A comprehensive look at financial volatility prediction by economic variables

Charlotte Christiansen, Maik Schmeling, Andreas Schrimpf

Institutions: Aarhus University, Leibniz University of Hanover, Bank for International Settlements

Published on: 01 Sep 2012 - Journal of Applied Econometrics (John Wiley & Sons, Ltd)

Topics: Forward volatility, Volatility swap, Volatility smile, Implied volatility and Volatility risk premium

BANK FOR INTERNATIONAL SETTLEMENTS

BIS Working Papers

No 374

A Comprehensive Look at Financial Volatility Prediction by Economic Variables

by Charlotte Christiansen, Maik Schmeling and Andreas Schrimpf

Monetary and Economic Department

March 2012

JEL classification: G12, G15, G17, C53

Keywords: Realized volatility; Forecasting; Data-rich modeling; Bayesian model averaging; Model uncertainty

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)

ISSN 1682-7678 (online)

A Comprehensive Look at Financial Volatility Prediction by Economic Variables*

Charlotte Christiansen ‡ **Maik Schmeling** § **Andreas Schrimpf** **

This draft: March 5, 2012

Abstract

We investigate if asset return volatility is predictable by macroeconomic and financial variables and shed light on the economic drivers of financial volatility. Our approach is distinct due to its comprehensiveness: First, we employ a data-rich forecast methodology to handle a large set of potential predictors in a Bayesian Model Averaging approach, and, second, we take a look at multiple asset classes (equities, foreign exchange, bonds, and commodities) over long time spans. We find that proxies for credit risk and funding (il)liquidity consistently show up as common predictors of volatility across asset classes. Variables capturing time-varying risk premia also perform well as predictors of volatility. While forecasts by macro-finance augmented models also achieve forecasting gains out-of-sample relative to autoregressive benchmarks, the performance varies across asset classes and over time.

JEL-Classification: G12, G15, G17, C53

Keywords: Realized volatility; Forecasting; Data-rich modeling; Bayesian model averaging; Model uncertainty

*We thank Tim Bollerslev (the co-editor), two anonymous referees, Eric Ghysels, Philippe Mueller, Peter Schotman, Christian Speck (DGF discussant), Grigory Vilkov, Robert Vermeulen (EFA discussant), Christian Upper and participants at the Arne Ryde Workshop in Financial Economics in Lund (2011), the European Finance Association Meeting in Stockhom (2011), the ESEM in Oslo (2011), and German Finance Association Meeting in Regensburg (2011) for helpful comments and discussions. Christiansen and Schrimpf acknowledge support from CREATES funded by the Danish National Research Foundation and from the Danish Council for Independent Research, Social Sciences. Schmeling gratefully acknowledges financial support by the German Research Foundation (DFG). The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank for International Settlements.

‡CREATES, Department of of Economics and Business, Business and Social Sciences, Aarhus University, Bartholins Alle 10, DK-8000 Aarhus C, Denmark, e-mail: `cchristiansen@creates.au.dk`, phone: +45 8716 5576.

§Department of Economics, Leibniz Universität Hannover, Königsworther Platz 1, D-30167 Hannover, Germany, e-mail: `schmeling@gif.uni-hannover.de`, phone: +49 511 762-8213.

**Bank for International Settlements (BIS), Monetary and Economic Department, Centralbahnplatz 2, 4002 Basel, Switzerland. Tel: +41 61 280 8942. Email: `andreas.schrimpf@bis.org`.

1. Introduction

Financial volatility is a crucial input for risk management, asset pricing, and portfolio management and it may exert important repercussions on the economy as a whole as evinced forcefully by the recent financial crisis. It is therefore of primary interest to learn more about the economic drivers of volatility in financial markets. In this paper, we empirically investigate whether information in a broad set of economic variables measuring financial and macro conditions is helpful in predicting future volatility. We provide a comprehensive analysis of volatility predictability for several major asset classes in a data-rich forecasting framework. While our main focus is on studying the determinants of equity market volatility drawing on a long-term dataset which covers more than 80 years of data, we also consider volatility in foreign exchange, bond, and commodity markets over shorter time spans.

