Econometric Estimation When the "True" Model Forecasts or Errors Are Observed

Eric T. Swanson
Federal Reserve Bank of San Francisco
eric.swanson@sf.frb.org

Abstract

Stochastic disturbance terms in an econometric model encompass two types of error: 1) specification error resulting from an econometric model that is simpler than the "true" economic model, and 2) stochastic innovations to the "true" economic model. It is standard practice to minimize these composite errors to estimate the econometric model. In many interesting cases, however, the "true" model forecasts or errors can be regarded as observed through futures markets, prediction markets, or surveys of professional forecasters. When the true model forecasts or errors are observed, econometric estimation can be improved by minimizing the distance from the econometric model's residuals to the true model errors, rather than to a vector of zeros. This paper derives the theory and applies the method to estimate simple time series models. The error-matching estimation method prescribed by this paper avoids overweighting large model errors that were unforecastable *ex ante*, and reduces standard errors substantially, by about 20–40% for the simple time series examples considered.

JEL Classification: C22, E37, E47, C21

Version 0.5
First Version: November 2000
Current Version: March 23, 2006

I thank Jonathan Wright, Ron Michener, Jon Faust, Alexei Onatski, Jim Hamilton, Andy Levin, Chris Sims, David Bowman, and seminar participants at the Federal Reserve Board, University of Virginia, and 2002 AEA Meetings for helpful discussions, comments, and suggestions. Dorrie Raymond and Joel Wesley provided valuable research assistance. The views expressed in this paper, and all errors and omissions, should be regarded as those solely of the author, and do not necessarily represent those of the individuals or groups listed above, the Federal Reserve System, or its Board of Governors.

1. Introduction

Suppose that a variable of interest, y_t, is determined by the relationship

$$y_t = g(X_t) + \varepsilon_t \tag{1}$$

where g is a function, X_t a vector of variables predetermined at time t, and ε_t a stochastic disturbance term, with $E[\varepsilon_t|X_t] = 0$.[1] We will refer to (1) as the *true* model determining y_t, although we only require that (1) encompass the econometrician's model in a sense that we specify below.

The econometrician, interested in this relationship, will often estimate a model of the form

$$y_t = f(X_t) + \eta_t \tag{2}$$

where f is a function possibly different from (e.g., simpler than) g, and η_t is a stochastic disturbance term.[2] We will refer to (2) as the *econometrician's* model for y_t.

It follows, of course, that the econometrician's disturbance η_t satisfies:

$$\eta_t \equiv (g(X_t) - f(X_t)) + \varepsilon_t \tag{3}$$

The econometrician's error term η_t encompasses both the specification error of the model choice f and the true model's stochastic disturbance ε_t.

There are many reasons why the econometrician may not be able to estimate (1) directly, and thus must settle for estimating (2). First, the exact functional form of g may not be known, or the true model from which g is derived may be very difficult to solve analytically. The econometrician is then forced to use a function f that serves as an approximation to g. Second, even if the functional form g is known, it may be computationally intractable to estimate it—this is often the case for nonlinear systems of equations with forward-looking variables, for instance. In that case, the econometrician must settle for a simpler approximating function f that can be estimated. Third, even

[1] We index all variables by t throughout the paper, and use time series regressions as examples. However, the methods of this paper can just as easily be applied to cross-sectional data. In addition, there is nothing that prevents y_t from being a vector, although for concreteness we will focus on the scalar case in the text.

[2] The econometrician may also choose a different set of right-hand-side variables than those that enter into the true model (1). We can allow for this possibility by redefining X_t to be the union of all the variables included in the true model (1) and the econometrician's model (2).

when the function g itself is known and estimable, limitations of the data or lack of data may prevent estimation of model (1) with all of the appropriate variables on the right-hand side. The econometrician must then settle for a more limited model, and argue (and hope) that the most important variables in the determination of y_t have been taken into account. Fourth, econometric estimates from a simple model of the form (2) are, it is sometimes argued, more robust to possible specification error and imperfections in the data than are more complicated econometric specifications—this argument is often used as a justification for LIML rather than FIML procedures, for example.

It is standard to estimate econometric models of the form (2) by minimizing the distance (in terms of some metric) of the disturbance terms η_t from a vector of zeros (e.g., ordinary least squares, nonlinear least squares, 2SLS, GMM, maximum likelihood). Intuitively, the best econometric model is the one that fits the available data as closely as possible. However, the theories that justify these methods assume no specific knowledge on the part of the econometrician about the stochastic error terms η_t or ε_t.

In this paper, we make the point that in many interesting cases, it may be more realistic to regard the data available to the econometrician as also including the "true" model forecasts, $g(X_t) = E[y|X_t, g]$, or the true model errors, ε_t. For example, Grossman (1989) provides conditions under which financial market prices aggregate the information in the economy to yield the conditional expectation $E[y_t|X_t, g]$ based on the *full* economywide information set $\{X_t, g\}$. Futures markets, prediction markets, and surveys of professional forecasters can all be regarded as potentially providing information about the true model forecast $g(X_t)$ that could be useful to the econometrician, where the "true" model here simply refers to a model g that encompasses the econometrician's model f.

There are several reasons why financial markets or even professional forecasters should be regarded as better able to compute the true model's expectation $g(X_t)$ than the econometrician. First, market participants may have access to a broader or richer collection of data for forecasting y_t than does the econometrician, since many data series are proprietary. Second, the market may have a better approximation of the true model than does the econometrician. For example, market participants may know their own preference parameters (such as risk aversion) or their own production and cost functions. More generally, even though market participants may not know the entire true model (and all relevant data) individually, each may know their own tiny piece of the true model, such as individual cost functions in a competitive market; the decentralized decisions of market participants then yield the correct prices and quantities to solve the aggregate problem, even though each market participant (and the econometrician) may be unable to solve the full model by herself. Third, market participants and professional forecasters may rely on a multitude of non-nested models in making their forecasts, using a substantial element of "judgment" (difficult to model econometrically) in deciding which models are the most appropriate under the given circumstances. Forecasts based on this algorithm, when judgment is well-informed, should outperform the forecasts of any one econometric model.[3] For all of these reasons, good measures of expectations from financial markets or professional forecasters may be much closer to the true model expectations than those otherwise obtainable by the econometrician.

The present paper asks the following question: *Given* information about the true model forecasts $g(X_t)$ or errors ε_t, how could estimation of the econometric model (2) be improved? The answer we propose is to minimize the distance between the econometric model's errors η_t and the true model errors ε_t, in contrast to the standard approach of minimizing the distance between the econometric model's errors η_t and a vector of zeros. This approach necessarily increases the precision of the estimates of the econometrician's model, and in practice the gains seem to be substantial, ranging from 20 to 40 percent in the simple time series exercises below.

A simple example illustrates the idea. Suppose that the true model for y_t is a linear function of a constant, an n-dimensional vector X_t, and a stochastic disturbance term ε_t:

$$y_t = \alpha + \beta X_t + \varepsilon_t \tag{4}$$

where ε_t satisfies the usual Gauss-Markov assumptions. Suppose also that the econometrician knows the form of the true model (4) and observes y_t and X_t for all t, with α and β to be estimated. Finally, assume that the true model errors ε_t are observed by the econometrician.

The usual approach to estimation of (4) is ordinary least squares, which provides consistent and unbiased estimates of the parameters α and β. However, this approach ignores the information that is available to the econometrician in the form of the error

[3] The Federal Reserve Board's forecasts are well described by this algorithm and dominate the forecasts of any one simple model (Sims, 2002, Romer and Romer, 2000).

terms ε_t. The econometrician can gain much greater efficiency by matching the estimated model's residuals to the observed errors ε_t, rather than simply minimizing the size of the model's residuals. Indeed, in this extremely simple example, where the econometrician's model f is the same as the true model g, the true parameters can be estimated perfectly with only $n + 1$ observations!

The point is less extreme, but qualitatively very similar, if we assume that the econometrician is not able to estimate the true model (4) for one of the reasons cited earlier. Intuitively, minimization of the econometric model's residuals (the OLS algorithm) ignores the fact that many shocks that hit the system over the estimation period were large and completely unforecastable within the framework of the model—e.g., terrorist attacks, major hurricanes, stock market crashes, international currency crises, or sudden movements in oil prices. By purging the data of such shocks, one can achieve much more precise estimates of the econometric model.

Of course, one might ask why the econometrician would be interested in the "wrong" model f at all, when she possesses data on the true model's forecasts, $g(X_t)$. There are many reasons. First, the econometrician may be interested in more than just forecasting. For example, the parameter values of the model f may themselves be of interest and shed light on structural features of the economy. Similarly, the econometrician may be interested in hypothesis tests regarding the parameters of the function f. Second, the econometrician may be interested in counterfactual simulations—what would have happened to the U.S. economy if oil prices or terrorists, for example, had not behaved as they did? The true model forecasts $g(X_t)$ provide only actual forecasts, not counterfactual forecasts. Third, even from a forecasting point of view, the true model forecasts $g(X_t)$ may not be available in a timely manner or may not extend into the future as far the econometrician would like. For example, the popular Money Market Services survey of professional forecasters only provides forecasts for statistical releases over the next two weeks, and the econometrician may want to forecast these statistics either several months into the future, or may want to forecast the next month's release more than two weeks prior to its release. The econometrician's model f then fills in the void when the desired "true" model forecast data $g(X_t)$ are not available.

There is a large and growing range of potential applications of the methods in this paper to the macroeconomics and finance literatures more broadly, as the availability of high-quality financial market and forecast data has become more widespread. Although there has been a growing body of work that uses high-frequency financial market data to help identify VARs (e.g., Rudebusch (1998), Brunner (2000), Cochrane and Piazzesi (2002), and Faust, Swanson, and Wright (2003)), these papers have not taken advantage of the information in these futures and forecast series for the *estimation* of those models. Recently, a number of authors estimating macro-finance models of the term structure have also begun incorporating survey-based measures of expectations for the future path of interest rates, output, and inflation into the model (e.g., Bernanke, Reinhart, Sack (2004), Kim and Orphanides (2005), Kim and Wright (2005), Chun (2005)). Again, relatively little attention in these papers is devoted to the question of how estimation of the model ought to change given the information available in the forecast series. Finally, the advent of internet-based "prediction markets" in the past few years has led to an increasing use of data from these markets to investigate political and economic questions (e.g., Wolfers and Zitzewitz (2004), Gürkaynak and Wolfers (2005)). As longer time series of prediction market data become available, interest will naturally turn from forecasting to the estimation of more structural political and economic models, and the question as to how to use the information contained in the prediction market data to efficiently estimate these models will become relevant.

