Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

What Does Financial Volatility Tell Us About Macroeconomic Fluctuations?

Marcelle Chauvet, Zeynep Senyuz, and Emre Yoldas

2013-61

What Does Financial Volatility Tell Us About Macroeconomic Fluctuations?*

First version: October 2010
This version: August 2013

Marcelle Chauvet[a]　　Zeynep Senyuz[b]　　Emre Yoldas[c]

Abstract

This paper provides an extensive analysis of the predictive ability of financial volatility measures for economic activity. We construct monthly measures of stock and bond market volatility from daily returns and model volatility as composed of a long-run component that is common across all series, and a set of idiosyncratic short-run components. Based on powerful in-sample predictive ability tests, we find that the stock volatility measures and the common factor significantly improve short-term forecasts of conventional financial indicators. A real-time out of sample assessment yields a similar conclusion under the assumption of noisy revisions in macroeconomic data. In a non-linear extension of the dynamic factor model for volatility series, we identify three regimes that describe the joint volatility dynamics: low, intermediate and high-volatility. We also find that the non-linear model performs remarkably well in tracking the Great Recession of 2007-2009 in real-time.

Keywords*:* Financial Volatility, Real-time Data, Predictive Ability Tests, Dynamic Factor Model, Markov Switching.

JEL Classification*:* C32, E32, E44

---

* We would like to thank William Bassett, John Driscoll, participants of the Applied Time Series Workshop at the Federal Reserve Bank of St. Louis, European Central Bank Workshop on the Role of Nonlinear Methods in Empirical Macroeconomics and Forecasting, Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, Conference on Real-Time Data Analysis at the Federal Reserve Bank of Philadelphia/CIRANO, International Symposium on Forecasting, Joint Statistical Meetings, and seminar participants at the Federal Reserve Board, Cornerstone Research, Indiana University, and The University of Queensland for useful comments. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, its staff or any other person associated with the Federal Reserve System.
[a] Department of Economics, University of California, Riverside. E-mail: chauvet@ucr.edu
[b] Federal Reserve Board, Division of Monetary Affairs. E-mail: zeynep.senyuz@frb.gov
[c] Federal Reserve Board, Division of Monetary Affairs, 20th and C Streets NW, Washington, DC 20551, Phone: +1 202 973 7302 E-mail: emre.yoldas@frb.gov (Corresponding author)

1. Introduction

The predictive ability of financial variables such as term spreads, stock returns and credit spreads for economic activity has been extensively studied, see for example Fama (1990), Estrella and Mishkin (1998), Ang et al. (2006), and Gilchrist and Zakrajsek (2012), among others. Recently, using the information content of second moment of stock returns for predicting economic activity has also attracted attention. For example, Fornari and Mele (2009) study the predictive ability of aggregate stock market volatility combined with other commonly used financial indicators. Bakhsi et al. (2011) provide a similar analysis for the forward variances extracted from option portfolios while Allen et al. (2012) construct an aggregate systemic risk measure that has predictive value for economic downturns.

The relation between financial volatility and economic activity has solid theoretical foundations. Schwert (1989a, 1989b) shows that measures of return volatility proxy for the uncertainty surrounding future cash flows and discount rates according to the standard present value model of stock prices. According to the rational valuation framework of Mele (2007), investors require higher returns during relatively bad times and increases in risk premia are larger in magnitude than declines in good times, leading to counter-cyclical return volatility. Bloom (2009) investigates the impact of shocks to economic uncertainty under stochastically evolving business conditions and finds that uncertainty shocks generate sharp recessions and recoveries. Christiano et al. (2010), Arellano et al. (2011) and Bloom et al. (2012) study uncertainty shocks in the context of dynamic stochastic general equilibrium models and argue that models incorporating uncertainty shocks provide a better fit in matching business cycle dynamics.

In this paper, we analyze the predictive value of various financial volatility measures as a proxy for anticipated real uncertainty. We consider the volatility of a broadly defined stock market portfolio as well as an aggregated volatility measure from industry portfolios. Moreover, we include a measure of volatility from the Treasury bond market to capture potential uncertainty arising from monetary policy, future interest rates, and term premia. We use daily data to construct the financial volatility measures at monthly frequency from January 1985 to June 2012.[1] We then use a variety of linear and nonlinear methods to assess the predictive power of these volatility measures for macroeconomic activity using both revised and real-time data.

Our paper makes several important contributions to the literature that focuses on the relation between real activity and volatility of financial asset returns. First, we consider multiple volatility measures in a unified framework. Second, we explicitly consider components of volatility that have potentially different linkages with the underlying economic fundamentals. Third, in our assessment of predictive ability of the volatility measures, we rely on recently developed in-sample tests that focus on finite-sample predictive ability, which also allows us to analyze sub-samples. Moreover, we implement out-of-sample tests that explicitly incorporate real-time data revisions. Finally, we identify distinct states of volatility and shed light on their relationship with macroeconomic activity.

We first show that the aforementioned volatility measures are significantly correlated with growth in industrial production and non-farm employment. These volatility measures also exhibit strong co-movement, which we capture in a dynamic factor framework motivated by the ICAPM model of Adrian and Rosenberg (2008). Decomposing volatility allows us to filter out the noisy short-run component and identify the long-run component of volatility that is presumably more strongly tied to underlying economic fundamentals. We then implement the in-sample predictive ability tests of Clark and McCracken (2012) for individual volatility measures as well as the common long-run component volatilities. We find that stock volatility measures and the common factor perform similar to conventional financial indicators, namely the term spread, the credit spread and the return on a broadly defined stock portfolio. Moreover, the volatility measures improve forecasts from conventional indicators, especially over relatively short forecast horizons. Test results obtained from dividing the sample into approximately equal

[1] Ex-post sample variances that are computed from higher frequency return data as lower frequency volatility measures have been extensively used in the empirical finance literature, see for example the early work of Poterba and Summers (1986) and French, Schwert and Stambaugh (1987). More recently, Andersen et al. (2001a, 2001b, 2003, and 2005) showed the empirical success of realized volatility for measuring and modeling underlying return variability.

sub-samples reveal that volatility measures, similar to conventional financial indicators, have predictive power only over the period from 1998 to 2012, which is characterized by several episodes of global and domestic financial turmoil.

We also provide an out-of-sample analysis with real-time data. We adopt the framework proposed by Clark and McCracken (2009), which take into account effects of data revisions on the distribution of test statistics. We find that mean squared forecast error for industrial production and employment growth from models using volatility measures are not significantly different from those using conventional financial indicators under noisy data revisions, especially over short forecast horizons. This result holds using both initial and final data releases to evaluate forecasts produced in real-time. According to the nested model comparisons, the stock volatility measures and the common factor significantly improve short-term forecasts from conventional financial indicators, assuming presence of noise in data revisions.

When we allow for nonlinear dynamics in the common factor of volatility measures, we find that three-regimes are necessary to adequately capture the joint volatility dynamics: low, intermediate, and high volatility. The regime classification from this model implies that the expansionary periods during which the economy performs well are also the periods that exhibit low financial volatility. The intermediate volatility regime is associated with either episodes of uncertainty during economic expansions or mild recessions. The high-volatility regime typically leads NBER recessions and periods of economic slowdowns slightly and prevails during the entire Great Recession. As a robustness check, we also estimate the model recursively and calculate regime probabilities as one would do in real-time. We find that all our findings from the full-sample hold, and that the model performs remarkably well in tracking the most recent recession in real-time.

The rest of the paper is organized as follows. The next section explains construction of the volatility measures. Section 3 introduces the data set and provides some preliminary analysis. Section 4 contains a comprehensive analysis of the predictive power of various volatility measures for macroeconomic activity in the context of predictive regressions as well as Markov-switching dynamic factor models. Section 5 concludes.

2. Volatility Measures

We construct three measures of realized volatility: volatility of a broad stock market portfolio, an aggregated measure of volatility from industry portfolios, and a bond market volatility measure from spot Treasury yields.

Let r_{ms} denote the daily excess return over the risk free rate for the value-weighted market portfolio, where s denotes the trading days in a given month, which is indexed by t. Our log-transformed realized volatility measure for the market portfolio, RVM, is defined as follows

$$(1) \qquad RVM_t = \frac{1}{2}\ln\left(\sum_{s\in t}^{n_t} r_{ms}^2\right), \quad t = 1, \dots, T,$$

where $\ln(.)$ is the natural-logarithm function, n_t denotes the number of trading days in month t, and T denotes the total number of months in the sample.[2]

Following Campbell et al. (2001), we also consider a stock volatility measure which is obtained by aggregating information from industry portfolios. Let r_{is} denote the daily value-weighted return of all firms in industry i and define $e_{is} = r_{is} - r_{ms}$. Then our second stock volatility measure, RVI, is given by

$$(2) \qquad RVI_t = \frac{1}{2}\ln\left(\sum_{i=1}^{m_i} w_{it} \sum_{s\in t}^{n_t} e_{is}^2\right), \qquad t = 1, \dots, T,$$

where w_{it} is the weight of industry i with respect to market capitalization, and m_i denotes the total number of industries. This definition of volatility stands between the systemic volatility as measured by the volatility of the market portfolio, and the idiosyncratic firm level volatility. Campbell et al. (2001) document strong correlation between this measure and GDP growth.

[2] A realized volatility measure taking into account the first order autocorrelation in daily returns can be calculated similarly. We also consider this alternative in our calculations and find that the results are qualitatively similar.

Our last volatility measure is obtained from the Treasury bond market. Let y_s denote the continuously compounded yield of the 10-year zero coupon T-bond. The daily bond return is given by $r_{bs} = 10(y_{s-1} - y_s)$. We then construct the bond market volatility measure, RVB, based on this daily return as follows

$$(3) \qquad RVB_t = \frac{1}{2}\ln\left(\sum_{s\in t}^{n_t} r_{bs}^2\right), \qquad t = 1, \ldots, T.$$

The realized volatility approach provides directly observable return volatility measures, which are fully nonparametric and incorporate the inherent information in the higher-frequency data. We exploit these properties to understand the extent of the relation between financial return volatility and aggregate economic activity.

3. Data and Preliminary Analysis

The daily stock returns are retrieved from Kenneth French's online Data Library. We consider 48 industries in our data set. The bond data are obtained from Gurkaynak et al. (2007).[3] We consider the growth in the U.S. industrial production index and non-agricultural payroll employment as measures of macroeconomic activity at the monthly frequency.[4] The conventional financial indicators that we consider are the term spread, the difference between the 10-year Treasury note yield and the 3-month Treasury bill yield, the credit spread, the difference between Moody's seasoned BAA and AAA corporate bond yield indices, and the return on the value weighted NYSE portfolio.[5]

[3] For a complete list of the industries and classification procedures refer to Kennett French's data library http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Gurkaynak et al. (2007) dataset can be downloaded from the Board of Governors of the Federal Reserve System at http://www.federalreserve.gov/pubs/feds/2006/index.html.

[4] Industrial production series is obtained from the Board of Governors of the Federal Reserve System (http://www.federalreserve.gov/releases/G17/default.htm) and payroll data are from the Bureau of Labor Statistics (http://www.bls.gov/data/#employment).

[5] Treasury and corporate yield data are obtained from the economic database of the Federal Reserve Bank of St. Louis (FRED®) available at http://research.stlouisfed.org/fred2/ and the value-weighted market return is from Kennett French's data library (see Footnote 3 above).