Economic theory suggests that variables capturing time-varying risk premia are primary candidates for understanding and forecasting volatility (Mele, 2007). This implies that return predictors from the extant literature (e.g. valuation ratios for equities, yield spreads for bonds, interest rate differentials for foreign exchange etc.) qualify as promising volatility predictors as well. However, countercyclical risk premia in financial markets do not mechanically imply countercyclical return volatility (Mele, 2007). Thus, it is worthwhile to investigate sources of volatility predictability separately from return predictability. In addition, several return predictors also have a direct theoretical link to volatility forecasting. For example, structural credit risk models such as Merton (1974) imply that equity volatility increases when leverage increases, so that higher market-wide leverage should be positively related to future stock return volatility.

Using information in financial and macroeconomic variables to forecast volatility in financial markets is not entirely new to the literature but is far from having received the same attention as the predictability of asset returns (see e.g. Goyal and Welch, 2003, 2008; Cochrane and Piazzesi, 2005; Ang and Bekaert, 2007; Ludvigson and Ng, 2009; Lustig, Roussanov, and Verdelhan, 2010, for recent contributions to return predictability). Moreover, while there is a vast econometric literature on pure time-series modeling and forecasting of volatility (see e.g. Andersen, Bollerslev, Christoffersen, and Diebold, 2006, for a recent survey), the empirical literature on the economic drivers of volatility is still fairly scarce.[1]

The seminal paper on the economic determinants of equity market volatility is Schwert (1989). While his findings point towards countercyclical movements of stock market volatility, the link between volatility and economic activity is not found to be very strong from a statistical perspective. On a more positive note, Engle, Ghysels, and Sohn (2008) analyze the effect of inflation and industrial production growth on daily stock return volatility, considering each macroeconomic variable separately. They find that macro fundamentals do indeed matter for stock return volatility. Diebold and Yilmaz (2010) consider a broad set of 40 international equity markets and find that stock market volatility is cross-sectionally related to fundamental macroeconomic volatility as measured by GDP volatility. In a recent paper, Paye (2012) studies equity volatility predictability by macroeconomic and financial variables. His results suggest meaningful and encouraging links between several economic variables and stock market volatility, whereas improvements in terms of out-of-sample forecast accuracy are found to be fairly modest.

It is rather difficult, however, to draw general conclusions from the extant literature on financial volatility predictability by macroeconomic and financial state variables. Most authors employ different sample periods, different forecasting models and methods, different predictors, different forecast evaluation criteria, and almost exclusively focus on stock market volatility. Another important aspect is that "model uncertainty" is neglected in this branch of the literature. Model uncertainty is the ex ante uncertainty of an economic agent with regard to the right choice of macro-finance variables that are best suited for volatility prediction. While theory provides some motivation for why some economic variables might qualify as predictors, it offers little guidance on which specific variable (or particular combination of variables) should enter the forecasting model for volatility. In this paper, we consider a Bayesian model averaging framework

[1]Modeling and forecasting time-varying volatility has its foundations in the class of (G)ARCH models (Engle, 1982; Bollerslev, 1986). More recently, the literature has expanded substantially drawing on the concept of realized volatility and high-frequency modeling (See, e.g. Andersen, Bollerslev, Diebold, and Labys, 2003). This literature is typically interested in high-frequency movements of volatility and time series aspects, while this paper is mainly interested in low frequency variation and its link to macroeconomic and financial conditions.

which takes account of this model uncertainty.[2] In essence, the approach we take is to let the data speak about the usefulness of specific predictors.

The recent paper by Paye (2012) comes closest to ours in that it also models realized volatility in a predictive regressions setting. In this paper we go beyond Paye's work in several regards. First, we study a larger set of macro-finance predictors (38 as opposed to 13). Second, we go beyond equity market volatility to consider other major asset classes such as foreign exchange, bond markets and commodities for which (to the best of our knowledge) the economic determinants of volatility have not been systematically analyzed before. Third and most important, we explicitly consider the effects of model uncertainty when dealing with a large number of potential predictors. We do this by a Bayesian model averaging (BMA) approach which has been shown to be an adequate tool for stock return and exchange rate predictability (see e.g. Avramov, 2002; Wright, 2008) and we show its virtues for studying the economic determinants of volatility as well.[3] This allows us also to go beyond naïve forecast combination approaches and to consider optimal Bayesian foreacsts for out-of-sample forecasting. In a nutshell, it is the comprehensive approach, both in terms of scope as well as in terms of the applied econometric methodology, which is the unique feature of this paper.

