Forecasting through the Rearview Mirror: Data Revisions and Bond Return Predictability

Eric Ghysels
Division of Finance, Kenan-Flagler Business School, Department of Economics, University of North Carolina at Chapel Hill and CEPR

Casidhe Horan
University of Michigan Ross School of Business

Emanuel Moench
Deutsche Bundesbank

A previous literature has documented that bond returns are predicted by macroeconomic information not contained in yields contemporaneously. That literature has mostly relied on final revised, rather than real time macroeconomic data. We show that the use of real time data substantially reduces the predictive power of macro variables for future bond returns as well as the implied countercyclicality of term premiums. We discuss potential interpretations of our results. (*JEL* G10, G12)

Received January 26, 2014; editorial decision June 16, 2017 by Editor Geert Bekaert.

A recent literature has documented that macroeconomic factors have strong predictive power for future bond yields over and above the information contained in yields themselves. Moreover, term premiums resulting from models that include macroeconomic information are found to be strongly countercyclical, in line with standard asset pricing theory.

Macroeconomic data are typically subject to future revisions and released with delay. In this paper, we analyze the extent to which the use of final versus real time macroeconomic data affects the quantitative assessments of the role

We thank first, and foremost, Geert Bekaert, the Editor. We thank three anonymous referees for their many constructive comments and insights. We also thank our discussants Michael Halling and Lars Lochstoer; seminar participants at the WFA and the AFA; and Thomas Gilbert and Jonathan Wright for useful comments and discussions. We are most grateful to Thomas Gilbert for sharing market consensus expectations data with us. Weiling Liu provided excellent research assistance. The views expressed in this paper are those of the authors and do not necessarily represent those of the Deutsche Bundesbank or the Eurosystem. The first author thanks the Federal Reserve Bank of New York for its financial support. Send correspondence to Eric Ghysels, Department of Finance, Kenan-Flagler Business School, Campus Box 3490, McColl Building, Chapel Hill, NC 27599-3490; telephone: (919) 962-9610. E-mail: eghysels@unc.edu.

doi:10.1093/rfs/hhx098 Advance Access publication August 22, 2017

of macroeconomic information in the term structure of U.S. government bond yields. We find that a sizable fraction of the predictive information about future bond returns contained in final macro data is carried by data revisions. Real time macroeconomic data contain some useful information about future bond returns, but their predictive power is economically and statistically much lower than that of final data.

We first document these results for nonfarm payroll employment, one of the key macroeconomic indicators that is closely followed by market participants and has been shown to be an important driver of bond yields in event study analyses (see, e.g., Jones, Lamont, and Lumsdaine 1998; Fleming and Remolona 1999; Gürkaynak, Sack, and Swanson 2005). We then show that our findings apply to other macroeconomic predictor series that have been used in the extant literature. Specifically, these are real gross domestic product (GDP) growth, which has been used for example by Wright (2011); the Chicago Fed National Activity Index (CFNAI), which has been employed in Joslin, Priebsch, and Singleton (2014), industrial production which has been used in Cooper and Priestley (2009); and a principal component of a large panel of macroeconomic time series in the spirit of Ludvigson and Ng (2009).

A quantitative example may best serve to highlight the economic significance of our results. Adding revised nonfarm payroll growth to a predictive regression of the excess holding return on a 2-year Treasury on the lagged Cochrane and Piazzesi (2005) return forecasting factor increases the adjusted R^2 from 19% to 25%. Replacing the final data with that available in real time reduces the adjusted R^2 to 21%. Moreover, the slope coefficient pertaining to nonfarm payroll becomes insignificant. Performing a variance decomposition, which characterizes the relative contribution of each component to the predictive power of the regression, we find that the Cochrane and Piazzesi (2005) return forecasting factor takes a 78% share of the predictive power whereas final nonfarm payroll captures the remaining 22%. In contrast, the predictive share drops from 22% to 7% with real time data, which is about the same magnitude as that captured by the revision component.

We consider alternative proxies for the information set potentially available to investors in real time, and find very similar results. Moreover, following the approach in Ludvigson and Ng (2009), we find that term premiums implied by a simple vector autoregressive model including real time nonfarm payroll data are considerably less countercyclical than term premiums obtained using the final revised counterpart. Hence, some of the previously documented cyclicality of bond premiums appears to be driven by data revisions. In line with this finding, a mean-variance investor endowed with final macroeconomic data would obtain a higher Sharpe ratio than one using real time data. Finally, we show that on days of macroeconomic news announcements investors react not only to the surprise component of the announcement but also to revisions to prior releases. Yet, the bond market's reaction to macroeconomic news does not anticipate future revisions. This is in contrast to Gilbert (2008), who studies

S&P 500 returns on days of payroll announcements and finds that the stock market's reaction to payroll announcements is correlated with future revisions, controlling for market expectations of the release.

What can we learn from these results? We discuss several potential interpretations. Assuming that revised macroeconomic data capture the state of the business cycle more precisely than real time data, our finding that the predictability of bond returns is stronger when revised data are used is consistent with standard asset pricing theory implying time-varying risk premiums. Our finding that revisions are correlated with bond returns could thus simply reflect the fact that both risk premiums and revisions have a cyclical component or, alternatively, that data revisions proxy for uncertainty about the state of the economy and thus correlate with risk premiums. A potentially complementary interpretation relates to the role of data revisions in investors' expectations about monetary policy. Orphanides (2001) has shown that the prescription of a Taylor-type policy rule for the federal funds rate based on real time data can substantially differ from one based on revised data. Bond yields reflect investors' expectations about future short rates as well as term premiums. If investors use data revisions to update their beliefs about the state of the economy and thus the Federal Reserve's likely path of policy, bond returns should be correlated with data revisions, as we empirically document. Finally, the documented difference in predictability of bond returns using final versus real time macro data could be consistent with informational frictions playing some role in explaining bond return dynamics. At a more general level, our findings suggest that different assumptions about the information set available to investors may lead to different assessments of the driving forces behind bond yields.

1. Why Revisions May Matter

In this section, we document that final revised nonfarm payroll (NFP, henceforth) growth data available to an econometrician substantially differ from real time observations available to investors. In other words, data revisions are a sizable component of NFP growth and may therefore matter for their ability to predict bond returns. We start by introducing some notation, and then move on to a quantitative assessment of the relative importance and the cyclical properties of real time observations, data revisions, and final revised data, respectively.

1.1 Notation

To fix ideas it is useful to introduce some notation. Let X_t denote a macroeconomic time series that will be used to predict future bond returns. Additionally, assume that X_t is published with delay and subject to future revisions. For our purpose, we shall use a double index, namely $X_{t|\tilde{t}}$ where $\tilde{t}$ is the time stamp of the information release pertaining to time period t. For example, $X_{t|T}$ denotes the final data collected for calendar month t at the end of a sample of size T.[1] For simplicity, we assume that the data $X_{t|T}$ are no longer subject to future revisions and can therefore be considered as final.[2]

In practice, since almost all macroeconomic time series are released with a 1- or 2-month lag, the release of new information in month t typically pertains to period $t-1$ or $t-2$, that is, $X_{t-1|t}$ or $X_{t-2|t}$. For ease of notation, we will uniformly refer to the real time observation available in t as $X_{t-1|t}$ and hence assume a one period publication lag. In our empirical analysis, however, we keep track of the information in each calendar month as it was available to investors at the end of that month.

We can decompose the final revised observation into two components:

$$X_{t|T} = X_{t-1|t} + \nu_{t|T}, \tag{1}$$

where

$$\begin{aligned} \nu_{t|T} &= (X_{t|T} - X_{t-1|t}) = \nu_{t|T}^{rev} + \nu_{t|T}^{pl} \\ \nu_{t|T}^{rev} &= X_{t|T} - X_{t|t+1} \\ \nu_{t|T}^{pl} &= X_{t|t+1} - X_{t-1|t} \end{aligned} \tag{2}$$

contains two elements: a component denoted by $\nu_{t|T}^{rev}$ that is purely related to future revisions of the initial announcement, $X_{t|T} - X_{t|t+1}$, and one denoted by $\nu_{t|T}^{pl}$, or publication lag, that captures the fact that macroeconomic data are released with a lag, $X_{t|t+1} - X_{t-1|t}$.

1.2 How important are data revisions?

Table 1 reports the summary statistics of real time observations, revisions and final data for NFP. Over the full 1964-2014 sample, monthly changes in NFP averaged 135K according to the final data, whereas the real time release was, on average, only 113K, a difference of 22K, or almost 20%, of the real time announcement. During NBER expansions $\nu_{t|T}$ averaged 32K, whereas during recessions the average is -37K, or almost 30%, of the real time release. Hence, real time data understate the final figure during expansions and also on average for the full sample, but overstate the final number during recessions. In other words, revisions are cyclical: the good news is even better ex post during expansions but the bad news during recessions is ex post much worse with the final data. The bulk of $\nu_{t|T}$ is attributable to the revision component, that is, $\nu_{t|T}^{rev}$, for the full sample as well as NBER expansions. In contrast, publication delays, that is, $\nu_{t|T}^{pl}$, are slightly more important during recessions.

[1] Similarly for calendar quarter. In our empirical work, we consider one quarterly series, that is, real GDP growth.

[2] In practice, macroeconomic time series are revised in benchmark revisions that incorporate information from additional sources even multiple years after the first release, but these revisions mostly tend to affect the level, and not the monthly growth rate, of these time series.

Table 1
Summary statistics of nonfarm payroll data

	Mean	SD	Min	Max
		NFP 1964-2014		
$X_{t\|T}$	135.69	207.71	−823.00	1115.00
$X_{t-1\|t}$	113.75	199.62	−674.00	733.00
$\nu_{t\|T}$	21.93	185.39	−795.00	1526.00
$\nu_{t\|T}^{pl}$	0.11	184.05	−898.00	1144.00
$\nu_{t\|T}^{rev}$	21.82	108.28	−441.00	382.00
		NFP 1964-2014: NBER expansions		
$X_{t\|T}$	188.51	149.42	−329.00	1115.00
$X_{t-1\|t}$	156.36	159.77	−467.00	733.00
$\nu_{t\|T}$	32.16	184.95	−795.00	1526.00
$\nu_{t\|T}^{pl}$	5.34	183.62	−898.00	1144.00
$\nu_{t\|T}^{rev}$	26.81	103.19	−441.00	382.00
		NFP 1964-2014: NBER recessions		
$X_{t\|T}$	−170.68	232.81	−823.00	306.00
$X_{t-1\|t}$	−133.33	227.50	−674.00	383.00
$\nu_{t\|T}$	−37.34	177.61	−529.00	409.00
$\nu_{t\|T}^{pl}$	−30.22	184.66	−477.00	434.00
$\nu_{t\|T}^{rev}$	−7.12	130.99	−325.00	335.00

This table reports the summary statistics of real time data, revisions and final data for nonfarm payroll (NFP) from 1964 to 2014. Three sample configurations cover, respectively, the full sample, NBER expansions, and recessions. The final revised data is labeled $X_{t|T}$. This can be decomposed into two components: $X_{t|T} = X_{t-1|t} + \nu_{t|T}$, defined in equations (1) and (2).

Revisions are also quite volatile, as shown by their standard deviations reported in Table 1. While final revised and real time data are similarly volatile, their difference $\nu_{t|T}$ also has a standard deviation of about the same order of magnitude. This is consistent with a negative correlation between the real time data and the combined publication lag and revision component. More generally, these numbers show a considerable amount of uncertainty about the true state of the economy in real time.

The relative importance of revisions in the variability of revised nonfarm payroll data and their large variability is consistent with the prior literature. For example, Aruoba (2008) 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. Moreover, he documents that revisions feature substantial degrees of serial correlation. This is confirmed by Croushore (2011), who further provides a comprehensive review of the literature on data revisions and real time data.

