

Ghysels, Eric; Horan, Casidhe; Moench, Emanuel

Working Paper

Forecasting through the rear-view mirror: Data revisions and bond return predictability

Staff Report, No. 581

Provided in Cooperation with:

Federal Reserve Bank of New York

Suggested Citation: Ghysels, Eric; Horan, Casidhe; Moench, Emanuel (2012) : Forecasting through the rear-view mirror: Data revisions and bond return predictability, Staff Report, No. 581, Federal Reserve Bank of New York, New York, NY

This Version is available at:

<https://hdl.handle.net/10419/93668>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Federal Reserve Bank of New York
Staff Reports

Forecasting through the Rear-View Mirror: Data Revisions and Bond Return Predictability

Eric Ghysels
Casidhe Horan
Emanuel Moench

Staff Report No. 581
November 2012



This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Forecasting through the Rear-View Mirror: Data Revisions and Bond Return Predictability

Eric Ghysels, Casidhe Horan, and Emanuel Moench

Federal Reserve Bank of New York Staff Reports, no. 581

November 2012

JEL classification: G10, G12

Abstract

Real-time macroeconomic data reflect the information available to market participants, whereas final data—containing revisions and released with a delay—overstate the information set available to them. We document that the in-sample and out-of-sample Treasury return predictability is significantly diminished when real-time as opposed to revised macroeconomic data are used. In fact, much of the predictive information in macroeconomic time series is due to the data revision and publication lag components.

Key words: return predictability, real-time data, dynamic factor models

Ghysels: University of North Carolina at Chapel Hill (e-mail: eghysels@unc.edu). Horan: University of Michigan (e-mail: casidheh@umich.edu). Moench: Federal Reserve Bank of New York (e-mail: emanuel.moench@ny.frb.org). Eric Ghysels thanks the Federal Reserve Bank of New York for its financial support. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

The question whether financial markets efficiently process all available information is important for a wide range of issues in economics and finance. According to the strong form of the efficient market hypothesis, asset prices reflect all publicly available information and hence at any given point in time all information necessary to predict future asset prices should be contained in current asset prices. Several recent papers have either implicitly or explicitly challenged that view for the US Treasury bond market. Ang and Piazzesi (2003) find that models with macro factors forecast better than models with only unobservable factors. Moench (2008) documents that an affine model with factors extracted from a large cross-section of macroeconomic variables has superior predictive ability for Treasury yields relative to models which use information contained in Treasury yields alone. Ludvigson and Ng (2009) (LN1 henceforth) argue that factors extracted from large macroeconomic data sets carry information useful for predicting Treasury bond returns beyond the information contained in the Treasury yield curve itself. Duffee (2011) uses filtering to uncover some factors not (or only weakly) spanned by yields and shows that their dynamics are correlated with some of the macro factors in LN1. Joslin, Priebsch, and Singleton (2010) show that a real growth and an inflation factor, both orthogonal to the first three principal components of the Treasury yield curve, drive much of the variation in Treasury term premia.

These results can be interpreted as casting doubt on the validity of the efficient markets hypothesis. Of course, when testing whether information not incorporated in contemporaneous asset prices is useful for predicting future prices, it is important to correctly specify the information set available to investors in real time. This is especially true when macroeconomic time series are used as predictors since these are typically published with a lag and subject to revisions.

In this paper, we revisit the evidence in favor of bond return predictability using macroeconomic factors extracted from real-time data. Broadly speaking there are two key issues with the data being used to construct macroeconomic factors. These two issues are: (1) publication delays and (2) data revisions. Regarding the former, publication delays in monthly economic time series range from just a few days to over a month. Ignoring the second issue, suppose, for a moment, that there are no data revisions. The presence of publication delays then means that a predictive regression involving month t macro factors based on final releases is de facto a regression with $t + 1$ information since month t data is not available to market participants until one month later. The misalignment due to publication lags may therefore imply spurious predictability patterns. The second issue with using revised data in predictive regressions is that data revisions might be informative for future bond returns. We document that revisions represent a non-trivial share of the total variance in many macroeconomic time series. We further show that revision errors feature sizeable degrees of serial and cross-sectional correlation.

The combination of publication delays and serially as well as cross-sectionally correlated revision components implies that the factors extracted from a set of revised macroeconomic data likely do not span the same space as the factors extracted from the real-time data. Consequently, the predictive information contained in macroeconomic time series should be re-assessed on the basis of factors extracted from real-time as opposed to revised data if we want to examine expected bond returns with the information available to market participants.

LN1 find that factors extracted from a panel of macroeconomic variables predict excess returns on 2-,3-,4-, and 5-year bonds, controlling for information contained in the term structure. Our results show that, when we explicitly take into account the real-time nature of macroeconomic information, a sizeable fraction - but not all - of the in-sample predictive power of macroeconomic variables for future Treasury bond returns is explained by the data revision and publication lag components. We further document that the additional predictive information of factors extracted from revised macroeconomic data largely disappears in a truly real-time out-of-sample forecast exercise. Our results thus indicate that the efficient market hypothesis is not easily dismissed.

We also examine whether the source of mis-specification is primarily due to the mis-alignment caused by publication delays, or the presence of revision errors. We isolate each component in our analysis and find that both are important.

Our research follows in the footsteps of some recent papers documenting the potential impact of mis-specification of information sets on the estimation of asset pricing models. For example, Christoffersen, Ghysels, and Swanson (2002) re-examine macroeconomic mimicking portfolios, designed to maximize correlation with macroeconomic news. They find that when final data are used, as in Chen, Roll, and Ross (1986), as opposed to real-time series, one obtains very different tracking portfolios and loadings when projected on equity returns. Similarly, Gilbert (2011) shows that an empirical relation exists between stock returns on macroeconomic news announcement days and the future revisions of the released data. The results suggest that revisions do matter, i.e. that investors care about the final revised value of a macroeconomic series. So far this line of research has focused on equity markets, but fixed income markets - where arguably macroeconomic news may matter more - have not been studied.

Our paper also relates to a broader literature on the usefulness of real-time data. For example, Orphanides (2001) stresses the importance of examining monetary policy decisions with vintage data rather than final revised series. Moreover, our real-time data is based on outstanding research efforts undertaken at the Philadelphia Federal Reserve, making such data are publicly available and easy to use (see in particular Croushore and Stark (2001)). Appraising forecasting performance, particularly of macroeconomic series, with final versus real-time data has also been discussed extensively in the literature, see e.g. Stark and

Croushore (2002), Orphanides and van Norden (2002), Koenig, Dolmas, and Piger (2003), Croushore (2006), among many others.

The paper is organized as follows. In Section 2, we provide a detailed overview of the real-time data set that we construct to analyze the question at hand and discuss the time series properties of the revision components. In Section 3, we lay out the econometric framework used in the paper and document the main empirical findings in Section 4. Section 5 concludes.

2 Data

We construct one-year excess holding period returns for 2-, 3-, 4-, and 5-year U.S. Treasury bonds provided in the Fama-Bliss data set of the Center for Research in Securities Prices (CRSP). These data have been used e.g. by Cochrane and Piazzesi (2005) who document that a single linear combination of forward rates captures almost all of the predictable variation of one-year excess returns on bonds of different maturities. Hence, this particular linear combination (henceforth the CP factor) summarizes the predictive information in the yield curve for future bond returns. We therefore use the CP factor to proxy for the predictive information for future bond returns contained in current market prices.

To study the additional predictive content of macroeconomic factors, we construct a panel of time series that are published at a monthly frequency and for which the initial release numbers as well as the release dates are available for a sufficiently long time span. Unfortunately the availability of real-time data is only sparse before the 1980s. We find that a data set covering 68 economic time series and spanning the sample period 1982m3-2011m12 maximizes the total number of time series and cross-sectional observations. These series include measures of industrial production, employment, housing indicators, personal income, price indices and the money stock. We obtain these data from the Archival Federal Reserve Economic Database (ALFRED) at the Federal Reserve Bank of St. Louis. Appendix A provides a detailed account of our real-time data set, along with the original data source, the first date of real-time data availability, as well as the transformations that we apply to the series in order to ensure stationarity.

Our data set broadly covers the same economic categories as the one used by LN1. However, LN1 also include in their panel a number of financial time series and extract factors from the joint data set. We are interested in the predictive content of macroeconomic data in real-time above and beyond the information already contained in market prices. We therefore restrict ourselves to a data set covering only macroeconomic time series that are potentially subject to publication lags and data revisions.

2.1 Publication Lags and Data Revisions

Investors who aim to use macroeconomic information to predict bond returns in real-time face two issues. First, most macroeconomic time series are released with a lag, and second, they are typically subject to subsequent revisions. Both issues are important when assessing the usefulness of macro data for predicting bond returns ex-post. In this section, we set the stage for our analysis by documenting that publication lags are widespread and that data revisions capture a sizeable share of the total variation of macroeconomic time series. We define revisions as the simple difference between the last available vintage and the initial announcement of a given time series in a given calendar month.

The first row in Table 1 provides summary statistics across the 68 time series in our panel for the number of days between the last day of the calendar month and the publication date of the first release for a given variable. The table shows that the mean and the median series in our data set are on average published about 13 days after the end of the calendar month that they measure. While some series are published with an average lag of only about five days, for others it takes more than one month, on average, before they are released. This underscores that the information set available to investors in real-time may differ quite substantially from the one available to the econometrician in an ex-post analysis of bond return predictability.

We now turn to the importance of data revisions. The second row in Table 1 provides the same summary statistics for the ratio of sample variances of the revision component and the final revised series across variables. Here, we do not take into account the publication lag, but simply compare the revision as the difference between the first release and the final revised vintage of a variable for a given month. The average of the variance ratios across the series in our panel amounts to 0.68, indicating that revision components make up for a sizeable fraction of the total variation in final revised macroeconomic time series. Hence, the information content of the final and first vintages of macroeconomic time series is potentially quite different.