The remainder of the paper proceeds as follows. Section two presents the problem under consideration, the assumptions, and the method in more detail. Section three works through two simple time series examples to demonstrate the practical benefits of the method. Section four provides additional discussion and extensions of the method. Section five concludes.

2. Estimation Assumptions and Methods

We assume that the econometrician is unable to estimate the "true" model (1), and must instead settle for estimating the econometric model (2). Model (1) encompasses model (2) in the sense that:

$$F(y_t|X_t, g, f) = F(y_t|X_t, g) \tag{5}$$

where $F(\cdot|X_t, g, f)$ denotes the distribution function of y_t conditional on X_t and the functional forms g and f, and $F(\cdot|X_t, g)$ denotes the distribution function of y_t conditional on

X_t and g alone. Equation (5) is the key condition that defines what it means for a model to be a "true" model in this paper. Indeed, there is no requirement that the observable forecast data, below, are necessarily generated directly by the true data generating process. It is only required that the observable forecast data be generated by a model that encompasses the econometrician's model.

Second, we assume that the econometrician's model (2) is of interest, despite being a "wrong" model, for one or more of the reasons discussed in the Introduction. In general, the fact that the econometrician is estimating model (2) to begin with should be evidence enough that (2) is considered by the econometrician to be of interest.

Finally, we assume that the econometrician possesses a set of forecast data $g(X_t)$ or forecast errors ε_t from the "true" model (1).

Estimation of models such as (2) is generally well understood under the usual assumption that the econometrician possesses no outside information about the true model errors ε_t. The present paper asks: How should estimation of (2) change when the econometrician in fact possesses such additional outside information?

2.1 Previous Approaches in the Literature

A small but growing number of authors have recognized the availability of data on ε_t through futures markets and other market-related measures of expectations. The paper that is most closely related to the present one is Rudebusch (1998), which explicitly compares residuals from a monetary policy VAR to measures of monetary policy surprises derived from the federal funds futures market. Although the two measures of monetary policy "shocks" are correlated, Rudebusch focuses on the discrepancies between the two to point out the shortcomings of the traditional VAR approach. However, Rudebusch never addresses how estimation of the VAR could be improved given the additional data at hand.

Brunner (1995, 2000), Bagliano and Favero (1998, 1999), and Christiano, Eichenbaum, and Evans (2000) do begin to address this question by augmenting the VAR to include fed funds futures errors (or equivalently, fed funds futures forecasts, since the federal funds rate is already included in the VAR) as one of the variables. The idea behind this approach is the following: the futures market forecasts are derived from a large number of variables, many of which are outside the econometrician's model. Thus, these forecasts provide a one-dimensional proxy for all of the variables that the econometrician has omitted. By augmenting the VAR to include these forecasts, or forecast errors, the econometrician can regain at least some of the information that he has excluded, or does not have available.

While this ε-augmented VAR is a reasonable approach, there is no reason to believe that it is the optimal way for the econometrician to incorporate the "true model" expectations information. For example, the VAR-augmenting method cannot be applied to cross-sectional data, as it prescribes using lags of the ε_t to help forecast future values of the variables of interest.[4] Surely, one would think, true model errors would be informative for regressions in the cross-sectional as well as time series domain.

The ε-augmented VAR approach is also unsatisfying because it is very difficult to put a structural interpretation on the estimated coefficients on the lagged ε_t. For example, in a canonical monetary policy VAR, the estimated coefficients on output and inflation in the Federal funds rate equation can be interpreted as the Federal Reserve's reduced-form responses to those variables. How is one to interpret the coefficients on the ε's? Each error ε_t, particularly the larger ones, tends to correspond to a fairly unique event, such as a terrorist attack, Russia/LTCM-induced "credit crunch", etc. Estimated coefficients on lagged values of the ε_t thus partly represent the Fed's average response to these "outside-the-box" events. But because these events are unique, it is hard to interpret just what this "average response" is capturing. Thus, while the ε-augmented VAR approach might be reasonable from a forecasting perspective, the difficulty in assigning a structural interpretation to the coefficients on the lagged errors ε makes the model much less useful for other purposes.

Finally, even for forecasting, the ε-augmented VAR approach leaves much to be desired. Situations in which the true model errors are available to the econometrician are also typically the same situations in which the true model forecasts are available to the econometrician in real time as well. For example, futures prices and surveys of market expectations are readily available from various sources well before the data are released. If the econometrician's only interest is in forecasting these data series, then the market-based "true" model forecasts will do so much more efficiently than any variation of the econometrician's model, by our assumption that the "true" model is an encompassing

[4] It is also not clear how the VAR-augmenting method should be applied to nonlinear econometric models (2).

model.

2.2 Matching Econometric Residuals to "True" Model Errors

The method proposed in this paper is to match the residuals from the econometrician's model (2) to the "true" model errors ε_t from (1). In contrast to the ε-augmented VAR approach, this method can easily be applied to models of cross-sectional data, to econometric models (2) that are nonlinear, and yields a model with coefficients that are useful for applications other than pure forecasting.[5]

Intuitively, one can think of the proposed error-matching method as a grid search over the entire parameter space for model (2), where the method advocates choosing the values for the parameters that lead to the best match between the implied econometric residuals η_t and the true model errors ε_t. The term "best match" here refers to minimum distance—i.e., the sum (over t) of the squared differences between the η_t and the ε_t in the scalar case.

In practice, it is actually much easier to implement error-matching than by the brute force grid search procedure above. In fact, it can be implemented by a single least squares regression. For example, when the econometric model (2) is linear, so that $y_t = X_t\beta + \eta_t$, ordinary least squares yields the mathematical projection of the vector y onto the column space of X, denoted $\mathcal{S}(X)$. This is the closest point in $\mathcal{S}(X)$ to y, or equivalently, the closest point in $y - \mathcal{S}(X)$ to 0 (which follows because distance is translation-invariant). To match the true model errors, as proposed above, we want to find the closest point in $y-\mathcal{S}(X)$ to ε, the vector of true model errors, instead of the closest point in $y-\mathcal{S}(X)$ to 0. This is identical to the closest point in $\mathcal{S}(X)$ to $y-\varepsilon$. *Put simply, OLS regression of $y-\varepsilon$ onto X produces the desired error-matched estimate of β.* We will denote this estimator by $\hat{\beta}^{EM}$.

When the econometric model (2) is nonlinear, so that $y_t = f(X_t;\beta) + \eta_t$, the same argument applies, although the space $\mathcal{S}(X)$ is no longer the column space of X, but rather the (nonlinear) range that the function f and column vectors of X imply as β varies over all values in its domain. In other words, NLS regression of $y - \varepsilon$ onto X produces the desired error-matched estimate of β for the nonlinear case.

[5] See White (1984) for a broader and more detailed discussion of the estimation and application of "wrong" econometric models.

Estimation of $\hat{\beta}^{EM}$ is thus computationally no more difficult than OLS or NLS, depending on whether the econometrician's model (2) is linear or nonlinear.

2.3 Equivalent Formulations of the Algorithm

It is evident from the discussion above that the true model error-matching procedure is equivalent to regressing (by OLS or NLS) the true model *forecasts,* $g(X_t)$, onto $f(X_t)$.[6] Intuitively, this makes a great deal of sense. The econometrician cannot possibly hope to explain the portion of y_t that is unpredictable even with the true model. Thus, the econometrician is better off focusing model (2) on what it can hope to explain, which is the systematically varying portion of y_t—in other words, $g(X_t)$, the true model forecast.

One can also think of the error-matching method as minimizing the *specification* error of the econometrician's model (2), instead of the total error of that model. By running the error-matched regressions instead of the standard regression, the econometrician purges the unforecastable noise term ε_t from the econometric residuals η_t in (3). All that remains is the specification error, which the error-matching algorithm subsequently minimizes.

2.4 Gains From Matching "True" Model Errors

Intuitively, by purging the unforecastable noise ε_t from the dependent variable of the econometric model, we can significantly enhance the precision of the econometrician's estimates. The econometrician's point estimate $\hat{\beta}^{EM}$ will generally differ from the usual OLS or NLS $\hat{\beta}$ as well, but in expected value they are the same,[7] so the difference is not systematic.

More formally, the true model errors ε_t in (1) are orthogonal to X_t, and hence to $g(X_t)$ and to $f(X_t)$. From equation (3), we then have $\text{Cov}(\eta_t, \varepsilon_t) = \sigma_\varepsilon^2$, the variance of ε_t, and hence $\text{Var}(\eta_t - \varepsilon_t) = \sigma_\eta^2 - \sigma_\varepsilon^2$. In other words, *the residuals of the error-matched regression, $\eta_t - \varepsilon_t$, necessarily have lower variance than the residuals from the usual OLS or NLS regression, η_t.*

---

[6] Note that, in the case of a VAR, this is *not* the same as estimating a VAR on the true model forecasts. To estimate an autoregression on y, the error-matching method prescribes regressing $y_t - \varepsilon_t$ onto $y_{t-1}, y_{t-2}, \ldots$ In contrast, a VAR on the true model forecasts prescribes regressing $y_t - \varepsilon_t$ onto $(y_{t-1} - \varepsilon_{t-1}), (y_{t-2} - \varepsilon_{t-2}), \ldots$, which will yield very different results.

[7] This follows from the orthogonality discussed below. Still, I should probably prove it more explicitly.

