and the effective federal funds rate (1979:10 to 1998:04), are available from the Federal Reserve Bank of St. Louis and the Board of Governors of the Federal Reserve, respectively. The target federal funds rate is the Board of Governors' announced target for the overnight rate on interbank loans - which is revealed following each Federal Open Market Committee (FOMC) meeting. Generally FOMC meetings are held 8 times a year, except in special circumstances. We constructed a monthly dataset for this series which corresponds to the timing of our IP and CPI datasets. To do this, we assign the appropriate federal funds rates to each month, given the release dates of the IP and CPI data, which are often the 15th of each month, and almost always between the 14th and 17th. Therefore, we align each announced change in the target federal funds rate with the latest months' release of IP and CPI data that was available to the Board of Governors at the time of their decision. The notation that we use reflects these timing issues. If we denote some particular target federal funds rate as r_t, the preliminary data release for this observation is denoted $X_{t-1}(\tau)$, where $\tau = (t-2)$, i.e. at the time of the Board's decision, t, the most recent information available - released at time $(t-1)$ - pertains to the calendar date $(t-2)$. In this way, it is the availability of information that defines our "calendar time". The effective federal funds rate is the actual overnight rate on interbank loans. Again, we construct a monthly dataset for this series that corresponds to our real-time datasets. In this case, for each month we calculate the average of the effective federal funds rate for the 4 weeks following the release date of IP and CPI data.

In order to construct operational real-time measures of inflation and of the output gap, we begin by assuming that our data are "final" after 23 revisions, so that our variables are assumed not to change appreciably beyond the 24th release. In our analysis we consider first and twelfth differences of CPI. In particular, we form:
$\pi_t(\tau-1) = 1200(log(CPI_t(\tau-1)) - log(CPI_t(\tau-2))))$ or
$\pi_t(\tau-1) = 100(log(CPI_t(\tau-1)) - log(CPI_t(\tau-13))))$,
where $\tau = t, t-1, t-2, ...$ We consider annualized monthly inflation because this corresponds to the standard data transformation used in many empirical studies. On the other hand, twelfth differenced CPI data are used in order to facilitate a comparison of real-time policy simulation outcomes based on either seasonally adjusted or seasonally unadjusted data. As τ varies, we essentially construct two new inflation series for each release of CPI data. Our output gap variable is the deviation of industrial production from its loglinear trend. Industrial production trend regressions are estimated using samples of 6-12 years of data, and as our empirical results were found to be qualitatively similar in all cases, we use 9 years of data for output gap calculations hereafter. In addition, we use real-time data in all trend regressions, and include seasonal dummy variables when constructing gap estimates based on seasonally unadjusted data. Thus, we construct the output gap as:
$y_t(\tau-1) = 100(log(IP_t(\tau-1)) - log(\widehat{IP}_t(\tau-1)))$,
where $\widehat{IP}_t(\tau-1)$ is the forecast of trend output based on information available at time $t-1$, and $\tau = t, t-1, t-2, ...$

4 Empirical Findings

In order to facilitate the comparison of our rules, we report four different types of results. In particular, we examine (1) ex-ante mean square forecast error and mean absolute percentage error, (2) turning point predictability, (3) parameter estimates, and (4) model specification. As dicussed above, our approach is to simulate a real-time policy setting environment, thereby mimicing the behavior of private agents forming expectations in real-time, or the behavior of government policy-makers. To do this we consider ex-ante policy evaluation periods of 50 and 100 months. In addition to our two policy evaluation periods, we examine seasonally adjusted and unadjusted data, output gap measures based on 24^{th} vintage data and real-time data and the target and effective federal funds rate. Thus there are 32 different permutations of data which can be used in the construction of tables associated with (1) - (4) above.

In order to streamline the presentation of our findings, we omit various results that are either uninteresting, or comparable to other findings which we do report. First, we report results for a 50 month evaluation period with the effective federal funds rate and a real-time output gap measure. Results for the 100 month evaluation period, the target federal funds rate, and the 24^{th} vintage based output gap measure are qualitatively similar and are available upon request from the authors. Second, we report results only for Rules 1-5. Although ex-ante mean square error results, for example, are always better when lagged values of the policy instrument are used, our findings *across* rules, windows, and vintages remain qualitatively similar, and hence tabulated results for Rules 6-9 are omitted for the sake of brevity. Third, we omit tabular evidence based on first differenced data. As noted above, first differences were only used to examine seasonally adjusted data. We omit these results because our findings based on twelfth differences are superior to comparable results based on first differences. This finding is attributable to the fact that first differenced real-time data are considerably more noisy than real-time data based on twelfth differences. This is because the former type of data involve first and second vintages while the latter involve first and twelfth vintages, and twelfth vintage data are relatively more accurate than second vintage data.[6] Finally, our evidence based on absolute percentage error loss measures does not differ from the evidence which we offer based on mean square error loss measures, and so is omitted for the sake of brevity.

[6]This observation has been made previously by Maravall and Pierce (1983) in the context of the noisiness of preliminary seasonally adjusted data.