2. Predictive Bond Return Regressions

In this section, we study the extent to which the predictability of bond returns with macro data depends on the specific macroeconomic information set being employed. A number of recent papers have used simple linear regressions to analyze the predictive power for future bond returns contained in macroeconomic variables over and above the information captured by the yield curve itself (see, e.g. Ludvigson and Ng 2009, 2011; Cooper and Priestley 2009; Cieslak and Povala 2015). Other papers instead have used macroeconomic factors in affine term structure models and have shown that the resulting term premiums have different properties than those emanating from models using only yield curve factors (see, e.g., Wright (2011); Joslin, Priebsch, and Singleton (2014)). Because risk premium dynamics implied by affine term structure models are largely determined by the underlying vector autoregression (VAR) in the model factors (Joslin, Le, and Singleton (2013), the impact different macroeconomic information sets have on the ability to predict bond returns can be assessed without estimating a fully specified affine model. While we focus on simple predictive regressions of excess bond returns on lagged predictor variables in this section, we also will study the implications of different macroeconomic information sets in a VAR setup in Section 3.

The typical finding reported in the literature using predictive regressions is that macroeconomic data significantly add predictive power beyond the Cochrane and Piazzesi (2005) return forecasting factor (the CP factor henceforth), a linear combination of forward rates constructed to predict returns on Treasury notes. These analyses are commonly based on annual excess holding period returns computed based on the Fama-Bliss zero-coupon bond yields from the Center for Research in Securities Prices (CRSP). Here, we follow this literature and consider the following linear regressions:

$$rx_{t+12}^{(n)} = \alpha_n + \beta_n' bm_t + \gamma_n^z Z_t + e_{t+12}^{(n)} \qquad (3)$$

where $rx_{t+12}^{(n)}$ denotes the 1-year excess holding period return on an n-year bond, bm_t is a vector capturing the benchmark yield curve information available at time t, and Z_t is a macroeconomic predictor variable. In this section, Z_t will be NFP as observed at different points in time. In Section 5, we will also consider industrial production (IP) growth, GDP growth, CFNAI and the LN factor for the purpose of documenting robustness with respect to other commonly used return predictor variables.

Before proceeding, a few words about the timing of data releases and returns are in order. Typically, the month t - 1 NFP is published the first Friday of month t, at 8:30 am EST. Returns are computed end-of-month. Therefore, month t returns are matched with end-of-month information sets captured by Z_t. This means that when using real time data to predict returns in $t+12$ one only has information for month t - 1 (that is, $X_{t-1|t}$) available. This is in contrast to the standard literature which uses $X_{t|T}$ to predict bond returns.

We consider several benchmark model specifications so as to capture the predictive information contained in the yield curve in period t. The first is the CP factor which we update through December 2014, the end of our sample period. Besides the CP factor we have two additional specifications for the benchmark bm_t. The first is motivated from the CP factor but avoids estimating factor loadings. Specifically, we observe that the loadings of the CP factor have a tent-shaped form with a peak at the 3- to 4-year forward. One can thus well approximate the predictive information for future bond returns contained in current market prices by using two yield spreads: (1) between the 3-year and the 1-year yield (labeled "S31") and (2) between the 5-year and the 4-year yield (labeled "S54"). Using observable spreads instead of an estimated linear combination as a regressor has the advantage of avoiding potential errors-in-variables and look-ahead biases. The second simply uses the first three principal components of Treasury yields as predictors following a large literature going back to at least Litterman and Scheinkman (1991).

This leaves us with three benchmark models: (1) the CP factor, (2) the pair of spreads S31 and S54 - henceforth referred to simply as spreads or S31+S54, and (3) the first three PCs of the Treasury yield curve. Hence, we have benchmark models with one (CP), two (spreads) and three (PCs) regressors. We focus on the results with the CP factor and the two spreads in the remainder of this section. The Online Appendix Section OA.2 provides results with the three principal components.

2.1 Predicting returns with different information sets

We estimate Equation (3) for different regressors Z_t capturing various macroeconomic information sets: (a) final revised data, $Z_t = X_{t|T}$, (b) first releases for month or quarter t (typically observed in month or quarter $t+1$) $Z_t = X_{t|t+1}$, (c) real time data, $Z_t = X_{t-1|t}$, (d) the publication lag component $Z_t = \nu^{pl}_{t|T}$, and (e) revisions $Z_t = \nu^{rev}_{t|T}$.

Table 2 provides the results for both the CP factor and the two spreads as benchmarks. For each benchmark model there are two panels with results for the 2-year and 5-year maturities, respectively. The first column covers the predictive regressions with only yield information, that is, the benchmark models with either CP factor or the two spreads. For the 2-year maturity this amounts to a 19% (CP factor) or 17% (two spreads) adjusted R^2 and similarly 23% or 20% for the 5-year maturity. Adding final NFP data as a regressor in the second column always yields significant slope estimates. Moreover, the adjusted R^2 increases considerably, with the difference being somewhat stronger for the 2-year maturity where it rises from 19% (17%) to 25% (25%) compared to the CP factor (two spreads) benchmark. Interestingly, when only final NFP data and no yield curve information is used in the regression, the coefficient remains statistically significant but the R^2 sharply drops from 25% to 7% for the 2-year maturity and 26% to 5% for the 5-year maturity. This shows that compared to information in the yield curve, revised NFP data only explain a small share of the variance in bond returns.

Replacing the final releases with the first releases in the next column reduces the slope coefficients and lowers their statistical significance for both benchmark models and bond maturities considered. More importantly, in all cases we observe a reduction in the adjusted R^2. Replacing the first release by

Table 2
Bond return predictability - in-sample results

				1-year excess holding return on 2-year Treasury - CP factor					
CP	43.25***	41.86***		42.18***	42.26***		43.31***	43.30***	42.01***
Final		−0.20***	−0.22**						
First				−0.15**					
Real time					−0.11	−0.14*			−0.18**
Pub lag							−0.04*		−0.18***
Revisions								−0.23**	−0.29***
Adj. R^2	0.19	0.25	0.07	0.22	0.21	0.02	0.19	0.21	0.25
				1-year excess holding return on 5-year Treasury - CP factor					
CP	148.24***	144.68***		145.74***	145.91***		148.39***	148.39***	145.36***
Final		−0.51***	−0.58**						
First				−0.35*					
Real time					−0.27	−0.37			−0.43*
Pub lag							−0.10		−0.45***
Revisions								−0.69***	−0.85***
Adj. R^2	0.23	0.26	0.05	0.24	0.23	0.02	0.23	0.24	0.26

(*continued*)

Table 2
Continued

	1-year excess holding return on 2-year Treasury - Two spreads								
S31	1.63***	1.67***		1.64***	1.63***		1.64***	1.66***	1.68***
S54	−3.15***	−3.79***		−3.55***	−3.55***		−3.12***	−3.26***	−3.81***
Final		−0.24***	−0.22**						
First				−0.18***					
Real time					−0.15*	−0.14*			−0.23**
Pub lag							−0.04**		−0.21***
Revisions								−0.25***	−0.33***
Adj. R^2	0.17	0.25	0.07	0.21	0.20	0.02	0.17	0.20	0.25
	1-year excess holding return on 5-year Treasury - Two spreads								
S31	5.22***	5.31***		5.24***	5.21***		5.23***	5.30***	5.34***
S54	−6.27	−7.81**		−7.14*	−7.07*		−6.21	−6.60*	−7.75**
Final		−0.57***	−0.58**						
First				−0.39*					
Real time					−0.29	−0.37			−0.48*
Pub lag							−0.12**		−0.51***
Revisions								−0.75***	−0.94***
								[−2.96]	[−3.53]
Adj. R^2	0.20	0.24	0.05	0.21	0.20	0.02	0.20	0.22	0.24

This table reports in-sample predictive regression results using different macroeconomic predictors Z_t in equation (3), namely (a) final revised data, (b) first releases for t (typically published in month/quarter $t+1$) (c) real time data, (d) the publication lag component and (e) revisions. The data series is NFP from 1964 to 2014. These analyses are based on annual excess holding period returns computed from the Fama-Bliss zero-coupon bond yield. The benchmark bm_t uses the Cochrane and Piazzesi 2005 return forecasting factor (CP factor) or two yield spreads: (1) between the 3-year and 1-year yield (S31) and (2) between the 5-year and 4-year yield (S54). All regressions include a constant not reported in the tables. t-statistics are based on Newey and West 1987 standard errors with a maximum lag length of 18 months. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% level, respectively.

the real time observations in the next column suggests even less predictability both in terms of the adjusted R^2 and the magnitude and significance of the slope coefficients. The striking difference with respect to final data suggests that information either captured by the publication lag or the revision components accounts for some of the predictable information in final data. Dropping the CP factor and considering only real time NFP data shows that the adjusted R^2 drops even more strongly than in the case of final revised data, from 21% (23%) to 2% (2%) for the 2- and 5-year maturities, respectively. This shows that real time NFP data explain almost none of the variation in excess bond returns.

The next two columns separate the real time data from the revision component. The results show that it is clearly the revision component which carries the bulk of predictive content. In all cases, the coefficient of the revision component is strongly statistically significant and the adjusted R^2 are as large or larger than those implied by the real time data.

The final columns in each panel of Table 2 combine the benchmark regressors with the real time, publication lag and revision into a single regression. In all cases, the real time regressor is the least important predictor while both the publication lag and revision components enter strongly significantly. Moreover, the adjusted R^2 jump back to the levels observed when using final data. In sum, these results suggest that the predictive ability of macroeconomic information is much more pronounced when final revised as opposed to real time data are used. Moreover, important predictive information appears to be contained in the revision components.

2.2 Variance decomposition

The regression results reported in the previous section show that final revised data are stronger predictors for bond returns than real time data with the revision component itself carrying important predictive information. This section sheds some light on the relative contributions of each component to the predictive relationships.

In Appendix Section A.1, we provide the details of a variance decomposition which allows us to trace out the incremental predictive information contained in the various macroeconomic series - final or real time, publication lag and revisions - using the Frisch-Waugh-Lovell theorem.[3] Omitting the details here, the basic idea is to regress bond returns first on the benchmark yield curve predictor(s) bm_t and to use the residuals to form a regression with $X_{t|T}$ or $X_{t-1|t}$. For real time data, the residuals of the latter provide the input for a regression using the publication lag component, $\nu_{t|T}^{pl}$, and finally, the residuals of this regression are regressed on the revision component, $\nu_{t|T}^{rev}$. This yields a

[3] See, for instance, pages 19-24 of Davidson and MacKinnon (1993).