Since the right-hand side variables X_t are the same for both the error-matched and traditional regressions, it follows that the variance of $\hat{\beta}^{EM}$ is lower than that of $\hat{\beta}$, by the same ratio as the reduction in the variance of the error-matched residuals. This is clearly true for linear econometric models (2), since the variance of $\hat{\beta}$ is $\sigma_\eta^2(X'X)^{-1}$ while that of $\hat{\beta}^{EM}$ is $(\sigma_\eta^2 - \sigma_\varepsilon^2)(X'X)^{-1}$. It is not necessarily the case for nonlinear models (2) in small samples, since $\partial f/\partial\beta$ is not necessarily the same when evaluated at $\hat{\beta}^{EM}$ as at $\hat{\beta}^{NLS}$. However, the variance of $\hat{\beta}^{EM}$ will be lower than that of $\hat{\beta}^{NLS}$ asymptotically, as both estimators are consistent.

This increase in precision is analogous to the example in the Introduction, where the econometrician could estimate β perfectly, using the error-matching method, after only a small number of observations.

Note that the econometrician's model (2) is still the "wrong" model, in the sense that it is not the same as, and will never be the same as, the "true" model (1). Still, it was assumed that the econometrician's model is itself of interest (else why would the econometrician estimate it?), and the error-matching method does a much better job of estimating this model than does the usual procedure of OLS or NLS. We will also provide evidence below that the error-matching method does a much better job of producing forecasts than does the ε-augmented VAR approach (used by previous authors) in most cases of interest.

2.5 Market Expectations as "True" Model Forecasts

As noted previously, it is not necessary that market participants possess literally the *true* data generating process. All that is required is that the market's model encompasses the econometrician's model—i.e., that the market efficiently incorporates all of the variables that are in the econometrician's model. Note that, by definition, an efficient market satisfies this assumption, although we do not require full market efficiency here but only efficiency with respect to all of the variables in the econometrician's model, which is a weaker assumption.

In practice, the econometrician should test whether the assumption of forecast error orthogonality is justified for any given series of forecasts and a proposed econometric model (2). As a first step, the linear orthogonality conditions

$$y_t - \hat{y}_t = a + X_t b + \nu_t \tag{6}$$

should be verified, where $\hat{y}_t$ denotes the market forecast data for y_t, and the joint restriction $a = 0$, $b = 0$ (where b is a vector) is to be tested.

3. Two Examples: The Federal Funds Rate and Nonfarm Payrolls

Two examples help to illustrate the method and benefits of of the true model error-matching method in practice. For simplicity, we focus on simple, univariate time series models; some extensions to VARs are presented in the next section.

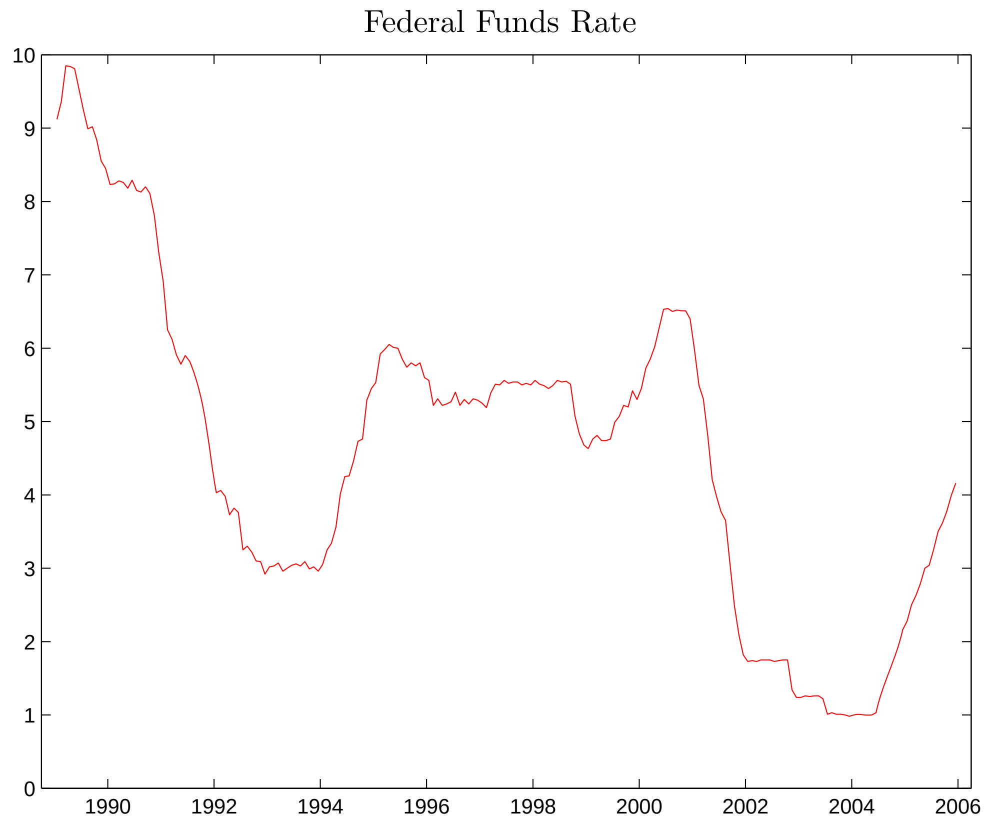
We apply the error-matching method to help estimate univariate autoregressions on two U.S. time series: the monthly federal funds rate and monthly employment, or nonfarm payrolls. These two statistics are often cited as the most timely and informative for the outlook for the U.S. economy and U.S. monetary policy; for example, the release dates for those two statistics lead to the biggest moves in financial markets (Ederington and Lee (1993), Fleming and Remolona (1997), Rudebusch (1998)). These two series provide serve as good examples for the error-matching method: first, forecasting these two series is of interest, owing to the series' importance, and second, both series are very widely and intensively forecast by financial markets and professional forecasters.

3.1 Data

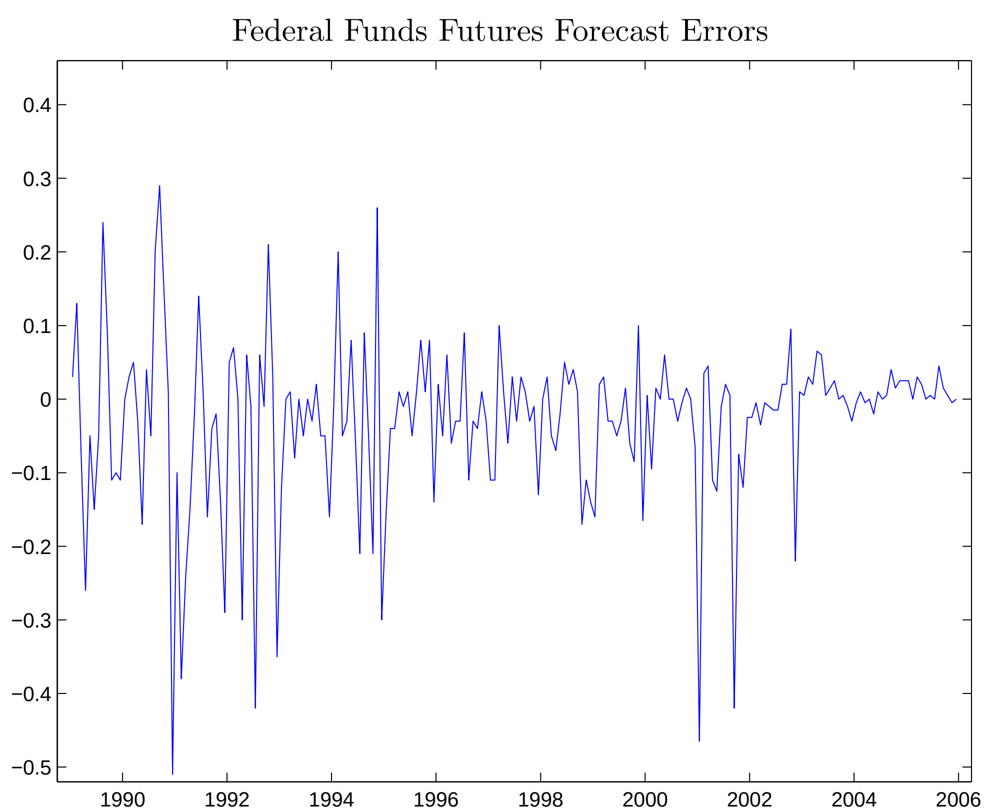
For forecasts of the federal funds rate, we use federal funds futures data obtained from the Federal Reserve Board. These futures contracts have been traded on the Chicago Board of Trade exchange since October 1988 and settle based on the average federal funds rate that is realized for a given calendar month, specified in the contract. The futures-based forecasts thus correspond exactly to our month-average federal funds rate series, above. We restrict attention in this example to the one-month-ahead forecast, which we take to be the federal funds futures rate for month t as quoted on the last day of month $t - 1$. In theory, this rate corresponds to the market expectation of the average funds rate for month t, given all information available at the end of month $t - 1$.[8]

[8] Piazzesi and Swanson (2004) investigate the importance of risk premia for fed funds futures. Although they find strong evidence of risk premia at longer horizons, they found no evidence of time-varying risk premia for the one-month-ahead federal funds futures contract, which supports our use of that contract here as a measure of market expectations. Allowing for a constant risk premium in this contract has no discernible effect on our results.

Figure 1: Federal Funds Rate and Forecast Errors, 1989–2005



(a) Month-average federal funds rate. Sample: Jan 1989 to Dec 2005, in percent.



(b) Federal funds futures forecast errors. Sample: Jan 1989 to Dec 2005, in percent. Month-average federal funds rate less corresponding federal funds futures rate as quoted on last day of previous month. See text for details.

Figure 1 plots the monthly average federal funds rate from 1989 to 2005 in the top panel and the one-month-ahead federal funds futures-based forecast errors in the bottom panel. It is worth emphasizing that many of the forecast errors in the bottom panel are quite large, even at this relatively short, one-month-ahead forecast horizon. For example, the terrorist attacks in September 2001 led the Fed to ease monetary policy by much more than markets had been expecting at the end of August, leading to a forecast error of about −42 basis points (bp). Similarly, the surprisingly rapid deterioration in the economic outlook in January 2001, December 1991, and May 1992 all led the Fed to ease policy by substantially more than the markets had expected, leading to forecast errors in each case of −40 to −50 bp.