Tables 2 through 8 summarize our main empirical findings. In order to get a feel for the data which we use in our analysis, and in particular the importance of the revision process, we begin by providing various summary statistics which are presented in Table 2a. In this table, we report on two types of data revisions, called fixed and increasing width revisions. The fixed revisions are constructed as $X_{t+i}(t) - X_{t+i-1}(t)$, and the increasing width revisions are $X_{t+i}(t) - X_t(t)$, for $i = 1, \ldots, 11, 18$, and 24. Table 2a contains three panels, corresponding to adjusted and unadjusted industrial production as well as adjusted CPI, for which mean, variance, skewness and kurtosis figures are reported.[7] Observe that the mean of fixed length revisions is significantly different from zero at a 95 percent level of confidence, for numerous vintages of data. This suggests there is systematic bias in revisions of our variables, and such information could in principle be used to increase the accuracy of preliminary releases. Notice also that based on increasing width revisions, the difference between "final" (i.e. i=24) and initial releases of data has mean bias which is significantly different from zero. This implies that a statistically significant correction could be made to all releases of the variables, prior to their final release. Finally, the skewness and kurtosis statistics reported in Table 2a suggest that data revisions are characterized by clear departures from normality. This may be due to the presence of outliers in the revision process, which implies that real-time policy setting based on recent releases of data may result in policy decisions which are quite different from those which would have been made, had we known the "final" data. Thus, the data revision process may be quite important for policy setting. A comprehensive analysis of the data revision process for our real-time datasets which includes plots of revision autocorrelation functions and discussion of the pattern of outliers in the data, for example, is given in Swanson, Ghysels and Callan (1997). As we use inflation and output gap variables in our subsequent analysis, we also present basic statistics which are analogous to those given in Table 2a for π and y (see Table 2b). The results based on Table 2b are largely similar to those discussed above for the raw data, with the notable exception that the incidence of significant mean revisions given in the second column of the table is reduced. However, note that the absolute magnitude of these revisions is actually quite large, suggesting that this finding may be due to the small sample sizes used in the construction of the statistics.

The rest of our empirical findings are presented in Tables 3 through 8. Our findings with

[7] Table 2 does not contain summary statistics for revisions in seasonally unadjusted CPI data, as revisions occur infrequently, and primarily in conjunction with base year benchmark revisions.

regard to data uncertainty are reported in Tables 3-4 for seasonally adjusted data and Tables 6-7 for seasonally unadjusted data. Tables 5 (adjusted data) and 8 (unadjusted data) summarize our findings with regard to parameter and model uncertainty. In Tables 3 and 6 we report mean square errors (MSEs) associated with 50 real-time policy decisions.[8] The first column of the table gives the vintage of the data. The calendar date of the data used in all rules is $t-1$, except for Rule 5 in which a mixture of calendar dates is used (see above). The second column in the table reports MSEs for Taylor's Rule (Rule 1). The next three panels report MSEs for Rules 2-5, where response coefficients and/or rule specifications vary, and where the "window" of observations used by policy setters is fixed, increasing, or nonexistent. Tables 4 and 7 are presented in the same way, except that so-called "Confusion Rates" are reported rather than MSEs. The confusion rate indicates the proportion of times that our policy decision correctly predicts the direction of change in realized interest rates. Thus, a value of 0.50 corresponds to a policy rule which captures directional changes in interest rates so poorly that were we to flip a coin, we could do equally well.[9] Finally, Tables 5 and 8 report the average and standard deviation of the response coefficients associated with the target variables when Rule 2 is the policy tool. As in the previous tables, the first column of these tables reports the vintage of data used to implement the rule.

Based on the results in Tables 3-8, our findings can be summarized as follows:

Vintage matters

This finding is supported by noting that MSE values in Tables 3 and 6 are dependent upon vintage. In Tables 3 and 6, lowest MSE values are boldfaced for each rule. For instance, in Table 3, Rule 1, the MSE varies from 1.439 to 4.447 depending on vintage. Moreover, MSEs are not monotonically increasing as vintage increases. Thus, it is not clear whether *only* using "final" data produces optimal forecast models. In fact, notice that the lowest MSEs in Tables 3 and 6 are associated with a *Fixed Window* of data, and occur for relatively recent vintages. For example, for Rule 4

[8] In the tables, the end of our sample is May, 1996, so that we do not exhaust our entire dataset, which ends in March, 1998. The reason for this is that we also carried out ex-post policy simulations (assuming that *finalized* data were known) in order to assess the accuracy of our ex-ante results (see below).

[9] A case could be made for using the target rather than effective federal funds rate when simulating confusion rates. Unlike the federal funds rate, the target rate contains a small number of discrete level shifts which are directly associated with policy decisions, while the effective rate is a smoothed version of the target rate which might also be impacted by market forces.

with the $Fixed\ Window$ of data, the lowest MSE vintage is $t-11$ for adjusted data and $t-10$ for unadjusted data. In addition, notice that for the same rule and the $Increasing\ Window$ of data, the analogous lowest MSE vintages are $t-2$ and $t-3$. Thus, we have evidence that preliminary data are useful.[10] It is important to note that the "MSE-best" vintage varies across rules. This is expected, given that our rules exploit the real-time information set differently. For example, it should be expected that the MSE-best vintage for Rule 2 is higher than the MSE-best vintage for Rule 3. This is indeed the case for all of our results, and follows because Rule 3 uses all information from vintage m back, say, while Rule 2 only uses information for a given vintage.