Our results are in line with earlier findings reported in Aruoba (2008) who documents the empirical properties of revisions to major macroeconomic variables and also finds that they are large relative to the variation in the original variables and feature substantial degrees of serial correlation. Regarding the latter, we find similar results. Looking at the summary statistics of the AR(1) coefficients for the revision components across the 68 variables shown in the third row of Table 1, we see that the average coefficient equals -0.44 suggesting that there is a substantial amount of serial correlation in these revision errors. In particular, most revision components are negatively correlated indicating that a positive revision in one month is often followed by a negative revision in the next month. To the extent that these revisions are systematic and make up for a substantial share of the variance in the final revised data, it is therefore not unlikely that their serial correlation contributes to the

predictive power for bond returns found in that data.

Revision errors are also cross-sectionally correlated, as shown by the fourth row in Table 1. This row documents the cross-sectional distribution of sample R^2 's obtained from regressing the 68 individual data revision components onto their first principal component. The cross-sectional mean amounts to 9% with a maximum of 67.3%, indicating that there is quite a degree of co-movement of revisions across macroeconomic variables. In line with this number, we find that the first principal component extracted from the revision components explains 9% of their total variation while the first three principal components explain 23%. Hence, data revision components are not only serially but also cross-sectionally correlated. This implies that the factors extracted from real-time data do not necessarily span the same space as the factors extracted from revised data. The finding that revisions are cross-sectionally correlated is consistent with Swanson, Ghysels, and Callan (1999) who document non-trivial multivariate dependence patterns in data revisions.

Overall, these summary statistics imply that bond return predictions based on macroeconomic factors might be quite different depending on whether real-time or revised data are used. In the sections that follow, we quantify this difference.

To conclude, it should also be noted that the construction of real-time factors and comparing them to final-data factors is of independent interest. The approach taken here is most closely related to the Chicago Fed National Activity Index (CFNAI) which is a weighted average of 85 monthly indicators of national economic activity.¹ The CFNAI is a principal component, and therefore is a weighted average of the 85 (transformed) economic indicators. Hence, it provides a single summary measure of a common factor in these national economic data. The CFNAI is revised with each monthly release because the underlying monthly data are released with varying degrees of delay.

3 Econometric Framework

We revisit the question whether macroeconomic variables carry information useful for predicting excess bond returns that is not subsumed by the cross-section of contemporaneous bond yields. Let Z_t denote a $K \times 1$ vector of conditioning variables obtained from the cross-section of bond yields in period t . These could be e.g. individual forward rates as in Fama and Bliss (1987), the return forecasting factor from Cochrane and Piazzesi (2005) which is obtained as a linear combination of five forward rates, or principal components of yields as in Joslin, Singleton, and Zhu (2011) or Adrian, Crump, and Moench (2012). It is standard

¹The economic indicators used for the CFNAI are drawn from four broad categories of data: (1) production and income (23 series), (2) employment, unemployment, and hours (24 series), (3) personal consumption and housing (15 series), and (4) sales, orders, and inventories (23 series). For a detailed discussion, see http://www.chicagofed.org/digital_assets/publications/cfnai/background/cfnai_background.pdf

practice in the literature to establish the predictive power of these conditioning variables by estimating the regression

$$rx_{t+12}^{(n)} = c_n + \beta'_n Z_t + \epsilon_{t+12}^{(n)}, \quad t = 1, \dots, T, \quad (1)$$

via OLS and then assessing the statistical and economic significance of the estimated coefficients $\hat{\beta}_n$ for holding period excess returns on bonds of different maturities. In principal, one can easily assess whether macroeconomic information is useful for predicting excess bond returns beyond the information contained in the yield curve factors Z_t by simply augmenting regression (1) with a set of $N \times 1$ macroeconomic predictor variables x_t :

$$rx_{t+12}^{(n)} = c_n + \gamma'_n x_t + \beta'_n Z_t + \epsilon_{t+12}^{(n)}, \quad t = 1, \dots, T. \quad (2)$$

However, as discussed in LN1, this regression quickly runs into degrees of freedom problems if N is large relative to T . As a remedy to the curse of dimensionality problem, LN1 consider the possibility that the macroeconomic variables x_{it} have a factor structure, i.e.

$$x_{it} = \lambda'_i F_t + e_{it}, \quad (3)$$

where F_t is an $r \times 1$ dimensional vector of common factors, λ'_i is a $r \times 1$ vector of factor loadings and where e_{it} denotes an idiosyncratic component. In matrix notation this becomes

$$X = F\Lambda + E,$$

where X is a $T \times N$ matrix of data observations, F is a $T \times r$ matrix of latent factor observations, Λ is a $r \times N$ matrix of factor loadings and E is a $T \times N$ matrix of idiosyncratic components. The number of common factors r is assumed to be much smaller than the number of cross-sectional elements N which implies that the space spanned by the variables in x is well approximated by the space spanned by the factors F . Hence, one can feasibly estimate the regression

$$rx_{t+12}^{(n)} = c_n + \gamma'_n f_t + \beta'_n Z_t + \epsilon_{t+12}^{(n)}, \quad (4)$$

where $f_t \subset F_t$ is the subset of pervasive factors that is important for predicting excess returns $rx_{t+12}^{(n)}$.

As the true common factors are not observed, one needs to replace F_t by estimates \hat{F}_t , which in practice are obtained via standard principal components techniques. In particular, if one allows for k factors in the estimation then under the restriction that $\Lambda'\Lambda/N = I_k$, Stock and Watson (2002) have shown that the factor loading matrix $\hat{\Lambda}$ can be consistently estimated by \sqrt{N} times the eigenvectors corresponding to the k largest eigenvalues of the matrix $X'X$. The corresponding factor estimates are given by $\hat{F} = X\hat{\Lambda}'/N$. Hence, the estimated factors represent cross-sectional averages of all variables in the panel where the weights are chosen such as to minimize the sum of squared idiosyncratic components E . LN1

obtain \hat{f}_t by constructing various subsets as well as non-linear functions of \hat{F}_t , and then using the Bayesian information criterion (BIC) to select the optimal model specification. Given this optimal set of factors extracted from a large panel of revised macroeconomic data, they then assess the statistical and economic significance of the estimated coefficients $\hat{\gamma}_n$ and conclude that macroeconomic factors indeed carry substantial information about future excess bond returns that is not spanned by the contemporaneous cross-section of bond yields.

Ludvigson and Ng (2011) (LN2 henceforth) extend the analysis in LN1 in a number of ways. Most importantly for our analysis, they use both an in-sample and an out-of-sample BIC criterion to select the best linear combination of factors. According to their out-of-sample BIC criterion, a small set of macroeconomic factors and nonlinear functions thereof feature significant predictive power for excess bond returns over the 1964-2007 sample period. In our empirical analysis, we use this set of factors as a benchmark for our own analysis.

Since the results in LN1 and LN2 are based on a set of revised macroeconomic data, a natural question that arises is whether investors could have exploited the macroeconomic information in real time. To formalize this, we denote x_{it} the final revised observation in period t for variable i . Further let \tilde{x}_{it} denote the reading on variable i that was available to investors in period t . In practice \tilde{x}_{it} may often be the time t release of x_i pertaining to period $t - 1$.² To the extent that our data set covers the information available to investors in real-time, even if they used (linear) projections, they wouldn't have a richer information set since the principal components that we extract span the dynamic factors which contain lags of the static factors.

More specifically, we can decompose

$$x_{it} = \tilde{x}_{it} + \nu_{it}, \quad (5)$$

where ν_{it} potentially contains two elements: a component that is purely related to data revisions and one that captures the fact that macroeconomic data are typically released with a lag. For most of the paper, we will consider the two components jointly as we are primarily interested in the predictive power of \tilde{x}_{it} with respect to $rx_{t+12}^{(n)}$. However, we will also shed some light on which of the two components in ν_{it} is more informative about future excess returns.

We assume that the real-time components of the macroeconomic data have a factor structure:

$$\tilde{x}_{it} = \tilde{\lambda}'_i F_t^{RT} + \tilde{e}_{it}, \quad (6)$$

²Alternatively, \tilde{x}_{it} may be some (linear) projection of the final release x_{it} onto the information at time t . Note, for example, that the initial CFNAI release includes *projected* monthly values for approximately one-third of the 85 series. In the following month's release when these missing data become available, correcting the projection error becomes a source of revision in the CFNAI. In our analysis we do not engage in projections of missing series, instead we simply keep track of the releases of the data in the construction of factors.

where F_t^{RT} denote the \tilde{r} factors that are common to the real-time components of the data and \tilde{e}_{it} the corresponding idiosyncratic components. In what follows, we estimate variants of the regression

$$rx_{t+12}^{(n)} = \tilde{c}_n + \tilde{\gamma}'_n f_t^{RT} + \tilde{\beta}'_n Z_t + \tilde{\epsilon}_{t+12}^{(n)}, \quad (7)$$

where $f_t^{RT} \subset F_t^{RT}$. We then assess whether the coefficients $\tilde{\gamma}'_n$ are statistically and economically significant and how the goodness of fit of the real-time predictive regression (7) compares to that of the regression (4) which uses revised data. We further perform a real-time out-of-sample exercise in which we extract the common factors month by month, then re-estimate equation (7) and use the estimated coefficients to predict excess returns one year ahead with the macroeconomic information available to investors in real-time.

4 Empirical Results

This section summarizes the empirical results of the paper. We start by comparing the factor spaces spanned by the principal components extracted from a larger panel of revised macro series as well as financial time series used in LN2. We find that the first factor extracted from our revised data set is highly correlated with the first factor in their data set which is the main driving force of the bond return predictability in their longer sample. We continue by assessing the first factor's in-sample predictive power when extracted from revised versus real-time data. We further study whether the revision components feature predictive power for bond returns. We then analyze if the real-time factor has significant predictive power for bond returns out-of-sample. Finally, we decompose the root-mean squared forecast error of the revised factor prediction model into the real-time and revision components in order to understand what drives the superior predictive power of revised data.