The full sample runs from January 1985 to June 2012. There are three main reasons for our choice of the starting date of the sample. The first is the change in broader U.S. monetary policy beginning from early 1980s. Second, there is a structural break in the volatility of output growth in the first half of 1984, documented by McConnell and Perez-Quiroz (2000). Third, capital markets have played an increasingly important role in financial intermediation beginning from mid-1980s.

Figure 1 plots the three realized volatility measures as described above. The volatility series are individually quite noisy, but a somewhat similar pattern can be observed; volatilities are generally higher during recessions and lower during expansions. The aggregated industry volatility moves closely with the market volatility. Notice that both series increased considerably from mid-1990s to early 2000s. This was followed by a very low volatility period that lasted until the beginning of the recent financial crisis in 2007-2008. The bond market measure is fairly correlated with the stock market measures, although its dynamics were noticeably different during the second half of 1990s.

We start with a simple correlation analysis to get some insight into the relationship between these financial volatility measures and the macroeconomic aggregates. Table 1 summarizes the results. Over the full-sample period, all volatility series are negatively correlated with the growth in industrial production, but the reported point estimates are significant only for the stock market measures according to the HAC p-values (Panel A). When we divide the sample into approximately equal two sub-periods, we observe that from January 1985 to December 1996 the correlation estimates are near-zero and insignificant while the period from January 1997 to June 2012 is characterized by significant and strongly negative correlations. Similar observations apply to the correlations between volatility series and the employment growth in the full-sample as well as the sub-samples (Panel B). One notable difference is that the correlations between the market and bond volatility series with employment growth are larger in magnitude compared to those between the same volatility series and industrial production growth. Overall, the correlation estimates in Table 1 suggest that there is a negative association between financial volatility measures and macroeconomic activity over the full-sample that is mostly driven by the dynamics in the second sub-sample, a period characterized with episodes of financial distress and the severe recession that followed the financial crisis of 2007-2008.

The similarities in the dynamics of the volatility measures and their relationship to macroeconomic aggregates suggest that combining information from these series could improve our understanding of how financial volatility and the macroeconomy are related. Therefore, we consider a simple common factor specification that filters out the inherent noise in individual volatility series and summarizes the common information that is more likely to be correlated with economic fundamentals. This approach is also motivated by the ICAPM model of Adrian and Rosenberg (2008), which decomposes aggregate stock return volatility into a relatively persistent long-term component and a transitory short-term component.

Let $y_t = (RVM_t, RVI_t, RVB_t)'$, then a simple dynamic factor model of volatility dynamics can be represented as follows:

$$(4) \qquad y_{i,t} = \lambda_i VF_t + u_{i,t}, \quad u_{i,t} = \phi_i u_{i,t-1} + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim NID(0, \sigma_i^2),$$

$$(5) \qquad VF_t = \alpha + \psi VF_{t-1} + \epsilon_t, \qquad \epsilon_t \sim NID(0, \tau^2).$$

Hence, VF represents the common factor and λs denote the loadings, which show the degree of correlation between individual volatility series and the common factor. The common factor is assumed to be uncorrelated with idiosyncratic terms at all leads and lags to ensure identification. The model is estimated using the Kalman filter, and the maximum likelihood estimates are reported in Table 2. Over the full sample, the extracted volatility factor is highly persistent and has an autoregressive coefficient estimate of around 0.8. All factor loadings are positive and highly significant indicating positive correlation between the individual volatility measures and the extracted factor. The volatility of the market portfolio has the largest loading on the common factor and also the least persistent idiosyncratic component. The other two volatility measures, especially the one obtained from industry portfolios, have more persistent idiosyncratic components and smaller factor loadings. Therefore, the parameter estimates suggest that the extracted common factor is essentially a systemic volatility measure. The estimates show some variability over the two sub-samples but the results are qualitatively very similar. Figure 2 plots the extracted common volatility factor.

4. Predicting Macroeconomic Activity using Financial Volatility

We employ two different approaches to assess predictive ability of the financial volatility measures for macroeconomic aggregates. The first is the in-sample testing framework proposed by Clark and McCracken (2012). This methodology is designed to evaluate the marginal significance of a variable – or a set of variables – by taking into account the trade-off between the signal provided by the variable(s) of interest and the noise introduced by parameter estimation in a finite sample. Our second evaluation method is based on the out-of-sample testing framework of Clark and McCracken (2009), which takes into account effects of data revisions on the asymptotic distribution of test statistics. We also conduct an event-timing analysis and estimate a Markov switching version of the dynamic factor model of realized volatilities outlined above and analyze the relationship between volatility and economic cycles.

4.1 In-sample Tests

We start our predictive analysis by augmenting simple autoregressive models of industrial production and payroll employment growth with a financial predictor and apply the in-sample tests of Clark and McCracken (2012) to test the significance of the corresponding predictive coefficient. The predictive model is given by,

$$g_{t+h} = \beta_0 + \beta_1 g_t + (T^{-1/2}\beta_2)x_t + u_{t+h}, \tag{6}$$

where g_{t+h} is the cumulative growth in the macro aggregate under consideration from time t to $t+h$, T is the sample size, and x_t denotes the financial predictor. The local-to-zero specification reflects the inherent trade-off between the signal provided by x_t and the noise introduced by imprecise estimation of its coefficient in a finite sample. The Clark and McCracken (2012) statistic tests the null hypothesis that these two factors offset each other and including x_t does not provide any significant gains in squared error loss. Under the alternative, signal dominates noise and the marginal predictor is useful.[6] This specification informs us whether the predictor is useful for projecting macroeconomic aggregates over the entire sample, as well as the sub-samples.

We also consider an alternative version of the test in which the conventional financial predictors are assumed to have non-zero predictive coefficients in the population and the volatility measures are treated as marginal predictors. Formally, we have

$$(7)\quad g_{t+h} = \beta_0 + \beta_1 g_t + \beta_2 x_{1,t} + (T^{-1/2}\beta_3)x_{2,t} + u_{t+h},$$

where $x_{1,t}$ is a conventional financial predictor and $x_{2,t}$ is a volatility measure. Note that this is a fairly conservative approach to assess the marginal predictive content of financial volatility measures since we are assuming that the baseline predictor has predictive power in *population*.

Table 3 reports the results for predicting industrial production growth over the full-sample. We report both the asymptotic p-values based on non-central normal distribution and the bootstrap p-values, which are obtained from a fixed-design wild bootstrap.[7] Both stock volatility measures are significant predictors of industrial production growth for the one-month horizon according to the asymptotic and bootstrap p-values. The predictive power of stock volatility measures are reflected in the common volatility factor as well. Among the conventional financial predictors, only the default spread provides useful information beyond the lagged industrial production growth for the one-month horizon. For longer forecast horizons, the stock volatility measures are typically significant according to the asymptotic p-values, but not according to the more conservative bootstrap p-values. Only the return on the market portfolio is found to be a significant predictor of industrial production growth for horizons beyond one-month. When we augment conventional predictors with the volatility measures, we find that the stock market based

[6] See the technical appendix for more details on this testing framework.

[7] We are grateful to Todd Clark and Michael McCracken for making their RATS program for the bootstrap procedure available.

volatility measures and the common factor usually provide a useful supplementary signal according to the asymptotic p-values. The bootstrapped p-values indicate significance only for the one and three-month horizons.

The in-sample test results for predicting industrial production growth are strikingly different across the two sub-samples (Tables 4 and 5). Over the first period, which runs from January 1985 to December 1996, only the term spread is significant for the one and three-month horizons. All other financial predictors, including volatility measures, do not provide any additional information over lagged growth for any of the forecast horizons. Moreover, none of the volatility measures can improve upon the term spread in instances when the latter is significant. Turning to the second sub-sample, the period from January 1997 to June 2012, provides an appreciably different perspective regarding the predictive content of conventional financial predictors and volatility measures. The results are qualitatively similar to those obtained from the full-sample, but the case for the marginal information content of stock volatility measures is strengthened.

The test results for payroll employment growth are reported in Tables 6 – 8. Over the full-sample period, stock volatility measures and the extracted factor are significant for all horizons according to the asymptotic p-values. Bootstrap p-values also support this conclusion with the exception of the 12-month horizon. Among the conventional predictors, return on the market portfolio is always significant under both testing schemes. The default spread helps improve short term forecasts while the term spread is significant only for the 12-month horizon. In terms of the marginal predictive power, the stock market measures and the common factor have the potential to improve employment growth forecasts in most instances. The sub-sample results are similar to the case of industrial production growth in the sense that no variable can systematically improve simple autoregressive forecasts in the first sub-sample. Moreover, the results from the second sub-sample closely resemble those from the full-sample.

4.2 Out-of-sample Tests in Real-time

We now turn to an out-of sample evaluation based on real-time data. This analysis allows us to assess relative and marginal predictive power of the financial volatility measures in a realistic setting. We evaluate predictive performance with respect to mean squared error (MSE) loss function using Diebold and Mariano (1995) type tests by taking into account the effects of real-rime data revisions on the distribution of test statistics. Specifically, we adopt the real-time out-of-sample testing framework of Clark and McCracken (2009) and consider two types of comparisons. [8] The first focuses on non-nested models that use either a conventional predictor or a volatility measure, while the second evaluates nested models in which volatility measures are treated as marginal predictors. Lagged growth rates of industrial production and employment are included in both cases. We evaluate real-time forecasts with respect to both *initial* and *final* data releases. The former is relevant as it is of interest to practitioners and policy makers who try to anticipate the initial data release, while the latter informs us about the ability to predict an ultimately more accurate measure of economic activity in real-time. We adopt a recursive forecasting scheme and evaluate the forecasts over the period from January 1996 to June 2012.[9]

The non-nested model comparison results for industrial production growth based on initial data release are reported in Table 9. The stock volatility measures and the common factor perform better than both the term spread and the market return over the one-month horizon, leading to reductions of up to 9% in the MSE. Moreover, they perform only slightly worse than the default spread. The bond volatility outperforms the term spread but provides a somewhat larger MSE than the default spread and the market return. The loss differences are significant in some instances when volatility measures perform better under the assumption of no noise in data revisions, but this no longer holds if one assumes that data revisions are driven by both news and noise. For the three-month horizon, all volatility measures beat the default spread substantially. However, they perform worse than the term spread and about the same as the market return on average. Assuming noise in data revisions imply that loss differences between the volatility measures and the default spread are typically significant at conventional levels. In other cases, differences are generally insignificant under both noisy and purely news driven data

[8] See the technical appendix for a detailed exposition of this testing framework.

[9] We have also implemented a rolling forecasting scheme and found that the results are qualitatively similar. Those results are available from the corresponding author upon request.

revisions. The results for the six-month horizon are broadly in line with those from the three-month horizon. Over the 12-month horizon, loss differences provide a similar depiction, but they are typically insignificant under noisy revisions. When we use the final release to evaluate forecasts that are produced using real-time data, we find that the results closely resemble those obtained using the first release (see Table 10).

Tables 11 and 12 report the non-nested model comparison results for employment growth. Overall, the results are in line with those for the industrial production growth with one exception: the term spread typically beats volatility measures for horizons beyond 3-months and the loss differences are significant regardless of the assumption one makes about possible noise in data revisions.