Adaptive least squares learning yields improved results

Based on both seasonally adjusted and unadjusted data, the lowest MSE value for Rule 1 is higher than the lowest MSE values (except for Rule 2) when the $Fixed$ and/or $Increasing\ Window$ of data used (see Tables 3 and 6). This suggests that adaptive model formation is useful. Also, in the $No\ Window$ case, where response coefficients are not updated, MSEs are actually worse than in Rule 1, where no estimation is done whatsoever. Thus, one might conclude based on this finding that "calibration" (i.e. Rule 1) is better than naive estimation (i.e. the $No\ Window$ case), but worse than adaptive least squares learning (i.e. the $Fixed$ and $Increasing\ Window$ cases).

Dynamic information sets are useful

In Tables 3 and 6, when comparing $Fixed$ and $Increasing\ Window$ cases, Rules 3 and 4 always "MSE-dominate" Rule 2. In fact, MSE-best values are reduced by more than 50% in the $Fixed\ Window$ case when Rules 2 and 3 are compared. For example, based on unadjusted data (Table 6), the MSE-best value is 0.534 for Rule 2 and 0.149 for Rule 3, corresponding to a 72% increase in precision when dynamic information sets are used.[11] This finding provides evidence that policy rules based on distributed lag polynomials of target variables may outperform simpler rules (see e.g. Rotemberg and Woodford (1997,1998) for further evidence).

Further evidence that dynamic information sets are useful is given in Figure 1, where the ex-ante performance of Rules 1-4 for a $Fixed\ Window$ of data is graphically illustrated. Even casual

[10] This does not mean, that final data are not useful (except in the case of Rule 2), but rather that the dataset which is "MSE-best" must contain preliminary data, while it *may* also contain earlier vintages, including final data.

[11] Note also that the lower MSE in this case is obtained based on the use of an earlier vintage of data.

observation of the graphs is sufficient to see that Rules 3 and 4 perform better than Rule 2, and Rules 2-4 perform better than Rule 1. One way to check whether the graphical evidence provided in Figure 1 is indicative of our policy rules being useful in practice (e.g. in the construction of forecasts and to establish monetary authority and hence monetary policy credibility) is to compare the ex-ante performance of the rules with analogous ex-post performance. By ex-post performance, we mean that only final revised data are used in the rules. We implement our ex-post analysis by using data released in April, 1998. Thus, all data in our ex-post sample prior to May, 1996 is final, while newer data has been revised fewer than 23 times, and hence is not final, according to our definition. Notice that this type of ex-post analysis is currently the norm rather than the exception in empirical economics, as data which have been revised many times are routinely downloaded from CITIBASE, for example, with no regard to the fact that these data, although representative of information at time t, say, were actually not available until time $t+i$ for i large. Use of datasets of this sort clearly invalidates any claim that later empirical findings are representative of the real-time flow of events in the economy. Nevertheless, a comparison of our ex-ante policy simulation results with an ex-post policy simulation should yield evidence concerning the usefulness of ex-ante rules. Such evidence is given in Figure 2, where it is apparent the performance of our ex-ante rules is essentially as good as an analogous set of ex-post rules where finalized data are assumed known. This result holds across rules and window specifications. (Results for windows other than the *Fixed Window* reported on in Figure 2 are available upon request from the authors.) One reason for the seemingly excellent ex-ante performance of the rules relative to their ex-post counterparts is that adaptive least squares is used, so that response coefficients are allowed to "adjust" to the type of data used.

Real-time specification flexibility pays off

Rule 4 is the MSE-best rule based on both adjusted and unadjusted data. Thus, learning is useful in two respects. First, updating response coefficients is better than fixing coefficients (*No Window* versus *Fixed* and *Increasing Window* cases). Second, updating the order of the lag polynomial (model uncertainty) in real-time is better than fixing the specification (Rule 4 versus Rules 1,2,3, and 5).

Revision bias is not as important as the use of real-time data

This finding is based upon a comparison of MSE-best values for Rules 3 and 5 in Tables 3 and 6, where it is clear that Rule 3 is MSE-best. Notice that Rule 3 is closest to Rule 5 as neither incorporates specification flexibility. Thus, Rule 5 directly augments Rule 3 by including the history of data revision in addition to different vintages of data. One caveat to this finding is that we cannot actually claim that the use of real-time data is more important than revision bias, rather we can say that given real-time data, revision history adds nothing in terms of MSE.

Seasonally unadjusted data are better

This may be surprising as it is often argued that seasonal adjustment filters extract the "relevant" component of the data. If a time series is viewed as the sum of different unobserved components, then it is argued that the seasonal component is "irrelevant" (see e.g. Ghysels (1994)). However, what is often not recognized is that seasonal adjustment entails more frequent and larger data revisions than are associated with unadjusted data (see e.g. Maravall and Pierce (1983)). Since all of our data are annual differences, the numerical entries in Tables 3 and 6 are directly comparable. Even cursory examination of these tables reveals that the MSE-best unadjusted rules dominate the corresponding MSE-best adjusted rules. For example, the *overall* MSE-best value in all of our tables is 0.124, which obtains when unadjusted data, Rule 4, and the *Fixed Window* of data are used. The corresponding entry based on adjusted data is 0.244, which roughly twice that based on unadjusted data. A further assessment of the noisiness of preliminary seasonally adjusted data is obtained when one examines Rule 2, which does not incorporate lagged polynomial information sets, and which is clearly our simplest adaptive rule. In the *Fixed Window* case, a MSE of 1.630 obtains with vintage $t-2$ when unadjusted data are used. The analogous figure based on adjusted data is 2.033. Furthermore, the MSE based on adjusted data does not decrease to 1.630 until vintage $t-8$, suggesting that many revisions are necessary in order to smooth out the noisiness associated with seasonally adjusted data. Thus, using unadjusted data in policy decisions is MSE-preferable. Moreover, unadjusted data are directly available.[12]

Confusion rate findings based on seasonally adjusted data are confused

Notice that all of our findings with respect to rule and data window selection and based on unadjusted data are the same when either MSE or confusion rate loss is used (see Tables 6 and 7).