Recognizing the size of these futures-market-based forecast errors, it would seem unfair to expect a simple time series model to fit these episodes when futures markets could not, even with billions of dollars on the line as an incentive and even with a vast array of economic indicators on hand as additional explanatory variables. This observation motivates estimating the time series model by trying to match the futures market's errors as closely as possible, rather than trying to completely minimize the model's prediction errors even over these unforecastable episodes.

Futures contracts for nonfarm payrolls are not available over most of our sample, although they have been introduced very recently by Deutsche Bank on a very limited basis—see Gürkaynak and Wolfers (2006) for details. Thus, instead of futures rates for nonfarm payrolls, we use the median forecast from a large survey of professional forecasters taken by Money Market Services. Money Market Services (MMS) publishes a weekly survey of market participants' expectations for each of the major statistical releases for the upcoming two weeks, and data are available going back to about 1985. Although these forecasts are not determined on an open exchange in the way that futures market forecasts are, they do have a significant market nonetheless and sell for several thousand dollars per year per customer, so their information content is perceived by the market to be valuable. Moreover, previous researchers (e.g., Balduzzi et al. (1996), Andersen et al. (2003), Brunner (2000)) have found the forecasts to pass standard tests of efficiency, which we replicate in Table 1, described below.

Figure 2 presents the nonfarm payrolls data in the top panel, graphed in first-differences, and the MMS median one-week-ahead survey forecast errors in the bottom panel.

Table 1: Forecast Correlations and Orthogonality Tests

Forecast	Data	correlation	Forecast Error Orthogonality Test F-stat	(p-val)
Federal Funds Futures	Federal Funds Rate	.80	0.60	(.834)
MMS Survey	Nonfarm Payrolls	.81	1.13	(.336)

Again, the forecast errors are substantial, amounting to two-tenths of a percent or more even just a single week before the statistical release, suggesting that fitting the model to match the true model errors may lead to significant gains in efficiency, relative to trying to completely minimize the model's residuals.

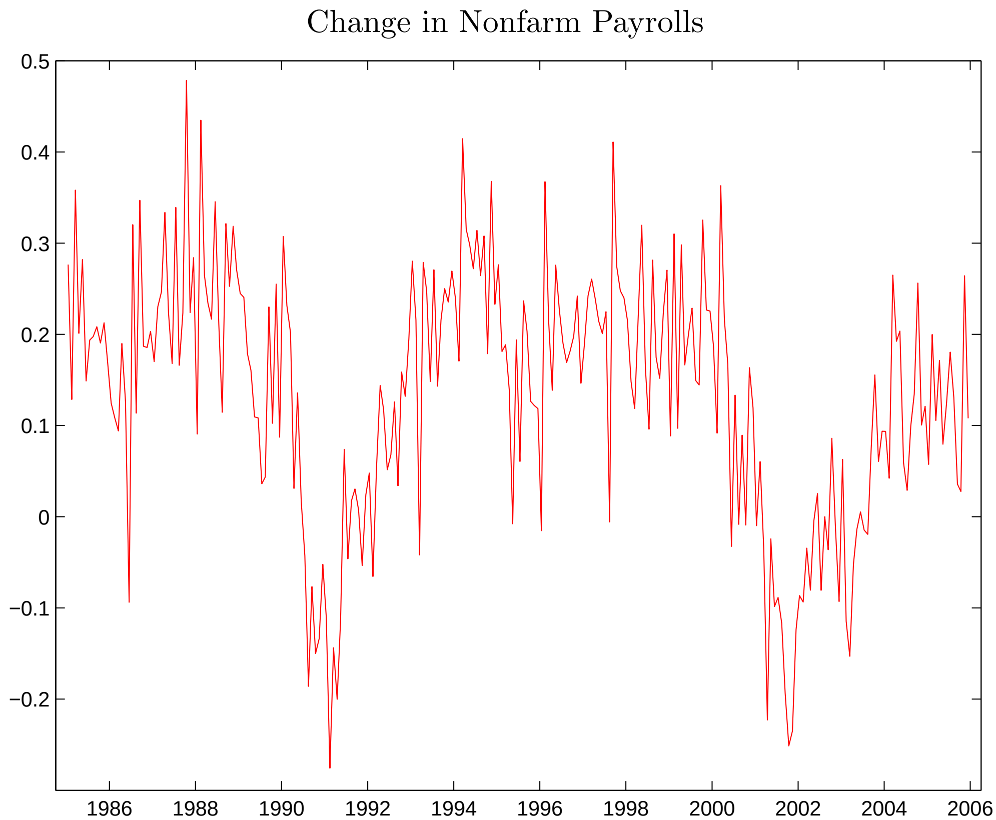
In Table 1, we test the efficiency of the forecasts with respect to the actual data as outlined in Section 2.5. The third column reports the raw correlation between the forecasts and the actual realizations of the series , which are high (for both series, this correlation is for the month-to-month changes, since nonfarm payrolls has a time trend and the fed funds rate has also been close to nonstationary). The last two columns of the table present results from these tests, corresponding to equation (6) with twelve lags of the regressors on the right-hand side. The forecasts easily pass the efficiency tests at standard levels of significance.

3.2 Estimation of Univariate Autoregressions

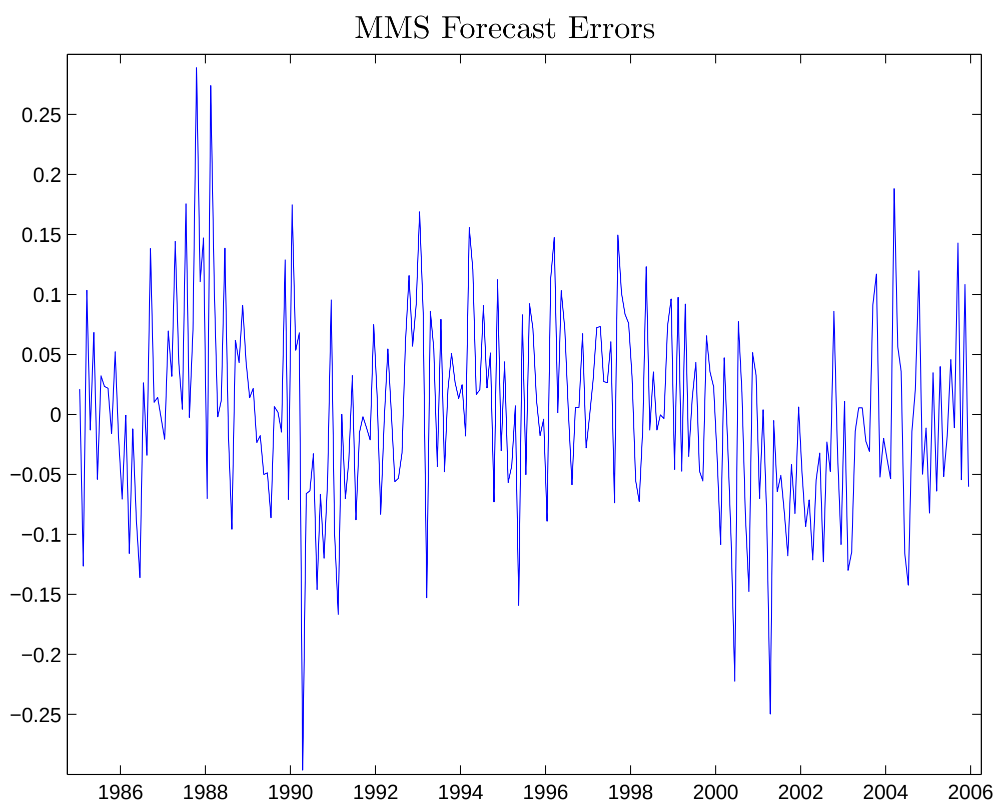
We now turn to estimating simple univariate autoregressions on each of these series. Although univariate autoregressions are extremely simple models, they are nonetheless often used as a benchmark forecast in many applications and their simplicty is an advantage when we are using them as an example.

For both the federal funds rate and nonfarm payrolls, we follow the suggestion of Sims, Stock, and Watson (1990) and estimate the regression on the levels of the data (log-levels for nonfarm payrolls). We begin with an autoregression on twelve monthly lags and drop the last lag if it is statistically insignificant, repeating this procedure until the last lag is significant. This process leads us to adopt an AR(4) specification for the federal funds rate and an AR(10) for nonfarm payrolls. (For OLS estimation of the nonfarm payrolls autoregression, this process would lead us to a shorter lag length, but we hold the lag length at 10 to maintain comparability to the error-matching estimation results.)

Figure 2: Nonfarm Payrolls and Forecast Errors, 1985–2005



(a) Change in nonfarm payrolls from previous month. Sample: Jan 1985 to Dec 2005, in percent.



(b) MMS survey forecast errors. Sample: Jan 1985 to Dec 2005, in percent. Change in nonfarm payrolls from previous month, less Money Market Services median forecast made one week previously. See text for details.