Nested model evaluation results for industrial production growth based on initial data release are reported in Table 13. For the one-month forecast horizon, the volatility measures usually improve on the performance of smaller models containing conventional predictors. For example, the common factor provides an 8% reduction in MSE over the model containing only lagged industrial production growth and the term spread with an associated p-value of 0.03. A similar result is obtained from the comparison with the market return. A slight improvement over the default spread is observed but the difference is marginally insignificant, with a p-value of 0.11. For longer horizons, adding the volatility measures does not improve forecasting performance and the MSE deteriorates somewhat. Evaluations based on final data release yield a similar conclusion (see Table 14).

When we add the volatility factor to predictive regressions for employment growth and use initial data to evaluate forecasts, we observe that MSE is reduced by 5% to 7% for the one-month horizon and that the gains are highly statistically significant (see Table 15). There are also improvements in MSE over the default spread and the market return for the three-month horizon. Beyond three months, the volatility measures can improve only the forecasts based on the default spread. Results are qualitatively identical when we use final data to evaluate the forecasts (see Table 16).

4.3 Event-timing Analysis with Regime Switching

The recent financial crisis and economic recession have revived widespread interest in predicting business cycle turning points rather than just focusing on point forecasts. Therefore, we estimate a nonlinear version of the dynamic factor model outlined in Section 2 and explore the relationship between volatility regimes and different states of the economy. Specifically, we allow for both the drift and the variance of the common factor to switch across different regimes according to an unobservable Markov process. Thus, the transition equation for the factor volatility (5) is now replaced with the following equation:

$$(5') \qquad VF_t = \alpha_{S_t} + \psi VF_{t-1} + \epsilon_t, \qquad \epsilon_t \sim ND(0, \tau_{S_t}^2),$$

where S_t is the state variable that governs the regimes, which evolves according to a first order Markov process with transition probabilities given by $p_{ij} = \Pr[S_t = j | S_{t-1} = i]$ where $i, j \in \{0, 1, \ldots, M-1\}$.[10] Hence, there are M different regimes characterized by different levels and variances of the volatility factor. As a result, we have $\alpha_{S_t} = \alpha_i 1(S_t = i)$ and $\tau_{S_t}^2 = \tau_i^2 1(S_t = i)$, $i = 0,1, \ldots, M-1$, where $1(.)$ represents the standard indicator function.

We consider two specifications featuring two and three states respectively. The two state model ($M = 2$) corresponds to the conventional bull/bear taxonomy of the financial markets, while the three state model ($M = 3$) introduces an additional state to provide a possibly more realistic approximation to the true underlying data generating process. In determining the optimal number of regimes, we follow Guidolin and Timmermann (2006) and rely on the Davies (1987) upper bound for the p-value of the likelihood ratio test statistic as well as the Hannan-Quinn information criterion. We find that both approaches strongly favor the three regime specification.

[10] This model is closely related to the model proposed in Chauvet (1998) for business cycle analysis.

Table 17 presents the maximum-likelihood estimates of the nonlinear common factor model of volatilities.[11] The model distinguishes between three different levels of volatility: high, intermediate, and low. The expected value of the common factor in the high-volatility regime is about twice as large as that in the low-volatility regime, while the intermediate-volatility regime is approximately midway between the two regimes in terms of the expected level of the common factor. The variance of the common factor takes its largest value in the high-volatility regime, suggesting that the *volatility of volatility* is strongly positively correlated with the level of volatility. The transition probability estimates reveal that the persistence of each regime is inversely correlated with the level of volatility in the corresponding regime. In particular, the high-volatility regime has an expected duration of about 5 months compared to 40 months for the low-volatility regime. The intermediate-volatility regime has an expected duration of 8 months. As in the case of the linear model, all volatility series are positively correlated with the common factor with statistically significant factor loadings. The idiosyncratic persistence parameters are comparable to those obtained under the linearity assumption.

The smoothed probabilities of regimes from the nonlinear dynamic factor model of volatilities are plotted in Figure 3. The low-volatility regime is strongly correlated with periods of robust economic expansion. The intermediate volatility regime is associated with either episodes of uncertainty during economic expansions or recessions. The 1990-1991 and 2001 recessions, which are usually regarded as mild recessions compared to the other post-war US recessions, as well as the prevailing uncertainty following the 2001 recession are classified as intermediate volatility regimes. The high-volatility regime typically leads the NBER recessions and periods of economic slowdowns slightly and prevails during the entire Great Recession.

According to the smoothed probabilities, the second half of 1980s was associated with low and intermediate-volatility regimes, with a brief interruption in October 1987 due to the Black Monday. The common factor enters into the high-volatility regime at the onset

[11] The models are estimated via numerical optimization of the likelihood function. In particular, we combine a nonlinear discrete version of the Kalman filter with Hamilton's (1989) filter using Kim's (1994) approximate maximum likelihood method.

of the 1990-1991 recession and then shifts to the intermediate-volatility regime for the rest of the recession. Following the end of the recession, the low-volatility regime dominates until 1997.

The period from early 1997 to mid-2003 is characterized by switches between intermediate and high volatility regimes. The Asian crisis of 1997 appears to have resulted in moderate volatility in U.S. asset markets. Starting from mid-1998, the probabilities of the high-volatility regime rise above 0.5 and remain high until the end of the first quarter of 1999. The notable event of this period was the 1998 Russian debt crisis and the LTCM default, which led to fears of a U.S. recession and transitory weakening of the economic conditions. The probabilities of the high-volatility regime rise again in early 2000 as the tech-bubble burst and remain elevated until the beginning of the 2001 recession. The smoothed regime probabilities classify the second half of 2002 as a high-volatility period. Over that time frame, U.S. economic activity slowed down amid corporate accounting scandals, declining confidence levels among households and businesses, and falling stock prices. A double-dip recession was averted according to the NBER Business Cycle Dating Committee, but the committee's announcement that the recession had ended in 2001 was not made until mid-2003. Interestingly, this timing coincides with the end of the high-volatility regime. The low-volatility regime persisted for four years beginning from mid-2003.

There is a steep increase in the probabilities of the high-volatility regime in the summer of 2007, when first signs of distress in the financial markets due to housing market problems made headlines. The probabilities remain elevated until June 2009, the end of the Great Recession according to the NBER. The sluggish recovery following the crisis is mostly captured by the intermediate volatility regime, with a brief shift to a high volatility regime from August to November 2011 as the economy slowed down. The U.S. government debt ceiling crisis and the increased stress in short-term U.S. Dollar funding markets contributed to elevated uncertainty during this time. The decline in economic activity and increased uncertainty were quite concerning that it led the Federal Reserve to announce that the federal funds rate would stay exceptionally low at least two more years, a first in its history.

As a robustness check, we also estimate the model recursively and calculate regime probabilities as one would do in real-time. We first estimate the model using data up to January 1997 and calculate regime probabilities for that month. We then expand the sample each subsequent month, re-estimate the model and calculate the corresponding regime probabilities for the remaining part of the sample. Figure 4 plots the recursive filtered probabilities of volatility regimes. The real-time performance of the dynamic factor model of volatilities is strikingly similar to that in the full-sample summarized above.

5. Concluding Remarks

We construct monthly measures of financial volatility from daily returns on market and industry portfolios as well as Treasury bonds and analyze the predictive value of such volatility measures for economic activity using both real time and revised data.

We model log realized volatility as composed of a long-run component that is common across all measures and transitory idiosyncratic components in a dynamic factor framework. We find that the stock market volatility measures as well as the common volatility factor help predict growth in industrial production and employment according to the in-sample tests of Clark and McCracken (2012). Moreover, these volatility measures improve forecasts from conventional financial indicators such as the term spread, the credit spread and the return on a broadly defined stock portfolio, especially over short forecast horizons. Our out-of-sample analysis with real-time data, which takes into account effects of data revisions, also shows that the stock volatility measures and the common factor significantly improve short-term forecasts from conventional financial indicators.

By estimating a nonlinear version of the dynamic factor model, we identify three distinct states of volatility dynamics. We find that economic expansions are usually characterized by the low-volatility regime. The intermediate-volatility regime characterizes episodes of increased uncertainty during economic expansions and mild recessions, whereas the high-volatility regime slightly leads the NBER recessions economic slowdowns. This model performs remarkably well in tracking the Great Recession in real-time.

References

Adrian, T. and J. Rosenberg, 2008, "Stock Returns and Volatility: Pricing the Long-Run and Short-Run Components of Market Risk," *The Journal of Finance*, 63, 2997-3030.

Allen, L., T. Bali, and Y. Tang, 2012, "Does Systemic Risk in the Financial Sector Predict Future Economic Downturns?," *Review of Financial Studies*, 25, 3000-3036.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens, 2001a, "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys, 2001b, "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42-55.

Andersen, T. G., T. Bollerslev, F.X. Diebold and P. Labys, 2003, "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579-625.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and J. Wu, 2005, "A Framework for Exploring the Macroeconomic Determinants of Systematic Risk," *American Economic Review Papers and Proceedings*, 95, 398-404.

Ang, A., M. Piazzesi, and M. Wei, 2006, "What does the Yield Curve Tell us about GDP Growth," *Journal of Econometrics*, 131, 1163-1212.

Arellano, C., Y. Bai, and P. Kehoe, 2012, Financial Markets and Fluctuations in Uncertainty, Federal Reserve Bank of Minneapolis Research Department Staff Report.

Bakshi, G., G. Panayotov, and G. Skoulakis, 2011, "Improving the Predictability of Real Economic Activity and Asset Returns with Forward Variances Inferred from Option Portfolios," *Journal of Financial Economics*, 100, 475-495.

Bloom, N., 2009, "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623-685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksen, and Stephen J. Terry, 2012, "Really Uncertain Business Cycles," Mimeo, Stanford.

Campbell., J., M. Lettau, B.G. Malkiel, and Y. Xu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *The Journal of Finance*, 56, 1-44.

Chauvet, M. 1998, "An Econometric Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching," *International Economic Review* 39, 969-996.

Christiano, L., R. Motto, and M. Rostagno, 2010, "Financial Factors in Economic Fluctuations," ECB Working Paper No. 1192.

Clark, T.E. and M.W. McCracken, 2012, "In-sample Tests of Predictive Ability: A new Approach," *Journal of Econometrics*, 170, 1-14.

Clark, T.E. and M.W. McCracken, 2009, "Tests of Equal Predictive Ability with Real-Time Data," *Journal of Business and Economic Statistics*, 27, 441-454.

Davies, R.B., 1987, Hypothesis Testing when a Nuisance Parameter is Present Only under the Alternatives, *Biometrika*, 74, 33-43.

Diebold, F.X. and R.S. Mariano, 1995, "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253-265.

Estrella, A. and F.S. Mishkin, 1998, "Predicting U.S. recessions: Financial Variables as Leading Indicators," *The Review of Economics and Statistics*, 80, 45-61.

Fama, E.F., 1990, "Stock Returns, Expected Returns, and Real Activity," *Journal of Finance*, 45, 1089-1108.

Fornari, F. and A. Mele, 2009, "Financial Volatility and Economic Activity," *mimeo*, London School of Economics.

French, K.R., G.W. Schwert, and R.F. Stambaugh, 1987, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3-29.

Gilchrist, S. and E. Zakrajsek, 2012, "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102-4, 1692-1720.

Guidolin, M. and A. Timmermann, 2006, An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns, *Journal of Applied Econometrics*, 21, 1-22.

Gurkaynak, R.S., B. Sack, and J.H. Wright, 2007, "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics,* 24, 2291-2304.