[12] It is also worth noting that the Sims (1974) and Wallis (1974) results on estimation with filtered data do not apply in our context, as interest rates are unfiltered so that seasonal adjustment impacts response coefficient estimates.

However, analogous findings based on adjusted data and confusion rate loss do not agree with comparable MSE results (see Tables 3 and 4). In particular, the "confusion rates-best" models (i.e. boldfaced entries) associated with Rule 1 in Table 4 are often lower that those associated with other rules in Table 4, regardless of which data window is used. This is clearly not the case with unadjusted data. If these sorts of conflicting findings occurred regardless of data adjustment, one might be led to believe that either (a) the choice of loss function is crucial or (b) our empirical findings are not robust. Indeed, there is much evidence in the economic forecasting literature that (a) is important (see e.g. Granger (1969) and Leitch and Tanner (1991)). However, given that our results are robust to evaluation period and output gap specification, it seems more likely that the problem lies with the use of seasonally adjusted data. For example, it is known that filters of various types alter the comovements between economic series (see e.g. Canova (1998)). In summary, our findings are consistent with a conjecture that turning point prediction is more sensitive to filtering than is MSE-based prediction. This is sensible if one believes that unusual (or one-time) events drive turning points, as these types of events tend to manifest themselves as observational outliers, which trigger smoothing corrections in standard adjustment programs (e.g. X-11 and X-12, see Findley, Monsell, Bell, Otto and Cheng (1998) and Ghysels, Granger and Siklos (1996)).

Response coefficients are weakly sensitive to the choice of vintage

Our results with respect to response coefficient sensitivity are gathered in Tables 5 (adjusted data) and 8 (unadjusted data). The layout of these tables is slightly different from the other tables. We take Rule 2 as a starting point and select the MSE-best vintage for each data window (called m_{opt}). We then form rules as follows:

$$R_t = \hat{a}_0 + \hat{b}_1 \pi_{t-m_{opt}+j}(t - m_{opt}) + \hat{c}_1 y_{t-m_{opt}+j}(t - m_{opt}), \tag{2}$$

where j = 1,..., 24. We use this rule instead of Rule 2 as it allows us separate out the impact of parameter uncertainty from that of specification uncertainty - both are mixed together if one examines parameter evolution based on Rule 2. In particular, direct use of Rule 2 mixes parameter and specification uncertainty because the calendar lags used in the rule depend on the vintage. The mean of the response coefficients given in Tables 5 and 8 varies smoothly, relatively monotonically, and with little variability across j. Note that output response coefficients do not depend critically on the data window for seasonally unadjusted data (see Table 8), but are more sensitive to data window when adjusted data are used (see Table 5). Inflation response coefficients, however, appear to depend critically on the data window used, again with higher variability across window being associated with seasonally adjusted data. Thus, while parameter uncertainty is not prevalent, response coefficients depend on the data window used, and hence on rule design.

Patience pays off

Rules and models based soly on preliminary data do not minimize mean square forecast error (MSE) risk. For example, our MSE-best rule (Rule 4, $Fixed\ Window$, unadjusted data) performs much better when one waits for 9 before using new data (i.e. a 58% MSE reduction, from 0.298 to 0.124). This suggests that preliminary data should be used with care by both policy-makers and private agents, and such data should perhaps be "down-weighted" when used to revise and update models and coefficient estimates.

5 Conclusions

In this paper we have examined model and data uncertainty within the context of monetary policy rules. Particular emphasis is placed on two related but different issues, namely (i) model uncertainty viewed through the lens of parameter and specification uncertainty, and (ii) the availability and timing of data with which to examine and implement policy rules. In order to carry out our analysis we built a number of large real-time datasets and carried out a series of experiments within the context of a real-time policy setting environment. Within the context of linear policy forecast models we find that data vintage (or release) is important, and adaptive least squares learning based methods are preferable to simpler model formation strategies. In addition, noise produced by seasonal adjustment filtering is prohibitively large when measured in a variety of different ways, prompting us to conclude that the use of seasonally unadjusted data is preferable to the use of seasonally adjusted data. Finally, it appears to be in the best interest of forecasters to wait until some of the data uncertainty associated with preliminary data has been removed by the revision process before updating adaptive forecast models, particularly if the objective is to forecast "final" data.

In summary, we believe that the importance of using real-time data has not yet been fully recognized in mainstream empirical economics, and many empirical techniques rely too heavily on the presumption that economic data are final and readily available. The construction of monetary policy rules is only one example, although a very important one, of how data and model uncertainty become relevant when we attempt to gather together empirical findings which are representative of the real-time flow of events in the economy.

6 References

Amato, J. and N.R. Swanson, (1999), The Real-time (In)Significance of M2, Working Paper, Texas A&M University.