Table 2: Federal Funds Rate and Nonfarm Payrolls Autoregressions

Federal Funds Rate, AR(4)

	(a) OLS Estimation			(b) Error-Matching Estimation		
	coeff	(std err)	[t-stat]	coeff	(std err)	[t-stat]
constant:	0.052	(0.0244)	[2.15]	0.057	(0.0201)	[2.85]
lag 1:	1.384	(0.0689)	[20.07]	1.375	(0.0568)	[24.22]
2:	−0.191	(0.1185)	[−1.61]	−0.217	(0.0976)	[−2.22]
3:	−0.068	(0.1185)	[−0.58]	−0.046	(0.0976)	[−0.47]
4:	−0.137	(0.0687)	[−2.00]	−0.120	(0.0566)	[−2.12]
SSR:		4.43			3.01	
ratio of std errs:		·			0.82	
ratio of SSR:		·			0.68	

Nonfarm Payrolls, AR(10)

	(c) OLS Estimation			(d) Error-Matching Estimation		
	coeff	(std err)	[t-stat]	coeff	(std err)	[t-stat]
constant:	0.835	(0.7154)	[1.17]	−0.026	(0.4104)	[−0.06]
lag 1:	1.187	(0.0636)	[18.66]	1.167	(0.0365)	[31.96]
2:	0.122	(0.0988)	[1.23]	0.048	(0.0567)	[0.84]
3:	−0.152	(0.1000)	[−1.52]	−0.038	(0.0574)	[−0.66]
4:	−0.030	(0.1004)	[−0.30]	−0.115	(0.0576)	[−2.00]
5:	−0.072	(0.1006)	[−0.72]	0.029	(0.0577)	[0.50]
6:	0.019	(0.1006)	[0.19]	0.033	(0.0577)	[0.57]
7:	−0.047	(0.1006)	[−0.47]	−0.038	(0.0577)	[−0.66]
8:	−0.015	(0.0999)	[−0.15]	−0.001	(0.0573)	[−0.02]
9:	−0.119	(0.0997)	[−1.20]	−0.107	(0.0572)	[−1.88]
10:	0.106	(0.0639)	[1.67]	0.081	(0.0366)	[2.21]
SSR:		2.10			0.69	
ratio of std errs:		·			0.57	
ratio of SSR:		·			0.33	

Sample period: Jan 1989 to Dec 2005 for federal funds rate, Jan 1985 to Dec 2005 for nonfarm payrolls, at monthly frequency. Autoregressions are in log-levels. Lag length was selected by beginning with 12 monthly lags and iteratively dropping the last lag until it was statistically significant. "Error-Matching Estimation" denotes estimation by matching the model's residuals to the "true" model errors from the forecast series presented earlier. See text for details.

Estimation results for these two autoregressions are presented in Table 2. Columns (a) and (c) present the results from estimating the two models by standard OLS. The results for both series are in line with those of many monthly time series: the coefficient on the first lag is large, positive, and highly statistically significant, while the coefficients on the other lags are smaller and often insignificant. The estimated autoregressions are close to unit-root processes.

In columns (b) and (d) of the table, we turn to estimating the autoregressions by matching the model residuals to the "true" model forecast errors as closely as possible. As predicted, the error-matching method yields a significantly more accurate projection of the data onto the variables of interest (in this case, lags of the endogenous variables): the standard errors around the coefficient estimates are reduced by almost 20 percent relative to OLS in the federal funds rate autoregression, and are reduced by almost 45 percent in the nonfarm payrolls autoregression.

This improvement in precision is enough to change the conclusion of hypothesis tests in some cases. For example, the last (tenth) lag of the nonfarm payrolls autoregression is not statistically significant at the 5 percent level using OLS, while it easily clears this threshhold using the error-matched estimates.[9]

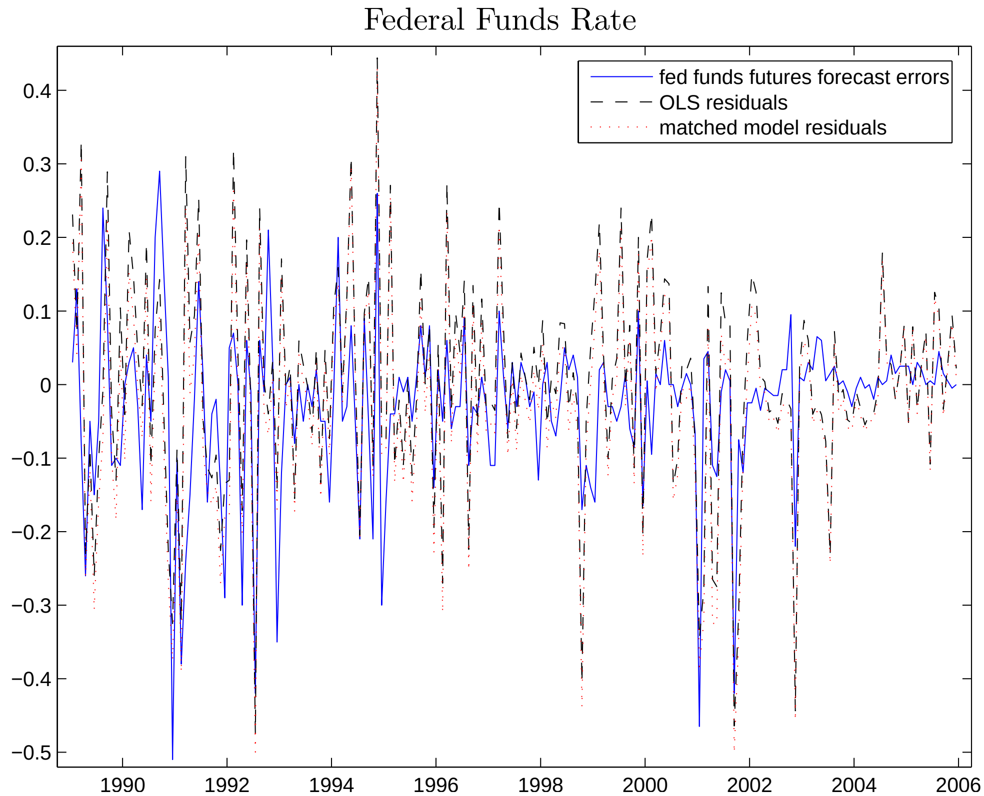
How well do the model residuals using each of the two estimation procedures (OLS and error-matching) match the "true" model errors in these examples? Figure 3 graphs all three sets of residuals—the "true" model errors from the forecast data series (previously plotted in Figures 1 and 2), and the two sets of model residuals obtained by each estimation method (OLS and error-matching).

The improvement in fit in Figure 3 using the error-matching procedure is actually quite subtle. To some extent, this is to be expected: both methods are consistent, thus both methods should yield very similar point estimates for β, and consequently both methods should yield very similar estimates for the model's residuals η_t. Indeed, the point estimates in Table 2 were similar across the two estimation methods.

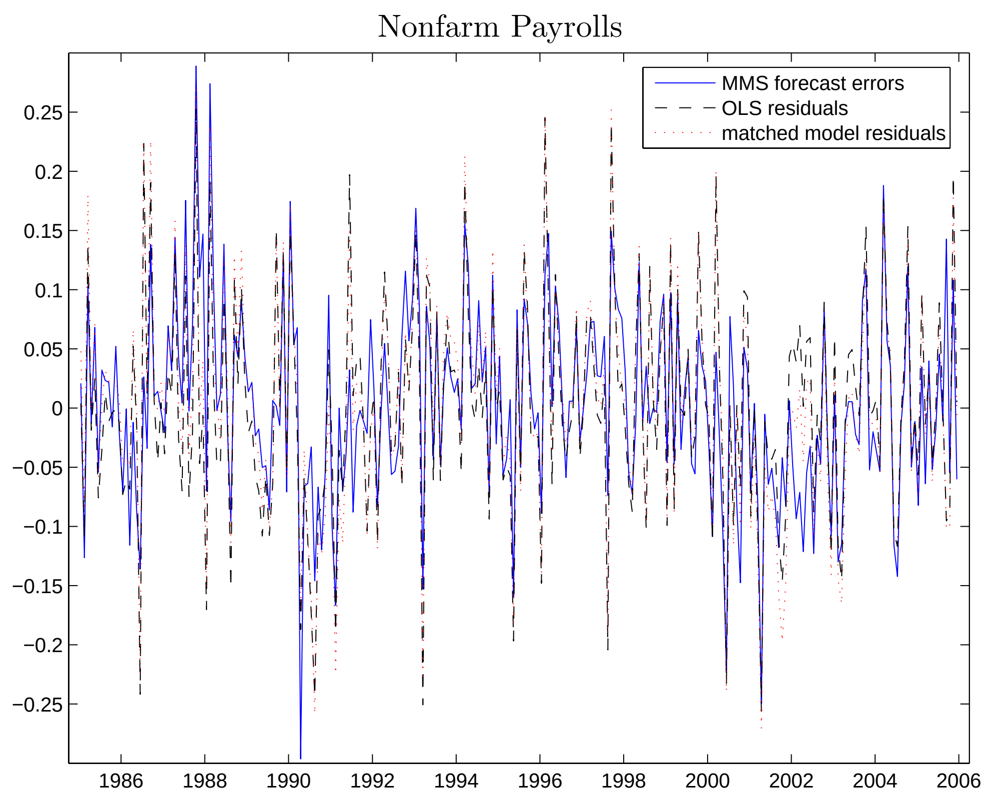
Nonetheless, the error-matched model shocks do match the "true" model errors more closely, but the improvement in fit is relatively small (e.g., RMSE of .0577 as opposed

[9] Continuing to drop the last lag from the OLS-estimated nonfarm payrolls autoregression would eventually lead us to adopt an AR(5) specification, with the fifth lag having a value of .015. As noted previously, we have kept the lag length of this autoregression fixed at 10 to maintain comparability across the two estimation procedures.

Figure 3: "True" Model Errors and Econometric Model Residuals



(a) Federal funds futures forecast errors vs. econometric model residuals. The blue line corresponds to "true" model errors ε_t, from futures market forecasts. The dashed black line corresponds to residuals from AR(4) model estimated by OLS. The dotted red line corresponds to residuals from AR(4) model estimated by matching residuals to "true" model errors. See text for details.



(b) Nonfarm payrolls forecast errors vs. econometric model residuals. The blue line corresponds to "true" model errors ε_t, from MMS forecasts. The dashed black line corresponds to residuals from AR(10) model estimated by OLS. The dotted red line corresponds to residuals from AR(10) model estimated by matching residuals to "true" model errors. See text for details.

.0524 for OLS in the nonfarm payrolls model, an improvement of about 10increase in statistical precision of the estimates in Table 2 comes not from changes in the parameter estimate β, but rather from the fact that the "true" model errors often miss in the same direction as the baseline model. Thus, a significant fraction of the baseline model's errors were unforecastable even using the "true" model, which makes the model estimation errors correspondingly smaller.

More importantly, error-matching estimation avoids over-weighting the very large errors that were associated with surprises that were unforecastable ex ante. Least squares will tend to weight these large errors very highly in the estimation, while the error-matching approach will downweight these errors to the extent that they were also made by the "true" forecasting model as well.