Understanding volatility movements is important since it is a consequential input for investment and asset allocation decisions. Moreover, understanding the macroeconomic causes of financial market volatility is interesting in itself since it helps to uncover linkages between price movements in financial markets and underlying risk factors or business cycle state variables. This is even more important since there is a growing body of evidence showing that risks associated with volatility are priced in option, stock, bond, and foreign exchange markets (e.g. Ang, Hodrick, Xing, and Zhang, 2006, Da and Schaumburg, 2009, Menkhoff, Sarno, Schmeling, and Schrimpf, 2012, Christiansen, Ranaldo, and Söderlind, 2011 among others). Volatility-based measures have also been shown to predict future stock market returns (see, e.g., Bollerslev,

[2]BMA is originally due to Leamer (1978). For treatments of model uncertainty in financial forecasting setups similar to this paper, see e.g. Avramov (2002), Cremers (2002), or Wright (2008); Faust, Gilchrist, Wright, and Zakrajsek (2011) recently use a BMA approach to predict U.S. business cycle fluctuations by credit spreads.

[3]Independent contemporaneous work by Cakmakli and van Dijk (2010) also considers the issue of stock return and volatility predictability based on many (macro-)economic variables. Unlike the Bayesian Model Averaging framework considered here, the authors extract information from macroeconomic series by dynamic factor models.

Tauchen, and Zhou, 2009). Furthermore, recent evidence in Mele (2008) and Fornari and Mele (2010) shows that stock market volatility is informative about future business cycle fluctuations so that a better understanding of the driving forces of financial volatility is important for policy makers and monetary authorities. In the same vein, Chauvet, Senyuz, and Yoldas (2010) find that financial volatility helps predict future economic activity when they consider volatility measures that cover both stock and bond markets.

Our results indicate that economic predictor variables add significant explanatory power in our forecasting exercises. Importantly, these results hold when controlling for the informational content in lagged asset return volatility. It is also important to point out that our main forecasting approach shows the strongest predictive ability for predictors that are associated with time-varying risk premia, leverage, or financial distress. Default spreads, for instance, stand out as useful predictors not only for equity market volatility but also for other asset classes. Moreover, the most robust predictors include valuation ratios (e.g. dividend yields in case of equity volatility), a measure of interest rate differentials vis à vis the U.S. (in case of foreign exchange volatility), and a measure of funding market (il)liquidity and heightened counterparty credit risk (the TED spread as in Brunnermeier, Nagel, and Pedersen, 2009) which matter for several asset classes.

In a nutshell, our results therefore suggest that there are economically meaningful relations between variables measuring financial conditions and future volatility of different asset classes. Purely macroeconomic variables (as opposed to financial variables) hardly show up as important predictors of financial volatility. These results are fairly robust to a number of additional checks and methodological variations. Our findings on out-of-sample predictability show that including macro-finance predictors are able to enhance the forecast performance relative to simple autoregressive benchmarks in particular in the case of forecast combination methods. These improvements do not hold for all asset classes and sample periods uniformly, however, and, as in the case of return predictability (Goyal and Welch, 2008; Timmermann, 2008), forecast performance can vary strongly over time.

The remaining part of the paper is structured as follows. Section 2 describes the data, Section 3 details the econometric framework, Section 4 presents the empirical results, and Section 5 concludes. Various additional results and details are delegated to an Online Appendix.

2. Data and Volatility Measurement

We base our main analysis on monthly observations of macroeconomic and financial variables and realized volatilities computed from daily return observations. We focus on a long sample for the U.S. equity market (represented by the S&P500), which runs from December 1926 to December 2010, a total of 1,009 monthly observations.[4] In addition, we also investigate a shorter sample for all four asset classes which covers the period from January 1983 to December 2010. The starting point of the shorter sample is guided by having a comprehensive and common dataset for all asset classes, both in terms of predictors and dependent variables. Thus, in the case of the short-term multi-asset class sample we have 366 monthly time series observations for each variable.