Anderson, E., L.P. Hansen and T.J. Sargent, (1997), Robustness and Risk-Sensitivity in General Equilibrium, Working Paper, University of Chicago and Hoover Institution.

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

Bernanke, B. and I. Mihov, (1998), Measuring Monetary Policy, Quarterly Journal of Economics, 113, 869-902.

Bray, M., (1982), Learning Estimation and the Stability of Rational Expectations, Journal of Economic Theory, 26, 318-339.

Bray, M. and D. Kreps, (1987), Rational Learning and Rational Expectations, in G. Feiwel eds. Arrows and the Ascent of Modern Economic theory, New York University Press, New York.

Bryant, R.C., P. Hooper and C. Mann (1993), Evaluating Policy Regimes: New Research in Empirical Macroeconomics, Brookings Institution, Washington, D.C.

Canova, F., (1998), Detrending and business cycle facts Journal of Monetary Economics, 41, 475-512.

Chung, H., (1990), Did Policymakers Really Believe in the Phillips Curve? An Econometric Test, Ph.D. Dissertation, University of Minnesota.

Croushore, D., (1993), Introducing: The Survey of Professional Forecasters, Federal Reserve Bank of Philadelphia Business Review, November-December, 3-15.

Croushore, D. and T. Stark (1999): "A Real-Time Data Set for Macroeconomists", Working Paper, Federal Reserve Bank of Philadelphia.

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.

Leitch G. and J.E. Tanner, (1991), Economic Forecast Evaluation: Profit Versus The Conventional Error Measures, American Economic Review, 81, 580-590.

Estrella, A. and F. Mishkin, (1998), Rethinking the Role of NAIRU in Monetary Policy: Implication of Model Formation and Uncertainty, Discussion Paper, Columbia University.

Fair, R. and R. Shiller, (1990), Comparing Information in forecasts From Econometric Models, American Economic Review, 80, 375-389.

Findley, D. B. Monsell, W.R. Bell, M. Otto and B.-C. Chen, (1998), New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program Journal of Business and Economic Statistics, 16, 127-152.

Frankel J. and M. Chinn, (1995), The Stabilizing Properties of a Nominal GNP Rule, Journal of Money, Credit and Banking, 27, 318-334.

Friedman, M., (1968), The Role of Monetary Policy, American Economic Review, 58, 1-17.

Fuhrer, J., (1997), Inflation/Output Variance Tradeoffs and Optimal Monetary Policy, Journal of Money, Credit and Banking, 29, 214-234.

Fuhrer, J. and J. Moore, (1995), Monetary Policy Trade-Offs and the Correlation Between Nominal Interest Rates and Real Output, American Economic Review, 85, 219-239.

Ghysels, E., (1987), Seasonal Extraction in the Presence of Feedback, Journal of Business and Economic Statistics, 5, 191-194.

Ghysels, E., (1994), On the Economics and Econometrics of Seasonality in C.A. Sims (ed.) Advances in Econometrics - Sixth World Congress, Vol. I, 257-316, Cambridge University Press.

Ghysels, E. C.W.J. Granger and P. Siklos, (1996), Is Seasonal Adjustment a Linear Data Transformation Process?, Journal of Business and Economic Statistics, 14, 374-386.

Granger, C.W.J., (1969), Prediction With a Generalized Cost of Error Function, Operational Research Quarterly, 20, 451-468.

Granger, C.W.J. and M. Deutsch, (1992), Comments on the Evaluation of Policy Models, Journal of Policy Modelling, 14, 497-516.

Hansen, E., (1996), Price Level Versus Inflation Targets in an Open Economy With Overlapping Wage Contracts, Working Paper, Reserve Bank of New Zealand.

Hansen, L.P. and T.J. Sargent, (1995), Discounted Linear Exponential Quadratic Gaussian Control, IEEE Transactions on Automatic Control, 40, 968-971.

Hansen, L.P. and T.J. Sargent, (1998), Discounted Robust Filtering and Control in the Frequency Domain, Working Paper, University of Chicago and Hoover Institution.

Henderson, D. and W. McKibbin, (1993), A Comparison of Some Basic Monetary Policy Regimes for Open Economies: Implications of Different Degrees of Instrument Adjustment and Wage Persistence, Carnegie-Rochester Conference Series on Public Policy, 39, 221-317.

King, R.G. and A.L. Wolman, (1996), Inflation Targeting in a St. Louis Model of the 21st Century, The Federal Reserve Bank of St. Louis Review, 78, 83-107.

Kuan, C.-M. and H. White, (1994), Adaptive Learning with Nonlinear Dynamics Driven by Dependent Processes, Econometrica, 62, 1087-1114.

Leitch, G. and J.E. Tanner, (1991), Economic forecast Evaluation: Profits Versus the Conventional Error Measures, American Economic Review, 81, 580-590.

Levin, A., V. Wieland and J.C. Williams, (1998), Robustness of Simple Monetary Policy Rules Under Model Uncertainty, Working Paper, Federal Reserve Board.

Lucas, R.E. (1972), Econometric Testing of the Natural Rate Hypothesis, in: A. Eckstein eds., The Econometrics of Price Determination: Conference, October 30-31, 1970, Board of Governors of the Federal Reserve System, Washington.

Maravall, A. and D. Pierce, (1983), Preliminary-Data Error and Monetary Aggregate Targeting, Journal of Business and Economic Statistics, 1, 179-186.