Hamilton, J.D., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and Business Cycles," *Econometrica*, 57, 357-384.

Kim, C.J., 1994, "Dynamic Linear Models with Markov-Switching," *Journal of Econometrics*, 60, 1-22.

Mele, A., 2007, "Asymmetric Stock Market Volatility and the Cyclical Behavior of Expected Returns," *Journal of Financial Economics*, 86, 446–478.

Newey, W.K. and K.D. West, 1987, "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* , 55, 703-708.

McConnell, M.M. and G. Perez-Quiros, 2000, "Output Fluctuations in the United States: What Has Changed since the Early 1980's?," *American Economic Review*, 90, 1464-1476.

Poterba, J.M. and L. Summers, 1986, "The Persistence of Volatility and Stock Market Fluctuations," *The American Economic Review*, 76, 1142-1151.

Schwert, G.W., 1989a, "Business Cycles, Financial Crises and Stock Volatility," *Carnegie-Rochester Conference Series on Public Policy*, 31, 83-125.

Schwert, G.W., 1989b, "Why Does Stock Market Volatility Change Over Time?," *The Journal of Finance*, 44, 1115-1153.

West, K.D., 1996, "Asymptotic Inference about Predictive Ability," *Econometrica*, 64, 1067-1084.

Tables

Table 1: Correlations between Financial Volatility Measures
and Macroeconomic Aggregates

Panel A Industrial Production Growth			Panel B Employment Growth		
RVM	RVI	RVB	RVM	RVI	RVB
Full-sample			Full-sample		
-0.220	-0.284	-0.126	-0.440	-0.414	-0.202
(0.08)	(0.01)	(0.25)	(0.00)	(0.01)	(0.16)
Sub-sample I			Sub-sample I		
-0.029	0.005	0.042	-0.112	-0.169	0.227
0.83	0.97	(0.67)	(0.53)	(0.29)	(0.03)
Sub-sample II			Sub-sample II		
-0.281	-0.360	-0.225	-0.464	-0.364	-0.490
(0.04)	(0.00)	(0.10)	(0.00)	(0.04)	(0.00)

Notes: Simple correlation estimates are reported. The full sample runs from January 1985 to June 2012, the first sub-sample is from January 1985 to December 1996, and the second sub-sample is from January 1997 to June 2012. Asymptotic p-values in parentheses are based on Newey-West (1987) HAC standard deviations.

Table 2: Parameter Estimates for the Dynamic Factor Model of Volatilities

Parameter	Full-sample	Sub-sample I	Sub-sample II
α	1.102	1.859	1.445
	(0.00)	(0.00)	(0.00)
ψ	0.790	0.544	0.783
	(0.00)	(0.00)	(0.00)
λ_1	0.262	0.274	0.237
	(0.00)	(0.00)	(0.00)
λ_2	0.194	0.182	0.195
	(0.00)	(0.00)	(0.00)
λ_3	0.059	0.082	0.045
	(0.00)	(0.00)	(0.00)
ϕ_1	0.194	-0.031	0.410
	(0.04)	(0.86)	(0.00)
ϕ_2	0.982	0.990	0.972
	(0.00)	(0.00)	(0.00)
ϕ_3	0.537	0.459	0.575
	(0.00)	(0.00)	(0.00)
σ_1	0.041	0.039	0.041
	(0.00)	(0.00)	(0.00)
σ_2	0.004	0.003	0.004
	(0.01)	(0.07)	(0.09)
σ_3	0.064	0.067	0.059
	(0.00)	(0.00)	(0.00)

Notes: The full sample runs from January 1985 to June 2012, the first sub-sample is from January 1985 to December 1996, and the second sub-sample is from January 1997 to June 2012. Asymptotic p-values in parentheses are based on the standard deviations based on the inverse hessian, which is obtained through numerical calculation. The variance of the common factor is set to unity for identification (i.e. $\tau = 1$ in the model stated in Section 3).

Table 3: In-sample Tests for Industrial Production Growth
(January 1985 – June 2012)

Baseline Predictors	Tested Predictor	Forecast Horizon							
		1-month		3-month		6-month		12-month	
		p_N	p_B	p_N	p_B	p_N	p_B	p_N	p_B
AR	TERM	0.52	0.53	0.42	0.45	0.38	0.39	0.18	0.23
AR	DEF	0.00	0.00	0.06	0.27	0.21	0.48	0.51	0.56
AR	MKTR	0.25	0.31	0.02	0.09	0.02	0.15	0.01	0.08
AR	RVM	0.02	0.03	0.05	0.15	0.04	0.17	0.05	0.14
AR	RVI	0.00	0.00	0.05	0.17	0.10	0.30	0.13	0.26
AR	RVB	0.37	0.38	0.36	0.47	0.57	0.58	0.78	0.67
AR	VF	0.00	0.01	0.03	0.11	0.05	0.19	0.06	0.15
AR, TERM	RVM	0.02	0.03	0.05	0.14	0.04	0.17	0.04	0.14
AR, TERM	RVI	0.00	0.01	0.07	0.25	0.12	0.37	0.14	0.27
AR, TERM	RVB	0.18	0.20	0.13	0.21	0.24	0.34	0.42	0.45
AR, TERM	VF	0.00	0.01	0.03	0.13	0.04	0.19	0.06	0.21
AR, DEF	RVM	0.58	0.61	0.10	0.15	0.08	0.16	0.09	0.19
AR, DEF	RVI	0.01	0.01	0.05	0.18	0.11	0.34	0.18	0.37
AR, DEF	RVB	0.99	1.00	0.96	0.93	0.98	0.94	0.97	0.89
AR, DEF	VF	0.32	0.35	0.05	0.10	0.07	0.18	0.10	0.27
AR, MKTR	RVM	0.04	0.04	0.09	0.14	0.07	0.15	0.15	0.25
AR, MKTR	RVI	0.01	0.01	0.06	0.16	0.12	0.27	0.24	0.32
AR, MKTR	RVB	0.44	0.46	0.43	0.53	0.69	0.68	0.87	0.74
AR, MKTR	VF	0.01	0.01	0.03	0.10	0.06	0.14	0.11	0.23

Notes: AR stands for the autoregressive term, TERM is the term spread, DEF is the credit spread, MKTR is the return on the market portfolio, and p_N (p_B) is the asymptotic (bootstrap) p-value of the Clark and McCracken (2012) test. The p-values are associated with the tested predictor while baseline predictors are assumed to have nonzero coefficients in the population.

Table 4: In-sample Tests for Industrial Production Growth
(January 1985 – December 1996)

Baseline Predictors	Tested Predictor	Forecast Horizon							
		1-month		3-month		6-month		12-month	
		p_N	p_B	p_N	p_B	p_N	p_B	p_N	p_B
AR	TERM	0.09	0.10	0.06	0.08	0.22	0.31	0.49	0.58
AR	DEF	0.16	0.17	0.43	0.46	0.58	0.60	0.66	0.69
AR	MKTR	0.96	0.96	0.56	0.65	0.60	0.82	0.18	0.31
AR	RVM	0.81	0.80	0.59	0.69	0.57	0.70	0.55	0.59
AR	RVI	0.94	0.94	0.90	0.90	0.98	0.96	1.00	1.00
AR	RVB	0.96	0.96	0.86	0.82	0.98	0.94	0.84	0.70
AR	VF	0.51	0.55	0.45	0.68	0.54	0.73	0.55	0.64
AR, TERM	RVM	0.76	0.76	0.52	0.64	0.53	0.65	0.53	0.59
AR, TERM	RVI	0.85	0.86	0.82	0.85	0.95	0.94	1.00	0.99
AR, TERM	RVB	0.95	0.95	0.82	0.76	0.97	0.92	0.72	0.57
AR, TERM	VF	0.44	0.49	0.37	0.64	0.50	0.71	0.55	0.64
AR, DEF	RVM	0.99	0.99	0.78	0.78	0.69	0.76	0.61	0.58
AR, DEF	RVI	0.92	0.90	0.88	0.90	0.97	0.93	1.00	1.00
AR, DEF	RVB	0.99	0.99	0.96	0.91	1.00	0.99	0.96	0.85
AR, DEF	VF	0.90	0.91	0.60	0.71	0.63	0.77	0.60	0.61
AR, MKTR	RVM	0.69	0.73	0.67	0.73	0.64	0.70	0.68	0.73
AR, MKTR	RVI	0.91	0.91	0.94	0.94	1.00	0.98	1.00	1.00
AR, MKTR	RVB	0.95	0.95	0.90	0.89	0.99	0.97	0.88	0.77
AR, MKTR	VF	0.37	0.41	0.48	0.68	0.58	0.72	0.65	0.76

Notes: AR stands for the autoregressive term, TERM is the term spread, DEF is the credit spread, MKTR is the return on the market portfolio, and p_N (p_B) is the asymptotic (bootstrap) p-value of the Clark and McCracken (2012) test. The p-values are associated with the tested predictor while baseline predictors are assumed to have nonzero coefficients in the population.

Table 5: In-sample Tests for Industrial Production Growth
(January 1997 – June 2012)

Baseline Predictors	Tested Predictor	Forecast Horizon							
		1-month		3-month		6-month		12-month	
		p_N	p_B	p_N	p_B	p_N	p_B	p_N	p_B
AR	TERM	0.80	0.82	0.70	0.72	0.59	0.59	0.25	0.30
AR	DEF	0.00	0.00	0.15	0.38	0.38	0.62	0.70	0.62
AR	MKTR	0.20	0.26	0.02	0.07	0.01	0.10	0.00	0.04
AR	RVM	0.03	0.06	0.06	0.16	0.03	0.08	0.01	0.04
AR	RVI	0.00	0.00	0.03	0.09	0.02	0.08	0.01	0.05
AR	RVB	0.25	0.29	0.34	0.47	0.50	0.54	0.69	0.65
AR	VF	0.01	0.02	0.05	0.17	0.05	0.16	0.02	0.10
AR, TERM	RVM	0.02	0.04	0.05	0.14	0.02	0.07	0.00	0.04
AR, TERM	RVI	0.01	0.01	0.04	0.14	0.03	0.14	0.00	0.04
AR, TERM	RVB	0.11	0.14	0.09	0.21	0.11	0.20	0.26	0.47
AR, TERM	VF	0.01	0.01	0.04	0.13	0.03	0.15	0.02	0.09
AR, DEF	RVM	0.36	0.37	0.07	0.12	0.06	0.09	0.03	0.07
AR, DEF	RVI	0.01	0.01	0.01	0.06	0.02	0.08	0.03	0.11
AR, DEF	RVB	0.95	0.96	0.88	0.83	0.85	0.75	0.76	0.72
AR, DEF	VF	0.22	0.23	0.05	0.09	0.06	0.15	0.07	0.19
AR, MKTR	RVM	0.06	0.09	0.09	0.17	0.04	0.08	0.18	0.23
AR, MKTR	RVI	0.01	0.01	0.01	0.06	0.01	0.06	0.07	0.14
AR, MKTR	RVB	0.26	0.28	0.29	0.41	0.50	0.52	0.74	0.67
AR, MKTR	VF	0.02	0.02	0.03	0.08	0.04	0.12	0.06	0.15

Notes: AR stands for the autoregressive term, TERM is the term spread, DEF is the credit spread, MKTR is the return on the market portfolio, and p_N (p_B) is the asymptotic (bootstrap) p-value of the Clark and McCracken (2012) test. The p-values are associated with the tested predictor while baseline predictors are assumed to have nonzero coefficients in the population.