Table 3
Bond return predictability: Variance decomposition

	2-year Treasury		5-year Treasury	
CP factor benchmark model				
R^2	0.25	0.25	0.26	0.27
bm_t	0.78	0.82	0.86	0.89
$X_{t\|T}$	0.22		0.14	
$X_{t-1\|t}$		0.07		0.04
$v^{pl}_{t\|T}$		0.05		0.02
$v^{rev}_{t\|T}$		0.06		0.05
Two spreads benchmark model				
R^2	0.25	0.26	0.24	0.25
bm_t	0.69	0.76	0.82	0.87
$X_{t\|T}$	0.31		0.18	
$X_{t-1\|t}$		0.11		0.04
$v^{pl}_{t\|T}$		0.07		0.03
$v^{rev}_{t\|T}$		0.06		0.05

This table reports variance decompositions based on the regressions

$$r^{(n)}_{t+12} = \alpha_n + \beta'_n bm_t + \gamma^f_n X_{t|T} + e^{(n)}_{t+12}.$$

and

$$r^{(n)}_{t+12} = \alpha_n + \beta'_n bm_t + \gamma^{rt}_n X_{t-1|t} + \gamma^{pl}_n v^{pl}_{t|T} + \gamma^{rev}_n v^{rev}_{t|T} + e^{(n)}_{t+12}.$$

using equation (A5) provided in Appendix A.1. The regressors are nonfarm payroll data from 1964 to 2014. The dependent variables $r^{(n)}_{t+12}$ refer to annual excess holding period returns computed from the Fama-Bliss zero-coupon bond yields from the Center for Research in Securities Prices (CRSP). The benchmark bm_t uses the Cochrane and Piazzesi 2005 return forecasting factor (CP factor) or two yield spreads: (1) between the 3-year and the 1-year yield (labeled "S31") and (2) between the 5-year and the 4-year yield (labeled "S54").

decomposition of the variance of the original fitted regression reported in the final column of Table 2.[4]

The results appear in Table 3. There are two panels, each pertaining to a benchmark model, CP or two spreads. Each panel has results for the 2-year maturity appearing in the first two columns and the 5-year maturity in the remaining third and fourth columns. The first line reproduces the adjusted R^2 which appeared in Table 2. We learn from the next line that the CP factor takes a 78% (2-year) and 86% (5-year) share of the predictive power whereas final NFP takes the remaining 22% for the 2-year and 14% for the 5-year. For the two spreads benchmark model the share of variance explained by the final NFP data are somewhat higher, at 31% and 18%, respectively. The variance decompositions for the companion regression using the components of the final data appear in the second (2-year maturity) and fourth (5-year) columns, respectively. The share of the variance explained by the benchmark yield curve information increases in all four cases. More importantly, comparing the numbers for the final revised ($X_{t|T}$) with those of real time ($X_{t-1|t}$) NFP, we see that the predictive share is reduced dramatically for both benchmark models and

[4] The formula for the variance decomposition appears in Equation (A5).

bond maturities considered: 22% to 7%, 14% to 4%, 31% to 11%, and, finally, 18% to 4%. Instead, the publication lag ($\nu_{t|T}^{pl}$) and revision ($\nu_{t|T}^{rev}$) components represent a sizeable share of the predictive variation, in all cases superior to that of $X_{t-1|t}$ when combined. In sum, the contribution of NFP to predicting bond returns is sharply reduced when using real time data. Compounded with our finding that the predictive share of yield curve information increases in the joint regressions with real time data while the overall R^2 remained roughly the same, we therefore infer a vastly reduced role for macroeconomic information in predicting bond returns when real time data, as opposed to final data, are used.

2.3 Out-of-sample analysis

In this subsection we analyze whether the in-sample results carry over to an out-of-sample (henceforth OOS) setting. We use a 10-year training sample for the OOS exercises. Since the NFP sample starts in 1964, this means that we have the first OOS predictions in 1974. We run recursively (i.e., using expanding samples) variants of regression (3), adding one monthly observation at a time to produce out-of-sample predictions. Specifically, we regress the individual bond returns $rx_{t+12}^{(n)}$ on (1) a constant and the benchmark yield curve information (CP factor or the two yield spreads $S31$ and $S54$) and (2) a constant, the benchmark yield curve information and Z_t capturing different components of macroeconomic information releases: (a) final revised data, (b) real time data, and (c) revisions. For each of these specifications, we use the estimated regression coefficients to predict excess bond returns 12-months out and record the corresponding forecast errors. We then assess whether the macroeconomic information significantly improves forecast accuracy by computing the ratio of mean squared forecast errors (MSE) of the unrestricted models which add the macro information and the MSE of the restricted models (i.e., specifications that only use a constant or a constant and the yield curve factors).

To assess whether a given macroeconomic factor significantly improves predictability with respect to the benchmark model, we report the ENC-NEW test for nested models suggested by Clark and McCracken (2001). We further test for all specifications that are not based on the real time information whether they are significantly outperformed by the model using the final revised series. Specifically, we report the Harvey, Leybourne, and Newbold (1998) test for nonnested models which represents an extension of the Diebold and Mariano (2002) test to serially correlated and heteroscedastic forecast error series. Following Clark and McCracken (2001), we label this the ENC-T test.

The empirical findings are reported in Table 4. As before, the benchmark bm_t is either the CP factor or the two spreads S31 and S54. The first column reports the ratio MSE_u/MSE_r of mean squared forecast error variances from an (unrestricted) model that uses a macroeconomic time series as regressor versus a (restricted) model that does not. The first row in each panel shows the mean squared forecast error for predictions based on the restricted model (i.e.,

Table 4
Out-of-sample assessment of bond return predictability

	2-year Treasury			5-year Treasury		
	$\frac{MSE_u}{MSE_r}$	ENC-NEW	ENC-T	$\frac{MSE_u}{MSE_r}$	ENC-NEW	ENC-T
CP factor						
bm_t	3.43			43.30		
$X_{t\|T}$	0.92	40.68**		0.95	19.82**	
$X_{t-1\|t}$	0.95	19.39**	3.24***	0.98	7.09**	3.35***
$\nu_{t\|T}$	1.00	1.64*	5.58***	1.00	1.31	4.81***
Two spreads						
bm_t	3.14			36.64		
$X_{t\|T}$	0.95	43.47**		0.96	23.63**	
$X_{t-1\|t}$	1.00	10.63**	4.14***	1.00	4.03**	4.03***
$\nu_{t\|T}$	0.99	6.21**	4.35***	0.99	4.70**	3.92***

This table provides out-of-sample forecast results for predictive bond return regressions. The benchmark bm_t uses either the Cochrane and Piazzesi 2005 return forecasting factor or the two yield spreads: (1) between the 3-year and the 1-year yield (labeled "S31") and (2) between the 5-year and the 4-year yield (labeled "S54"). MSE_u/MSE_r denotes the ratio of mean squared forecast error variances from an (unrestricted) model that involves a macroeconomic regressor versus a (restricted) model that does not. The first row in each panel shows the mean squared forecast error for predictions based on the restricted model (that is, using yield curve information alone equivalent to setting $\gamma_n^z=0$). The other entries in each panel are ratios with respect to the first row entry. ENC-NEW denotes Clark and McCracken's (2001) ENC-NEW test of equal forecast accuracy for nested models with the critical values obtained from Table 1 in their paper. ENC-T is the Harvey, Leybourne, and Newbold (1998) test for equal forecast performance of all models relative to the one using the $X_{t|T}$ data. As shown by West (1996), this test is asymptotically normal for non-nested forecasts. The two horizontal panels present results for excess returns on 2- and 5-year bonds, respectively. A 10-year training sample is used for the OOS exercises. All predictive regressions are reestimated month by month and forecast errors from predictions of 1-year excess holding period returns are recorded. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% level, respectively.

using yield curve information alone equivalent to setting $\gamma_n^z=0$). The other entries in each panel are ratios with respect to the first row entry. The two horizontal panels present results for excess returns on 2- and 5-year bonds, respectively.

A first observation, of independent interest, is that while the CP factor benchmark model featured a higher adjusted R^2 in-sample fit (see Table 2), we note that the two spreads benchmark specification yields a smaller MSE out-of-sample. The difference is particularly pronounced for the 5-year Treasury, where the MSE drops by 20% from 43.30 to 36.64.

The MSE ratios with final data range between 0.92 and 0.96, all representing a significant drop according to the ENC-NEW test. Hence, final revised NFP significantly improve the out-of-sample predictive power for bond returns. With real time data the ratios are larger and range from 0.95 to 1.00. Despite some values equal to one – meaning that $MSE_u = MSE_r$ – the Clark and McCracken (2001) ENC-NEW test still rejects the null of equal forecasting ability, indicating that the real time data do improve the out-of-sample forecasts relative to the benchmark yield curve information somewhat.[5] In the case

[5] The fact that one rejects the null despite the MSE ratio being equal to one is a matter of forecast accuracy in finite samples (at estimated model parameters) versus in population (at the limiting, population values of model

of using two spreads as benchmark, adding revisions of NFP reduces the MSE ratios slightly and – according to the ENC-NEW test – significantly so. Consistent with the in-sample results, we thus find that adding real time NFP data improves the predictive ability relative to only using yield curve data by less than when final data are used. Moreover, revisions themselves carry important predictive information.

We assess whether the predictive ability of the model involving final data is superior to that using real time data or revisions by means of the ENC-T test, provided in the third and sixth column of Table 4. Under the null of no superior predictive ability coming from revisions, the real time forecasts encompass the final revised ones. Hence, the positive and significant test statistics that we observe imply that we can reject the null, namely that the revisions contain useful additional information over and above the real time data. In a similar vein, the corresponding test statistics for the revision components $\nu_{t|T}$ imply that the real time data carry predictive information beyond that contained in revisions.

In sum, the results in this subsection thus confirm our in-sample findings: while real time NFP data appear to contain some predictive information for future Treasury returns over and above that included in yields themselves, the degree of predictability is significantly larger when using final revised data.

2.4 Capturing market expectations

Market participants may anticipate some of the revisions and publications lags. To assess whether endowing investors with a larger information set changes our conclusions, we consider predictive regressions in which we replace the macroeconomic data with regression-based predictions using market information proxies, namely,

$$\begin{aligned} rx^{(n)}_{t+12} &= \alpha_n + \beta_n' bm_t + \gamma_n^x \left(\hat{b}' x_t\right) + e^{(n)}_{t+12} \\ X_{t|T} &= a + b' x_t + \eta_{t|T}, \end{aligned} \tag{4}$$

where $x_t = \{X_{t-1|t}, FIN\}$, which uses the real time observations as well as the following set of financial market indicators: the Treasury yields with maturities ranging from $n=1,\dots,5$ years, and the five equity risk factors MKT, SMB, HML, momentum, and short-term reversal from French's data library. While the former by construction capture all information embedded in bond yields in period t, the latter have been shown to span the main sources of risk in equity

parameters). The ENC-NEW test is a test of equal accuracy in population. In finite samples, the forecast error variance due to the sampling error in model parameters (with more of that in the unrestricted model than the restricted) will inflate the MSE of the unrestricted model relative to the restricted, and can cause the smaller model to have a lower MSE than the larger. With an infinitely large sample, one would estimate the model parameters precisely enough that the unrestricted would be more accurate than the restricted. Clark and McCracken (2013) provide a discussion of this. We thank Todd Clark for helpful feedback on this issue.

markets. Combined, these indicators should thus represent a good proxy for the information set available to market participants in real time.

We also consider survey forecasts of nonfarm payroll employment. Specifically, we use the consensus expectation for NFP news announcements from the Money Market Services database (like in Gilbert 2008) until 1999, and from Bloomberg starting in 2000 in order to measure investors' expectations of the payroll release at time t.[6] Since the survey data only starts in 1985, the results in this section are based on the 30 year sample from 1985 to 2014.[7]

Table 5 summarizes the results. The first two panels feature the CP factor as benchmark model, the bottom two panels refer to the two spreads benchmark. Comparing the real time entries with the regressions using market or survey forecasts, we note that the adjusted R^2 is largely unchanged. Specifically, for the 2-year maturity it takes on values of 18% and 19% for market and survey expectations, as compared to 19% for real time observations when the CP benchmark is used and 17% as compared to 15% when the two spreads are used as benchmark. In all cases, the adjusted R^2s are considerably smaller than those obtained using final NFP data.

When we jointly consider the survey expectations with the final NFP data in the second to last column of Table 5, the latter remain strongly statistically significant and the adjusted R^2s increase meaningfully. In contrast, the R^2s remain unchanged when we consider the survey data along with the real time NFP data in the final column of the table. Moreover, the coefficient on the real time data is insignificant for both bond maturities and benchmark models considered. Thus, adding the survey data drives out real time but not final data. These findings indicate that the final revised data imply considerably stronger predictive power for future bond returns than reasonable alternative proxies of the information set that investors may have in real time.