4.1 A Comparison of Factor Spaces

LN2 extract eight factors from their panel of 131 macroeconomic and financial time series and find that a subset of them best predicts excess bond returns in-sample. This subset comprises the first and eighth as well as the second power of the fifth and the third power of the first principal component. Their results further suggest that the first principal component, which heavily loads on business cycle indicators such as measures of production and employment, captures the bulk of the predictive power of the macro factors for bond returns. Since our data set captures fewer series than the one employed by LN2, it is not clear ex-ante whether the factors extracted from our panel of revised data similarly capture predictive information for bond returns. Table 2 reports the pairwise correlation coefficients between the first eight principal components extracted from our data set with the four factors which LN2 find to best predict bond returns in their sample. We denote the latter by f_{1t}^{LN} , f_{8t}^{LN} , $(f_{5t}^{LN})^2$, and

$(f_{1t}^{LN})^3$, respectively. Similarly, we label the principal components extracted from our revised and real-time data set as $f_{1t}^{RV}, \dots, f_{8t}^{RV}$ and $f_{1t}^{RT}, \dots, f_{8t}^{RT}$, respectively.

The upper panel of the table shows these correlations for the principal components extracted from our revised data, and the lower panel reports them for the components extracted from the real-time data. As the upper panel shows, the first principal component from LN2's data set and the first principal component from our revised data are strongly correlated with a pairwise correlation coefficient of 0.85. Table 2 also shows that none of the remaining factors extracted from our revised data set are much correlated with the set of principal components that LN2 find to have forecasting power for bond returns in their sample.

The second panel in Table 2 reports the pairwise correlation coefficients between the LN2 factors and the first eight principal components extracted from our real-time data set. The panel shows that most of the pairwise correlations are considerably lower for the real-time counterparts, consistent with our earlier finding that publication lags and data revisions represent a sizeable fraction of the variance in the revised data. Most importantly, the pairwise correlation coefficient between f_t^{LN} and the first principal component in our data set drops from 0.85 to 0.55 when using real-time data. This suggests that the predictive power for bond returns contained in the macro data might well be different when using real-time data. All pairwise correlations of higher order principal components with the LN2 factors are essentially zero. In our subsequent analysis we therefore focus on analyzing the predictive content of the first principal component extracted from both the revised and the real-time version of our panel of 68 monthly time series.

Figure 1 superimposes the first principal components extracted from our revised and real-time data, respectively. It shows the while both factors are highly correlated, they differ markedly in some periods. Most importantly, the real-time factor f_{1t}^{RT} visibly lags the revised factor f_{1t}^{RV} highlighting the importance of the lags with which macroeconomic data are published.

4.2 In-Sample Predictive Regressions

We investigate the in-sample predictive power of f_{1t}^{RV} and f_{1t}^{RT} with respect to the set of principal components that LN2 find to be the best forecasters of excess bond returns. We then turn to the question whether the revision components comprised in final revised data carry predictive information that is not captured by the real-time components of the data or contemporaneous bond yields.

Table 3 provides estimation results for equation (4) which compare the predictive power of the different sets of principal components for one-year excess holding period returns on Treasury notes with 2-, 3-, 4-, and 5-years to maturity, respectively, in the four panels of the table. The left-hand columns show results for the factors found by LN2 to be the best predictors of bond returns in their data set. We use the factors estimated by LN2 for the full

sample from January 1964 through December 2007, but use March 1982 as the start date for the regression to enable direct comparison with the factors extracted from our shorter data set. The right-hand side columns replace f_{1t}^{LN} and its cubed version $(f_{1t}^{LN})^3$ by the corresponding counterparts of f_{1t}^{RV} and f_{1t}^{RT} , respectively. This allows us to directly assess the individual significance (in terms of t -statistics based on Newey-West adjusted standard errors) as well as goodness of fit (as measured by the adjusted R^2) of the predictive bond regressions obtained using the three different information sets. Like LN1 and LN2, we use the CP factor to summarize the predictive information of contemporaneous bond yields for bond returns one year out.

The first two rows in each of the four panels in Table 3 show that f_{1t}^{LN} is a strongly significant predictor of the excess one-year holding period returns on Treasury notes. This finding also holds when controlling for the predictive information contained in the CP factor as well as the other LN2 factors. Moreover, the first principal component explains the bulk of the predictive power of the macro factors since the adjusted R^2 s are very similar across the two specifications. This is in line with the evidence in LN1 as well as LN2; both find that the first principal component is the single best predictor of excess returns in their samples. They interpret this finding as being consistent with models of countercyclical risk premia.

The third and fourth row in each panel document that the in-sample predictive power of f_{1t}^{RV} is comparable to that of f_{1t}^{LN} as both the point coefficients and the t -statistics and R^2 s are similar to those in the first two rows. However, looking at the last two rows in each panel, the picture is not the same when we consider the first principal component extracted from the real-time data set. In fact, for all four maturities the point coefficients on each f_{1t}^{RT} is substantially lower in absolute terms than the one obtained for its revised counterpart, f_{1t}^{RV} . Moreover, the factor remains a barely significant predictor of excess bond returns individually and completely loses its statistical significance when controlling for the remaining LN2 factors as well as its own cubed version. Hence, the results in this table show that the predictive power of macroeconomic factors for bond returns is substantially reduced when real-time data are used.

4.2.1 Publication Lags or Data Revisions?

Is this finding driven by the revision component in revised macroeconomic time series? To answer this question, we augment the predictive bond return regressions using the real-time factor f_{1t}^{RT} and the CP factor with two different measures of the revision component. The first, f_{1t}^{ν} , is simply the first principal component extracted from the 68 time series of revision components of the individual variables in our panel. The second applies the loadings corresponding to f_{1t}^{RV} , the first principal component extracted from the revised data set, to the individual revision components in order to obtain a measure of the revision component entailed by f_{1t}^{RV} . We label this component $f_{1t}^{\nu RV}$.

The left-hand panel of Table 4 shows the results from these regressions for the same sample period used above, 1982m3-2007m12. The right-hand side panels repeat the same regressions for the extended sample period 1982m3-2011m12, which covers the recent financial crisis period. We first discuss the results for the sample ending in 2007. The first row in each of the four left-hand side panels of Table 4 reproduces the last row in each of the panels in Table 3. The second row uses the first principal component of revision errors, f_{1t}^ν , as a regressor. The third row uses $f_{1t}^{\nu RV}$, the revision component comprised in f_{1t}^{RV} , as a regressor. The last two rows in each panel repeat these regressions controlling for f_{1t}^{RT} . The results from these regressions show that both f_{1t}^ν as well as $f_{1t}^{\nu RV}$ are highly significant predictors of excess bond returns for all four maturities that we consider. Indeed, the reported R^2 s of the regressions using only the revision components show that $f_{1t}^{\nu RV}$ explains a similar share of the variation in excess bond returns as the real-time macroeconomic factor. Moreover, in a joint regression with f_{1t}^{RT} , the coefficients on the revision components f_{1t}^ν and $f_{1t}^{\nu RV}$ remain strongly significantly different from zero. In fact, the significance of the real-time factor increases once the revision components are added as a regressor, indicating that the covariance between the real-time macroeconomic factor and the revision components is important. Importantly, the signs on the revision component coefficients are all positive, suggesting that bond returns tend to be higher following periods of large revisions to macroeconomic time series, which tend to occur around recessions.

Our findings remain intact when we re-estimate the same regressions over the longer sample ending in 2011, as shown in the right-hand side panels of Table 4. In particular, the coefficients on revision components are all highly statistically significant, suggesting that one is likely to overstate the extent to which Treasury returns are predictable by macroeconomic factors. However, a notable difference with respect to the shorter sample is that the variance explained by both the real-time factor as well as the revision components are generally lower than for the sample ending in 2007. This suggests that forces not captured by our set of macroeconomic indicators (such as large-scale asset purchases by the Fed) may have determined bond returns over the most recent sample period.

In the above regressions we have obtained f_{1t}^ν and $f_{1t}^{\nu RV}$ from the simple differences between the final revised observations of all series in a given month and the recorded real-time observations in the same month. Since most series are published with a substantial delay (see Table 1 above), these differences do not allow us to discriminate between the effects of actual data revisions and publication lags. In order to assess to what extent the publication lag alone can impact the predictive power of macroeconomic factors, we perform the following simple exercise. Instead of using the twelve-month lag of f_{1t}^{RV} as a predictor of current excess bond returns, we run regressions using 13-month and 14-month lagged observations of the revised factor, respectively. At a monthly sampling frequency, given the publication lags of up to 35 days for the variables in our data set, one would essentially have

to wait two months to get a first release of all series in the panel for the current calendar month. Hence, lagging the revised factor by fourteen instead of twelve months (as in the regressions in Table 3 above) gives us a sense of the real-time predictive content of the macroeconomic data absent any subsequent data revisions. Since most data series in our panel feature average publication lags less than 35 days, we also perform a similar analysis using the 13-month lag of the revised factor estimate f_{1t}^{RV} as a predictor of current excess bond returns, assuming implicitly that all revised data for a calendar month are available one month later. In these regressions we control for the twelve-month lag of the CP factor in order to take into account the predictive information contained in contemporaneous bond prices, which according to the efficient market hypothesis should reflect market participants' knowledge of the current state of the economy.

Table 5 shows the results of these regressions. The left-hand panel reports the regression results for the 1982-2007 sample and the right-hand columns show the corresponding results for the sample ending in 2011. We start by discussing the results for the sample excluding the financial crisis. The first row in the four horizontal panels on the left-hand side of the table shows results for a regression of contemporaneous bond returns on one-year lagged values of the CP factor and f_{1t}^{RV} and is identical to the fourth row in Table 3. As discussed above, the significance of the coefficient on the one-year lagged factor suggests that by using final revised data there appears to be significant predictive information in macroeconomic time series for one-year excess holding returns on bonds beyond that contained in current bond prices. The second and third row of each of the four horizontal panels of Table 5 lag the revised factor by one or two additional months, respectively. While still statistically significantly different from zero, the coefficients on the 13-month lagged factor observation are considerably smaller in magnitude and the adjusted R^2 s of the regression are also substantially lower. More importantly, we find that the coefficients on the 14-month lagged factor observation are all insignificant and the adjusted R^2 s of these regressions are still substantially lower. These results imply that - absent any data revisions - the first principal component extracted from our data set would not have any incremental predictive power for excess bond returns beyond the CP factor if one waited for all variables in the panel to be published. Hence, small changes in the information set available to investors can have quite drastic implications for the degree of excess return predictability associated with macroeconomic variables

The right-hand side panel of Table 5 repeats these same regressions for the longer sample including the recent financial crisis. While we still see the absolute values of the point coefficients as well as the predictive R^2 s decline substantially from the regression using the 12-month lag to the one using the 14-month lagged factor observation, the associated coefficients retain their statistical significance over the longer sample. This reinforces our finding above that there appears to be more predictive power of macroeconomic variables if one includes the financial crisis.