Table 6: In-sample Tests for Employment Growth
(January 1985 – June 2012)

Baseline Predictors	Tested Predictor	Forecast Horizon							
		1-month		3-month		6-month		12-month	
		p_N	p_B	p_N	p_B	p_N	p_B	p_N	p_B
AR	TERM	0.90	0.88	0.55	0.47	0.19	0.20	0.04	0.06
AR	DEF	0.00	0.00	0.01	0.05	0.10	0.23	0.58	0.54
AR	MKTR	0.02	0.02	0.01	0.02	0.00	0.05	0.00	0.03
AR	RVM	0.01	0.01	0.01	0.04	0.01	0.05	0.03	0.12
AR	RVI	0.00	0.00	0.01	0.04	0.02	0.11	0.04	0.16
AR	RVB	0.22	0.23	0.40	0.43	0.63	0.62	0.90	0.80
AR	VF	0.00	0.00	0.00	0.01	0.01	0.05	0.03	0.10
AR, TERM	RVM	0.01	0.01	0.01	0.03	0.01	0.06	0.02	0.08
AR, TERM	RVI	0.00	0.00	0.01	0.06	0.03	0.16	0.06	0.20
AR, TERM	RVB	0.22	0.25	0.28	0.35	0.33	0.36	0.47	0.47
AR, TERM	VF	0.00	0.00	0.00	0.01	0.01	0.07	0.03	0.13
AR, DEF	RVM	0.11	0.12	0.06	0.10	0.04	0.12	0.05	0.17
AR, DEF	RVI	0.00	0.00	0.01	0.04	0.02	0.11	0.04	0.17
AR, DEF	RVB	0.91	0.92	0.98	0.96	0.99	0.96	0.99	0.96
AR, DEF	VF	0.02	0.03	0.02	0.05	0.03	0.11	0.05	0.17
AR, MKTR	RVM	0.05	0.06	0.04	0.06	0.05	0.10	0.11	0.22
AR, MKTR	RVI	0.01	0.01	0.01	0.04	0.02	0.09	0.06	0.19
AR, MKTR	RVB	0.30	0.32	0.51	0.56	0.78	0.75	0.94	0.90
AR, MKTR	VF	0.00	0.00	0.00	0.01	0.02	0.07	0.06	0.16

Notes: AR stands for the autoregressive term, TERM is the term spread, DEF is the credit spread, MKTR is the return on the market portfolio, and p_N (p_B) is the asymptotic (bootstrap) p-value of the Clark and McCracken (2012) test. The p-values are associated with the tested predictor while baseline predictors are assumed to have nonzero coefficients in the population.

Table 7: In-sample Tests for Employment Growth
(January 1985 – December 1996)

Baseline Predictors	Tested Predictor	Forecast Horizon							
		1-month		3-month		6-month		12-month	
		p_N	p_B	p_N	p_B	p_N	p_B	p_N	p_B
AR	TERM	0.49	0.53	0.29	0.36	0.24	0.37	0.17	0.34
AR	DEF	0.54	0.54	0.53	0.56	0.68	0.64	0.84	0.77
AR	MKTR	0.46	0.46	0.20	0.23	0.10	0.13	0.07	0.15
AR	RVM	0.75	0.77	0.64	0.61	0.64	0.67	0.67	0.66
AR	RVI	0.34	0.36	0.70	0.70	0.92	0.85	1.00	0.95
AR	RVB	0.94	0.94	0.94	0.91	0.98	0.94	0.98	0.93
AR	VF	0.31	0.38	0.29	0.40	0.45	0.59	0.56	0.61
AR, TERM	RVM	0.74	0.75	0.60	0.53	0.56	0.53	0.57	0.55
AR, TERM	RVI	0.27	0.33	0.58	0.64	0.82	0.79	0.95	0.89
AR, TERM	RVB	0.94	0.93	0.93	0.87	0.97	0.91	0.95	0.88
AR, TERM	VF	0.30	0.36	0.22	0.33	0.35	0.47	0.48	0.49
AR, DEF	RVM	0.87	0.87	0.82	0.71	0.70	0.64	0.59	0.60
AR, DEF	RVI	0.29	0.34	0.65	0.68	0.91	0.82	1.00	0.96
AR, DEF	RVB	0.97	0.96	0.98	0.95	1.00	0.97	0.99	0.97
AR, DEF	VF	0.43	0.46	0.38	0.46	0.43	0.57	0.43	0.53
AR, MKTR	RVM	0.85	0.85	0.84	0.78	0.83	0.80	0.82	0.78
AR, MKTR	RVI	0.47	0.49	0.82	0.84	0.97	0.94	1.00	0.99
AR, MKTR	RVB	0.95	0.95	0.97	0.93	0.99	0.96	0.98	0.96
AR, MKTR	VF	0.46	0.53	0.49	0.58	0.61	0.72	0.69	0.73

Notes: AR stands for the autoregressive term, TERM is the term spread, DEF is the credit spread, MKTR is the return on the market portfolio, and p_N (p_B) is the asymptotic (bootstrap) p-value of the Clark and McCracken (2012) test. The p-values are associated with the tested predictor while baseline predictors are assumed to have nonzero coefficients in the population.

Table 8: In-sample Tests for Employment Growth
(January 1997 – June 2012)

Baseline Predictors	Tested Predictor	Forecast Horizon							
		1-month		3-month		6-month		12-month	
		p_N	p_B	p_N	p_B	p_N	p_B	p_N	p_B
AR	TERM	0.94	0.94	0.81	0.71	0.52	0.43	0.14	0.15
AR	DEF	0.00	0.01	0.03	0.18	0.10	0.35	0.37	0.44
AR	MKTR	0.02	0.04	0.00	0.02	0.01	0.07	0.00	0.03
AR	RVM	0.03	0.04	0.04	0.12	0.03	0.07	0.02	0.07
AR	RVI	0.03	0.04	0.01	0.04	0.01	0.07	0.02	0.11
AR	RVB	0.09	0.12	0.27	0.33	0.43	0.48	0.62	0.49
AR	VF	0.01	0.02	0.03	0.09	0.03	0.17	0.05	0.19
AR, TERM	RVM	0.03	0.05	0.05	0.10	0.02	0.07	0.01	0.04
AR, TERM	RVI	0.01	0.02	0.02	0.07	0.02	0.11	0.03	0.11
AR, TERM	RVB	0.07	0.10	0.21	0.31	0.25	0.36	0.29	0.42
AR, TERM	VF	0.01	0.01	0.03	0.10	0.03	0.16	0.03	0.15
AR, DEF	RVM	0.10	0.12	0.11	0.14	0.11	0.17	0.09	0.21
AR, DEF	RVI	0.03	0.05	0.02	0.07	0.02	0.09	0.03	0.12
AR, DEF	RVB	0.46	0.47	0.76	0.77	0.85	0.83	0.81	0.68
AR, DEF	VF	0.11	0.11	0.11	0.18	0.12	0.28	0.12	0.30
AR, MKTR	RVM	0.09	0.11	0.12	0.15	0.06	0.07	0.10	0.19
AR, MKTR	RVI	0.08	0.11	0.01	0.03	0.01	0.04	0.02	0.10
AR, MKTR	RVB	0.08	0.10	0.26	0.35	0.42	0.46	0.58	0.50
AR, MKTR	VF	0.02	0.03	0.03	0.06	0.03	0.11	0.06	0.17

Notes: AR stands for the autoregressive term, TERM is the term spread, DEF is the credit spread, MKTR is the return on the market portfolio, and p_N (p_B) is the asymptotic (bootstrap) p-value of the Clark and McCracken (2012) test. The p-values are associated with the tested predictor while baseline predictors are assumed to have nonzero coefficients in the population.

Table 9: Out-of-sample Tests for Industrial Production Growth: Non-nested Model Evaluation based on Initial Data Release

	Volatility Measure											
	RVM			RVI			RVB			VF		
Competing Predictor	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2
	1-month horizon											
TERM	0.94	0.31	0.13	0.93	0.38	0.09	0.98	0.37	0.22	0.91	0.21	0.08
DEF	1.03	0.66	0.65	1.03	0.75	0.72	1.07	0.38	0.43	1.01	0.93	0.92
MKTR	0.98	0.60	0.34	0.97	0.66	0.21	1.02	0.64	0.31	0.96	0.32	0.12
	3-month horizon											
TERM	1.07	0.53	0.51	1.18	0.34	0.12	1.08	0.29	0.22	1.07	0.61	0.59
DEF	0.63	0.02	0.13	0.69	0.22	0.24	0.63	0.01	0.18	0.63	0.01	0.11
MKTR	1.00	0.99	0.99	1.10	0.59	0.11	1.01	0.88	0.84	1.00	0.98	0.98
	6-month horizon											
TERM	1.17	0.31	0.33	1.45	0.09	0.02	1.09	0.32	0.31	1.21	0.32	0.28
DEF	0.67	0.07	0.14	0.83	0.54	0.53	0.62	0.04	0.19	0.69	0.07	0.14
MKTR	0.97	0.75	0.65	1.21	0.41	0.06	0.91	0.27	0.10	1.00	0.98	0.97
	12-month horizon											
TERM	1.15	0.48	0.42	1.63	0.10	0.07	1.08	0.56	0.52	1.20	0.45	0.32
DEF	0.83	0.47	0.09	1.18	0.59	0.40	0.78	0.23	0.08	0.86	0.59	0.17
MKTR	0.99	0.93	0.86	1.41	0.20	0.07	0.93	0.39	0.12	1.03	0.83	0.67

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the volatility based model to that from the competing model. p_1 (p_2) is the asymptotic p-value associated with the Clark and McCracken (2009) test for non-nested models under the assumption of noisy (purely news driven) data revisions.

Table 10: Out-of-sample Tests for Industrial Production Growth: Non-nested Model Evaluation based on Final Data Release

	Volatility Measure											
	RVM			RVI			RVB			VF		
Competing Predictor	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2
	1-month horizon											
TERM	0.95	0.35	0.18	0.94	0.37	0.11	0.98	0.33	0.23	0.93	0.28	0.13
DEF	1.02	0.73	0.74	1.01	0.90	0.90	1.05	0.49	0.53	1.00	0.99	0.99
MKTR	1.00	0.90	0.82	0.98	0.76	0.42	1.03	0.54	0.16	0.98	0.58	0.37
	3-month horizon											
TERM	1.07	0.54	0.54	1.16	0.38	0.17	1.08	0.27	0.23	1.07	0.61	0.62
DEF	0.64	0.02	0.14	0.69	0.18	0.24	0.64	0.01	0.19	0.64	0.01	0.11
MKTR	1.00	0.98	0.98	1.09	0.61	0.14	1.01	0.89	0.86	1.00	0.97	0.97
	6-month horizon											
TERM	1.18	0.28	0.31	1.45	0.09	0.02	1.10	0.28	0.28	1.22	0.28	0.26
DEF	0.67	0.08	0.14	0.83	0.52	0.52	0.63	0.04	0.19	0.70	0.08	0.15
MKTR	0.97	0.77	0.69	1.20	0.41	0.07	0.91	0.25	0.09	1.01	0.92	0.90
	12-month horizon											
TERM	1.21	0.34	0.25	1.70	0.08	0.05	1.12	0.40	0.32	1.26	0.35	0.20
DEF	0.83	0.49	0.11	1.17	0.61	0.42	0.77	0.22	0.08	0.87	0.61	0.20
MKTR	1.00	0.98	0.96	1.40	0.21	0.07	0.92	0.32	0.10	1.04	0.81	0.64

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the volatility based model to that from the competing model. p_1 (p_2) is the asymptotic p-value associated with the Clark and McCracken (2009) test for non-nested models under the assumption of noisy (purely news driven) data revisions.