3. Term Premiums and Mean-Variance Portfolios

So far, we have established that the predictability of bond returns is diminished when real time as opposed to final-revised nonfarm payroll data are used as predictor in addition to information embedded in the yield curve. In this section, we study the economic significance of this result. We first show that term premiums implied by a predictive model for excess bond returns that uses real time NFP are substantially less cyclical than term premiums implied by a model based on final revised NFP. We then document that a mean-variance investor endowed with real time NFP data would achieve lower Sharpe ratios than one

[6] The MMS consensus expectation is the median of about 40 economists' forecasts surveyed each Friday about the next week's macroeconomic announcements. The Bloomberg consensus is also the median of the surveyed economists who can submit their forecasts until the day before the payroll announcement. The number of forecasters ranged between 40 and 100 over the sample we consider.

[7] For completeness we report in Online Appendix Section OA.3 all the results for NFP covering only the 1985-2014 sample. The main findings remain unchanged as compared to the full sample 1964-2014.

Table 5
Bond return predictability with market expectations

	1-year excess holding return on 2-year Treasury									
CP	41.76***	42.55***	42.95***	42.20***	41.86***	40.74***	42.38***	42.72***	42.00***	42.62***
Final		−0.19***							−0.32***	
First			−0.15***							
Real time				−0.13**						−0.06
Pub lag					−0.02					
Revisions						−0.27***				
Mkt fcst							−0.15**			
Survey fcst								−0.18**	0.20*	−0.12*
Adj. R^2	0.15	0.24	0.20	0.19	0.15	0.19	0.18	0.19	0.25	0.19
	1-year excess holding return on 5-year Treasury									
CP	152.32***	154.53***	155.64***	153.67***	152.40***	149.49***	154.37***	155.11***	153.21***	154.73***
Final		−0.53***							−0.83***	
First			−0.41***							
Real time				−0.40**						−0.21
Pub lag					−0.02					
Revisions						−0.74**				
Mkt fcst							−0.50**			
Survey fcst								−0.51***	0.48	−0.31
Adj. R^2	0.17	0.23	0.20	0.20	0.17	0.20	0.20	0.20	0.23	0.20

(*continued*)

Table 5
Continued

	1-year excess holding return on 2-year Treasury									
S31	0.98**	1.12***	1.08**	1.03**	0.98**	0.99**	1.14***	1.06**	1.12***	1.06**
S54	−2.37**	−4.01***	−3.59***	−3.48***	−2.37**	−2.51***	−4.19***	−3.78***	−3.83***	−3.87***
Final		−0.27***							−0.35***	
First			−0.22***							
Real time				−0.20***						−0.09
Pub lag					−0.02					
Revisions						−0.30***				
Mkt fcst							−0.31***			
Survey fcst								−0.28**	0.14	−0.20**
Adj. R^2	0.09	0.23	0.16	0.15	0.08	0.14	0.17	0.17	0.24	0.17
	1-year excess holding return on 5-year Treasury									
S31	3.42*	3.71**	3.60**	3.51**	3.42*	3.45**	3.71**	3.56**	3.72**	3.56**
S54	−2.74	−6.18	−4.93	−4.80	−2.74	−3.10	−6.12	−5.13	−5.26	−5.34
Final		−0.57**							−0.98***	
First			−0.39*							
Real time				−0.38*						−0.22
Pub lag					−0.02					
Revisions						−0.82**				
Mkt fcst							−0.58*			
Survey fcst								−0.48	0.68*	−0.28
Adj. R^2	0.09	0.14	0.11	0.11	0.09	0.12	0.11	0.11	0.15	0.11

This table reports predictive regressions with market information proxies appearing in equation (4), where x_t uses the real time observations as well as: the Treasury yields with maturities ranging from $n = 1, \ldots, 5$ years, and the equity risk factors MKT, SMB, HML, momentum, and short-term reversal. The entries with *Mkt fcst* refer to the above market expectation models. In addition, we also use the median survey expectation for NFP announcements from the Money Market Services database (like in Gilbert 2008) until 1999, and from Bloomberg starting in 2000. The entries with *Survey fcst* refer to these survey forecasts. The sample period is 1985:02-2014:12. t-statistics are based on Newey and West (1987) standard errors with a maximum lag length of 18 months. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% level, respectively.

endowed with final revised NFP data, although in both cases the differences to an investor only using yield curve information would be rather small.

3.1 Term premiums

A number of previous studies have argued that yield curve models that embed macroeconomic information imply term premiums which are more countercyclical than models that exclusively rely on information contained in yields themselves (see, e.g., Ludvigson and Ng 2009; Wright 2011; Joslin, Priebsch, and Singleton 2014). In these analyses, the countercyclicality increases because the added final-revised macroeconomic variables are found to help predict future bond yields and have themselves a strong cyclical component. In the previous sections, we have found that real time macroeconomic variables are less powerful predictors of future bond yields and are also less cyclical than their final revised counterparts. We may thus expect that the use of real time data reduces the countercyclicality of the model-implied risk premiums.

We assess this conjecture by following the analysis in Ludvigson and Ng (2009). They use predictive return regressions to establish the importance of macroeconomic factors for the predictability of excess bond returns. They further document that the bond risk premiums implied by a model augmented with (final revised) macroeconomic factors are considerably more countercyclical than those implied by a model that does not incorporate macroeconomic information. As in Ludvigson and Ng (2009), we compute the term premium for a bond with n years to maturity as

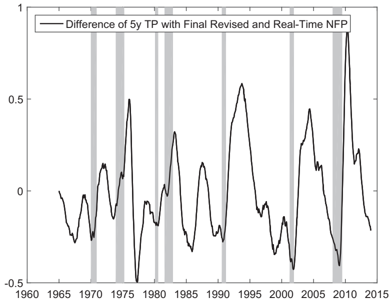
$$tp_t^{(n)} = \frac{1}{n}\left[\hat{E}_t(rx_{t+1}^{(n)}) + \hat{E}_t(rx_{t+2}^{(n-1)}) + \ldots \hat{E}_t(rx_{t+n-1}^{(2)})\right],$$

where we obtain the conditional expectations $\hat{E}_t(rx_{t+h}^{(n-h+1)})$ using h-year ahead predictions from a vector autoregression (VAR) that includes as variables the excess returns themselves as well as a set of predictor variables. Note that the above expression is equivalent to the difference of the yield of an n-year bond and the expectations-hypothesis value, that is, the average expected short rate over the life of the bond (see, e.g., Cochrane and Piazzesi 2009).

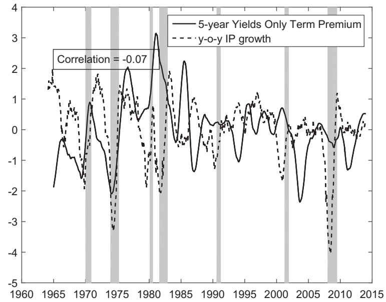
Specifically, in our benchmark specification the VAR includes observations on excess returns and the CP factor or the two spreads. We then compare this benchmark specification to one that adds final-revised or real time NFP data, respectively.[8]

The first panel of Figure 1 shows the difference in 5-year term premium implied by both the model using real time NFP and that using final-revised NFP as additional predictor. This difference is strongly countercyclical, rising around recessions and falling in expansions. Moreover, the quantitative difference

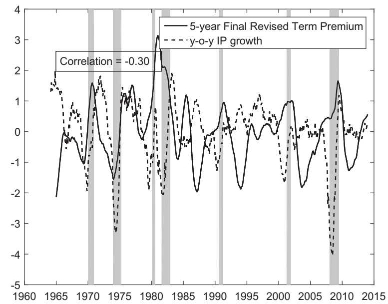
[8] Following Ludvigson and Ng (2009), we use a monthly VAR with $p = 12$ lags.



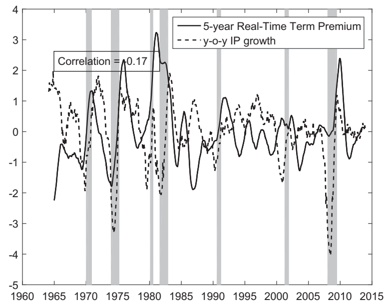
(a) **5-year term premium difference**



(b) **5-year term premium: CP only**



(c) **5-year term premium: CP & final revised NFP**



(d) **5-year term premium: CP & real time NFP**

Figure 1
The cyclicality of term premiums
This figure illustrates the cyclicality of term premiums implied by different specifications of the vector autoregression discussed in Section 3.1. The first panel shows the difference between the term premium implied by a VAR with the CP factor and real time nonfarm payroll employment and a similar VAR with final revised nonfarm payroll growth as predictors. The second panel provides a plot of the 12-month moving average of the term premium implied by a VAR using only past returns and the Cochrane and Piazzesi (2005) factor as predictors versus the 12-month moving average of the monthly growth rate of industrial production. The third panel shows the corresponding plot of the term premium implied by a VAR using past returns, the CP factor and final revised nonfarm payroll growth as predictors. The fourth panel shows the corresponding plot of the term premium implied by a VAR using past returns, the CP factor and real time nonfarm payroll growth as predictors.

can be sizable, reaching almost 1% in 2010. Following Ludvigson and Ng (2009), we visualize the cyclicality of the different term premium estimates by superimposing the growth rate of industrial production.[9] As a point of comparison, the second panel shows the term premium implied by a VAR using only yield curve information to predict returns. Consistent with

[9] Specifically, as in their analysis, we show the 12-month moving average of both the term premium and the monthly growth rate of industrial production.

Ludvigson and Ng (2009), this yield-only term premium is essentially acyclical. The third panel of Figure 1 shows the corresponding plot for the model adding final revised NFP as predictor. The implied term premium is considerably more countercyclical and features a correlation of -30% with IP growth. Finally, the fourth chart provides the corresponding plot for the term premium implied by adding real time NFP as predictor. As conjectured above, this real time term premium is considerably less countercyclical, the correlation with IP growth drops to only -17%.

In sum, the analysis in this section shows that the countercyclicality of term premiums is considerably weaker when real time data, as opposed to final revised macroeconomic data, are considered.

3.2 Mean-variance portfolios

We now highlight the economic significance of the differences in bond return predictability by comparing the optimal (in a mean-variance sense) portfolios of investors endowed with real time versus final revised macroeconomic information. Specifically, given the expected returns $E_t[rx_{t+1}]$ implied by the predictive regressions in (3), we compute portfolio weights (see, e.g., Carriero and Giacomini 2011) as:

$$\begin{aligned}\omega_t &= a + BE_t[rx_{t+1}],\\ \text{where } a &= \frac{\Sigma\iota}{\iota'\Sigma\iota},\\ \text{and } B &= \frac{1-\gamma}{\gamma}\left(\Sigma^{-1} - \frac{\Sigma^{-1}\iota\iota'\Sigma^{-1}}{\iota'\Sigma^{-1}\iota}\right),\end{aligned}$$

whereas γ denotes the coefficient of relative risk aversion, Σ the unconditional covariance matrix of excess returns rx, and ι a vector of ones with length equal to the number of bonds in the portfolio.[10]

Note that these portfolio weights add to one in each period and therefore rule out implicit leverage or underinvestment.

With the portfolio weights ω_t at hand, we compute excess returns on the optimal portfolio as $rx_t^* = \omega_{t-1} rx_t$ and compare their sample average, standard deviation as well as the Sharpe ratio across three specifications: the benchmark model using only yield curve information, the model adding the final revised NFP series as predictor, and the model adding real time NFP as predictor. We carry out this exercise for both in-sample as well as out-of-sample regressions.[11]

10 Following Carriero, Kapetanios, and Marcellino (2012), we set $\gamma = .5$. Varying γ does not affect the results qualitatively. Since our predictive model does not imply conditional variance dynamics, we assume $Var_t(rx_{t+1}) = Var(rx_{t+1}) = \Sigma$.