Maravall, A. and D. Pierce, (1986), The Transmission of Data Noise Into Policy Noise in U.S. Monetary Control, Econometrica, 54, 961-979.

Marcet, A. and J. Nicolini, (1997), Recurrent Hyper Inflations and Learning, Working Paper, Pompeu Fabra University.

Marcet, A. and T.J. Sargent, (1989a), Convergence of Least Squares Learning Mechanisms in Self Referential, Linear Stochastic Models, Journal of Economic Theory, 48, 337-368.

Marcet, A. and T.J. Sargent, (1989b), Convergence of Least Squares Learning in Environments With Hidden State Variables and Private Information, Journal of Political Economy, 97, 1306-1322.

Marimon, M., (1997), Learning From Learning in Economics, in: D. Kreps and K. Wallace eds. Advances in Economics and Econometrics: Theory and Applications, Volume 1, Cambridge University Press, Cambridge.

Matheny, K., (1996), Output Targeting and an Argument for Stabilization Policies, Working Paper, Purdue University.

McCallum, B.T., (1993), Discretion Versus Policy Rules in Practice, Two Critical Points: A Comment, in: Carnegie-Rochester Conference Series on Public Policy, 39, 215-220.

McCallum, B.T., (1997), Issues in the Design of Monetary Policy, Working Paper, Carnegie Mellon University.

Orphanides, A., (1997), Monetary Policy Rules Based on Real-Time Data, Working Paper, Board of Governors of the Federal Reserve System.

Phelps, E., (1967), Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time, Economica, 2, 22-44.

Rotemberg, J.J. and M. Woodford, (1997), An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy, NBER Macroeconomics Annual, 12, 297-346.

Rotemberg, J.J. and M. Woodford, (1998), Interest-Rate Rules in an Estimated Sticky Price Model, Discussion Paper, Harvard Business School.

Sargent, T.J., (1989), Two Models of Measurements and the Investment Accelerator, Journal of Political Economy, 97, 251-287.

Sargent, T.J., (1998), The Conquest of American Inflation, Marshall Lecture Series, The University of Cambridge.

Sargent, T.J. (1998), Discussion of "Policy Rules for Open Economies" by Laurence Ball, Working Paper, University of Chicago and Hoover Institution.

Sims, C.A. (1974), Seasonality in Regression, Journal of the American Statistical Association 69, 618-626.

Swanson, N.R., (1998), Money and Output Viewed Through A Rolling Window, Journal of Monetary Economics, 41, 455-474.

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: *Festschrift in Honor of Clive W.J. Granger*, eds. R. Engle and H. White, Oxford: Oxford University Press, 45-75.

Swanson, N.R. and H. White, (1995), A Model Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks, Journal of Business and Economic and Statistics, 13, 265-275.

Swanson, N.R. and H. White, (1997), A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks, Review of Economics and Statistics, 79, 540-550.

Taylor, J.B., (1979), Estimation and Control of an Econometric Model with Rational Expectations, Econometrica, 47, 1267-1286.

Taylor, J.B., (1993a), Discretion versus Policy Rules in Practice, Carnegie-Rochester Conference Series on Public Policy, 39, 195-214.

Taylor, J.B., (1993b), Macroeconomic Policy in the World Economy: From Econometric Design to Practical Operation, W.W. Norton, New York.

Trivellato, U. and E. Rettore, (1986), Preliminary-Data Errors and Their Impact on the Forecast Error of simultaneous-Equations Models, Journal of Business and Economic Statistics, 4, 445-453.

Wallis, K.F. (1974), Seasonal Adjustment and Relations Between Variables, Journal of the American Statistical Association 69, 18-31.

Woodford, M.M., (1990), Learning to Believe in Sunspots, Econometrica, 58, 277-308.

TABLE 1

Table 2a: Real-Time Dataset Summary Statistics

Seasonally Adjusted Industrial Production: 1979:2-1996:5

	FIXED WIDTH REVISIONS				*INCREASING WIDTH REVISIONS*			
	Mean	*Variance*	*Skewness*	*Kurtosis*	*Mean*	*Variance*	*Skewness*	*Kurtosis*
i=1	0.076	0.398	13.79	19.59	0.076	0.398	13.79	19.59
i=2	0.096	0.400	13.94	19.87	0.171*	0.848	9.776	9.846
i=3	0.023	0.410	14.08	20.04	0.226*	0.899	9.790	9.837
i=4	0.015*	0.475	14.11	20.21	0.221*	0.887	9.814	9.927
i=5	0.014*	0.409	14.11	20.93	0.218*	0.806	9.800	9.620
i=6	0.014*	0.469	14.10	20.69	0.224*	0.781	9.763	9.147
i=7	0.014*	0.486	14.04	20.51	0.231*	0.829	9.795	9.686
i=8	0.015*	0.442	14.06	20.98	0.226*	0.804	9.784	9.446
i=9	0.015*	0.407	14.07	20.12	0.232*	0.820	9.791	9.606
i=10	0.015*	0.429	14.08	20.33	0.227*	0.827	9.786	9.543
i=11	0.005	0.480	14.10	20.68	0.225*	0.801	9.788	9.537
i=12	0.014*	0.426	14.11	20.20	0.225*	0.803	9.811	9.894
i=18	0.015*	0.434	13.99	19.95	0.230*	0.840	9.770	9.448
i=24	0.014	0.444	13.65	19.37	0.220*	0.821	9.718	9.853