Some evidence of OLS overweighting the large forecast errors in the data can be seen in the model impulse response functions, presented in Figure 4. Solid blue lines denote the point estimates generated by the coefficients in Table 2, while the dotted red lines denote bootstrapped two-standard-error bands.[10]

When the impulse responses are generated by the OLS coefficient estimates, the model has a substantially greater tendency to mean-revert in these two examples. Intuitively, the large, unforecastable errors in the federal funds rate and nonfarm payrolls series, can be thought of as additional white noise that is pulling the OLS coefficient estimates back toward a white noise process. By contrast, the true model error-matching coefficient estimates downweight these large surprises and produce impulse responses that are more persistent.

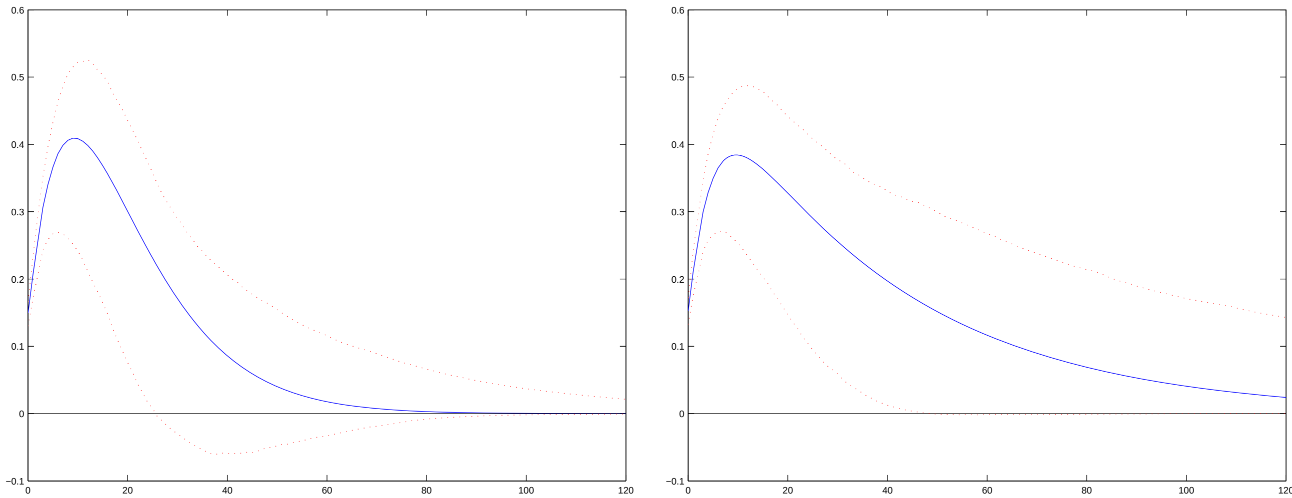
The precision of the impulse responses from the error-matched model estimates is also substantially greater.[11] Together with the more persistent point estimates, the result is

[10] The bootstrap standard error bands were computed as in Christiano, Eichenbaum, and Evans (2000). Five thousand synthetic time series were generated using the parameters in column (a) of Table 2, and drawing randomly from the residuals of that regression with replacement. For each synthetic time series, a new set of coefficients was estimated, and the corresponding impulse response function computed. The set of synthetic impulse response functions can be sorted at each time t by their sizes at that point in time. The dotted lines in Figure 2(a) present the 125th and 4875th largest such responses at each point in time.

[11] Bootstrap standard error bands for the error-matched regressions were computed as follows. One thousand synthetic time series were generated using the parameters in panel (b) (or (d) or (e))) of Table 2, drawing randomly from the residuals and "true" model errors of that regression with replacement—thus, if a given random draw picks the 42nd residual from the regression, the 42nd "true" model error is also taken and recorded, and used to help generate the synthetic time series. For each synthetic time series, a new set of coefficients was estimated using the error-matching method, matching the corresponding vector of "true" model errors that were recorded. Impulse response functions were then computed and sorted in

Figure 4: OLS and Error-Matched Model Impulse Response Functions

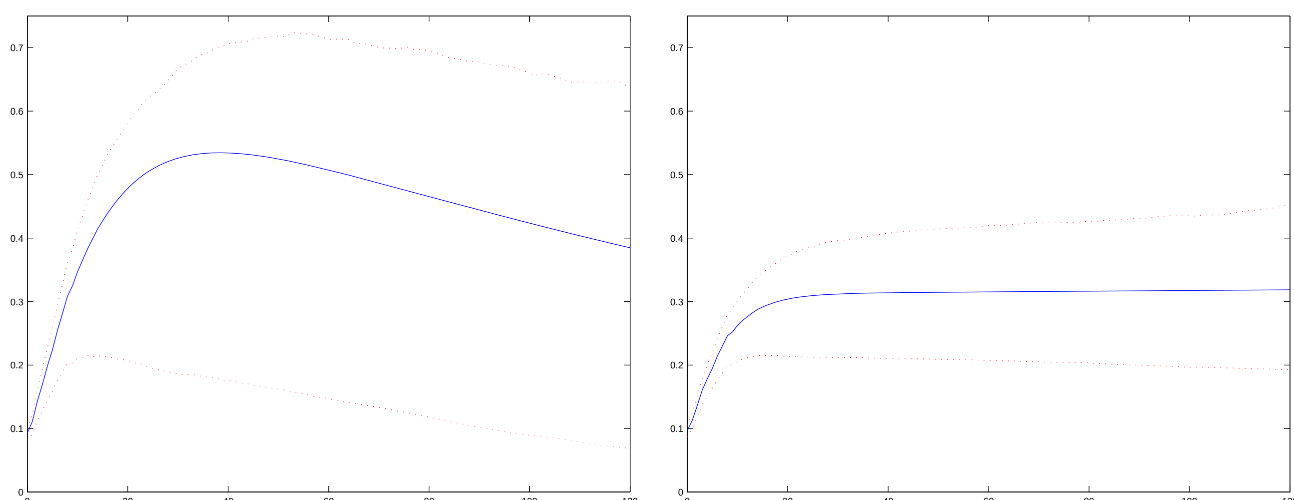
Federal Funds Rate



(a) OLS Estimation (b) Error-Matching Estimation

Impulse response functions from AR(4) model estimated on monthly federal funds rate data (in levels). "OLS" denotes estimation by OLS; "Matched Model" denotes estimation by matching model residuals to the "true" model forecast errors. The blue line depicts point estimates and dotted red lines depict ±two-standard-error bands, obtained from 5000 bootstrap simulations. See text for details.

Nonfarm Payrolls



(c) OLS Estimation (d) Error-Matching Estimation

Impulse response functions from AR(10) model estimated on monthly nonfarm payrolls data (in log levels). "OLS" denotes estimation by OLS; "Matched Model" denotes estimation by matching model residuals to the "true" model forecast errors. The blue line depicts point estimates and dotted red lines depict ±two-standard-error bands, obtained from 5000 bootstrap simulations. See text for details.

---

the usual way.

that the error-matched model's impulse response to a federal funds rate shock is statistically different from zero for fully three and a half years, compared to only two years for the model estimated by OLS. The statistical significance of the nonfarm payrolls impulse responses is also subsantially higher using the error-matching procedure.

3.3 Results from an ε-Augmented VAR

As an alternative to the true model error-matching approach above, one might consider the "ε-augmented VAR" approach followed by other authors in the literature. This method prescribes appending the observed "true" model errors to the econometrician's VAR (the method only applies to VARs, recall). Like any time series, the "true" model errors (TMEs) can be regressed on lags of themselves and lags of the other time series in the VAR, and stochastic innovations calculated. Of course, if the "true" model forecasts are efficient with respect to the variables in the VAR, then all of the coefficients in the TME equation should be insignificantly different from zero. As discussed in Section 2, however, the TMEs may still enter significantly into the regressions for the other variables of the VAR, as they proxy for variables that the econometrician has excluded from the model.

In the context of our simple Nonfarm Payrolls autoregression,[12] the ε-augmented VAR consists of two time series: Nonfarm Payrolls and the TMEs. The usual Cholesky factorization is assumed, with the TME series ordered last—this preserves the "Nonfarm Payrolls shock" as a single entity, rather than dividing it arbitrarily into two components. The results of this VAR are presented in Figure 4.[13]

The TME equation behaves exactly as predicted above: its impulse response functions (bottom panels, Figure 4) are economically and statistically indistinguishable from zero (note the scale of the axes), and its coefficients are individually and jointly statistically insignificant (the joint test has a p-value of .24).