2.1. Measuring Financial Volatility

The main variables of interest are volatilities of the different asset classes which serve as dependent variables in our predictive regressions. We compute the realized variance for asset class i in month t as the sum of squared intra-period (daily) returns: $\sum_{\tau=1}^{M_t} r_{i;t;\tau}^2$ where $r_{i;t;\tau}$ is the τth daily continuously compounded return in month t for asset i and M_t denotes the number of trading days during month t. In our empirical analysis, we define the realized volatility to be the log of the square root of the realized variance since it is better behaved (i.e. closer to normality) than the raw series:

$$RV_{i;t} = \ln \sqrt{\sum_{\tau=1}^{M_t} r_{i;t;\tau}^2}, \quad t = 1, ..., T. \tag{1}$$

Realized volatility is an accurate proxy for the true, but latent, integrated volatility as the number of intra-period observations becomes large (See, e.g. Andersen, Bollerslev, Diebold, and Labys, 2003). We then proceed by using realized volatilities computed this way as observable

[4] We are most grateful to G. William Schwert for providing these data.

dependent variables (See, e.g. Andersen, Bollerslev, Christoffersen, and Diebold, 2006, for a survey).

For the construction of realized stock market volatility ($RV_{S,t}$) we compute the realized volatility measure according to Eq. (1) based on daily returns on the S&P500. Realized bond market volatility ($RV_{B,t}$) is computed from returns on the 10-year Treasury note futures contract traded on the Chicago Board of Trade (CBOT).[5] In the same manner, we employ Standard & Poor's GSCI commodity index to construct our proxy for commodity market volatility ($RV_{C,t}$).[6]

Our construction of aggregate foreign exchange (FX) market volatility is somewhat less standard and draws on a portfolio approach for the global foreign exchange market as suggested in recent work by Lustig, Roussanov, and Verdelhan (2011), among others. Hence, we use a basket of currencies against the U.S. dollar for our main analyses. We do so to obtain an aggregate measure of foreign exchange volatility (from the perspective of a U.S. investor) similar to the aggregate stock market index, bond index, and commodity index we use for the other asset classes. For robustness, however, we also report results for four major individual exchange rates (German mark/euro, Japanese yen, British pound, and Swiss franc against the U.S. dollar) in the Online Appendix. To construct the aggregate FX volatility measure we form an equally weighted portfolio consisting of all currencies with available data at a given point in time.[7] For the aggregate FX portfolio we calculate the time series of the daily spot rate changes which are then used to construct realized FX volatility (denoted by $RV_{FX,t}$) according to Eq. (1).

[5]We use the daily closing price of the bond futures contract which is available from Datastream (mnemonic symbol is CTYCS00(PS)). This is the contract used by Fleming, Kirby, and Ostdiek (1998) who consider volatility linkages between equity and bond markets. The advantages of using futures data are that these contracts are highly liquid and that we can compute bond returns straight away without having to rely on return approximations based on yields.

[6]These data are available from Datastream. In place of the GSCI index, it may have been preferable to use data on the GSCI futures contract as this is actively traded at the CME (See e.g. Fong and See, 2001). Yet, the GSCI futures only start trading in 1992. Still, the correlation between the realized volatility for the GSCI index and the GSCI futures amounts to 0.97 during the period 1992-2009, so we deem it reasonable to use the GSCI index to obtain a longer time-series.

[7]The foreign exchange rates are available from Thomson Financial Datastream. We use the currencies of the following countries (all quoted against U.S. Dollar): Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Egypt, Euro area, Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Kuwait, Malaysia, Mexico, Netherlands, New Zealand, Norway, Philippines, Poland, Portugal, Russia, Saudi Arabia, Singapore, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, Ukraine, and United Kingdom. Not all currencies are available during the entire sample period, as some currencies enter or exit the sample.

[Insert Table 1 about here]

Table 1 shows summary statistics for the realized volatility series. The average volatilities for commodities and stocks are much larger than the average volatilities for foreign exchange and bonds. The same holds for the standard deviations of the realized volatility series. As is well known, realized volatility is highly persistent and we find this behavior for all four asset markets under investigation as indicated by the autocorrelation coefficients.