11 In the out-of-sample analysis we use a 10-year initial learning sample and obtain optimal portfolio weights based on 1-year-ahead return predictions using an expanding estimation window, see also Section 2.3.

Table 6 presents the results in four panels. The left panel shows the in-sample and the right panel the out-of-sample results. The top panel reports results using the CP factor as benchmark yield curve information, the bottom panel instead uses the two spreads to predict excess returns. The average excess return of the optimal portfolio when only the CP factor is used is 0.28%. This increases to 0.32% when one adds final revised NFP as a regressor and to 0.31% when real time NFP is employed. The volatility of optimal portfolio returns also increases but only slightly so when one adds macroeconomic information to predict returns.

As a consequence, the optimal portfolio's Sharpe ratio increases from 0.30 when only the CP factor is used to 0.34 and 0.32 when final revised and real time NFP are added as regressors, respectively. One can assess the statistical significance of the Sharpe ratio differential using the test proposed in Ledoit and Wolf (2008).[12] The test indicates that the differences, albeit small, are indeed statistically different from zero. The results are quantitatively similar for the out-of-sample analysis shown in the right panel of the table. The Sharpe ratio of the optimal portfolio increases from 0.29 using only the CP factor to 0.32 (0.30) when final revised (real time) NFP is added as regressor. However, the Ledoit-Wolf test indicates that these differences are not statistically significant.

Moving to the bottom panel, we see that average excess returns and Sharpe ratios are substantially larger when the two spreads instead of the CP factor are used as yield curve predictors. For example, the in-sample Sharpe ratio for the model using only the two spreads is 0.39, as compared with 0.28 when only the CP factor is used to predict excess returns. The Sharpe ratio increases from 0.39 to 0.43 (0.42) when adding final revised (real time) NFP in the in-sample regressions. These differences in Sharpe ratios are both statistically significant according to the Ledoit-Wolf test. In out-of-sample regressions, the Sharpe ratio also jumps from 0.39 to 0.42 when adding final revised NFP data, but only to 0.40 when adding real time NFP. The latter difference is not statistically significant according to the Ledoit-Wolf test while the former is. Hence, in contrast to final-revised NFP data, observing real time NFP does not significantly improve an investors' portfolio allocation. That said, from an economic point of view the improvement in bond return predictability from using macroeconomic data in addition to information embedded in the yield curve are rather small even when final revised data are used.

4. Announcement Effects

In the previous sections, we have documented that nonfarm payroll growth has stronger predictive power for bond returns when final revised as opposed to

[12] Ledoit and Wolf (2008) propose two tests, one based on HAC standard errors, and one using a resampling algorithm. We employ the former in order to adjust for serial correlation and heteroscedasticity in the return series.

Table 6
Mean-variance portfolio returns

	In-sample predictions					Out-of-sample predictions				
	$E[rx_t^*]$	$Std(rx_t^*)$	SR	Δ_{SR}	p-val	$E[rx_t^*]$	$Std(rx_t^*)$	SR	Δ_{SR}	p-val
	Benchmark: CP					**Benchmark: CP**				
CP	0.28	0.93	0.30			0.30	1.05	0.29		
CP + FR	0.32	0.96	0.34	0.03	0.04	0.34	1.07	0.32	0.02	0.17
CP + RT	0.31	0.95	0.32	0.02	0.06	0.32	1.06	0.30	0.01	0.45
	Benchmark: two spreads					**Benchmark: two spreads**				
Two spreads	0.38	0.98	0.39			0.54	1.37	0.39		
Two spreads + FR	0.45	1.03	0.43	0.04	0.01	0.57	1.37	0.42	0.02	0.05
Two spreads + RT	0.42	1.00	0.42	0.02	0.02	0.55	1.36	0.40	0.01	0.17

This table reports predictive regressions with market information proxies appearing in equation (4), where x_t uses the real time observations as well as: the Treasury yields with maturities ranging from $n = 1, \ldots, 5$ years, and the equity risk factors MKT, SMB, HML, momentum, and short-term reversal. In addition, we also use the median survey expectation for NFP announcements from the Money Market Services database (like in Gilbert 2008) until 1999, and from Bloomberg starting in 2000. The sample period is 1985:02-2014:12. t-statistics are based on Newey and West (1987) standard errors with a maximum lag length of 18 months. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% level, respectively.

real time data are used. Alternative market-based proxies for the information set available to investors give rise to the same result. Does this imply that bond investors cannot exploit the bond return predictability in real time? Since they operate in a data rich environment this question is difficult, if not impossible, for an econometrician to assess ex post.

However, investors' reaction to macroeconomic news potentially reveals what information they did not have before. In this section, we thus analyze announcement effects to nonfarm payroll releases to shed some light on two related questions. First, to what extent are data revisions anticipated? Second, do investors react to announcements of revisions?

To do so, we need to examine the reaction of the Treasury yield curve to NFP announcements, controlling for investors' predictions of the release. Our analysis is closely related to Gilbert (2008), who shows that the return on the S&P 500 index on days of payroll announcements predicts future revisions with a positive sign in expansions and a negative sign in recessions. This is interpreted as evidence that equity investors anticipate the final revised figure of payroll employment when reacting to the initial announcement. In the context of our findings thus far, his results could be interpreted in the following way. When observing a new macroeconomic data release investors update their beliefs, incorporate the new information into prices and at the same time anticipate the future revisions.

We assess whether similar effects might be at work in the Treasury market. While Jones, Lamont, and Lumsdaine (1998) and Fleming and Remolona (1999), among others, have documented the importance of NFP announcements for Treasuries, to the best of our knowledge, no prior paper has studied the relationship between payroll announcement returns in the Treasury market and data revisions.[13]

Following Gilbert (2008) we use the median market expectation for NFP news announcements from the Money Market Services database until 1999, and from Bloomberg starting in 2000 to characterize investors' time-t expectations of the payroll release. More specifically, we estimate the following announcement day regressions

$$\begin{aligned} \Delta y_t^{(n)} &= \alpha_n + \beta_n^S NFP_t^S + \beta_n^{r1}\left(NFP_{t-1|t+1} - NFP_{t-1|t}\right)\dots \\ &\quad + \beta_j^{r2}\left(NFP_{t-2|t+1} - NFP_{t-2|t}\right) \\ &\quad + \beta_j^{rf}\left(NFP_{t|T} - NFP_{t|t+1}\right) + \epsilon_{nt}, \end{aligned} \tag{5}$$

where $\Delta y_t^{(n)}$ denotes the daily change in zero-coupon yields on 2-, 5-, 7-, and 10-year Treasury notes from Gürkaynak, Sack, and Wright (2007) on

[13] The response of other asset markets to macroeconomic news announcements has been studied using high-frequency data in, e.g., Andersen et al. (2003) for foreign exchange markets, Faust et al. (2007) for interest rate and foreign exchange futures, and Kilian and Vega (2011) for commodity prices.

days of payroll announcements. NFP_t^S is the surprise component of the forecast for payroll growth, that is, the difference between the first release and the aforementioned consensus survey forecast; $NFP_{t-1|t+1}$ - $NFP_{t-1|t}$ and $NFP_{t-2|t+1}$ - $NFP_{t-2|t}$ refer to the revisions of the prior 2 months' observations released with the new announcement. Finally, $NFP_{t|T}$ - $NFP_{t|t+1}$ denotes the cumulative revision made after the first release.

Table 7 reports the results. The sample period is 1985:02-2014:12 because the MMS/Bloomberg survey only becomes available in February 1985. We report the full sample estimates in the upper panel and the bottom two panels repeat the regressions for expansion and recession samples separately. These have been determined by whether the announcement date falls into an NBER recession or expansion. For each maturity, the first line reports the parameter estimates and the second the corresponding t-statistics.

We start with the full sample results reported in the top panel. The second column shows that all maturities strongly react to the surprise component of payroll releases as the coefficients are highly statistically significant. As all coefficients are positive, we find that the yield curve shifts up significantly when the actual payroll release exceeds the market expectation. Interestingly, the coefficient on the revision to the prior month's release is statistically significant (third column), suggesting that investors incorporate that information into Treasury prices, over and above the surprise about the new monthly release. The revision to the release 2 months ago (fourth column) also has positive coefficients across all maturities, but these are not statistically significant in the case of the full sample. More importantly, however, the subsequent final revisions reported in the last column are essentially all zero. These results imply that, controlling for past revisions and contemporaneous announcement surprises, the yield curve reaction to payroll news does not anticipate future revisions.

For the full sample, this finding is in line with Gilbert (2008), who documents a statistically insignificant coefficient on future revisions in similar regressions using the daily change of the S&P 500 index as dependent variable. Splitting the sample into expansion and recession periods, however, Gilbert (2008) finds statistically significant and offsetting effects: stock returns react positively to future revisions in expansions and negatively in recessions. Looking at the middle and lower panels of Table 7, we do not see a significant impact of future revisions during either NBER expansions or recessions. While the coefficients on future revisions also have opposite signs in expansions versus recessions in our analysis, neither are found to be statistically significant. Hence, we cannot confirm the result by Gilbert (2008) for the bond market. The use of different dependent variables and a different sample period are likely potential explanations for this discrepancy.[14] Interestingly, looking at the

[14] There is an interesting similarity of the results in Gilbert (2008) with those in Boyd, Hu, and Jagannathan (2005) and Andersen et al. (Andersen, Bollerslev, Diebold, and Vega (2007)). Both papers report an asymmetry in the

Table 7
Treasury yield curve reaction to nonfarm payroll news

	cst	Surprise	$NFP_{t-1\|t+1}$ $-NFP_{t-1\|t}$	$NFP_{t-2\|t+1}$ $-NFP_{t-2\|t}$	$NFP_{t\|T}$ $-NFP_{t\|t+1}$	Adj. R^2
			Full sample (N = 353)			
2-year	0.00	0.05***	0.02***	0.01	0.00	0.32
	(0.48)	(10.69)	(2.63)	(0.70)	(0.25)	
5-year	0.00	0.05***	0.02**	0.02	−0.00	0.26
	(0.62)	(9.01)	(2.36)	(1.12)	(−0.23)	
7-year	0.00	0.04***	0.02**	0.02	−0.00	0.23
	(0.69)	(8.14)	(2.15)	(1.24)	(−0.35)	
10-year	0.00	0.04***	0.02**	0.02	−0.00	0.20
	(0.88)	(7.23)	(1.99)	(1.36)	(−0.49)	
			Expansions (N = 316)			
2-year	−0.00	0.06***	0.03***	0.02	0.01	0.33
	(−0.46)	(10.24)	(3.02)	(1.27)	(1.18)	
5-year	−0.00	0.06***	0.03***	0.03*	0.01	0.30
	(−0.69)	(9.36)	(2.91)	(1.93)	(1.08)	
7-year	−0.00	0.05***	0.03***	0.04**	0.01	0.28
	(−0.65)	(8.71)	(2.68)	(2.12)	(0.98)	
10-year	−0.00	0.04***	0.02**	0.04**	0.00	0.24
	(−0.44)	(7.92)	(2.41)	(2.26)	(0.81)	
			Recessions (N = 37)			
2-year	0.01	0.05***	−0.01	0.01	−0.01	0.29
	(0.51)	(3.40)	(−0.47)	(0.32)	(−0.94)	
5-year	0.01	0.03*	−0.01	0.00	−0.02	0.07
	(0.36)	(1.66)	(−0.51)	(0.08)	(−1.00)	
7-year	0.00	0.02	−0.01	−0.00	−0.02	−0.00
	(0.10)	(0.90)	(−0.38)	(−0.13)	(−1.00)	
10-year	−0.00	0.00	−0.00	−0.01	−0.01	−0.05
	(−0.10)	(0.22)	(−0.05)	(−0.32)	(−0.98)	