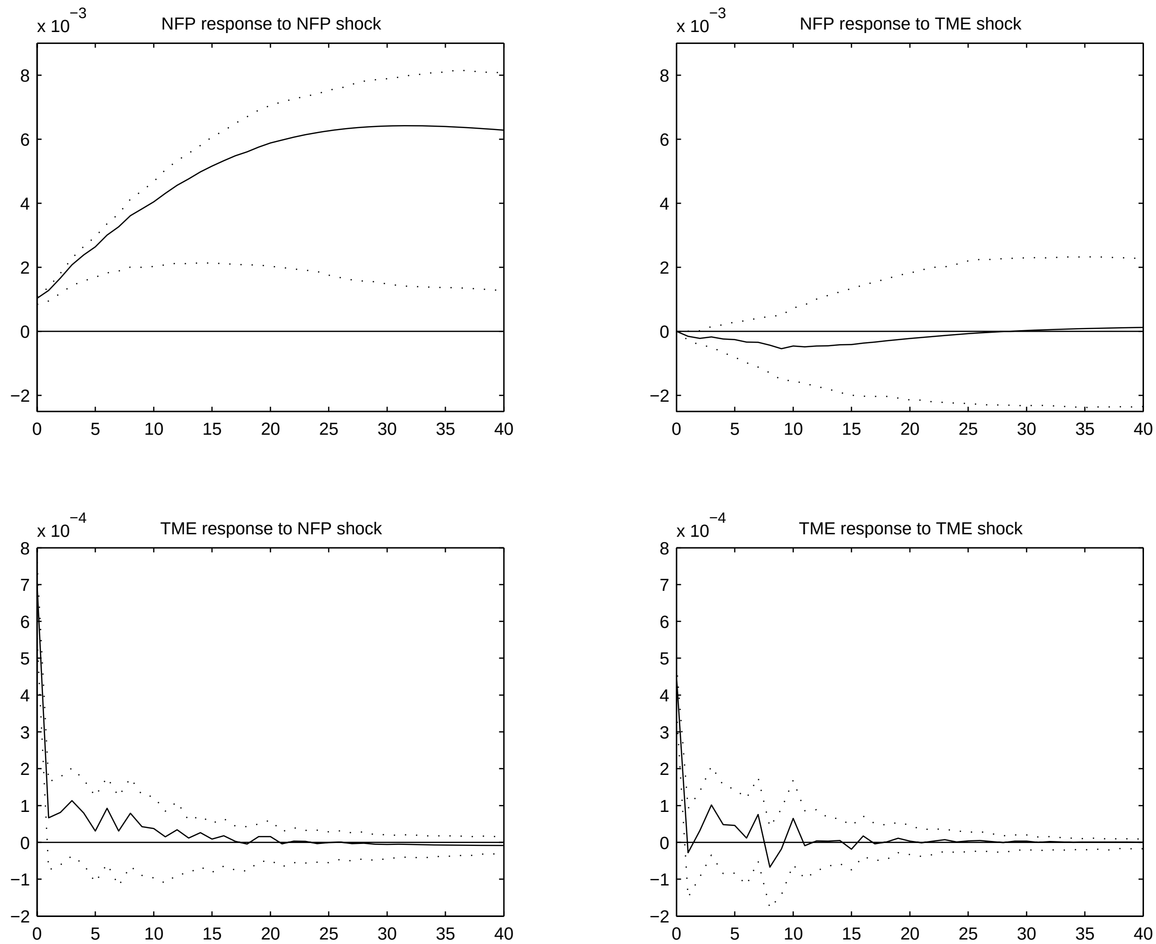
More importantly, the inclusion of the TMEs into the Nonfarm Payrolls equation—the approach favored by previous authors—yields essentially no benefit whatsoever. The response of Nonfarm Payrolls to a Nonfarm Payrolls shock (upper left panel of Figure 4) is

---

[12] Results for the federal funds rate autogression are qualitatively very similar.

[13] Figure 4 uses the MMS forecast data. Results using the Wrightson forecast data are very similar. The procedure used to calculate bootstrapped two-standard-error bands is the same as that in Christiano, Eichenbaum, and Evans (2000).

Figure 5: ε-Augmented VAR Impulse Responses



very similar to what we estimated using the standard (univariate autoregression) method earlier, both in terms of the point estimate of the impulse response, and in terms of the large error bands surrounding this point estimate. The impulse response of Nonfarm Payrolls to an innovation in the TME series (upper right panel of Figure 4) are essentially insignificant, and the coefficients on the TMEs in this equation are jointly indistinguishable from zero (p-value of .12).

The finding that the ε's make very little difference when incorporated into the VAR in this way is in agreement with the findings of other authors who have used this method (Christiano, Eichenbaum, and Evans (2000), Brunner (2000), Bagliano and Favero (1998)). While the ε's (or equivalently, the "true" model forecasts) do contain information about variables other than Nonfarm Payrolls, this information is far from unidimensional and is thus very difficult to extract from the single, one-dimensional series ε_t. As a result, the coefficients on the ε_t are generally small and statistically insignificant, and help very little in improving the model's forecasts and reducing the variance of its parameter estimates. Note that these observations are not specific to Nonfarm Payrolls, but rather hold for the ε-augmented VAR approach more generally.

The upper right-hand panel of Figure 4 also helps to illustrate another of the criticisms leveled at the ε-augmented VAR approach in Section 2. It depicts the response of Nonfarm Payrolls to a "TME shock." Moreover, this "TME shock" is not just any innovation in the forecast error series, but rather the component of such innovations that is orthogonal to the errors in the estimated NFP equation. Interpreting this "TME shock," not to mention the impulse responses to such a shock, is very difficult.

4. Discussion and Extensions

As demonstrated in section 2, the true model error-matching procedure will always perform better than the standard method, in terms of precision, when the econometrician possesses valid measures of the "true" model errors. However, the extent of this improvement can lie anywhere between zero (when the "true" model errors are identically zero) and infinity (when the "true" model parameters can be recovered with perfect precision). The amount of improvement depends upon the model and data under consideration.

Error-matching yields relatively little improvement in performance when the "true" model errors are small relative to the econometric model residuals. In some sense, the "true" model in this case is too good, in that it provides a nearly perfect, virtually deterministic explanation of the data of interest. The "true" model errors are then largely uninformative—indeed, the "true" model forecasts are essentially the same as the observed data—and thus help very little in estimating the econometric model. Paradoxically, then, the better is the "true" model relative to the econometric model, the less useful it is for helping us to estimate the econometric model.

Error-matching does very well when the "true" model is very close to the econometric model. In this case, the "true" model errors are extremely informative about the parameters of the econometrician's model—in the limit, when the "true" model and econometric model are the same, knowledge of the true model errors allows the econometrician to recover the "true" model with perfect precision, as in the example in the Introduction. In the case of our Nonfarm Payrolls example above, if the "true" model (market's model) for Nonfarm Payrolls had been a simple univariate AR(12) process, with fixed coefficients, then knowledge of the "true" model errors would have allowed us to recover the "true" (market's) model with only 13 observations. Would this have been interesting? The answer is yes if and only if the "true" (in this case, the market's, or equivalently, the econometrician's) model is itself of interest. We have assumed throughout this paper that the "true" model is one that is of interest.

Of course, if the "true" model is garbage, then the results produced by the error-matching algorithm will be garbage as well. Graphing the "true" model forecasts against the actual data, and first-stage orthogonality tests, will help to weed out potential forecasts that are clearly unhelpful, but the econometrician must still be wary. For example, if the econometrician's only measure of the "true" model errors is contaminated with a large degree of white noise, then these errors will still pass the orthogonality tests, yet trying to match the econometric model residuals to these noisy proxies for the true model errors could in fact yield much worse results than the standard method (OLS or NLS).

4.1 Futures Market Quotes and Risk Premia

For the reasons discussed above, the econometrician must be careful when using futures market quotes as measures of "true" model forecasts, as risk premia are an issue in essentially all futures markets. In the case of the federal funds futures market, the assumption of a small, constant risk premium seems to work very well, and is not rejected by the data, but in some markets this risk premium appears to vary over time. For example, the forward and future foreign currency markets are notoriously poor predictors of future exchange rates, and it is widely hypothesized that this failure is due in large part to time-varying risk premia.

If the risk premium "noise" in the forecast series is small relative to the "true" model forecast error, then it will typically be safe to ignore such noise. If, however, the non-expectational errors in the forecast series are substantial, then the econometrician must be very careful to avoid the "garbage" problem discussed above. If the noise in the true model error data can be restricted on theoretical grounds to only a few observations or subsample periods, then the econometrician can dummy out those true model errors in question, effectively treating them as zeros. In the limit, when all such true model errors are dummied out, this reduces the error-matching procedure to the standard method (OLS or NLS).

4.2 Combining the Error-Matched and ε-Augmented Approaches

Since our Nonfarm Payrolls example in Section 3 was a time series, an astute reader may have noticed that there is no reason why we cannot combine both the error-matching and ε-augmented approaches into a single procedure. This will allow us to gain the added precision that derives from matching the "true" model errors, while at the same time using lagged values of the ture model errors as proxies for variables that we have excluded from the regression. Note, however, that the results of this exercise will suffer from all the same criticisms we have leveled at the ε-augmented VAR approach in Section 2. For example, the resulting model is probably only useful for generating Nonfarm Payrolls forecasts, and these forecasts should in general be dominated by those that we could obtain from the market anyway.

This approach is going to yield the optimal forecasting VAR with the given data. There is no way to improve it further. I should probably prove this assertion.

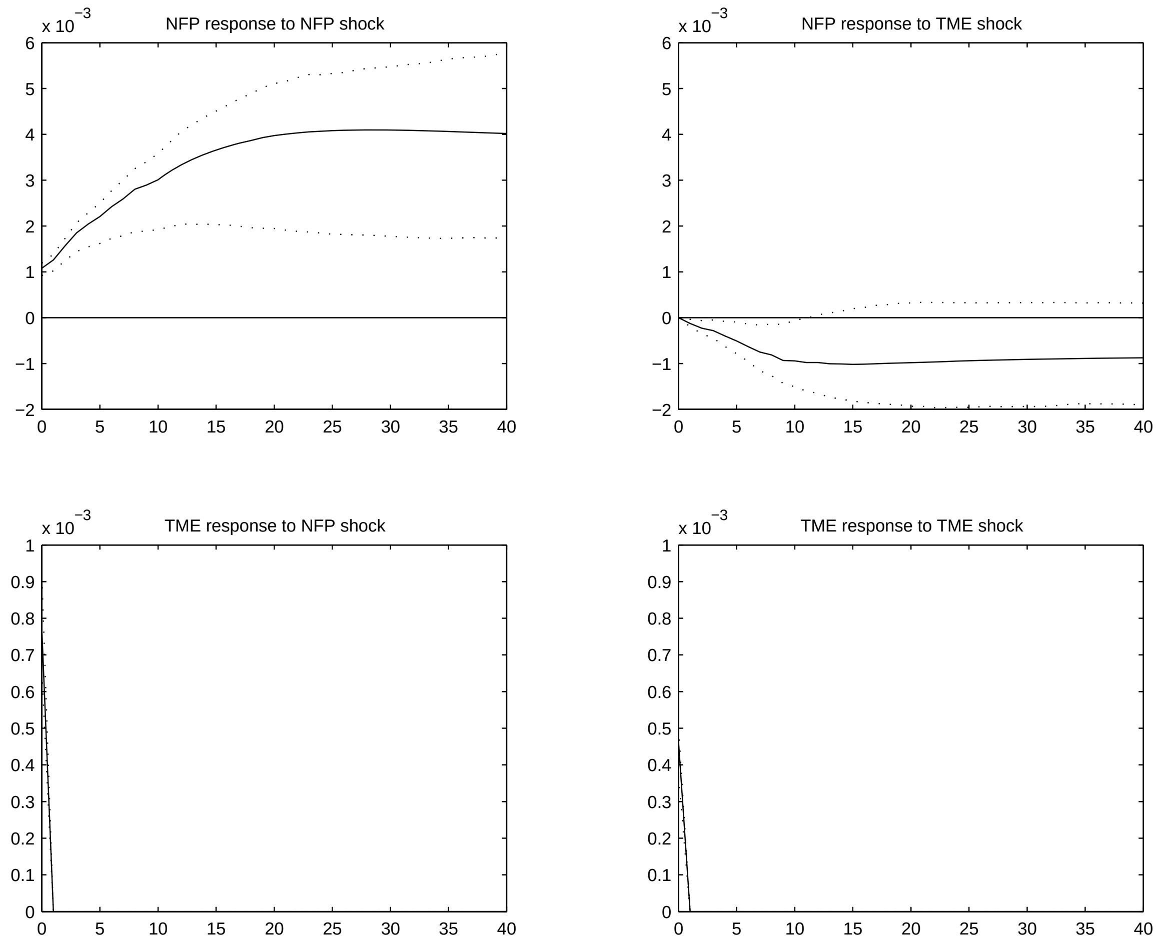
Figure 5 presents the results from such a error-matched, ε-augmented VAR.[14] Note that the econometrician observes not only the "true" model errors for the Nonfarm Payrolls series, but also the "true" model errors for the TME time series as well—in fact, the "true" model errors for the TME time series is simply the TME time series itself. This follows because, if one were to ask market participants (Wrightson, MMS), what their forecast for the TME series each period would be, they must answer "zero," unless they knowingly issue biased or inefficient forecasts for Nonfarm Payrolls. In other words, the "true" model forecasts for the TME series is identically zero every period, and hence the "true" model errors for the TME time series is that series itself.[15]