[Insert Figures 1 and 2 about here]

Figures 1 and 2 plot our realized volatility measures for the different asset classes. The time series are highly variable and they do not appear to follow an identical pattern across asset classes. This is also reflected in the pair-wise correlation coefficients that are reported in Panel B of Table 1 which are generally not very high in absolute terms, i.e. below 48%. Given this heterogeneous behavior, one may suspect that the volatility of different asset classes is at least partly driven by different economic variables.

2.2. Macroeconomic and Financial Predictors

Overall, we rely on a comprehensive set of 38 macroeconomic and financial predictive variables. The results for the long-term sample period draw upon a reduced set of 13 predictors (indicated in Table A.1) which are available as of December 1926. Since many of those variables are motivated via the risk-premium channel by Mele (2007), there is some overlap with the predictive variables used in the comprehensive study on stock return predictability by Goyal and Welch (2008). Table A.1 provides an overview and summary statistics of the predictors whereas the Online Appendix provides further details on data sources and construction.

The variables considered in this paper are motivated by theory, mostly focussing on the time-varying risk-premium channel emphasized by Mele (2007). This implies that specifically those variables that have been shown to be useful predictors of returns and hence drivers of risk premia should qualify as potential predictors. Given the scope of our study with its multi-asset class focus, we do not only consider popular forecasting variables for equity returns (such as valuation ratios, industrial production growth etc.) but we also take into account potential drivers of countercyclical risk premia in other markets. In addition, we also consider variables associated with market and funding (il)liquidity, heightened credit and counterparty risk, as one may suspect that a specific variable capturing risk premia or market conditions in one market is also influential for volatility in another market.[8] For instance, during the recent financial crisis stress in money markets quickly spilled over to affect market conditions and returns in other asset classes (Baba, 2009). Our tests allow for such a possibility.[9]

In the following, we provide a brief overview and some further details regarding the motivation behind some specific variables. For ease of exposition, we classify the predictive variables according to five broad economic categories:

(1) Equity Market Variables and Risk Factors: Our list of explanatory variables includes well-known equity market valuation ratios such as the dividend price ratio (D-P) and the earnings-price ratio (E-P), commonly considered in predictive regressions for stock returns (e.g. Campbell and Shiller, 1988; Goyal and Welch, 2008). We also include lagged equity market returns (MKT) to capture the well-established leverage effect (Black, 1976; Glosten, Jagannathan, and Runkle, 1993; Nelson, 1991), i.e. the finding that negative returns are associated with higher subsequent volatility. Other equity variables include the risk factors by Fama and French (1993), a short-term reversal factor (STR) which is related to market volatility and distress as analyzed in Nagel (2012). We also consider S&P500 turnover (TURN), which is commonly viewed as a proxy for difference in opinion (Scheinkman and Xiong, 2003; Baker and Wurgler, 2007) and hence potential uncertainty about future market valuations.

(2) Interest Rates, Spreads, and Bond Market Factors: Our set of bond market variables include variables for instance the T-bill rate (T-B) which has shown to be a useful predictor of equity excess returns (e.g. Ang and Bekaert, 2007). In addition, we include prominent bond return predictors such as the term spread (Campbell and Shiller, 1991) and the factor by Cochrane and Piazzesi (2005) based on the term structure of forward rates. These variables are intended

[8]We thank an anonymous referee for pointing this out.

[9]In unreported tests, we also tested if volatilities in one specific market Granger-cause volatilities in other markets, so-called volatility spillover effects (See e.g. Diebold and Yilmaz, 2009). We do not find evidence for these kinds of lead-lag relationships, which is most likely due to the low frequency of the data considered in this paper.

to capture the evolution of risk premia in bond markets.

(3) Foreign Exchange Variables and Risk Factors: We consider three foreign exchange specific forecasting variables. The average forward discount (AFD) – which measures interest rate differentials vis à vis the U.S. for a broad range of currencies – might be particularly useful given the findings on its ability to capture countercyclical FX risk premia in Lustig, Roussanov, and Verdelhan (2010). In addition, we also include the dollar risk factor (DOL) and a carry trade factor (C-T) from Lustig, Roussanov, and Verdelhan (2011) which have been shown to capture a large fraction of common FX return variation.