Seasonally Unadjusted Industrial Production: 1979:2-1996:5

	FIXED WIDTH REVISIONS				*INCREASING WIDTH REVISIONS*			
	Mean	*Variance*	*Skewness*	*Kurtosis*	*Mean*	*Variance*	*Skewness*	*Kurtosis*
i=1	-0.191*	4.444	-0.145	3.454	-0.191*	4.044	-0.145	3.454
i=2	-0.058*	0.110	1.347	28.31	-0.207*	4.126	-0.142	3.404
i=3	-0.016	0.078	-6.073	81.67	-0.216 *	4.390	-0.413	4.213
i=4	-0.025	0.357	-8.080	107.2	-0.146	4.247	-0.075	3.547
i=5	0.046*	0.288	13.07	180.4	-0.184 *	4.060	-0.154	3.435
i=6	0.007	0.017	12.33	169.2	-0.177 *	4.077	-0.153	3.417
i=7	0.014*	0.032	13.59	191.5	-0.176 *	4.087	-0.150	3.406
i=8	0.015*	0.031	12.10	179.0	-0.180 *	4.075	-0.150	3.421
i=9	0.012*	0.022	13.08	180.3	-0.179 *	4.071	-0.154	3.419
i=10	0.013*	0.026	12.43	167.4	-0.182 *	4.078	-0.161	3.435
i=11	0.009	0.018	12.75	174.5	-0.183 *	4.093	-0.167	3.454
i=12	0.009	0.026	11.54	156.7	-0.181 *	4.064	-0.151	3.437
i=18	<0.001	0.001	-12.56	99.04	-0.205 *	4.121	-0.163	3.408
i=24	-0.001	0.043	-13.73	194.0	-0.220 *	4.222	-0.232	3.596

Seasonally Adjusted Consumer Price Index: 1987:2-1996:5

	FIXED WIDTH REVISIONS				*INCREASING WIDTH REVISIONS*			
	Mean	*Variance*	*Skewness*	*Kurtosis*	*Mean*	*Variance*	*Skewness*	*Kurtosis*
i=1	-0.003*	<0.001	-3.880	26.33	-0.003*	<0.000	-3.880	26.33
i=2	-0.003*	<0.001	-2.972	21.18	-0.007*	0.001	-2.391	11.36
i=3	-0.004*	<0.001	-3.576	21.64	-0.006*	0.001	-2.505	12.43
i=4	-0.003*	<0.001	-3.839	25.53	-0.004*	<0.000	-3.274	18.60
i=5	-0.001*	<0.001	-7.105	52.06	-0.003*	0.001	-1.694	11.96
i=6	<0.001	<0.001	-0.551	18.39	-0.003*	0.001	-1.814	12.86
i=7	-0.001	<0.001	-0.743	22.04	<0.000	0.001	0.439	15.16
i=8	0.003*	<0.001	3.904	30.62	-0.001	0.001	-0.169	16.37
i=9	0.002*	<0.001	4.502	38.60	<0.000	0.001	0.325	16.32
i=10	0.003*	<0.001	4.560	36.19	-0.002	0.001	-0.410	13.74
i=11	0.001	0.001	0.241	23.66	-0.004*	0.001	-1.577	14.09
i=12	-0.001	0.001	-1.865	25.46	-0.006*	0.001	-3.363	16.47
i=18	-0.002*	<0.001	-3.018	21.14	-0.006*	0.001	-2.837	13.68
i=24	<0.001	<0.001	-5.155	22.52	-0.003*	<0.000	-3.880	26.33

Notes: * denotes a mean value that is significantly different from zero based on a 95% confidence interval constructed using a heteroskedasticity and autocorrelation consistent variance estimator. Fixed width revisions are constructed as $ln(X_{t+i-1}(t-2))-ln(X_{t+i-2}(t-2))$ and increasing width revisions are constructed as $ln(X_{t+i-1}(t-2))-ln(X_{t-1}(t-2))$, where X is either IP or CPI. In these definitions, the subscript refers to the release date of the data, while the bracketed index denotes the the date to which the release pertains (see Table 1). For example, the "$i-2$" rows in the table correspond to $ln(X_{t+1}(t-2))-ln(X_t(t-2))$ for the fixed width revisions. In this case, the second release for the period t is subtracted from the third release for period t. For increasing width revisions, the "$i=2$" rows correspond to $ln(X_{t+1}(t-2))-ln(X_{t-1}(t-2))$, so that the first release (or first available data) is subtracted from the third release for the period $t-2$. The sample period for which we present summary statistics is determined by our real-time simulation experiments. The period reported on for IP is longer than for CPI because additional IP data were needed in order to estimate trend lines for use in output gap construction. Results for our larger sample periods (from 1950 for IP and 1978 for CPI) are qualitatively similar, and are available upon request from the authors. The end period of the data is 1996:5, corresponding to the last interest rate observation used in our subsequent ex-ante analysis. This end date was used (as opposed to the actual end of our sample - 1998:3) in order to facillitate an ex-ante versus ex-post comparison of policy rule performance (see discussion in Section 5).