Table 11: Out-of-sample Tests for Employment Growth: Non-nested Model Evaluation based on Initial Data Release

	Volatility Measure											
	RVM			RVI			RVB			VF		
Competing Predictor	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2
	1-month horizon											
TERM	0.94	0.16	0.03	0.94	0.14	0.01	0.99	0.51	0.41	0.93	0.10	0.03
DEF	1.00	0.96	0.94	1.01	0.87	0.81	1.06	0.09	0.07	1.00	0.92	0.89
MKTR	0.97	0.45	0.12	0.98	0.61	0.28	1.03	0.42	0.07	0.97	0.41	0.16
	3-month horizon											
TERM	1.08	0.18	0.07	1.06	0.44	0.07	1.08	0.03	0.00	1.06	0.22	0.27
DEF	0.84	0.16	0.07	0.83	0.13	0.07	0.85	0.22	0.12	0.83	0.10	0.04
MKTR	1.00	0.97	0.96	0.98	0.83	0.60	1.01	0.77	0.74	0.99	0.70	0.75
	6-month horizon											
TERM	1.14	0.07	0.04	1.18	0.10	0.00	1.10	0.08	0.01	1.15	0.04	0.07
DEF	0.88	0.26	0.13	0.91	0.44	0.31	0.85	0.28	0.14	0.89	0.19	0.11
MKTR	1.02	0.63	0.57	1.05	0.54	0.05	0.99	0.71	0.57	1.03	0.52	0.63
	12-month horizon											
TERM	1.22	0.08	0.03	1.33	0.04	0.01	1.14	0.20	0.06	1.24	0.05	0.03
DEF	0.96	0.74	0.47	1.05	0.55	0.40	0.90	0.06	0.15	0.98	0.89	0.76
MKTR	1.03	0.63	0.51	1.12	0.27	0.00	0.96	0.43	0.12	1.05	0.23	0.43

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the volatility based model to that from the competing model. p_1 (p_2) is the asymptotic p-value associated with the Clark and McCracken (2009) test for non-nested models under the assumption of noisy (purely news driven) data revisions.

Table 12: Out-of-sample Tests for Employment Growth: Non-nested Model Evaluation based on Final Data Release

	Volatility Measure											
	RVM			RVI			RVB			VF		
Competing Predictor	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2	MSE Ratio	p_1	p_2
	1-month horizon											
TERM	0.92	0.09	0.02	0.92	0.06	0.00	0.98	0.34	0.15	0.91	0.07	0.02
DEF	1.02	0.61	0.46	1.02	0.74	0.62	1.10	0.06	0.04	1.01	0.78	0.69
MKTR	0.97	0.46	0.15	0.97	0.47	0.13	1.04	0.32	0.04	0.96	0.39	0.15
	3-month horizon											
TERM	1.09	0.13	0.04	1.06	0.47	0.09	1.09	0.03	0.00	1.07	0.19	0.23
DEF	0.84	0.18	0.09	0.82	0.13	0.07	0.84	0.22	0.12	0.83	0.11	0.04
MKTR	1.01	0.75	0.69	0.99	0.85	0.66	1.01	0.64	0.60	0.99	0.90	0.92
	6-month horizon											
TERM	1.14	0.08	0.04	1.18	0.11	0.00	1.11	0.07	0.01	1.15	0.04	0.07
DEF	0.88	0.25	0.13	0.91	0.43	0.30	0.85	0.28	0.14	0.89	0.19	0.11
MKTR	1.02	0.64	0.59	1.05	0.55	0.07	0.99	0.74	0.61	1.03	0.53	0.64
	12-month horizon											
TERM	1.25	0.06	0.02	1.36	0.03	0.01	1.16	0.18	0.04	1.28	0.03	0.02
DEF	0.97	0.78	0.55	1.05	0.53	0.39	0.90	0.06	0.15	0.99	0.94	0.86
MKTR	1.03	0.56	0.45	1.12	0.26	0.00	0.96	0.37	0.06	1.05	0.12	0.40

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the volatility based model to that from the competing model. p_1 (p_2) is the asymptotic p-value associated with the Clark and McCracken (2009) test for non-nested models under the assumption of noisy (purely news driven) data revisions.

Table 13: Out-of-sample Tests for Industrial Production Growth: Nested Model Evaluation based on Initial Data Release

	Volatility Measure							
	RVM		RVI		RVB		VF	
Baseline Predictor	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value
				1-month horizon				
TERM	0.95	0.10	0.95	0.18	0.97	0.07	0.92	0.03
DEF	0.99	0.36	0.96	0.08	1.00	0.57	0.98	0.11
MKTR	0.97	0.09	0.96	0.18	0.99	0.22	0.95	0.01
				3-month horizon				
TERM	1.00	0.15	1.12	0.84	1.03	0.81	1.01	0.64
DEF	1.00	0.55	1.10	1.00	0.99	0.03	0.98	0.01
MKTR	1.01	1.00	1.12	0.92	1.01	0.63	1.01	0.64
				6-month horizon				
TERM	1.06	0.90	1.32	1.00	1.03	0.82	1.10	0.91
DEF	1.01	0.89	1.24	1.00	0.99	0.01	1.02	1.00
MKTR	1.03	0.99	1.30	1.00	1.00	0.45	1.07	0.94
				12-month horizon				
TERM	1.00	0.49	1.43	1.00	1.01	0.59	1.05	0.67
DEF	1.04	1.00	1.41	1.00	0.96	0.00	1.07	0.96
MKTR	1.04	0.93	1.51	1.00	0.98	0.26	1.08	0.84

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the unrestricted model including a volatility measure to that from the restricted model. The reported p-value is associated with the Clark and McCracken (2009) test for nested models under the assumption noisy data revisions.

Table 14: Out-of-sample Tests for Industrial Production Growth: Nested Model Evaluation based on Final Data Release

	Volatility Measure							
	RVM		RVI		RVB		VF	
Baseline Predictor	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value
				1-month horizon				
TERM	0.96	0.12	0.96	0.19	0.98	0.07	0.94	0.05
DEF	0.99	0.35	0.96	0.06	1.00	0.42	0.98	0.12
MKTR	0.98	0.16	0.97	0.19	0.99	0.28	0.96	0.04
				3-month horizon				
TERM	1.01	0.74	1.11	0.83	1.03	0.88	1.01	1.00
DEF	1.00	0.49	1.09	1.00	1.00	0.22	0.98	0.03
MKTR	1.02	1.00	1.11	0.92	1.02	0.70	1.02	0.87
				6-month horizon				
TERM	1.06	0.92	1.31	1.00	1.02	0.81	1.11	0.94
DEF	1.01	0.94	1.24	1.00	0.99	0.00	1.02	1.00
MKTR	1.04	0.99	1.30	1.00	1.00	0.47	1.07	0.97
				12-month horizon				
TERM	1.03	0.71	1.46	1.00	1.00	0.53	1.07	0.73
DEF	1.05	1.00	1.41	1.00	0.96	0.00	1.08	0.96
MKTR	1.04	0.97	1.51	1.00	0.98	0.26	1.08	0.86

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the unrestricted model including a volatility measure to that from the restricted model. The reported p-value is associated with the Clark and McCracken (2009) test for nested models under the assumption noisy data revisions.

Table 15: Out-of-sample Tests for Employment Growth: Nested Model Evaluation based on Initial Data Release

	Volatility Measure							
	RVM		RVI		RVB		VF	
Baseline Predictor	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value
	1-month horizon							
TERM	0.94	0.01	0.95	0.02	0.99	0.17	0.93	0.00
DEF	0.96	0.03	0.94	0.00	1.00	0.04	0.96	0.00
MKTR	0.96	0.01	0.96	0.01	0.99	0.11	0.95	0.00
	3-month horizon							
TERM	1.01	0.92	1.00	0.49	1.02	0.94	1.00	0.49
DEF	0.97	0.05	1.00	0.38	0.99	0.00	0.94	0.00
MKTR	1.00	0.64	0.99	0.35	1.00	0.48	0.98	0.00
	6-month horizon							
TERM	1.04	1.00	1.08	1.00	1.03	0.90	1.05	0.99
DEF	0.98	0.09	1.07	1.00	0.99	0.00	0.97	0.00
MKTR	1.03	1.00	1.07	0.99	0.99	0.29	1.03	0.99
	12-month horizon							
TERM	1.05	1.00	1.16	1.00	1.03	0.92	1.07	0.95
DEF	0.99	0.22	1.14	1.00	0.99	0.01	1.00	0.41
MKTR	1.03	1.00	1.15	1.00	0.97	0.10	1.05	0.95

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the unrestricted model including a volatility measure to that from the restricted model. The reported p-value is associated with the Clark and McCracken (2009) test for nested models under the assumption noisy data revisions.

Table 16: Out-of-sample Tests for Employment Growth: Nested Model Evaluation based on Final Data Release

	Volatility Measure							
	RVM		RVI		RVB		VF	
Baseline Predictor	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value	MSE Ratio	p-value
	1-month horizon							
TERM	0.93	0.00	0.93	0.00	0.99	0.23	0.92	0.00
DEF	0.96	0.01	0.92	0.00	1.00	0.09	0.95	0.00
MKTR	0.96	0.01	0.95	0.00	1.00	0.33	0.95	0.00
	3-month horizon							
TERM	1.02	1.00	1.00	0.50	1.02	0.95	1.01	0.82
DEF	0.97	0.05	0.99	0.36	0.99	0.00	0.94	0.00
MKTR	1.01	0.94	0.99	0.36	1.00	0.52	0.99	0.00
	6-month horizon							
TERM	1.04	1.00	1.08	1.00	1.02	0.91	1.05	0.99
DEF	0.98	0.09	1.06	1.00	0.99	0.00	0.97	0.00
MKTR	1.03	1.00	1.07	0.99	0.99	0.29	1.03	0.99
	12-month horizon							
TERM	1.06	1.00	1.17	1.00	1.03	0.92	1.09	0.97
DEF	1.00	0.46	1.15	1.00	0.99	0.00	1.01	0.70
MKTR	1.04	1.00	1.16	1.00	0.97	0.09	1.06	0.97

Notes: The evaluation sample runs from January 1995 to June 2012. MSE ratio is the ratio of the mean squared error loss from the unrestricted model including a volatility measure to that from the restricted model. The reported p-value is associated with the Clark and McCracken (2009) test for nested models under the assumption noisy data revisions.