However, in summary, the results presented in this section show that the predictive content of macroeconomic factors is substantially reduced if we take into account the real-time nature of the data and that the stronger forecasting power of revised data can to a good extent be attributed to the incremental predictive information contained in data revision components.

4.3 Out-of-Sample Predictions

In this section we analyze whether the in-sample results discussed above carry over to a truly real-time out-of-sample setting. Specifically, we run the following exercise. Starting in January 1994, we extract the first principal component from the revised and real-time version of our data set of 68 macroeconomic variables using data from March 1982 through the month when the forecast is made. We also use bond price data up to the current month to re-estimate the CP return forecasting factor. We then run three regressions of the individual bond returns on (i) a constant (ii) a constant and the twelve month lag of the estimated principal component, and (iii) a constant, the lagged principal component as well as the lagged CP factor as regressors. For each of these specifications, we use the estimated regression coefficients to predict excess bond returns twelve months out and record the corresponding forecast error. We then assess whether the estimated macroeconomic factors significantly improve forecast accuracy by computing the ratio of mean squared forecast errors of the restricted models (i.e. specifications which only use a constant or a constant and the CP factor) divided by the unrestricted models which add the extracted first principal component from the respective data set (i.e. f_{1t}^{LN} , f_{1t}^{RV} , or f_{1t}^{RT}).

The left- and right-hand side panels of Table 6 summarize the results from this exercise for the two sample periods ending in 2007:12 and 2011:12, respectively. As before, the four horizontal panels show analogous regression results for 2-,3-,4-, and 5-year maturity Treasury notes. The first and fourth row in each of the four horizontal panels report the forecast results using f_{1t}^{LN} as the predictor variable. As we only have data for the LN2 data set through the end of 2007, we can only provide results for these specifications in the left-hand side panels of the table. The results for f_{1t}^{LN} confirm that the first principal component extracted from LN2's panel of revised macro as well as financial time series significantly reduces the mean squared out-of-sample forecast errors with respect to both the simple rolling mean prediction (rows 1, 7, 13, and 19 in the table) as well as the forecast model using a constant and the CP factor as regressors (rows 4, 10, 16, 22).

Comparing the out-of-sample predictions based on a constant and our first factor extracted from revised macro data, f_{1t}^{RV} , with those of a rolling mean model (rows 2, 8, 14, 20), we see that this factor reduces forecast errors by about 4% to 6% depending on the maturity. According to the ENC-NEW test suggested by Clark and McCracken (2001), these MSE reductions are statistically significant at the 5 percent level for all four bond

maturities that we consider. This result prevails when we control for the information contained in contemporaneous bond prices by adding the CP factor as an additional regressor to both forecasting models. Indeed, the improvement in forecast precision when adding f_{1t}^{RV} to the model that includes a constant and the CP factor is quite sizeable, ranging from 14% for the 5-year bond to 22% for the 2-year bond. Not surprisingly, these large reductions in mean-squared forecast errors are found to be significant according to the ENC-NEW statistic.

We next look at the out-of-sample forecast results using the real-time factor, f_{1t}^{RT} . The comparison between a simple rolling mean forecast as well as a model that includes both a constant and f_{1t}^{RT} (rows 3, 9, 15, 21) shows that adding the real-time factor actually results in larger out-of-sample forecast errors since the ratio of root mean squared errors of the unrestricted to the restricted model exceeds one for all maturities. Hence, by itself the first principal component extracted from our panel of real-time variables does not add any useful information for predicting bond returns. This is in sharp contrast to the results obtained both for f_{1t}^{LN} as well as f_{1t}^{RV} , which do show a significant forecast improvement.

Interestingly, the real-time macro factor does carry useful predictive information jointly with the CP factor (rows 6, 12, 18, 24). Indeed, we find that the reduction of mean squared forecast errors is economically and statistically significant when adding f_{1t}^{RT} to the model with a constant and the CP factor. This indicates that the interaction between contemporaneous bond prices and the information contained in macroeconomic variables may be important for predicting bond returns.

We find somewhat different out-of-sample forecast results for the sample period ending in 2011. Over the longer sample including the financial crisis neither our revised nor our real-time factor help improve forecasts relative to the rolling mean model. However, when adding the factors to the model consisting of a constant and the CP factor, both f_{1t}^{RV} and f_{1t}^{RT} reduce forecast errors in a sizeable and statistically significant way. In unreported results we see that the mean squared forecast error of the model that uses a constant and the CP factor is much larger than that of the rolling mean forecasts. Hence, the fact that we find a significant improvement in forecast results when adding the macroeconomic factors to the CP model is merely indicating of how poorly the CP factor has performed as a predictor of bond returns in the crisis, rather than documenting the usefulness of factors extracted from real-time macroeconomic data.

In sum, Table 6 documents that our in-sample results about the limited usefulness of real-time macroeconomic data for bond return prediction carry over to a pure out-of-sample setting. In the next subsection, we will provide a decomposition of the mean squared forecast errors and study why revised data provide better forecasts than real-time data.

4.3.1 Decomposition of the mean squared forecast error

When we examine out-of-sample performance, we would also like to know the contribution of each of the various sources of prediction error. In particular, we address the question, how do the mean squared forecast error (MSFE) implied by the revised data and the real-time factor relate to each other? In this section we propose such a decomposition, in the spirit of Theil's U statistic (see Theil (1961)).

Recall that our forecasts are based on variants of regression (7). In order to save on notation, we add the intercept to the regressors Z_t and denote it as \tilde{Z}_t . Then, ignoring the superscript denoting bond return maturity let

$$\begin{aligned}\widehat{rx}_{t+12}^{RV} &= (\tilde{\gamma}_t^{RV})' f_t^{RV} + (\tilde{\beta}_t^{RV})' \tilde{Z}_t \\ \text{and } \widehat{rx}_{t+12}^{RT} &= (\tilde{\gamma}_t^{RT})' f_t^{RT} + (\tilde{\beta}_t^{RT})' \tilde{Z}_t\end{aligned}\quad (8)$$

be the model-implied expected one-year excess holding period return where f_t^i , $i = RV$ and RT are the factors from our revised and real-time data set, respectively.³ Given k observations of out-of-sample predictions (again ignoring the variations of maturity), we have

$$\begin{aligned}MSFE_{RV} &\equiv \frac{1}{k} \sum_{t=1}^k (rx_k - \widehat{rx}_k^{RV})^2 \equiv \frac{1}{k} \sum_{t=1}^k (e_t^{RV})^2 \\ \text{and } MSFE_{RT} &\equiv \frac{1}{k} \sum_{t=1}^k (rx_k - \widehat{rx}_k^{RT})^2 \equiv \frac{1}{k} \sum_{t=1}^k (e_t^{RT})^2.\end{aligned}\quad (9)$$

Then we have the following relationship between $MSFE_{RV}$ and $MSFE_{RT}$:

$$\begin{aligned}MSFE_{RT} &\equiv MSFE_{RV} + \frac{1}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT})^2 + \frac{1}{k} \sum_{t=1}^k (\tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))^2 \\ &\quad + \frac{1}{k} \sum_{t=1}^k ((\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)^2 + \frac{2}{k} \sum_{t=1}^k (e_t^{RT} (\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT}) \\ &\quad + \frac{2}{k} \sum_{t=1}^k (e_t^{RT} \tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV})) + \frac{2}{k} \sum_{t=1}^k (e_t^{RT} (\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t) \\ &\quad + \frac{2}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT} \tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV})) \\ &\quad + \frac{2}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT} (\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t) \\ &\quad + \frac{2}{k} \sum_{t=1}^k (\tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}) (\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t).\end{aligned}\quad (10)$$

³Recall that in order to perform a real-time out-of-sample exercise in which we extract the common factors month by month, we re-estimate equation (7). Therefore all the parameters in the above equations are time-varying.

Many of the terms in the above equation have a straightforward interpretation, while others do not. For example $\frac{1}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT})^2$ measures the contribution due to the difference in estimated factor loadings, and conversely $\frac{1}{k} \sum_{t=1}^k (\tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))^2$ pertains to the difference in the factors. A priori we expect some of the components to be small, such as for example $\frac{1}{k} \sum_{t=1}^k ((\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)^2$, since the intercept estimates should be similar in both specifications.

In Table 7 we report the above decomposition for regressions using (i) an intercept, (ii) the one-year lag of the Cochrane and Piazzesi (2005) return forecasting factor CP_t and (iii) the one-year lagged observations of the first principal component extracted from our panel of 68 revised (f_{1t}^{RV}) or real-time macroeconomic time series (f_{1t}^{RT}). For each of the two sample periods, we present results for excess returns on 2-, 3-, 4-, and 5-year bonds. We also estimated these regressions without the CP factor and found similar results. Therefore the results reported in Table 7 may be considered robust with regards to model specification. The top and bottom rows display $MSFE_{RT}$ and $MSFE_{RV}$. The former ranges from 1.930 to 26.917 across all maturities and the two samples and the latter takes on values between 1.645 and 24.336. Which components are the main contributors to these differences? We find that those involving $(f_t^{RT} - f_t^{RV})$ are dominant. In particular, (1) $\frac{1}{k} \sum_{t=1}^k (\tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))^2$ ranges from 0.115 to 0.897, with all entries being highly statistically significant, and (2) $\frac{2}{k} \sum_{t=1}^k (e_t^{RT} \tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))$, taking on statistically significant negative values between -0.293 and -2.194 . Since $\tilde{\gamma}_t^{RV}$ is negative and features little variation, we can interpret this result as meaning that large forecast errors are concurrent with large revisions in the factors. The changes in loadings also have a significant impact, albeit less important than the contributions attributed to $(f_t^{RT} - f_t^{RV})$. For example, $\frac{2}{k} \sum_{t=1}^k (e_t^{RT} (\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT})$ takes on values between -0.033 and -0.953 , although they are only significant in the 1982m3-2011m12 sample. Finally, the contribution of $\frac{1}{k} \sum_{t=1}^k ((\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)^2$, which we expect to be small, is indeed taking on values between 0.003 and 0.017, although they appear to be statistically significant.