(4) Liquidity and Credit Risk Variables: To proxy for heightened credit risk we rely on the yield spread between BAA and AAA rated bonds (often labeled the default spread, DEF). Credit risk tends to be higher in situations where leverage rises, which – according to models such as Merton (1974) – should influence volatility. Furthermore, we include the TED spread (difference between the 3-month LIBOR rate and T-Bill rate), a common measure of funding (il)liquidity in interbank markets (e.g. Brunnermeier, Nagel, and Pedersen, 2009). We also consider an aggregate measure of bid-ask spreads in foreign exchange markets to proxy for FX market (il)liquidity (Menkhoff, Sarno, Schmeling, and Schrimpf, 2012) as well as the measure of stock market liquidity in equity markets by Pastor and Stambaugh (2003).

(5) Macroeconomic Variables: We also consider general macroeconomic variables, such as inflation and industrial production growth (either computed in monthly or annual growth rates). The latter variable is central in the recent return predictability of excess returns in bonds and foreign exchange (see e.g. Ludvigson and Ng, 2009; Lustig, Roussanov, and Verdelhan, 2010). Output-based measures have also been found to be successful predictors of equity returns (e.g. Cooper and Priestley, 2009). Including these variables also allows us to assess if macroeconomic conditions are causal (in a *post hoc ergo propter hoc* sense) for volatility or by contrast whether causality runs the other way as emphasized in papers such as Fornari and Mele (2010) or Chauvet, Senyuz, and Yoldas (2010).

3. Econometric Framework

We now outline our econometric approach.[10] Note that we use a univariate framework throughout the paper which aims at predicting financial volatility for each asset class separately. We use standard predictive regressions for the future realized volatility of asset i

$$RV_{i;t} = \alpha + \sum_{\ell=1}^{L} \rho_l RV_{i;t-\ell} + \beta_j' z_{j;t-1} + u_t, \tag{2}$$

where β_j denotes the k_j-dimensional vector of regression coefficients on the predictive variables and i indexes the asset type. The subscript j indicates that the composition of the vector of predictive variables $z_{j;t}$ depends on the particular model $\mathcal{M}_j$. As we have a large number of potentially relevant predictor variables, we investigate $j = 1, ..., 2^\kappa$ models, where κ denotes the overall number of predictive variables under consideration.

Since volatility is fairly persistent, it is important to include autoregressive terms in the predictive regression to investigate if there is additional predictive content of the macroeconomic and financial variables that goes beyond the information contained in the time-series history of volatility. We therefore also report results from fitting an autoregressive model for the RV series as the relevant benchmark case. While we largely focus on one autoregressive lag ($L = 1$), we also discuss models with higher order terms in Section 4.2 and report additional results in the Online Appendix.

Since the number of potential models is very large, it is computationally infeasible to evaluate all possible models analytically. With $\kappa = 38 + 1$ potentially useful predictive variables we have $2^{39} = 549,755,813,888$ possible model specifications. Given these considerations, we rely on two approaches in this paper. First, we make use of a Bayesian model averaging approach with a stochastic model search algorithm (MC^3). Second, we employ a model selection approach based on information criteria. We detail these two approaches next.

[10] See e.g. Avramov (2002) and Ludvigson and Ng (2009) for related approaches in the literature on stock return predictability and bond return predictability. Wright (2008) studies the predictability of exchange rates in a similar data-rich forecasting environment.

3.1. Bayesian Model Averaging and MC^3

Our baseline results draw on a Bayesian model averaging (BMA) approach. We briefly outline the approach in the following, while some further technical details are discussed in the Appendix. A particularly attractive feature of the BMA approach is that model uncertainty can be addressed in a coherent way. Results in the stock return predictability literature suggest that model uncertainty can be substantial when forecasting returns, see for instance Avramov (2002) or Schrimpf (2010). In our context, model uncertainty refers to a situation where it is not clear ex ante what the important predictive variables might be or which combination of variables may be useful for prediction purposes. Unlike the classical approach, BMA does not posit the existence of a true model and is therefore particularly suited to deal with a setup where model uncertainty plays a role. Moreover, the BMA approach can be used to obtain optimal weights for forecast combination (See e.g. Timmermann, 2006). To handle the large number of potential models we rely on Markov Chain Monte Carlo Model Composition (MC^3), a sampling approach drawing from the model space which is particularly suited for high-dimensional problems such as the one encountered here (See, e.g. Fernandez, Ley, and Steel, 2001; Koop, 2003). We outline this approach in section B.2 of the appendix. The results are based on 500,000 draws and a burn-in period of 50,000 draws.