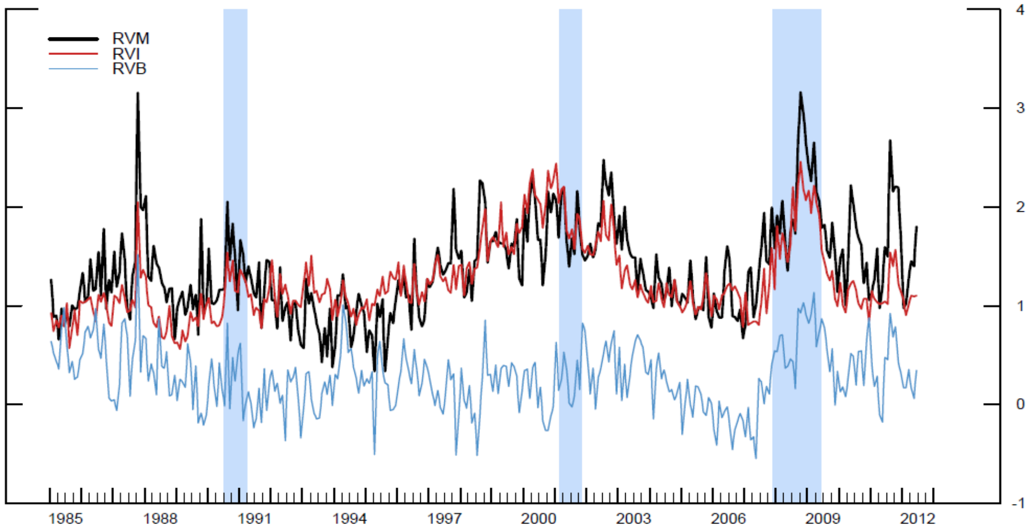
Table 17: Parameter Estimates of the Nonlinear Dynamic Factor Model of Volatilities

Parameter	Estimate	Parameter	Estimate
α_0	8.943	λ_1	0.366
	(0.00)		(0.00)
α_1	6.372	λ_2	0.323
	(0.00)		(0.00)
α_2	4.547	λ_3	0.123
	(0.00)		(0.00)
ψ	0.227	ϕ_1	0.242
	(0.01)		(0.00)
τ_0	1.918	ϕ_2	0.987
	(0.00)		(0.00)
τ_1	0.894	ϕ_3	0.535
	(0.00)		(0.00)
p_{00}	0.820	σ_1	0.435
	(0.00)		(0.00)
p_{11}	0.875	σ_2	0.136
	(0.00)		(0.00)
p_{22}	0.974	σ_3	0.780
	(0.00)		(0.00)

Notes: The sample runs from January 1985 to June 2012. Asymptotic p-values in parentheses are based on the standard deviations based on the inverse hessian, which is obtained through numerical calculation. The variance of the common factor in regime 2 is set to unity for identification (i.e. $\tau_2 = 1$ in the model stated in Section 4.3).

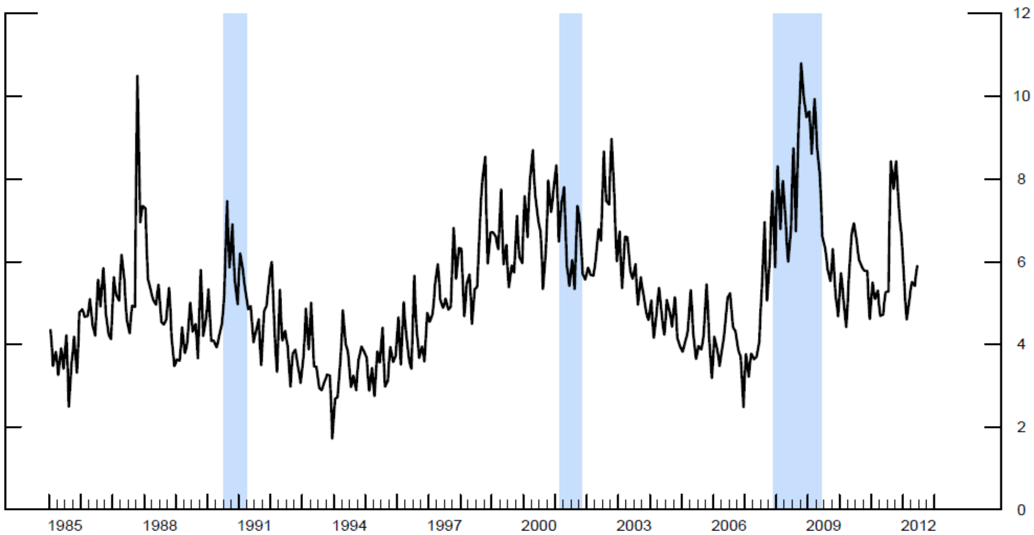
Figures

Figure 1: Volatility Measures



Notes: See the main text for variable definitions. The shaded areas represent recessions according to the business cycle dating committee of the NBER.

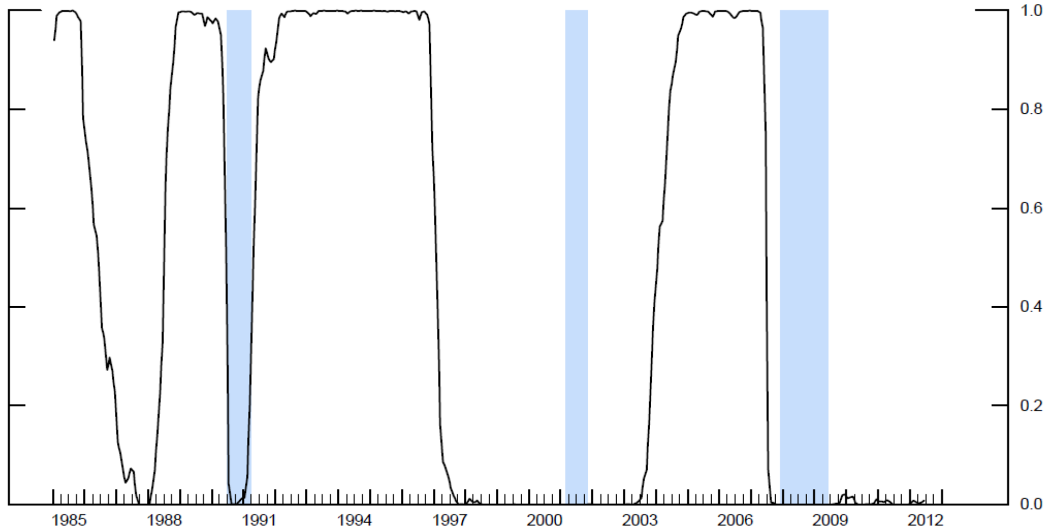
Figure 2: Common Volatility Factor



Notes: The shaded areas represent recessions according to the business cycle dating committee of the NBER.

Figure 3: Smoothed Probabilities of Volatility Regimes

Panel a: Low-volatility regime



Panel b: Intermediate-volatility regime

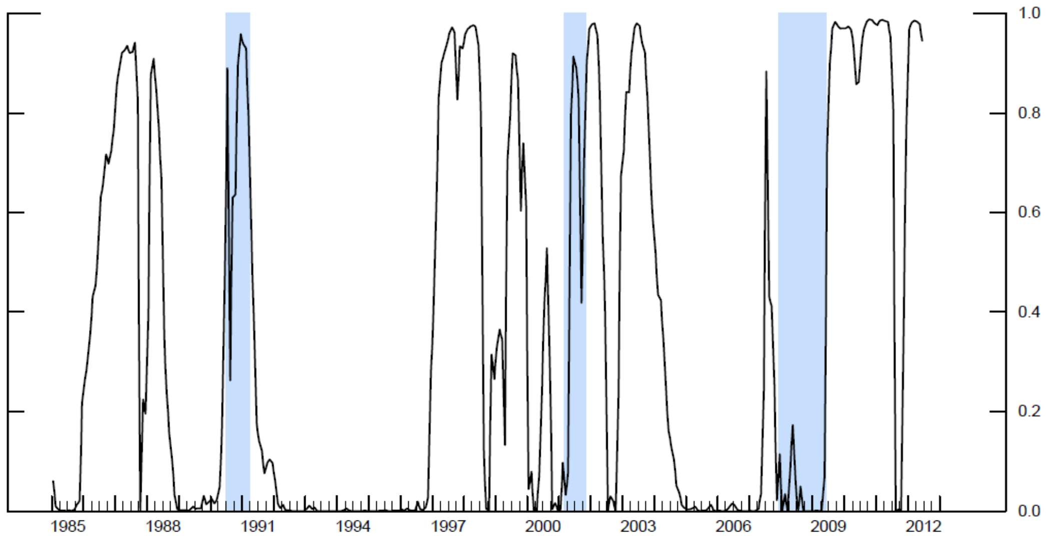
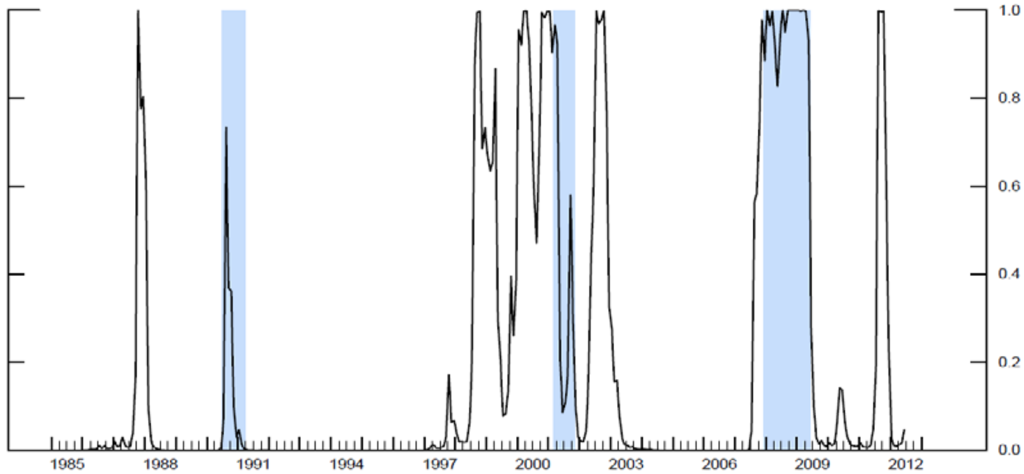


Figure 3: Volatility Regimes (Cont'd)

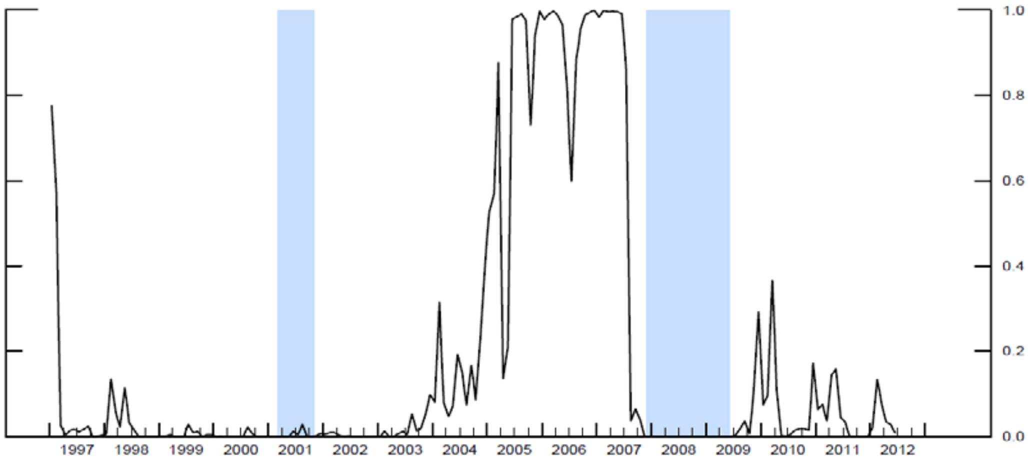
Panel c: High-volatility regime



Notes: Filtered and smoothed probabilities of regimes from the dynamic factor Markov-switching model of financial volatilities are reported. The sample runs from January 1985 to June 2012. See the main text for characterization of regimes.