The decomposition clearly tells us that the out-of-sample forecast performance is most significantly affected by the (mis-)specification of the factors. It further shows that the forecast errors implied by the revised data factors are lower than those implied by the real-time factors mainly because in times of large data revisions forecast errors are also large.

5 Conclusion

In this paper, we have revisited the predictive power of macroeconomic factors for bond returns taking into account the real-time nature of macroeconomic information. We have shown that a sizeable fraction of the in-sample predictive power of macroeconomic variables for future Treasury bond returns is explained by the data revision and publication

lag components. We have further documented that the additional predictive information of factors extracted from revised macroeconomic data largely disappears in a truly real-time out-of-sample forecast exercise. Moreover, we have shown that factors extracted from revised macroeconomic data imply smaller forecast errors than factors extracted from real-time data mainly because data revisions and forecast errors are positively correlated. In sum, our results underscore that the efficient market hypothesis is not easily dismissed.

Our results suggest a number of avenues for future research. First, as noted before, the factors extracted from a set of revised macroeconomic data do not span the same space as the factors extracted from the real-time data. The use of final data to extract factors is prevalent in the literature. Only a handful of papers, including Christoffersen, Ghysels, and Swanson (2002) and Gilbert (2011), acknowledge the importance of properly aligning data releases when studying the macroeconomic sources of return predictability. The construction of real-time factors is therefore of independent interest.

Second, given prior evidence on the predictive power of macroeconomic factors for conditional first and second moments of stock returns (e.g. Ludvigson and Ng (2007)) it would be interesting to see how much of this predictability ‘survives’ in real-time.

Finally, a few recent papers (see e.g. Joslin, Priebsch, and Singleton (2010), Duffee (2011), Wright (2011), among others) have argued that macroeconomic factors that are not spanned by the cross-section of Treasury yields contribute in a sizeable way to the dynamics of Treasury risk premia. These studies commonly use final revised macroeconomic data to capture the information set available to investors. Since our results have shown that the predictive content of macroeconomic factors for excess bond returns depends crucially on the publication lag and data revision components comprised in final revised macroeconomic variables, it would be interesting to analyze the importance of unspanned macroeconomic factors when taking into account the real-time nature of macroeconomic data. We leave this for future research.

References

- Adrian, T., R. Crump, and E. Moench, 2012, “Pricing the term structure with linear regressions,” *Journal of Financial Economics*, forthcoming.
- Ang, A., and M. Piazzesi, 2003, “A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables,” *Journal of Monetary Economics*, 50(4), 745–787.
- Aruoba, S., 2008, “Data revisions are not well behaved,” *Journal of Money, Credit and Banking*, 40(2-3), 319–340.
- Chen, N., R. Roll, and S. Ross, 1986, “Economic forces and the stock market,” *Journal of Business*, 59(3), 383–403.
- Christoffersen, P., E. Ghysels, and N. Swanson, 2002, “Let’s get “real” about using economic data,” *Journal of Empirical Finance*, 9(3), 343–360.
- Clark, T., and M. McCracken, 2001, “Tests of equal forecast accuracy and encompassing for nested models,” *Journal of Econometrics*, 105(1), 85–110.
- Cochrane, J., and M. Piazzesi, 2005, “Bond risk premia,” *American Economic Review*, 95(1), 138–160.
- Croushore, D., 2006, “Forecasting with real-time macroeconomic data,” vol. 1 of *Handbook of Economic Forecasting*, pp. 961–982. Elsevier.
- Croushore, D., and T. Stark, 2001, “A real-time data set for macroeconomists,” *Journal of Econometrics*, 105(1), 111–130.
- Duffee, G., 2011, “Information in (and not in) the term structure,” *Review of Financial Studies*, 24(9), 2895–2934.
- Fama, E., and R. Bliss, 1987, “The information in long-maturity forward rates,” *American Economic Review*, 77(4), 680–92.
- Gilbert, T., 2011, “Information aggregation around macroeconomic announcements: Revisions matter,” *Journal of Financial Economics*, 101(1), 114–131.
- Joslin, S., M. Pribsch, and K. Singleton, 2010, “Risk premiums in dynamic term structure models with unspanned macro risks,” Working Paper, Stanford University.
- Joslin, S., K. Singleton, and H. Zhu, 2011, “A new perspective on Gaussian dynamic term structure models,” *Review of Financial Studies*, 24(3), 926–970.

- Koenig, E., S. Dolmas, and J. Piger, 2003, "The use and abuse of real-time data in economic forecasting," *Review of Economics and Statistics*, 85(3), 618–628.
- Ludvigson, S., and S. Ng, 2007, "The empirical risk-return relation: A factor analysis approach," *Journal of Financial Economics*, 83(1), 171–222.
- Ludvigson, S., and S. Ng, 2009, "Macro factors in bond risk premia," *Review of Financial Studies*, 22(12), 5027–5067.
- Ludvigson, S., and S. Ng, 2011, "A factor analysis of bond risk premia," in *Handbook of Empirical Economics and Finance*, ed. by A. Ulah, and D. Giles, pp. 313–372. Chapman and Hall.
- Moench, E., 2008, "Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach," *Journal of Econometrics*, 146(1), 26–43.
- Orphanides, A., 2001, "Monetary policy rules based on real-time data," *American Economic Review*, 91(4), 964–985.
- Orphanides, A., and S. van Norden, 2002, "The unreliability of output-gap estimates in real time," *Review of Economics and Statistics*, 84(4), 569–583.
- Stark, T., and D. Croushore, 2002, "Forecasting with a real-time data set for macroeconomists," *Journal of Macroeconomics*, 24(4), 507–531.
- Stock, J., and M. Watson, 2002, "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, 97, 1167–1179.
- Swanson, N., 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 *Cointegration, Causality and Forecasting - A Festschrift in Honour of Clive W. J. Granger*, ed. by R. Engle, and H. White. Oxford University Press.
- Theil, H., 1961, *Economic forecasts and policy*. North-Holland Pub. Co.
- Wright, J., 2011, "Term premia and inflation uncertainty: Empirical evidence from an international panel dataset," *American Economic Review*, 101(4), 1514–1534.

Tables and Figures

Table 1: **Summary Statistics for Publication Lags and Revision Components**

This table provides summary statistics and regression R^2 's for the publication lag and revision components comprised in the 68 time series of our macroeconomic data set. For each of the statistics, the cross-sectional mean, standard deviation, median, min, and max are shown across the 68 variables in our panel of macroeconomic time series. 'Publication' Lag refers to the average number of days it takes from the end of the calendar month it measures until the first release of a series is published. 'Variance Ratio' denotes the ratio of the variance of the pure revision component (i.e. ignoring the publication lag) and the variance of the final revised series. AR(1) Coefficient of ν_{it} refers to the estimated first-order autoregressive coefficient associated with the revision component. The ' R^2 : ν_{it} on f_{1t}^ν ' is the R^2 obtained from a regression of the individual revision components on the first principal component extracted from all revision components, f_{1t}^ν . ' R^2 : x_{it} on f_{1t}^{RV} ' is the R^2 obtained from a regression of the individual final revised series on the first principal component extracted from our set of 68 final revised macroeconomic time series. ' R^2 : x_{it} on f_{1t}^{LN} ' denotes the R^2 obtained from a regression of the individual final revised series on the first principal component extracted from the data set used by Ludvigson and Ng (2009) which comprises both final revised macroeconomic as well as financial time series.

	Mean	StDev	Median	Min	Max
Publication Lag	12.749	9.182	12.971	4.842	35.106
Variance Ratio	0.677	0.722	0.412	0.091	4.685
AR(1) Coefficient of ν_{it}	-0.435	0.197	-0.489	-0.684	0.106
R^2 : ν_{it} on f_{1t}^ν	0.088	0.155	0.013	0.000	0.673
R^2 : x_{it} on f_{1t}^{RV}	0.151	0.240	0.011	0.000	0.845
R^2 : x_{it} on f_{1t}^{LN}	0.172	0.204	0.106	0.000	0.725

Table 2: **Correlation with Ludvigson and Ng (2011) Factors**

This table shows the pairwise correlations between the set of factors which Ludvigson and Ng (2011) find have the strongest predictive power for bond returns, and the first eight principal components extracted from our revised and real-time macroeconomic data set, respectively. The Ludvigson and Ng factors are labeled with a superscript LN , the principal components extracted from our revised data set are indicated by a superscript RV , and the principal components extracted from our real-time data set are labeled with a superscript RT . The LN2 factors have been extracted from data covering the sample period 1964-2007. The revised and real-time factors have been extracted from monthly data covering the period 1982:03-2011:12. The correlations have been computed over that sample period.