Table 2b: Vintages of the Output Gap and Inflation, Summary Statistics

Seasonally Adjusted Output Gap: 1990:2-1996:3

	FIXED WIDTH REVISIONS				*INCREASING WIDTH REVISIONS*			
	Mean	*Variance*	*Skewness*	*Kurtosis*	*Mean*	*Variance*	*Skewness*	*Kurtosis*
i=1	-8.675*	540.7	-0.830	5.743	-8.675*	540.7	-0.830	5.743
i=2	-5.631*	384.2	-1.109	9.198	-14.30*	1087	-0.526	4.543
i=3	-4.011*	187.0	-3.779	33.96	-18.31*	1267	-0.646	4.452
i=4	-0.192	122.2	-7.887	86.90	-18.50*	1398	-0.920	5.057
i=5	-0.389	139.0	-8.119	87.28	-18.89*	1625	-1.140	5.991
i=6	-0.873	87.56	-9.754	104.1	-19.77*	1783	-1.234	5.917
i=7	0.097	34.03	-3.703	58.49	-19.67*	1834	-1.193	5.645
i=8	0.162	54.95	-5.178	61.66	-19.51*	1946	-1.135	5.243
i=9	0.070	75.62	2.289	52.15	-19.44	2087	-1.002	4.802
i=10	-0.323	97.71	-0.336	35.42	-19.76	2169	-0.899	4.429
i=11	0.042	211.7	3.069	56.22	-19.72*	2347	-0.763	4.058
i=12	-1.281	151.9	-8.149	79.83	-21.00*	2440	-0.670	3.762
i=18	-0.556	28.02	-1.987	27.74	-25.89*	2632	-0.461	3.379
i=24	-2.474	527.2	-7.995	76.15	-33.06	3513	-0.852	4.130

Seasonally Unadjusted Output Gap: 1990:2-1996:3

	FIXED WIDTH REVISIONS				*INCREASING WIDTH REVISIONS*			
	Mean	*Variance*	*Skewness*	*Kurtosis*	*Mean*	*Variance*	*Skewness*	*Kurtosis*
i=1	-7.375*	480.5	0.630	6.688	-7.375*	480.5	0.630	6.688
i=2	-3.129	587.1	3.501	27.13	-10.50*	1095	1.293	8.607
i=3	-2.018	153.0	1.331	9.921	-12.52*	1165	1.103	7.251
i=4	-4.046*	3228	-10.51	114.8	-16.56*	4571	-6.713	65.89
i=5	6.848	3534	10.37	112.0	-9.721*	1243	1.071	6.916
i=6	1.187	132.8	9.861	103.5	-8.534*	1313	1.006	6.365
i=7	1.960	317.7	9.972	105.1	-6.574	1639	1.477	8.046
i=8	1.914	301.9	10.24	109.7	-4.660	1845	1.512	7.523
i=9	0.416	276.6	5.994	76.21	-4.245	2054	1.407	6.727
i=10	0.887	286.6	8.896	97.76	-3.358	2322	1.434	6.375
i=11	-0.123	328.4	2.660	56.88	-3.481	2729	0.974	5.800
i=12	0.503	337.4	7.622	84.59	-2.978	3125	1.118	5.899
i=18	-1.255	157.7	-10.62	116.0	-9.481	4296	0.444	4.910
i=24	-1.750	204.6	-8.509	77.65	-16.22	4938	0.402	4.271

Seasonally Adjusted Inflation: 1990:2-1996:3

	FIXED WIDTH REVISIONS				*INCREASING WIDTH REVISIONS*			
	Mean	*Variance*	*Skewness*	*Kurtosis*	*Mean*	*Variance*	*Skewness*	*Kurtosis*
i=1	-0.086	1.326	-2.231	31.21	-0.086	1.326	-2.231	31.21
i=2	-0.098	1.460	-2.551	30.02	-0.184	2.769	-1.552	15.14
i=3	-0.455*	3.049	-3.585	14.20	-0.639*	5.650	-1.563	7.050
i=4	-0.266*	1.913	-5.038	27.37	-0.906*	7.220	-1.374	5.282
i=5	0.110	1.575	3.104	36.79	-0.795*	8.997	-0.682	5.055
i=6	-0.110	3.284	-1.314	17.58	-0.906*	12.10	-0.374	3.949
i=7	-0.105	3.101	-1.505	19.79	-1.011*	15.01	-0.208	3.319
i=8	0.196	5.539	1.709	27.36	-0.815	20.95	0.194	3.930
i=9	0.124	3.171	3.810	36.92	-0.692	24.32	0.336	3.670
i=10	0.265	4.620	3.742	27.87	-0.427	29.31	0.429	3.317
i=11	0.297	13.741	3.280	32.00	-0.130	43.31	0.795	4.867
i=12	0.026	8.354	-1.008	23.69	-0.104	51.67	0.545	4.025
i=18	0.102	1.279	10.82	119.0	-0.432	64.37	0.189	3.428
i=24	-0.015	0.029	-10.82	119.0	-0.262	64.00	0.148	3.442

Notes: See notes to Table 2a. Fixed width revisions are constructed as $(\hat{X}_{t+i-1}(t-2) - \hat{X}_{t+i-2}(t-2)) * 100$ and increasing width revisions are constructed as $(\hat{X}_{t+i-1}(t-2) - \hat{X}_{t-1}(t-2)) * 100$, where $\hat{X}$ is either the output gap or inflation. The sample period reported corresponds to the period used in the 50 month policy simulation reported on in subsequent tables.