In the Bayesian framework it is common to derive posterior probabilities $p(\mathcal{M}_j|D)$ for a particular model, where different models are defined by inclusion or exclusion of specific explanatory variables. These posterior model probabilities, which reflect the usefulness of a particular model after having seen the data D, are used in the BMA framework as weights in a composite model:

$$E[\beta|D] = \sum_{j=1}^{2^\kappa} p(\mathcal{M}_j|D)\beta_j|D, \tag{3}$$

where $\beta_j|D$ denotes the posterior mean of the predictive coefficients in the jth model. Likewise, combined forecasts of BMA can be obtained by weighting the forecasts of the individual models by the corresponding posterior model probabilities. Thus, in line with the Bayesian tradition, the data allow us to learn by updating our belief about the quality of a particular model. The posterior model probability is given by

$$p(\mathcal{M}_j|D) = \frac{p(D|M_j)p(\mathcal{M}_j)}{\Sigma_{i=1}^{2^\kappa} p(D|M_i)p(\mathcal{M}_i)}, \tag{4}$$

where $p(D|M_j)$ is the marginal likelihood and $p(\mathcal{M}_j)$ denotes the prior probability of model j (as determined by inclusion and exclusion of specific predictive variables). The expression for the marginal likelihood is obtained as

$$p(D|M_j) = \int p(D|M_j, \beta_j)p(\beta_j|\mathcal{M}_j)d\beta_j, \tag{5}$$

where $p(\beta_j|\mathcal{M}_j)$ refers to the prior on the parameters of model j and $p(D|M_j, \beta_j)$ is the likelihood of the model.[11]

3.2. Model Selection Based on Information Criteria

For completeness, we provide comparisons to a classical model selection approach that neglects model uncertainty. Given the large amount of predictors, some standard pretesting is necessary before estimating and evaluating the different models. To this end, we reduce the initial set of potential predictors by only considering variables with a t-statistic greater than two in absolute value in a predictive regression containing the respective macro-finance predictor and the lagged dependent variable. In this way, we end up with a smaller set of predictors such that an analytical evaluation of all models is computationally feasible. This is a common approach and is also used by e.g. Ludvigson and Ng (2009) in the context of bond return predictability. For each of the different model specifications, the Schwarz Information Criterion (BIC) is computed. Then the models are ranked according to the BIC. The BIC favors models that provide a good fit while at the same time penalizing highly parameterized models. Our tables report estimation results for the three best model specifications according to the BIC and we report coefficients, Newey and West (1987) HAC standard errors with optimal lag length selection by Andrews (1991), and the adjusted R^2.

[11] We focus on a 1-month forecasting horizon in the paper since the Bayesian approach does not allow for longer horizons with overlapping observations. However, we also consider longer horizons (based on quarterly data) with detailed results reported in the Online Appendix (Table IA.8). These results do not yield much additional insight.

3.3. Out-of-Sample Forecast Evaluation

We also evaluate how forecasting models augmented by macro-finance predictors perform in an out-of-sample context. As a general rule, we always evaluate the out-of-sample performance of the forecast against the benchmark forecast of an autoregressive model. We basically employ the same procedure as in our in-sample tests but we now estimate our models recursively and evaluate the resulting out-of-sample forecasts. More specifically, we start with an initialization period, estimate predictive regressions in the same way as above to produce the first out-of-sample forecast. We then expand the estimation window and repeat the above steps to obtain out-of-sample forecasts for the next period and continue in this way until we reach the end of the sample period. In the following, we denote the forecast by the macro-finance augmented model by $f^M_{i,t+1}$ and the forecast of the autoregressive benchmark model by $f^B_{i,t+1}$.