The results in Figure 5 are very much in line with our expectations. The impulse response functions for the TME series (bottom row) are identically zero except on impact, and are estimated with perfect precision, since by subtracting off the "true" model errors

[14] Figure 5 uses the MMS forecast errors. Results using the Wrightson forecasts errors are very similar.

[15] Of course, this is equivalent to simply imposing that all of the coefficients in the second equation of the ε-augmented VAR are zero.

Figure 6: ε-Augmented and TME-Matching Methods Combined



for that series we are left with nothing to explain. This is a special case of one of the points made above: since the "true" model for the TMEs is the same as the econometric model (it is an AR, with all zero coefficients), the TME-matching method recovers the true model perfectly after only finitely many observations.

The response of Nonfarm Payrolls to a Nonfarm Payrolls shock (upper left panel, Figure 5) is very similar to what we estimated in Section 3 using the univariate error-matching procedure. The precision of the estimate is also no better than what we obtained with the univariate error-matching approach in Figure 2c. Thus, again, the augmentation of the VAR to include the ε's yields essentially no improvement over the unaugmented version of the model. In practice, the ε's serve as very poor proxies for the variables omitted from the model, even though in theory they are potentially useful in this respect. This is not to say, however, that the ε's are *uninformative* in practice—indeed, the results of the (univariate) error-matched approach clearly demonstrate that the information contained in the ε's is significant.

4.3 Extensions

The most obvious generalization of the error-matching method, as presented in this paper, is to the estimation of VARs instead of univariate autoregressions. The direct generalization of the estimation procedure to this case is trivial, and improvements in the precision of the estimates of the reduced-form model are comparable to what can be obtained in the univariate case. However, the availability of high-frequency (weekly or daily) measures of "true" model errors also raises new possibilities for identification of the structural shocks to the model. Faust, Swanson, and Wright (2001) consider these issues in detail, and their application to a monetary policy VAR.

More generally, the benefits of error-matching will be most pronounced in applications where the marginal value of increased precision of the estimates is very high. Empirical tests for the presence of unit roots seem to be one such example, since low power against the null in these tests is very often a problem.

It should also be emphasized that there is no theoretical reason why the error-matching method of this paper couldn't be applied equally well to cross-sectional or panel data, despite the present paper's emphasis on time series examples. The only limitation in this respect seems to be the availability of good forecast data for commonly-used cross-sectional or panel datasets. Such forecasts do exist, as in corporate finance, where earnings and other statistics are widely forecast for large cross-sections of corporations, but the unbiasedness and efficiency of these forecasts is often somewhat suspect (reference?). Perhaps their efficiency has been improving, making these data more useful going forward, or perhaps a certain subset of these forecasts stands out as being more efficient than the others, in which case the econometrician could restrict attention to this subset of forecasts. Forecasts of GDP for a wide cross-section of countries are also avaible, although again, the unbiasedness and efficiency of these forecasts have been somewhat questionable empirically (reference?). Still, the potential benefits to empirical cross-country panel studies of growth are large if these shortcomings of the data could be overcome.

5. Conclusions

Forecast and futures market data are potentially very useful for improving the estimation of standard econometric models. Few previous studies have incorporated these data into their analysis, and those that have done so have generally not asked the question of how such forecast data *should* be incorporated from the viewpoint of econometric estimation. This paper has asked that question, and found an answer very different from the previous literature.

The method prescribed in this paper is that of matching the econometric model residuals to the market-based forecast errors instead of simply minimizing the econometric model's residuals. Under the assumption that the market's forecasts are efficient with respect to all of the variables in the econometrician's model, this approach necessarily yields an improvement in the precision of the econometrician's estimates. In practice, this improvement can be substantial, ranging from about 20 to 40 percent for the simple time series examples considered. Moreover, the error-matching method avoids overweighting large model errors that were unforecastable *ex ante* even with the "true" model.

There are potentially a large number of applications of the method going forward. Recent work on VARs, macro-finance models of the yield curve, and prediction markets have all begun incorporating financial market and forecast data into the model. Ongoing work in these areas should provide new opportunities to test the usefulness of the error-matching estimation algorithm.

References

Andersen, Torben, Tim Bollerslev, Frank Deibold, and Clara Vega (2003). "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review* 93, 38–62.

Bagliano, Fabio and Carlo Favero (1998). "Measuring Monetary Policy with VAR Models: An Evaluation," *European Economic Review* 42, 1069–1112.

Bagliano, Fabio and Carlo Favero (1999). "Information from Financial Markets and VAR Measures of Monetary Policy," *European Economic Review* 43, 825–37.

Balduzzi, Pierluigi, Edwin Elton, and T. Clifton Green (2001). "Economic News and Bond Prices: Evidence from the U.S. Treasury Market," *Journal of Financial and Quantitative Analysis* 36, 523–43.

Bernanke, Ben, Vincent Reinhart, and Brian Sack (2004). "Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment," *Brookings Papers on Economic Activity* 2, 1–78.

Brunner, Allan (1996). "Using Measures of Expectations to Identify the Effects of a Monetary Policy Shock," *Federal Reserve Board International Finance Discussion Paper* 537.

Brunner, Allan (2000). "On the Derivation of Monetary Policy Shocks: Should We Throw the VAR Out with the Bath Water?" *Journal of Money, Credit, and Banking* 32, 254–79.

Chun, Albert (2005). "Expectations, Bond Yields, and Monetary Policy," unpublished manuscript, Stanford University.

Christiano, Lawrence, Martin Eichenbaum and Charles Evans (2000). "Monetary Policy Shocks: What Have We Learned and To What End?" *Handbook of Monetary Economics,* ed. John Taylor and Michael Woodford.

Cochrane, John and Monika Piazzesi (2002). "The Fed and Interest Rates: A High-Frequency Identification," *American Economic Review Papers and Proceedings* 92, 90–95.

Ederington, Louis and Jae Ha Lee (1993). "How Markets Process Information: News Releases and Volatility," *Journal of Finance* 48, pp. 1161–91.

Faust, Jon, Eric Swanson, and Jonathan Wright (2004). "Identifying VARs Based on High-Frequency Futures Data," *Journal of Monetary Economics* 51, 1107–31.

Fleming, Michael and Eli Remolona (1997). "What Moves the Bond Market?" *Federal Reserve Bank of New York Economic Policy Review,* 31–50.

Grossman, Sanford (1989). *The Informational Role of Prices* (Cambridge, MA: MIT Press).

Gürkaynak, Refet and Justin Wolfers (2006). "Macroeconomic Derivatives: An Initial Analysis of Market-Based Macro Forecasts, Uncertainty, and Risk," *Journal of Economic Literature* forthcoming.

Kim, Don and Athanasios Orphanides (2005). "Term Structure Estimation with Survey Data on Interest Rate Forecasts," *Federal Reserve Board Finance and Economics Discussion Series* 2005–48.

KIM, DON AND JONATHAN WRIGHT (2005). "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates," *Federal Reserve Board Finance and Economics Discussion Series* 2005–33.

KUTTNER, KENNETH (2001). "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market," *Journal of Monetary Economics* 47, 523-44.

PIAZZESI, MONIKA AND ERIC SWANSON (2004). "Futures Rates as Risk-Adjusted Forecasts of Monetary Policy," *NBER Working Paper* 10547.

ROMER, CHRISTINA AND DAVID ROMER (2000). "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review* 90, 429–57.

RUDEBUSCH, GLENN (1998). "Do Measures of Monetary Policy in a VAR Make Sense?" *International Economic Review* 39, 907–41.

SIMS, CHRISTOPHER (2002). "The Role of Models and Probabilities in the Monetary Policy Process." *Brookings Papers on Economic Activity* 2, 1–62.

SIMS, CHRISTOPHER, JAMES STOCK, AND MARK WATSON (1990). "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica* 58, 113–44.

WHITE, HALBERT (1994). *Estimation, Inference, and Specification Analysis* (New York: Cambridge University Press).

WOLFERS, JUSTIN AND ERIC ZITZEWITZ (2005). "Prediction Markets," *Journal of Economic Perspectives* 18, 107–26.