	f_1^{LN}	f_8^{LN}	$(f_5^{LN})^2$	$(f_1^{LN})^3$
Revised Factors				
f_1^{RV}	0.853	0.140	-0.323	0.606
f_2^{RV}	0.032	-0.101	-0.051	0.034
f_3^{RV}	-0.038	-0.102	0.170	-0.047
f_4^{RV}	-0.011	-0.141	-0.001	0.001
f_5^{RV}	-0.013	-0.165	0.029	0.005
f_6^{RV}	0.034	-0.168	0.019	0.050
f_7^{RV}	0.059	-0.277	-0.053	0.020
f_8^{RV}	-0.003	0.071	0.024	-0.039
Real-Time Factors				
f_1^{RT}	0.552	0.128	-0.179	0.397
f_2^{RT}	-0.140	0.050	-0.034	-0.055
f_3^{RT}	0.055	-0.051	0.025	0.048
f_4^{RT}	0.065	-0.014	-0.050	0.088
f_5^{RT}	-0.057	-0.064	0.066	-0.094
f_6^{RT}	0.030	0.059	-0.097	0.046
f_7^{RT}	-0.032	0.050	0.043	0.027
f_8^{RT}	-0.030	0.080	-0.018	0.002

Table 3: In-Sample Bond Return Regressions

This table shows regression results for predictive bond return regressions of the form

$$rx_{t+12}^{(n)} = \tilde{c}_n + \tilde{\gamma}'_n F_t^I + \tilde{\beta}'_n CP_t + \tilde{\epsilon}_{t+1}^{(n)},$$

where $rx_{t+12}^{(n)}$ denotes the one-year excess holding period return on an n -year bond, CP_t denotes the one-year lag of the Cochrane and Piazzesi (2005) return forecasting factor, and F_t^I denotes the one-year lagged observations of a vector of factors extracted from one of three large macroeconomic data sets. Factors with the superscript LN have been extracted from the Ludvigson and Ng (2011) panel of revised macroeconomic and financial variables. Factors with superscript RV have been extracted from our panel of 68 revised macroeconomic time series. Factors with superscript RT are from the real-time counterpart of that data set. The t -statistics in parentheses are based on Newey-West standard errors with a maximum lag of 18 months. The four panels present results for excess returns on 2-, 3-, 4-, and 5-year bonds, respectively. The sample period is 1982:03-2007:12.

f_{1t}^{LN}	f_{8t}^{LN}	$(f_{5t}^{LN})^2$	$(f_{1t}^{LN})^3$	CP	f_{1t}^{RV}	$(f_{1t}^{RV})^3$	f_{1t}^{RT}	$(f_{1t}^{RT})^3$	\bar{R}^2
2-YEAR									
-0.620 (-3.431)	-0.161 (-1.964)	0.056 (0.809)	0.051 (2.703)	0.413 (4.769)					0.352
-0.436 (-3.088)				0.376 (3.702)					0.333
	-0.127 (-1.516)	0.039 (0.578)		0.423 (4.540)	-0.484 (-2.656)	0.020 (1.206)			0.338
	-0.151 (-1.490)	0.105 (1.285)		0.421 (4.719)			-0.219 (-1.100)	-0.004 (-0.124)	0.297
				0.406 (4.035)	-0.439 (-3.018)				0.335
				0.412 (4.369)			-0.272 (-2.165)		0.288
3-YEAR									
-1.039 (-3.392)	-0.408 (-2.414)	0.167 (1.280)	0.088 (2.420)	0.773 (4.559)					0.337
-0.757 (-3.484)				0.697 (3.541)					0.310
	-0.347 (-1.998)	0.113 (0.909)		0.790 (4.315)	-0.846 (-2.868)	0.019 (0.660)			0.338
	-0.404 (-1.957)	0.237 (1.586)		0.786 (4.482)			-0.312 (-0.876)	-0.037 (-0.679)	0.301
				0.750 (3.792)	-0.860 (-3.904)				0.329
				0.761 (4.116)			-0.547 (-2.516)		0.280
4-YEAR									
-1.270 (-3.069)	-0.608 (-2.357)	0.242 (1.307)	0.136 (2.489)	1.163 (4.955)					0.342
-0.830 (-2.847)				1.048 (3.766)					0.310
	-0.543 (-2.036)	0.131 (0.704)		1.175 (-2.606)	-1.046 (0.575)	0.025			0.347
	-0.623 (-2.014)	0.290 (1.366)		1.166 (4.722)			-0.288 (-0.567)	-0.057 (-0.728)	0.316
				1.107 (3.913)	-1.069 (-3.676)				0.336
				1.120 (4.219)			-0.631 (-2.063)		0.294
5-YEAR									
-1.181 (-2.253)	-0.716 (-2.254)	0.373 (1.637)	0.144 (1.984)	1.367 (4.580)					0.296
-0.766 (-2.265)				1.243 (3.543)					0.267
	-0.654 (-2.007)	0.218 (0.976)		1.374 (4.222)	-1.062 (-2.278)	0.022 (0.438)			0.305
	-0.752 (-2.027)	0.383 (1.519)		1.361 (4.340)			-0.182 (-0.294)	-0.085 (-0.846)	0.284
				1.299 (3.634)	-1.143 (-3.366)				0.294
				1.312 (3.890)			-0.662 (-1.796)		0.262

Table 4: Revision Error Regressions

This table shows regression results for predictive bond return regressions of the form

$$rx_{t+12}^{(n)} = \tilde{c}_n + \tilde{\gamma}'_n f_{1t}^{RT} + \tilde{\delta}'_n f_{1t}^* + \tilde{\beta}'_n CP_t + \tilde{\epsilon}_{t+1}^{(n)},$$

where $rx_{t+12}^{(n)}$ denotes the one-year excess holding period return on an n -year bond, CP_t denotes the one-year lag of the Cochrane and Piazzesi (2005) return forecasting factor, and f_{1t}^{RT} denotes the one-year lagged observation of the first principal component extracted from our real-time macroeconomic data set. f_{1t}^* represents either of two different measures of data revisions: f_{1t}^ν is the first principal component extracted from all individual data revision components, and $f_{1t}^{\nu RV}$ measures the revision component entailed by the first principal component extracted from the revised version of our data set. The t -statistics in parentheses are based on Newey-West standard errors with a maximum lag of 18 months. The two vertical panels present results for the sample periods 1982 : 03 – 2007 : 12 and 1982 : 03 – 2011 : 12, respectively.

1982m3-2007m12					1982m3-2011m12				
CP	f_{1t}^{RT}	f_{1t}^ν	$f_{1t}^{\nu RV}$	\bar{R}^2	CP	f_{1t}^{RT}	f_{1t}^ν	$f_{1t}^{\nu RV}$	\bar{R}^2
2-YEAR					2-YEAR				
0.412 (4.369)	-0.272 (-2.165)			0.288	0.577 (3.746)	-0.368 (-3.749)			0.222
0.396 (4.446)		0.165 (2.686)		0.269	0.495 (3.135)		0.080 (3.118)		0.183
0.396 (4.442)			0.185 (3.073)	0.272	0.498 (3.185)			0.218 (3.345)	0.187
0.406 (4.061)	-0.450 (-2.810)	0.377 (3.196)		0.329	0.592 (3.906)	-0.464 (-4.057)	0.133 (4.913)		0.263
0.407 (4.053)	-0.462 (-2.867)		0.402 (3.476)	0.336	0.602 (4.010)	-0.488 (-4.019)		0.367 (5.186)	0.274
3-YEAR					3-YEAR				
0.761 (4.116)	-0.547 (-2.516)			0.280	0.995 (3.193)	-0.740 (-4.116)			0.192
0.730 (4.172)		0.311 (3.124)		0.257	0.829 (2.568)		0.149 (3.594)		0.145
0.730 (4.162)			0.355 (3.663)	0.260	0.835 (2.610)			0.409 (3.730)	0.150
0.750 (3.830)	-0.892 (-3.548)	0.732 (4.251)		0.324	1.023 (3.315)	-0.923 (-4.393)	0.253 (5.562)		0.234
0.753 (3.817)	-0.918 (-3.648)		0.787 (4.702)	0.332	1.042 (3.406)	-0.970 (-4.323)		0.705 (5.792)	0.246
4-YEAR					4-YEAR				
1.120 (4.219)	-0.631 (-2.063)			0.294	1.463 (3.234)	-0.941 (-3.817)			0.198
1.080 (4.263)		0.445 (3.765)		0.283	1.252 (2.693)		0.194 (3.924)		0.160
1.080 (4.254)			0.499 (4.189)	0.285	1.259 (2.730)			0.530 (3.844)	0.164
1.105 (3.951)	-1.081 (-3.112)	0.955 (4.547)		0.332	1.500 (3.326)	-1.178 (-4.135)	0.327 (5.252)		0.232
1.108 (3.937)	-1.113 (-3.193)		1.024 (5.010)	0.338	1.524 (3.403)	-1.237 (-4.058)		0.907 (5.304)	0.241
5-YEAR					5-YEAR				
1.312 (3.890)	-0.662 (-1.796)			0.262	1.635 (2.823)	-1.070 (-3.509)			0.161
1.270 (3.906)		0.494 (3.250)		0.254	1.396 (2.346)		0.217 (3.434)		0.129
1.270 (3.894)			0.559 (3.768)	0.257	1.404 (2.375)			0.593 (3.389)	0.132
1.297 (3.672)	-1.150 (-2.777)	1.037 (4.174)		0.290	1.677 (2.900)	-1.338 (-3.917)	0.369 (4.914)		0.188
1.300 (3.656)	-1.189 (-2.878)		1.119 (4.674)	0.296	1.705 (2.963)	-1.404 (-3.879)		1.021 (4.972)	0.196

Table 5: Publication Lag Regressions

This table shows regression results for predictive bond return regressions of the form

$$rx_{t+12}^{(n)} = \tilde{c}_n + \tilde{\gamma}'_n f_{1t-j}^{RV} + \tilde{\beta}'_n CP_t + \tilde{\epsilon}_{t+1}^{(n)},$$

where $rx_{t+12}^{(n)}$ denotes the one-year excess holding period return on an n -year bond, CP_t denotes the one-year lag of the Cochrane and Piazzesi (2005) return forecasting factor, and f_{1t-j}^{RV} denotes the j -months lagged observation of the first principal component extracted from our revised macroeconomic data set. We report results for $j = 0, \dots, 2$. The t -statistics in parentheses are based on Newey-West standard errors with a maximum lag of 18 months. The two vertical panels present results for the sample periods 1982 : 03 – 2007 : 12 and 1982 : 03 – 2011 : 12, respectively.