Figure 4: Recursive Real-time Probabilities of the Volatility Regimes

Panel a: Low-volatility regime



Panel b: Intermediate-volatility regime

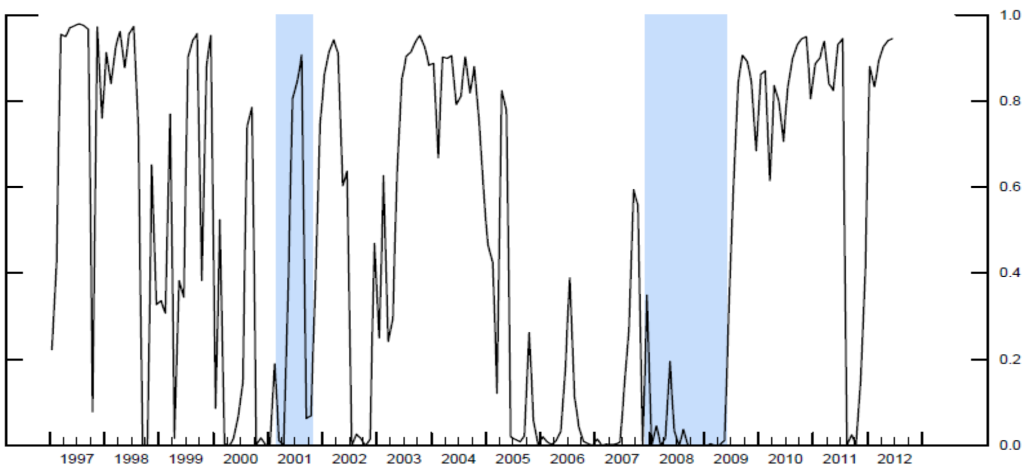
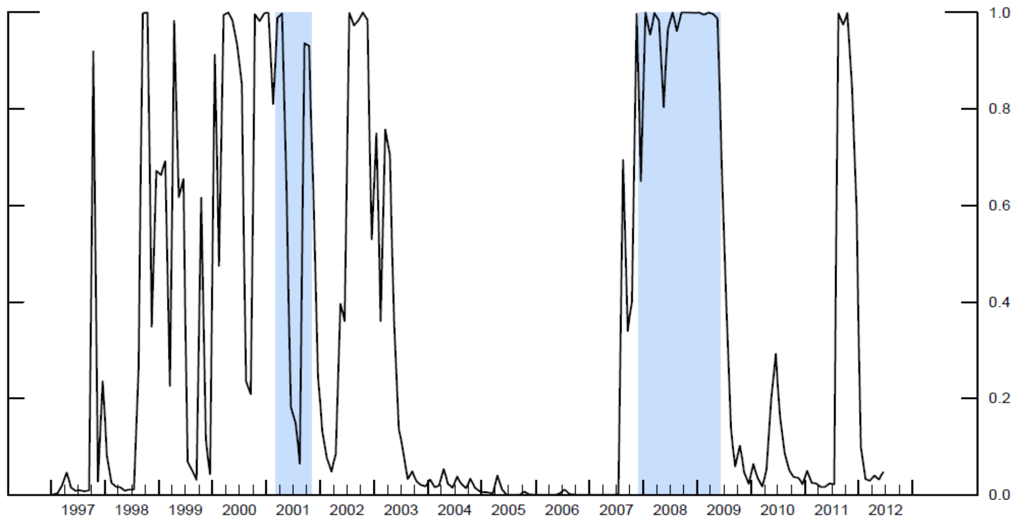


Figure 4: Recursive Real-time Probabilities of the Volatility Regimes (Cont'd)

Panel c: High-volatility regime



Technical Appendix

A. In-sample Tests of Clark and McCracken (2012)

The tests are based on the following DGP:

$$(A.1) \quad y_{T,t+\tau} = x'_{T,2,t}\beta_{2,T} + u_{T,t+\tau} = x'_{T,1,t}\beta_{1,T} + x'_{T,22,t}(T^{-1/2}\beta_{22}) + v_{T,t+\tau},$$

where $x_{T,1,t}$ is $k_1 \times 1$, $x_{T,22,t}$ is $k_2 \times 1$, and $Ex_{T,2,t}u_{T,t+\tau} = 0$, for all $t = 1, \ldots, T$. The models can be characterized *weakly nested* because the unrestricted model is the true model, but as the sample size grows large, the DGP converges to the smaller model. As argued by CM, this setup captures the practical reality that the predictive content of the marginal predictor(s) can be low. Notice that the dependent variable as well as the predictors and the error DEF depend on T, the forecast origin. The sequence $\{y_{T,t}, x'_{T,2,t}\}_{t=1}^{T}$ is used to generate τ–step forecasts from the restricted and unrestricted models, $x'_{T,1,t}\beta_{1,T}$ and $x'_{T,2,t}\beta_{2,T}$ respectively. The parameters are estimated via ordinary least squares. Let $\hat{u}^2_{T,i,T+\tau}$, $i = 1, 2$ denote the loss associated with the τ–step forecasts form the restricted and the unrestricted models, respectively. For our applications we will focus on the case where $k_2 = 1$. In addition, let $B_2(T) = \left(T^{-1}\sum_{t=1}^{T-\tau} x_{T,2,t}x'_{T,2,t}\right)^{-1}$, $J_2 = \left(0_{1\times k_1}, 1\right)'$, and $t(T) = \frac{T^{1/2}J'_2\hat{\beta}_{2,T}}{\left(J'_2B_2(T)V(T)B_2(T)J_2\right)^{1/2}}$, where $V(T)$ is the nonparametric kernel-based estimator of the relevant long-run covariance matrix. Hence, $t(T)$ is the HAC-robust t-statistic for testing the null hypothesis that $\lim_{T\to\infty} T.E\left(u^2_{1,T+\tau} - u^2_{2,T+\tau}\right) = 0$. Note that under the null, the increase in the mean square error due to the omitted variable bias is exactly offset by the decrease arising from imprecise estimation of the marginal predictive coefficients. Under certain regularity conditions, it can be shown that $t(T) \to N(\text{sign}(\beta_{22}), 1)$. Therefore, as long as we have a priori information regarding the direction of the predictive relationship, standard normal critical values can be used for inference. Moreover, for improved finite sample performance CM propose a fixed-design wild bootstrap procedure that imposes the null hypothesis via restricted least squares.

B. Out-of-sample Tests of Clark and McCracken (2009)

This out-of-sample evaluation framework is based on real-time data and assumes that the observables are subject to revisions over a finite number of periods, r. Let $y_s(t)$ denote the value of the time t vintage of the observation s realization of y, where $t \geq s$. When the revision process is complete – or when the series is not subject to revisions – we simply use y_s to denote the time s value of y. The sample of observations, $\{\{y_s(t), x_s'(t)\}_{s=1}^{t}\}_{t=R}^{\bar{T}}$ include a scalar random variable $y_s(t)$ – to be predicted – and a $(k \times 1)$ vector of predictors, $x_s(t)$. In case of nested models we have $x_s(t) = x_{2,s}(t) = (x_{1,s}'(t), x_{22,s}'(t))'$, with $x_{i,s}(t)$ being the $(k_i \times 1)$ vector of predictors associated with model i. When models are non-nested $x_{1,s}(t)$ and $x_{2,s}(t)$ denote distinct sub-vectors of $x_s(t)$.

For each forecast origin $t = R, \ldots, T \equiv R + P - \tau$ we predict $y_{t+\tau}(t')$, where τ is the forecast horizon and t' is the vintage used for evaluating the forecasts, and $r' = t' - (t + \tau)$ is the vintage horizon so that $\bar{T} = T + \tau + r'$. Let $\hat{u}_{i,t+\tau}(t') = y_{t+\tau}(t') - x_{i,t}'(t)\hat{\beta}_{i,t}$ for $i = 1,2$ denote the sequence of forecast errors from models 1 and 2, where $\hat{\beta}_{i,t}$ denotes the recursive OLS estimator of the pseudo parameter vector $\beta_i^\star$ from data vintage t. Let $h_{i,t+\tau}(t') = \left(y_{t+\tau}(t') - x_{i,t}'(t)\beta_i^\star\right)x_{i,t}(t)$, $h_{i,s+\tau} = \left(y_{s+\tau} - x_{i,s}'\beta_i^\star\right)x_{i,s}$, $B_i = \left(Ex_{i,s}x_{i,s}'\right)^{-1}$, and $d_{t+\tau}(t') = u_{1,t+\tau}^2(t') - u_{2,t+\tau}^2(t')$. Furthermore, in case of non-nested models, let $h_{t+\tau}(t') = \left(h_{1,t+\tau}'(t'), h_{2,t+\tau}'(t')\right)'$ and $h_{s+\tau} = \left(h_{1,s+\tau}', h_{2,s+\tau}'\right)'$ and for nested models let $h_{t+\tau}(t') = h_{2,t+\tau}(t')$ and $h_{s+\tau} = h_{2,s+\tau}$.

The test statistic is based on the average square loss differential as in Diebold and Mariano (1995). In particular, CM show that in case of non-nested models $P^{1/2}\bar{d} \rightarrow N(0, \Omega)$ where $\Omega = S_{dd} + 2\left(1 - \pi^{-1}\ln(1+\pi)\right)(FBS_{dh} + FBS_{hh}F')$, S_{dd} is the long-run variance of $d_{t+\tau}(t')$, S_{hh} is the long-run variance of $h_{t+\tau}$, S_{dh} is the long-run covariance between $d_{t+\tau}(t')$ and $h_{t+\tau}$, $\pi = \lim_{R,P\rightarrow\infty} P/R > 0$, $F = 2\big(-Eu_{1,t+\tau}(t')x_{1,t}'(t)$, $Eu_{2,t+\tau}(t')x_{2,t}'(t)\big)$, and B is a block-diagonal matrix with block-diagonal elements B_1 and B_2. The expression for the asymptotic variance matrix is very similar to that provided in West (1996). However, with unrevised data and the quadratic loss function for estimation and forecast comparison we would have $F = 0$ since $Eu_{i,t+\tau}x'_{i,t} = 0.$ However, with revised data $E\big(y_{t+\tau}(t') - x'_{i,t}(t)\beta_i^\star\big)x_{i,t}(t) = 0$ does not hold due to the presence of measurement noise inherent in real-time data. Therefore, one needs to estimate the additional DEFs to obtain an accurately sized test-statistic.

In case of nested models, CM show that $P^{1/2}\bar{d} \to N(0,\Omega)$ where $\Omega = S_{dd} + 2\big(1 - \pi^{-1}\ln(1+\pi)\big)F(-JB_1J' + B_2)S_{hh}(-JB_1J' + B_2)F'$, $J' = \big(I_{k_1\times k_1}, 0_{k_1\times(k_2-k_1)}\big)$ and the other DEFs are defined above. This result stands in sharp contrast to those presented in Clark and McCracken (2005) and McCracken (2007), which show that asymptotic distributions are non-standard for nested model comparisons using unrevised data.

The tests are made operational by replacing the long-run variance and covariance DEFs with consistent estimators. CM use standard kernel-based HAC estimators as in Newey and West (1987) and set the bandwidth equal to 2τ, which allows for noise in data revisions to generate serial correlation even in one-step ahead forecast errors.