This table shows OLS results for the announcement day regressions of the form

$$\begin{aligned}\Delta y_t^{(n)} &= \alpha_n + \beta_n^S NFP_t^S + \beta_n^{r1}\left(NFP_{t-1|t+1} - NFP_{t-1|t}\right)\ldots \\ &\quad + \beta_j^{r2}\left(NFP_{t-2|t+1} - NFP_{t-2|t}\right) + \beta_j^{rf}\left(NFP_{t|T} - NFP_{t|t+1}\right) + \epsilon_{nt}.\end{aligned}$$

$\Delta y_t^{(n)}$ denote the daily change in zero-coupon yields on 2-, 5-, 7-, and 10-year Treasury notes from Gurkaynak-Sack-Wright (2007) on days of nonfarm payroll announcements. NFP_t^S is the surprise component of the forecast for payroll growth, that is, the simple difference between the first release and the consensus survey forecast. We obtain the latter from the Money Market Services database until 1999 and from Bloomberg after 1999; $NFP_{t-1|t+1} - NFP_{t-1|t}$ and $NFP_{t-2|t+1} - NFP_{t-2|t}$ refer to the revisions of the prior two months' observations released with the new announcement. $NFP_{t|T} - NFP_{t|t+1}$ denotes the cumulative revision made after the first release. The sample period is 1985:02-2014:12. For each maturity, the first line reports the parameter estimates the second the t-statistics. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% level, respectively.

subsample pertaining to business cycle expansions, we find that revisions to the NFP release two months ago are also associated with statistically significant

response of stock markets to positive and negative news in business cycle expansions versus recessions. More specifically, Boyd, Hu, and Jagannathan (2005) document that news of rising unemployment typically moves equity prices up in expansions and down in recessions. They explain this finding with the information content of unemployment news about future short rates. Neither Andersen, Bollerslev, Diebold, and Vega (2007) nor Boyd, Hu, and Jagannathan (2005) study the role of data revisions.

coefficients. This reinforces our full sample result that bond investors price in information contained in past revisions.

In summary, the announcement day regressions show that information about revisions to past releases appear to be incorporated into bond prices on announcement days whereas future revisions do not seem to be priced in. While we defer an interpretation of the ensemble of our empirical results to Section 6, these findings appear to suggest that investors do not fully anticipate future revisions. An alternative interpretation could be that the publication of large data revisions increases investors' uncertainty about macroeconomic conditions and thus leads to lower bond prices and higher yields.

5. Robustness

In this section we document the robustness of our results with respect to alternative macroeconomic predictor variables previously considered in the literature. Specifically, we examine the evidence obtained from: (1) the growth rate of industrial production (henceforth denoted IP), (2) the growth rate of real gross domestic product (denoted GDP), (3) the Chicago Fed National Activity Index (or CFNAI) and (4) a principal component extracted from a large panel of macroeconomic data, in the spirit of Ludvigson and Ng (2009) and Ludvigson and Ng (2011). These series have been used in various studies on bond return predictability. More specifically, IP was used by Cooper and Priestley (2009), GDP growth appears in the study by Wright (2011), whereas Joslin, Priebsch, and Singleton (2014) use the CFNAI, and Ludvigson and Ng (2009) and Ludvigson and Ng (2011) document that principal components extracted from a large macroeconomic data panel improve bond return predictions.[15]

In addition to analyzing a broader set of series, we also report results for different samples. All series are seasonally adjusted and with the exception of quarterly GDP growth, all series are observed monthly. The availability of real time data varies across series and imposes different start dates for these analyses: IP growth is available from 1964 onwards, real GDP starts in 1992, CFNAI in 2001, and the real time factor in the spirit of Ludvigson and Ng (2009) is constructed using data from March 1982 onwards. We focus on a sample ending in December 2014. In addition, we also report results for the sample ending in December 2007, that is, not including the Great Recession. The Online Appendix reports details of the robustness results.

---

15 Cooper and Priestley (2009) use the cyclical component of IP, obtained as the difference between the log of IP and a fitted quadratic trend. When comparing the predictive ability of this cyclical component for bond returns with that of monthly IP growth, we find the latter to be superior in our sample. The discussion paper version of Joslin, Priebsch, and Singleton (2014) also used monthly IP growth and reported similar results to those using the CFNAI. Cieslak and Povala (2015) propose to predict bond returns using a factor that represents a linear combination of expected inflation and bond yields. They measure expected inflation as a simple discounted moving average of past CPI inflation. CPI inflation is revised only to the extent that the seasonal adjustment factors get revised. Accordingly, Cieslak and Povala (2015) interpret their factor essentially as real time data, and we therefore do not include it in our robustness analysis.

5.1 Data revisions

Table A.1 in Appendix section A.2 reports the data revision summary statistics of real time data, revisions and final data for (1) monthly IP growth from 1964 to 2014; (2) real GDP growth from 1992 to 2014; (3) the Chicago Fed National Activity Index (CFNAI), which runs from 2001 to 2014, and, finally the (4) Ludvigson and Ng factor available from 1982 to 2014.[16] As for NFP, we show these results for the full sample, as well as for NBER expansions and recessions. For the short CFNAI sample there are no separate expansion/recession results reported.

A picture similar to that of NFP emerges for IP, which covers the same sample period. On average IP grew by .22% per month (which amounts to 2.64% per year) according to the final data, but only .16% based on real time data with the difference almost a third of the final number. Interestingly, during NBER recessions $\nu_{t|T}$ is zero because $\nu_{t|T}^{rev}$ and $\nu_{t|T}^{pl}$ largely offset each other. As for NFP, standard deviations for $\nu_{t|T}$ (and its components) are large, namely .83% for the full sample and even 1.13% during recessions. Moving on to GDP growth, the next series reported in Table A.1, we observe somewhat smaller differences albeit still with large standard deviations which are equal to the average quarterly growth for the full sample and NBER expansions and even larger during recessions. As for the GDP figures, we note that the combined publication lag/revision components for CFNAI are relatively small on average, but have large standard deviations.

One may wonder whether the important role that revisions and publication lags play in the dynamics of the four series reported thus far extends to a larger set of macroeconomic variables. We address this issue by also examining factors extracted from a panel of macroeconomic time series in the spirit of Ludvigson and Ng (2009) and Ludvigson and Ng (2011). While our data set broadly covers the same economic categories, we restrict ourselves to macroeconomic time series that are potentially subject to publication lags and data revisions and for which real time data are available. In the Online Appendix Section OA.1 we provide the details of the resulting panel of 60 series and show that, despite the differences in coverage, the first principal component (PC) extracted from its final revised version is highly correlated with the first PC obtained by Ludvigson and Ng (2009) and Ludvigson and Ng (2011). Crucially, this first PC has the strongest predictive power for future bond returns in their analysis. In contrast, the first PC extracted from the corresponding panel of real time observations shows a much weaker correlation with the LN factor. This indicates not only that the revision and publication lag components are important for the panel of 60 macroeconomic time series, but also that they have a systematic component which is picked up by the estimated factor.

[16] The Online Appendix provides details about the construction of the LN factor.

Looking at the combined revision and publication lag components, we find that, on average, across all series, the variance of ν_t is about 81% of the variance of the final revised series, indicating that revisions make up for a sizable fraction of the total variation in final revised macroeconomic time series. For the median series, this ratio still amounts to a sizable 31%. For purpose of comparison, it is worth recalling that the same ratio for NFP reported earlier is 20% (based on the full sample results reported in Table 1), indicating that NFP is by no means an extreme example as far as revisions go.[17]

The LN factor is centered around zero and scaled to have standard deviation equal to one, as is indicated in the full sample panel of Table A.1. What is remarkable is that the standard deviation of $\nu_{t|T}$ in the full sample is almost equal to one as well (.94 to be precise) and that the standard deviation of $\nu_{t|T}^{pl}$ for the LN factor is even larger, namely equal to 1.20. This finding clearly shows that the results reported for individual series extend to a large cross-section of macroeconomic data.[18]

In Online Appendix Sections OA.4 through OA.6 we report summary statistics and regression results for a sample ending in 2007 and thus excluding the Great Recession for each of the series. To conserve space, we will only comment on the NFP case here. Over the 1964-2007 sample monthly changes in NFP averaged 153K according to the final data, while the real time release was, on average, only 127K, a difference of 25K or again almost 20%, of the real time release, similar to the finding in the full sample. We also find again that real time data understate/overstate the final figures during expansions/recessions. The bulk of $\nu_{t|T}$ is attributable to the revision component, that is, $\nu_{t|T}^{rev}$, for the full sample as well as NBER expansions. In contrast, publication delays, that is, $\nu_{t|T}^{pl}$, are more important during recessions. What is most remarkable in our view, however, is again the sheer magnitude of the standard deviations of these components. Very similar results are obtained for the other series considered.

5.2 Predicting returns with different information sets

We now turn our attention to in-sample predictive regressions involving different macroeconomic series (Table OA.4 of the Online Appendix provides the details). The CP and two spreads panels for IP cover the same sample as NFP and are therefore identical for the yields-only regressions. For GDP and CFNAI the samples differ and thus yield different regression fits. Specifically, the shorter GDP and CFNAI samples have much lower predictive fit with the yields only benchmarks. For the 2-year maturity and regardless of the benchmark model, it appears that adding final macroeconomic data regressors

[17] These 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. See also Croushore (2011) for an analysis of data revisions across variables and an overview of the relevant literature.

[18] Online Appendix Section OA.1.1 offers a more detailed discussion of the data revisions.

always results in significant slope estimates. For the 5-year maturity the results are mixed, as we find that final data do not improve predictability in the cases of IP (CP factor benchmark) and CFNAI (CP factor and two spreads benchmarks).

The increments in adjusted R^2 due to adding final macroeconomic data are sometimes substantial, particularly in the shorter samples of GDP and CFNAI and in the case of the LN factor as predictor. Replacing the final releases with real time observations results in considerably lower adjusted R^2s for all variables, except CFNAI, where it remains the same or increases slightly. The declines are particularly pronounced for GDP and the LN factor. Moreover, as for NFP many of the slope coefficients become insignificant when moving from final revised to real time data. Finally, when regressing bond returns jointly on the lagged real time data, publication lag and revision components, the real time regressor is either not statistically significant - in the case of IP (both maturities and benchmarks), GDP (5-year) and CFNAI (5-year) – or considerably smaller in magnitude and less strongly significant than the other regressors. This is also in line with the evidence reported for NFP. Dropping the benchmark yield curve predictors and using only the macroeconomic predictors in the regressions shows for most variables that although the final data yield a larger R^2 than their real time counterparts, the share of bond return variance explained by macroeconomic data alone is small in general.

Next, we exclude the Great Recession from the sample, with results reported in the Online Appendix Sections OA.4 through OA.7. We find that the in-sample fits improve dramatically. Hence, the Great Recession diminished the role played by macroeconomic data in predicting bond returns. This is perhaps not surprising, because in the period from 2008 to 2014 the Federal Reserve implemented a number of unconventional policies directly aimed at lowering and stabilizing U.S. bond yields. Focusing on the relative importance of final revised versus real time macroeconomic information releases, a pattern similar to that reported in Table 2 emerges. Specifically, for all variables the adjusted R^2s drop considerably, and the slope coefficients also decline and in many cases become insignificant when moving from final revised to real time data. Hence, cutting the sample short by excluding the Great Recession does not alter our conclusions.