1982m3-2007m12					1982m3-2011m12				
CP_t	f_{1t}^{RV}	f_{1t-1}^{RV}	f_{1t-2}^{RV}	\bar{R}^2	CP_t	f_{1t}^{RV}	f_{1t-1}^{RV}	f_{1t-2}^{RV}	\bar{R}^2
2-YEAR					2-YEAR				
0.406	-0.439			0.335	0.598	-0.501			0.271
(4.035)	(-3.018)				(3.957)	(-4.611)			
0.396		-0.354		0.298	0.581		-0.433		0.246
(3.891)		(-2.114)			(3.724)		(-3.866)		
0.363			-0.288	0.245	0.534			-0.381	0.207
(3.349)			(-1.496)		(3.225)			(-3.068)	
3-YEAR					3-YEAR				
0.750	-0.860			0.329	1.031	-0.983			0.242
(3.792)	(-3.904)				(3.330)	(-5.068)			
0.724		-0.711		0.288	0.991		-0.865		0.214
(3.560)		(-2.606)			(3.089)		(-4.355)		
0.674			-0.607	0.243	0.924			-0.795	0.186
(3.160)			(-1.898)		(2.780)			(-3.682)	
4-YEAR					4-YEAR				
1.107	-1.069			0.336	1.510	-1.257			0.239
(3.913)	(-3.676)				(3.340)	(-4.763)			
1.060		-0.863		0.293	1.402		-1.095		0.201
(3.543)		(-2.323)			(2.874)		(-4.011)		
0.984			-0.728	0.247	1.297			-1.015	0.173
(3.087)			(-1.671)		(2.554)			(-3.391)	
5-YEAR					5-YEAR				
1.299	-1.143			0.294	1.687	-1.420			0.194
(3.634)	(-3.366)				(2.901)	(-4.574)			
1.292		-0.910		0.276	1.647		-1.265		0.179
(3.548)		(-2.113)			(2.757)		(-3.944)		
1.212			-0.762	0.237	1.525			-1.182	0.154
(3.121)			(-1.506)		(2.452)			(-3.312)	

Table 6: **Out of Sample Forecast Comparisons**

This table reports out-of-sample forecast results for regressions of the form

$$rx_{t+12}^{(n)} = \tilde{c}_n + \tilde{\gamma}'_n f_{1t}^I + \tilde{\beta}'_n Z_t + \epsilon_{t+1}^{(n)},$$

where $rx_{t+12}^{(n)}$ denotes the one-year excess holding period return on an n -year bond, CP_t denotes the one-year lag of the Cochrane and Piazzesi (2005) return forecasting factor, and f_{1t}^I denotes the one-year lagged observations of a vector of factors extracted from one of three large macroeconomic data sets. Factors with the superscript LN have been extracted from the Ludvigson and Ng (2011) panel of revised macroeconomic and financial variables. Factors with superscript RV have been extracted from our panel of 68 revised macroeconomic time series. Factors with superscript RT are from the real-time counterpart of that data set. $\frac{MSE_u}{MSE_r}$ denotes the ratio of mean squared forecast error variances from an (unrestricted) model that uses a macroeconomic factor as regressor versus a (restricted) model that does not. ENC-NEW denotes Clark and McCracken (2001)'s ENC-NEW test of equal forecast accuracy. We highlight in bold the ENC-NEW statistics that are statistically significant at least at the 10% level. The four horizontal panels present results for excess returns on 2-, 3-, 4-, and 5-year bonds, respectively. The two vertical panels present results for the sample periods 1982 : 03 – 2007 : 12 and 1982 : 03 – 2011 : 12, respectively.

		1982m3-2007m12		1982m3-2011m12	
Row	Comparison	$\frac{MSE_u}{MSE_r}$	ENC-NEW	$\frac{MSE_u}{MSE_r}$	ENC-NEW
2-YEAR					
1	f_{1t}^{LN} vs. const	0.854	16.840		
2	f_{1t}^{RV} vs. const	0.954	8.433	1.060	12.189
3	f_{1t}^{RT} vs. const	1.013	0.236	1.027	1.208
4	f_{1t}^{LN} + CP vs. const + CP	0.863	15.153		
5	f_{1t}^{RV} + CP vs. const + CP	0.787	26.185	0.741	55.028
6	f_{1t}^{RT} + CP vs. const + CP	0.923	9.385	0.875	21.101
3-YEAR					
7	f_{1t}^{LN} vs. const	0.887	13.514		
8	f_{1t}^{RV} vs. const	0.941	9.878	1.036	16.680
9	f_{1t}^{RT} vs. const	1.008	0.604	1.010	3.431
10	f_{1t}^{LN} + CP vs. const + CP	0.908	10.344		
11	f_{1t}^{RV} + CP vs. const + CP	0.801	23.959	0.758	51.128
12	f_{1t}^{RT} + CP vs. const + CP	0.927	8.617	0.878	20.338
4-YEAR					
13	f_{1t}^{LN} vs. const	0.924	9.107		
14	f_{1t}^{RV} vs. const	0.956	7.643	1.046	11.814
15	f_{1t}^{RT} vs. const	1.012	-0.224	1.008	1.802
16	f_{1t}^{LN} + CP vs. const + CP	0.944	6.284		
17	f_{1t}^{RV} + CP vs. const + CP	0.825	19.966	0.792	40.569
18	f_{1t}^{RT} + CP vs. const + CP	0.938	6.870	0.895	16.232
5-YEAR					
19	f_{1t}^{LN} vs. const	0.959	5.026		
20	f_{1t}^{RV} vs. const	0.967	5.505	1.0172	10.589
21	f_{1t}^{RT} vs. const	1.014	-0.618	1.0006	1.717
22	f_{1t}^{LN} + CP vs. const + CP	0.974	3.064		
23	f_{1t}^{RV} + CP vs. const + CP	0.861	14.904	0.8234	31.798
24	f_{1t}^{RT} + CP vs. const + CP	0.952	5.170	0.9108	13.039

Table 7: MSFE Decomposition

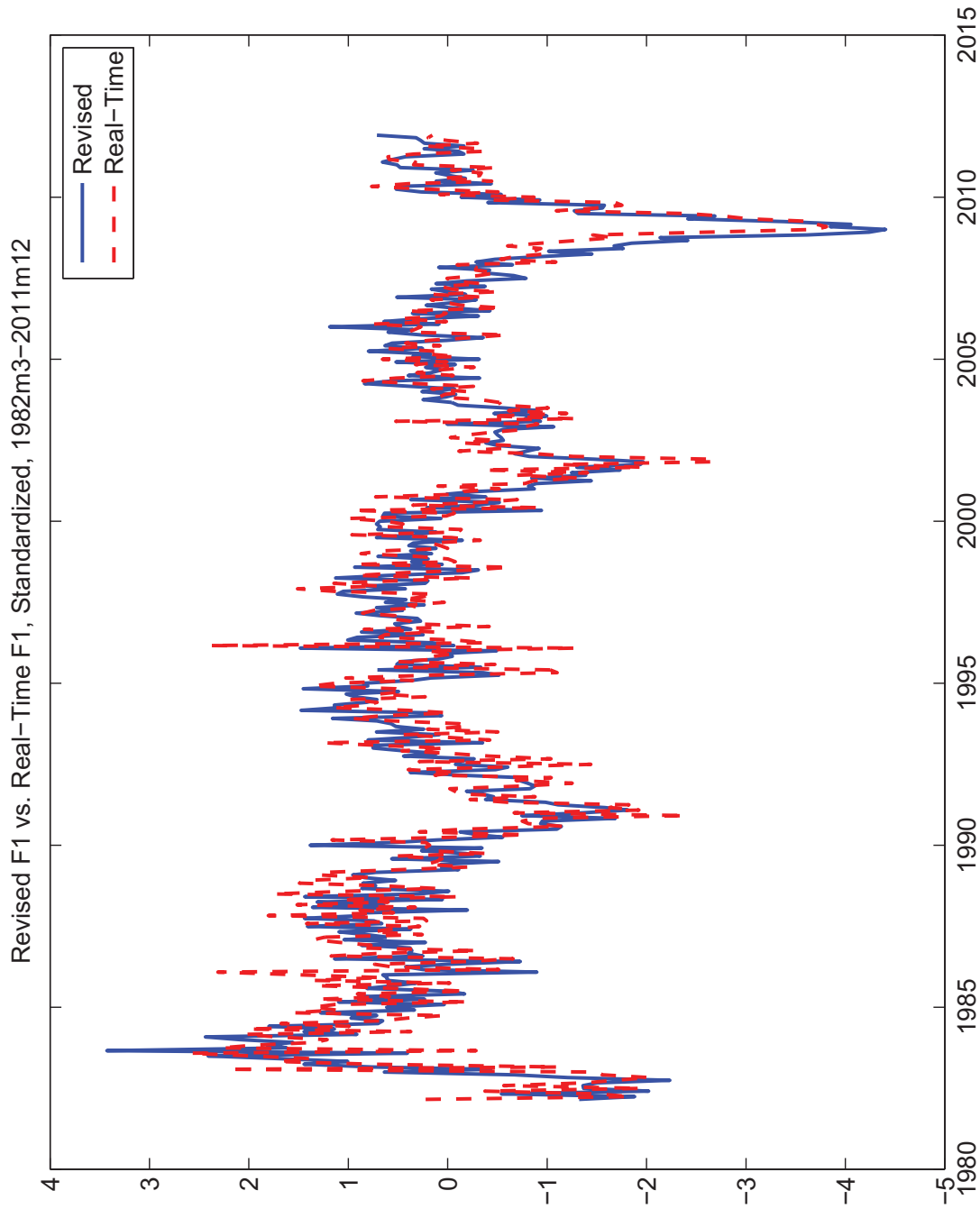
This table reports forecast error variance decompositions for the mean squared forecast error (MSFE) of the predictions

$$\begin{aligned}\widehat{rx}_{t+12}^{RV,(n)} &= (\tilde{\gamma}_t^{RV})' f_t^{RV} + (\tilde{\beta}_t^{RV})' \tilde{Z}_t \\ \widehat{rx}_{t+12}^{RT,(n)} &= (\tilde{\gamma}_t^{RT})' f_t^{RT} + (\tilde{\beta}_t^{RT})' \tilde{Z}_t\end{aligned}$$

yielding $MSFE_{RV}$ and $MSFE_{RT}$, respectively. The regressions involve $rx_{t+12}^{(n)}$, the one-year excess holding period return on an n -year bond. The regressors \tilde{Z}_t include an intercept and CP_t , which denotes the one-year lag of the Cochrane and Piazzesi (2005) return forecasting factor. The factor f_{1t}^{RV} represents the one-year lagged observations of the first principal component extracted from our panel of 68 revised macroeconomic time series, and f_{1t}^{RT} represents the first principal component extracted from the real-time counterpart of that data set. For each of the two sample periods, we present results for excess returns on 2-, 3-, 4-, and 5-year bonds.

Component	1982m3-2007m12				1982m3-2011m12			
	2y	3y	4y	5y	2y	3y	4y	5y
$MSFE_{RT}$								
1 $\frac{1}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT})^2$	1.930	7.454	14.375	21.992	2.176	8.408	17.258	26.917
2 $\frac{1}{k} \sum_{t=1}^k (\tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))^2$	0.002	0.009	0.022	0.024	0.011	0.043	0.092	0.113
	(2.387)	(2.454)	(2.611)	(2.694)	(1.455)	(1.483)	(1.543)	(1.497)
3 $\frac{1}{k} \sum_{t=1}^k ((\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)^2$	0.115	0.427	0.659	0.728	0.135	0.505	0.790	0.897
	(5.240)	(5.223)	(5.189)	(5.152)	(4.866)	(4.791)	(4.692)	(4.544)
4 $\frac{2}{k} \sum_{t=1}^k (e_t^{RT} (\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT})$	0.003	0.011	0.015	0.016	0.003	0.011	0.015	0.017
	(4.497)	(4.541)	(4.447)	(4.469)	(5.190)	(5.278)	(5.244)	(5.306)
5 $\frac{2}{k} \sum_{t=1}^k (e_t^{RT} \tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))$	-0.033	-0.145	-0.339	-0.427	-0.069	-0.281	-0.669	-0.953
	(-1.564)	(-1.313)	(-1.260)	(-1.282)	(-2.200)	(-2.229)	(-2.297)	(-2.321)
6 $\frac{2}{k} \sum_{t=1}^k (e_t^{RT} (\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)$	-0.293	-0.993	-1.514	-1.747	-0.359	-1.212	-1.838	-2.194
	(-2.920)	(-2.675)	(-2.397)	(-2.254)	(-3.120)	(-2.912)	(-2.571)	(-2.292)
7 $\frac{2}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT} \tilde{\gamma}_t^{RV} (f_t^{RT} - f_t^{RV}))$	-0.080	-0.311	-0.530	-0.651	-0.080	-0.304	-0.502	-0.624
	(-1.964)	(-2.172)	(-2.337)	(-2.291)	(-2.445)	(-2.628)	(-2.668)	(-2.586)
8 $\frac{2}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT} (\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)$	-0.006	-0.032	-0.076	-0.083	0.011	0.034	0.050	0.065
	(-2.262)	(-3.647)	(-3.893)	(-3.905)	(0.654)	(0.542)	(0.429)	(0.464)
9 $\frac{2}{k} \sum_{t=1}^k ((\tilde{\gamma}_t^{RT} - \tilde{\gamma}_t^{RV}) f_t^{RT} (\tilde{\beta}_t^{RT} - \tilde{\beta}_t^{RV}) \tilde{Z}_t)$	0.002	0.009	0.017	0.018	0.006	0.022	0.039	0.045
	(1.808)	(1.868)	(1.920)	(1.916)	0.006	0.022	0.039	0.045
	0.003	0.010	0.015	0.015	0.009	0.032	0.047	0.053
	(0.703)	(0.661)	(0.663)	(0.632)	(1.479)	(1.451)	(1.463)	(1.454)
$MSFE_{RV}$	1.645	6.439	12.643	19.885	1.843	7.258	15.282	24.336

Figure 1: Factor Comparison - Revised versus Real-time



This figure plots the first principal component extracted from a panel of 68 revised macroeconomic time series, f_{it}^{RV} , along with the first principal component extracted from the corresponding set of real-time data, f_{it}^{RT} . The sample period is 1982:03-2011:12.

A Data Appendix

This appendix lists the 68 macroeconomic time series used in our analysis. TCode provides the transformation applied to each series with TCode = 1 denoting simple monthly difference and TCode = 2 the month-over-month growth rate. Start Date provides the first real-time observation for a given series. All series have been obtained from the Archival Federal Reserve Data base (ALFRED) at the St. Louis Fed.

Mnemonic	Variable Description	TCode	Start Date
AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees: Manufacturing	2	11/1/1964
AWHNONAG	Average Weekly Hours Of Production And Nonsupervisory Employees: Total private	2	5/1/1970
AWOTMAN	Average Weekly Overtime Hours of Production and Nonsupervisory Employees: Manufacturing	2	8/1/1966
CE16OV	Civilian Employment	2	12/1/1964
CLF16OV	Civilian Labor Force	2	11/1/1964
CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items	2	6/1/1972
CURRDD	Currency Component of M1 Plus Demand Deposits	2	11/1/1964
CURRSL	Currency Component of M1	2	11/1/1964
DEMDEPSL	Demand Deposits at Commercial Banks	2	9/1/1964
DMANEMP	All Employees: Durable goods	2	11/1/1964
DSPI	Disposable Personal Income	2	1/1/1980
DSPIC96	Real Disposable Personal Income	2	2/1/1980
HOUST	Housing Starts: Total: New Privately Owned Housing Units Started	2	12/1/1964
HOUST1F	Privately Owned Housing Starts: 1-Unit Structures	2	2/1/1972
HOUST2F	Housing Starts: 2-4 Units	2	2/1/1973
INDPRO	Industrial Production Index	2	11/1/1964
M1SL	M1 Money Stock	2	12/1/1979
M2SL	M2 Money Stock	2	12/1/1979
MANEMP	All Employees: Manufacturing	2	11/1/1964
NDMANEMP	All Employees: Nondurable goods	2	11/1/1964
OCDSL	Other Checkable Deposits	2	2/1/1981
PAYEMS	All Employees: Total nonfarm	2	11/1/1964
PCE	Personal Consumption Expenditures	2	12/1/1979
PCEC96	Real Personal Consumption Expenditures	2	3/1/1980
PCEDG	Personal Consumption Expenditures: Durable Goods	2	12/1/1979
PCEDGC96	Real Personal Consumption Expenditures: Durable Goods	2	3/1/1980
PCEND	Personal Consumption Expenditures: Nondurable Goods	2	12/1/1979
PCENDC96	Real Personal Consumption Expenditures: Nondurable Goods	2	3/1/1980
PCES	Personal Consumption Expenditures: Services	2	12/1/1979
PCESC96	Real Personal Consumption Expenditures: Services	2	3/1/1980
PPCGEF	Producer Price Index: Finished Consumer Goods Excluding Foods	2	1/1/1982
PI	Personal Income	2	2/1/1966
PPICFF	Producer Price Index: Crude Foodstuffs & Feedstuffs	2	1/1/1982
PPICPE	Producer Price Index: Finished Goods: Capital Equipment	2	1/1/1978
PPICRM	Producer Price Index: Crude Materials for Further Processing	2	3/1/1978
PPIFCF	Producer Price Index: Finished Consumer Foods	2	1/1/1982
PPIFGS	Producer Price Index: Finished Goods	2	1/1/1982
PPIIFF	Producer Price Index: Intermediate Foods & Feeds	2	1/1/1982
PPIITM	Producer Price Index: Intermediate Materials: Supplies & Components	2	3/1/1978
SAVINGSL	Savings Deposits - Total	2	12/1/1979
SRVPRD	All Employees: Service-Providing Industries	2	9/1/1971
STDCBSL	Small Time Deposits at Commercial Banks	2	12/1/1979
STDLSL	Small Time Deposits - Total	2	12/1/1979
STDITI	Small Time Deposits at Thrift Institutions	2	12/1/1979
SVGCBSL	Savings Deposits at Commercial Banks	2	12/1/1979
SVGTI	Savings Deposits at Thrift Institutions	2	12/1/1979
SVSTCBSL	Savings and Small Time Deposits at Commercial Banks	2	12/1/1979
SVSTSL	Savings and Small Time Deposits - Total	2	12/1/1979
TCDSL	Total Checkable Deposits	2	3/1/1981
UEMP15OV	Civilians Unemployed - 15 Weeks & Over	2	11/1/1964
UEMP15T26	Civilians Unemployed for 15-26 Weeks	2	1/1/1982
UEMP27OV	Civilians Unemployed for 27 Weeks and Over	2	1/1/1966
UEMP5TO14	Civilians Unemployed for 5-14 Weeks	2	11/1/1964
UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	2	11/1/1964
UEMPMEAN	Average (Mean) Duration of Unemployment	2	1/1/1972
UEMPMED	Median Duration of Unemployment	2	1/1/1982
UNEMPLOY	Unemployed	2	12/1/1964
UNRATE	Civilian Unemployment Rate	1	2/1/1960
USCONS	All Employees: Construction	2	12/1/1964
USFIRE	All Employees: Financial Activities	2	12/1/1964
USGOOD	All Employees: Goods-Producing Industries	2	9/1/1971
USGOVT	All Employees: Government	2	12/1/1964
USMINE	All Employees: Mining and logging	2	12/1/1964
USPRIV	All Employees: Total Private Industries	2	8/1/1971
USSERV	All Employees: Other Services	2	12/1/1964
USTPU	All Employees: Trade, Transportation & Utilities	2	12/1/1964
USTRADE	All Employees: Retail Trade	2	12/1/1964
USWTRADE	All Employees: Wholesale Trade	2	12/1/1964