5.3 Variance decompositions

The variance decomposition results reported in Table 3 revealed that the contribution of NFP in the prediction of bond returns is reduced in half when the real time data are used. How does that compare to the other series we examine? Online Appendix Table AO.5 offers the results. In terms of relative contributions of final release data, both GDP growth and the CFNAI are the most impressive, at least for the 2-year maturity. GDP growth accounts for a 94% (CP factor) or 81% (two spreads) and CFNAI for a 40% or 71% share of predictive power in their relatively short samples. Replacing $X_{t|T}$ with real time data implies a drop from 94 % to 23% with the CP factor benchmark and from

81 to 21% with the two spreads benchmark for GDP. For IP and the LN factor, we find results similar to those of NFP, using real time instead of final revised data the predictive power drops by half or more compared to the final data. For the CFNAI there is no drop in predictive power, however. Hence, with the exception of CFNAI for which only a very short real time sample is available, our finding that real time data account for a substantially smaller share of bond return predictability than final data also holds for other macroeconomic series previously considered in the literature. Removing the Great Recession from the sample, as reported in the Online Appendix, yields very similar results.

5.4 Out-of-sample analysis

We again use a ten year training sample for the OOS exercises, beginning with the first month/quarter of the available samples, namely 1964:01 for IP, 1992:Q1 for GDP and 1982:03 for the LN factor. Unfortunately, since the sample for CFNAI only covers fourteen years of data we cannot include it in the OOS exercise. We then reestimate the prediction model with an expanding sample and collect the resulting 1-year-ahead bond return forecasts. Online Appendix Table OA.6 reports the details.

Recall that with NFP the CP factor benchmark model featured a higher in-sample fit, whereas the specification using two spreads yielded a smaller mean squared forecast error out-of-sample. This is also the case for IP, GDP and the LN factor. The MSE ratios with final data in some cases feature substantial drops relative to a model using only yield curve information, for example close to 40% for the LN factor with the CP benchmark. In all cases the drop is significant according to the ENC-NEW test. With the exception of GDP for the 2-year maturity with the CP benchmark, the ratios are all closer to one when using real time data. More importantly, in two cases the ratio is no longer significantly different from one, that is, real time data do not improve out-of-sample predictions beyond the benchmark model and in two other cases the statistic is only significant at the 10 % level. The ENC-T statistics also tell us that the final data uniformly feature significantly higher predictive power compared to real time data and/or revisions. Hence, our finding that real time data are substantially less powerful in predicting bond returns out-of-sample than final data also applies to the other macroeconomic predictor variables considered in the previous literature.

5.5 Capturing market expectations

For NFP we found that accounting for broader information sets via auxiliary regressions does little to alter the conclusions we obtained with the real time and first releases. According to the findings reported in Online Appendix Table OA.7 this result also holds for the additional series considered. Moreover, in many cases we find insignificant slope coefficients when market expectation proxies are used as predictors.

5.6 Mean-variance portfolios

We repeat the portfolio analysis reported in subsection 3.2, where we examined the economic significance of the differences in bond return predictability by comparing the optimal portfolios of investors endowed with real time versus final revised NFP data. Using the same setting we re-examine these findings for IP growth, GDP growth and the LN factor. Recall that with NFP as predictor the Sharpe ratio of a mean-variance portfolio increased slightly relative to the yields only benchmarks, but that the increase was lower when real time data was used. While the differences in Sharpe ratios were small, the Ledoit and Wolf (2008) test indicated statistical significance both using in-sample and out-of-sample regression predictions.

The findings are mixed (they are reported in Online Appendix Table OA.8). For the LN factor, we find stronger results than for NFP. Looking at the in-sample results, the Sharpe ratio increases from 0.53 for the CP benchmark (0.67 for the two spreads benchmark) to 0.58 (0.74) adding final data but only to 0.56 (0.71) when adding real time data. The differences in Sharpe ratios are found to be statistically significant at the 10% level for the final revised but not for the real time data in the case of the CP benchmark, and are found to be statistically significant at the 5% level in the case of the two spreads benchmark. The out-of-sample predictions imply even stronger Sharpe ratio differentials which are found to be statistically significant relative to both the CP and the two spreads benchmarks. The increases in Sharpe ratios involving final data are significant, whereas for the real time data the increase is not significant vis-à-vis the CP prediction model in-sample.

For the other series, the results are more nuanced. For IP growth, the Sharpe ratio differentials are small in and out-of-sample for both benchmarks and for final revised and real time data alike. For GDP growth, the Sharpe ratio differentials coming from both final revised and real time data are quite sizable, regardless of which benchmark model is used and more so in the out-of-sample analysis. Specifically, the increase in Sharpe ratios relative to both benchmarks is found to be highly statistically significant in the out-of-sample analysis when final revised data but not when real time data are added. In sum, the evidence for other macroeconomic predictor series supports our result for NFP that using final data leads to superior mean-variance portfolio performance than when real time data are used.

6. Discussion and Concluding Remarks

The results in the previous sections have shown that the predictability of bond returns using macroeconomic information is considerably weaker when real time data, as opposed to revised data, are employed. This is the case not only for nonfarm payroll growth, our lead example, but also for various other predictor variables used in previous studies. Moreover, the term premium implied by a model taking into account final macro data is substantially more countercyclical than that using only real time data.

How can we interpret these results? Standard asset pricing theory implies that risk premiums are time varying because investors expect higher compensation for risky payoffs in bad or more uncertain states of the world. Since revisions reflect statistical agencies' gradual incorporation of additional and more timely data, it is reasonable to assume that revised data capture the state of the economy more precisely than real time data. Our finding that the predictability of bond returns is stronger when revised data are used is therefore consistent with standard asset pricing theory implying time-varying risk premiums. In the same vein, our result that revisions are correlated with bond returns could simply reflect that both risk premiums and revisions have a cyclical component, or that data revisions proxy for economic uncertainty and therefore comove with expected returns.

A complementary interpretation has a role for data revisions in investors' expectations about monetary policy. An extensive literature going back to Taylor (1993) has documented that U.S. monetary policy is well characterized by simple rules linking the federal funds rate to measures of the output gap and inflation. Hence, investors' assessment of current and future macroeconomic conditions represent a key input to their expectations about the future path of short rates. However, as shown by Orphanides (2001), the prescription of a Taylor-type policy rule for the federal funds rate based on real time data can substantially differ from one based on revised data. To the extent that investors incorporate data revisions when they update their beliefs about the state of the economy and likely path of monetary policy, bond returns should reflect these changes and therefore be correlated with revisions, as we find in the data.

The results in Section 4 seem to be at odds with the notion that real time investors fully anticipate future revisions, however. An alternative interpretation of our findings could thus be that the difference in predictability of bond returns using final versus real time macro data captures informational frictions. A related argument has recently been made by Cieslak (2016), who documents that survey expectations of the real federal funds rate feature serially correlated forecast errors which induce a predictable component in bond returns. Although Cieslak (2016) abstracts from data revisions and distinguishes between survey forecasts made by investors in real time versus ex post forecasts of an econometrician using full sample information, her analysis also points to the role of informational frictions in bond return dynamics. More research is needed to parse out the driving forces of bond yields taking into account investors' information sets.

Appendix 1: Variance Decomposition

To trace out the incremental predictive information contained in the various macroeconomic information sets, we would be tempted to run the regression:

$$r_{t+12}^{(n)} = \alpha_n + \beta_n bm_t + \gamma_n^{rt} X_{t-1|t} + \gamma_n^{pl} \nu_{t|T}^{pl} + \gamma_n^{rev} \nu_{t|T}^{rev} + e_{t+12}^{(n)}. \tag{A1}$$

More specifically, the order in which the regressors appear in the above regression follows the flow of information: $X_{t-1|t}$ is released first, then the publication lag $\nu_{t|T}^{pl}$ is realized and $\nu_{t|T}^{rev}$ is available last. Unfortunately, the regressors in the above equation are not mutually orthogonal so that a unique variance decomposition is not feasible. However, using the classic Frisch-Waugh-Lovell theorem (see, e.g., Davidson and MacKinnon 1993, p 19-24), one can sequentially add regressors in the order of the information flow and obtain a unique variance decomposition. The setup for the Frisch-Waugh-Lovell (FWL) theorem is a classical regression:

$$y = X\beta + X_\star \beta_\star + \epsilon$$

and a consistent estimator for $\beta_\star$ can be obtained via the regression:

$$My = MX_\star \beta_\star + \epsilon \tag{A2}$$

where $P = X(X'X)^{-1}X'$, and $M = (I - P)$, which amounts to projecting y on the space spanned by X and compute its residuals, then project $X_\star$ onto the space spanned by X yielding residuals $MX_\star$, and finally regressing My on $MX_\star$.

The variance decomposition shares many features with the well known Cholesky decomposition used in VAR model impulse response analysis. There is an important difference between standard applications of Cholesky factorizations and the variance decomposition, however. The order variables appear in a VAR model may greatly affect impulse response analysis and in typical applications there is no natural order. In our variance decomposition there is a natural order. As we mention right below equation (A1), the order in which the regressors appear in the regression follows the information flow: $X_{t-1|t}$ is released first, then the publication lag $\nu_{t|T}^{pl}$ is realized and $\nu_{t|T}^{rev}$ is available last.[19] The variance decomposition consists of running the following set of regressions:

$$\begin{aligned} X_{t-1|t} &= \delta_n^0 + \delta_n^1 bm_t + u_{rt,t+12}^{(n)} \\ \nu_{t|T}^{pl} &= \delta_n^0 + \delta_n^1 bm_t + \delta_n^2 X_{t-1|t} + u_{pl,t+12}^{(n)} \\ \nu_{t|T}^{rev} &= \delta_n^0 + \delta_n^1 bm_t + \delta_n^2 X_{t-1|t} + \delta_n^3 \nu_{t|T}^{pl} + u_{rev,t+12}^{(n)} \end{aligned} \tag{A3}$$

and using the residuals $u_{rt,t+12}^{(n)}$, $u_{pl,t+12}^{(n)}$ and $u_{rev,t+12}^{(n)}$ to run the regressions:

$$\begin{array}{lll} r_{t+12}^{(n)} = \alpha_n + \beta_n bm_t + e_{bm,t+12}^{(n)} & \text{yielding:} & R_{bm,n}^2 \\ & \text{and} & Var(P_1 r_{t+12}^{(n)}) \\ e_{bm,t+12}^{(n)} = \gamma_n^{rt} u_{rt,t+12}^{(n)} + e_{rt,t+12}^{(n)} & \text{yielding:} & R_{rt,n}^2 \\ & \text{and} & Var(P_{2|1} M_1 r_{t+12}^{(n)}) \\ e_{rt,t+12}^{(n)} = \gamma_n^{pl} u_{pl,t+12}^{(n)} + e_{pl,t+12}^{(n)} & \text{yielding:} & R_{pl,n}^2 \\ & \text{and} & Var(P_{3|1-2} M_{1-2} r_{t+12}^{(n)}) \\ e_{pl,t+12}^{(n)} = \gamma_n^{rev} u_{rev,t+12}^{(n)} + e_{rev,t+12}^{(n)} & \text{yielding:} & R_{rev,n}^2 \\ & \text{and} & Var(P_{4|1-3} M_{1-3} r_{t+12}^{(n)}) \end{array} \tag{A4}$$

with the following projection matrices: (a) $P_1 = X_a(X_a'X_a)^{-1}X_a'$ with $X_a = (\iota\ [bm_t^{(n)}]_{t=1}^T)$ and $M_1 = (I - P_1)$, (b) $P_{2|1} = M_1 X_b (X_b' M_1 X_b)^{-1} X_b' M_1$, with $X_b = ([x_{t-1|t}]_{t=1}^T)$ and $M_{1-2} = (I - P_{2|1})$, (c)

[19] Hence, the variance decomposition shares features with the use of Cholesky factorizations in mixed frequency VAR models, see Ghysels 2016, where high-frequency data are stacked in chronological order into low frequency vectors.

$P_{3|1-2} = M_{1-2}X_c(X_c' M_{1-2}X_c)^{-1}X_c' M_{1-2}$, with $X_c = ([\nu_{t|T}^{pl}]_{t=1}^{T})$ and $M_{1-3} = (I - P_{3|1-1})$, $P_{4|1-3} = M_{1-3}X_d(X_d' M_{1-3}X_d)^{-1}X_d' M_{1-3}$, with $X_d = ([\nu_{t|T}^{rev}]_{t=1}^{T})$ and $M_{1-3} = (I - P_{3|1-1})$.

Based on the above equation and denoting $SSR = Var(P_1 r_{t+12}^{(n)}) + Var(P_{2|1}M_1 r_{t+12}^{(n)}) + Var(P_{3|1-2}M_{1-2}r_{t+12}^{(n)}) + Var(P_{4|1-3}M_{1-3}r_{t+12}^{(n)})$ we can write:

$$\begin{aligned} 1 \quad = \quad & \frac{Var(P_1 r_{t+12}^{(n)})}{SSR} + \frac{Var(P_{2|1}M_1 r_{t+12}^{(n)})}{SSR} \\ & + \frac{Var(P_{3|1-2}M_{1-2}r_{t+12}^{(n)})}{SSR} + \frac{Var(P_{4|1-3}M_{1-3}r_{t+12}^{(n)})}{SSR} \end{aligned} \tag{A5}$$

Finally, the above analysis is also applied to a similar type of regression, but involving final data instead, namely

$$r_{t+12}^{(n)} = \alpha_n + \beta_n bm_t + \gamma_n^f X_{t|T} + e_{t+12}^{(n)}. \tag{A6}$$

Appendix 2: Robustness Analysis

Table A.1
Summary statistics: alternative predictors

	Mean	SD	Min	Max
		IP 1964-2014		
$X_{t\|T}$	0.22	0.75	−4.30	3.09
$X_{t-1\|t}$	0.16	0.71	−3.56	2.43
$\nu_{t\|T}$	0.06	0.83	−3.17	4.98
$\nu_{t\|T}^{pl}$	−0.00	0.76	−2.35	4.57
$\nu_{t\|T}^{rev}$	0.06	0.44	−1.49	2.18
		IP 1964-2014: NBER expansions		
$X_{t\|T}$	0.37	0.58	−1.90	3.09
$X_{t-1\|t}$	0.31	0.54	−1.72	2.43
$\nu_{t\|T}$	0.06	0.76	−2.04	4.81
$\nu_{t\|T}^{pl}$	0.01	0.70	−1.95	4.15
$\nu_{t\|T}^{rev}$	0.06	0.42	−1.17	2.13
		IP 1964-2014: NBER recessions		
$X_{t\|T}$	−0.67	0.94	−4.30	1.98
$X_{t-1\|t}$	−0.67	0.98	−3.56	1.58
$\nu_{t\|T}$	−0.00	1.13	−3.17	4.98
$\nu_{t\|T}^{pl}$	−0.05	1.02	−2.35	4.57
$\nu_{t\|T}^{rev}$	0.05	0.56	−1.49	2.18
		GDP 1992-2014		
$X_{t\|T}$	0.64	0.62	−2.11	1.89
$X_{t-1\|t}$	0.63	0.49	−1.57	1.74
$\nu_{t\|T}$	0.01	0.65	−2.05	1.62
$\nu_{t\|T}^{pl}$	0.01	0.53	−1.16	1.37
$\nu_{t\|T}^{rev}$	−0.01	0.39	−1.15	0.97

(*continued*)

Table A.1
Continued

	Mean	Std	Min	Max
		GDP 1992-2014: NBER expansions		
$X_{t\|T}$	0.77	0.46	−0.39	1.89
$X_{t-1\|t}$	0.72	0.39	−0.26	1.74
$\nu_{t\|T}$	0.05	0.60	−1.17	1.62
$\nu^{pl}_{t\|T}$	0.03	0.52	−1.16	1.37
$\nu^{rev}_{t\|T}$	0.02	0.37	−1.03	0.97
		GDP 1992-2014: NBER recessions		
$X_{t\|T}$	−0.34	0.82	−2.11	0.53
$X_{t-1\|t}$	0.01	0.71	−1.57	0.96
$\nu_{t\|T}$	−0.35	0.89	−2.05	1.44
$\nu^{pl}_{t\|T}$	−0.14	0.63	−0.90	1.32
$\nu^{rev}_{t\|T}$	−0.21	0.51	−1.15	0.35
		CFNAI 2001-2014		
$X_{t\|T}$	−0.33	0.85	−4.18	0.77
$X_{t-1\|t}$	−0.36	0.74	−3.48	0.75
$\nu_{t\|T}$	0.02	0.36	−1.78	0.97
$\nu^{pl}_{t\|T}$	0.01	0.25	−1.01	0.63
$\nu^{rev}_{t\|T}$	0.02	0.22	−0.88	0.39
		LN Factor 1982-2014		
$X_{t\|T}$	−0.00	1.01	−5.18	3.52
$X_{t-1\|t}$	−0.01	1.01	−3.73	2.62
$\nu_{t\|T}$	0.01	0.94	−4.12	2.70
$\nu^{pl}_{t\|T}$	0.00	1.21	−4.14	3.47
$\nu^{rev}_{t\|T}$	0.00	0.68	−3.08	1.59
		LN Factor 1982-2014: NBER expansions		
$X_{t\|T}$	0.24	0.70	−1.75	3.52
$X_{t-1\|t}$	0.16	0.87	−3.28	2.62
$\nu_{t\|T}$	0.08	0.91	−2.91	2.70
$\nu^{pl}_{t\|T}$	0.02	1.20	−4.14	3.47
$\nu^{rev}_{t\|T}$	0.06	0.63	−1.99	1.59
		LN Factor 1982-2014: NBER recessions		
$X_{t\|T}$	−1.85	1.13	−5.18	−0.31
$X_{t-1\|t}$	−1.29	1.05	−3.73	0.74
$\nu_{t\|T}$	−0.56	1.01	−4.12	1.37
$\nu^{pl}_{t\|T}$	−0.16	1.26	−2.81	2.81
$\nu^{rev}_{t\|T}$	−0.40	0.88	−3.08	1.07

This table reports the summary statistics of real time data, revisions and final data for (1) industrial production growth (IP) from 1964 to 2014; (2) GDP growth from 1992 to 2014; (3) the Chicago Fed National Activity Index (CFNAI), which runs from 2001 to 2014; and (4) a LN-type factor available from 1982 to 2014. There are three sample configurations, covering respectively the full sample, NBER expansions and recessions. For the short CFNAI sample, there are no expansion/recession results reported. The final revised data is labeled $X_{t|T}$. This can be decomposed into two components: $X_{t|T} = X_{t-1|t} + \nu_{t|T}$, where $X_{t-1|t}$ is the real time data and $\nu_{t|T} = (X_{t|T} - X_{t|t+1}) + (X_{t|t+1} - X_{t-1|t})$ which we write as: $\nu_{t|T} = \nu^{rev}_{t|T} + \nu^{pl}_{t|T}$ with $\nu^{rev}_{t|T}$ capturing the cumulative future revisions of the initial announcement, $X_{t|T} - X_{t|t+1}$, and $\nu^{pl}_{t|T}$ the publication lag component which captures the fact that macroeconomic data are released with a lag, $X_{t|t+1} - X_{t-1|t}$.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega. 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review* 93:38–62.

——-. 2007. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics* 73:251–77.

Aruoba, S. 2008. Data revisions are not well behaved. *Journal of Money, Credit and Banking* 40: 319–40.

Boyd, J. H., J. Hu, and R. Jagannathan. 2005. The stock market's reaction to unemployment news: Why bad news is usually good for stocks. *Journal of Finance* 60:649–72.

Carriero, A., and R. Giacomini. 2011. How useful are no-arbitrage restrictions for forecasting the term structure of interest rates?. *Journal of Econometrics* 164:21–34.

Carriero, A., G. Kapetanios, and M. Marcellino. 2012. Forecasting government bond yields with large Bayesian vector autoregressions. *Journal of Banking and Finance* 36:2026–47.

Cieslak, A. 2016. Short-rate expectations and unexpected returns in Treasury bonds. Working Paper.

Cieslak, A., and P. Povala. 2015. Expected returns in Treasury bonds. *Review of Financial Studies* 28: 2859–2901.

Clark, T., and M. McCracken. 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105:85–110.

——-. 2013. Advances in Forecast Evaluation. *Handbook of Economic Forecasting, Vol. 2, Part B*, eds. G. Elliott, and A. Timmermann, 1107–1201. New York: Elsevier.

Cochrane, J., and M. Piazzesi. 2005. Bond risk premia. *American Economic Review* 95, 138–60.

——-. 2009. Decomposing the yield curve. Working Paper.

Cooper, I., and R. Priestley. 2009. Time-varying risk premiums and the output gap. *Review of Financial Studies* 22:2801–33.

Croushore, D. 2011. Frontiers of real time data analysis. *Journal of Economic Literature* 49:72–100.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and inference in econometrics*. Oxford: Oxford University Press.

Diebold, F. X., and R. S. Mariano. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13:253–63.

Faust, J., J. H. Rogers, S.-Y. B. Wang, and J. H. Wright. 2007. The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Monetary Economics* 54:1051–68.

Fleming, M. J., and E. M. Remolona. 1999. Price formation and liquidity in the US Treasury market: The response to public information. *Journal of Finance* 54:1901–15.

Ghysels, E. 2016. Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics* 193: 294–314.

Gilbert, T. 2011. Information aggregation around macroeconomic announcements: Revisions matter. *Journal of Financial Economics* 101:114–31.

Gürkaynak, R. S., B. Sack, and E. Swanson. 2005. The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American Economic Review* 95:425–36.

Gürkaynak, R. S., B. Sack, and J. H. Wright. 2007, "The US Treasury yield curve: 1961 to the present. *Journal of Monetary Economics* 54:2291–2304.

Harvey, D. S., S. J. Leybourne, and P. Newbold. 1998, "Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16:254–59.

Jones, C. M., O. Lamont, and R. L. Lumsdaine. 1998, “Macroeconomic news and bond market volatility. *Journal of Financial Economics* 47:315–37.

Joslin, S., A. Le, and K. J. Singleton. 2013. Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. *Journal of Financial Economics* 109:604–22.

Joslin, S., M. Priebsch, and K. J. Singleton. 2014. Risk premiums in dynamic term structure models with unspanned macro risks. *Journal of Finance* 69:1197–1233.

Kilian, L., and C. Vega. 2011. Do energy prices respond to US macroeconomic news? A test of the hypothesis of predetermined energy prices. *Review of Economics and Statistics* 93:660–71.

Ledoit, O., and M. Wolf. 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15:850–59.

Litterman, R. B., and J. Scheinkman. 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1:54–61.

Ludvigson, S., and S. Ng. 2007. The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics* 83:171–222.

——-. 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22:5027–67.

——-. 2011. A factor analysis of bond risk premia. *Handbook of Empirical Economics and Finance*, eds. A. Ulah, and D. Giles, 313–72. New York: Chapman and Hall.

West, Kenneth D., and Whitney K. Newey. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–08.

Orphanides, A. 2001. Monetary policy rules based on real time data. *American Economic Review* 91:964–85.

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*, eds. R. Engle, and H. White. Oxford: Oxford University Press.

Taylor, J. B. 1993. Discretion versus policy rules in practice. in *Carnegie-Rochester Conference Series on Public Policy* 39:195–214.

West, K. D. 1996. Asymptotic inference about predictive ability. *Econometrica* 64:1067–84.

Wright, J. 2011. Term premia and inflation uncertainty: Empirical evidence from an international panel dataset. *American Economic Review* 101:1514–34.