We report Theil's U (TU) which is given by the root mean square error (RMSE) of our macro-finance augmented model relative to the RMSE of the benchmark model such that a value smaller than one indicates that the model beats the benchmark in terms of forecast accuracy. In addition, we report out-of-sample R^2s as in Campbell and Thompson (2008). The out-of-sample R^2 is computed as

$$R^2_{OOS} = 1 - \frac{\sum_{t=R}^{T-1}(RV_{i,t+1} - f^M_{i,t+1})^2}{\sum_{t=R}^{T-1}(RV_{i,t+1} - f^B_{i,t+1})^2} \tag{6}$$

where T denotes the overall sample size, and R is the initialization period.

Besides these purely descriptive forecast evaluation criteria, we provide bootstrap-based statistical inference in order to assess if models augmented by macro-finance predictors significantly outperform the benchmark forecast. Since the benchmark model is nested by the model of interest, the asymptotic test put forth by Clark and West (2007) may be used. However, the theoretical setup considered in Clark and West (2007) does not cover our case where the forecasts are generated by forecast combination and where a model search over a large amount of models is conducted. Hence, we rely on a bootstrap approach instead of asymptotic tests.[12] We

[12]We are grateful to Todd E. Clark for this suggestion. In a similar vein, Wright (2008) relies on a bootstrap approach to evaluate the out-of-sample accuracy of BMA generated forecasts.

provide a brief description of our bootstrap procedure in the Online Appendix. In addition, we follow Paye (2012) and report results from the asymptotically valid test by Giacomini and White (2006) which is computationally equivalent to the seminal test by Diebold and Mariano (1995). Finally, we report Mincer-Zarnowitz (MZ) tests of unbiased forecasts (by regressing actual realized volatility on a constant and the f^M volatility forecast) and report results for a Wald test of a zero intercept and unit slope coefficient. Our results are based on the GLS version of the Mincer-Zarnowitz test suggested by Patton and Sheppard (2009).

4. Empirical Results

We first present in-sample results before moving on to out-of-sample forecasting accuracy.

4.1. In-Sample Analysis

Bayesian Model Averaging. We first present our baseline results for the long-term sample of U.S. equity market volatility obtained by the BMA approach. The results are reported in Table 2 which presents the eight best predictor variables in terms of posterior probability of inclusion ($\pi|D$). The posterior probability of inclusion reflects the belief of how likely a variable is included in the model after having seen the data. We start from a prior probability of inclusion π of 0.5, which implies that every model is deemed equally likely a priori (Koop, 2003; Faust, Gilchrist, Wright, and Zakrajsek, 2011). Hence, if $\pi|D$ exceeds 0.5, our belief of the usefulness of a particular economic variable as a predictor of volatility has been revised upwards in the light of the evidence.[13] In addition, we report posterior means and standard deviations as well as Bayesian t-ratios. These t-ratios incorporate adjustments for model uncertainty and are thus not comparable to classical t-statistics. We indicate by 1, if a specific predictor variable appears in the top five models according to the posterior model probability $p(\mathcal{M}_j|D)$.

[Insert Table 2 about here]

[13] We also checked if alternative values for the prior probability of inclusion would alter our results, but did not find that our results are much affected by alternative hyperparameter choices.

As a first observation, we find that several macro-finance variables are included in the top five models and/or have a posterior probability exceeding 0.5 so that it is useful to rely on their informational content when forecasting volatility. As expected, the autoregressive component is important and explains a large fraction of the variability of volatility.

In terms of economic effects, it is noteworthy that the two most important economic predictors of equity market volatility are associated with the effects of leverage. Default spreads (DEF) tend to widen when firm leverage and counterparty credit risk increases, a situation which precipitates higher volatility. Moreover, low past equity returns are a useful and robust predictor of higher subsequent equity market volatility. This is in line with the "leverage effect" emphasized by Black (1976) and modeled explicitly by asymmetric GARCH models, e.g. Nelson (1991) and Glosten, Jagannathan, and Runkle (1993). The earnings-price (E-P) ratio is typically considered to be associated with time-varying risk premia, which are shown to be related to time-variation in equity market volatility in Mele (2007). The economic effect we find implies that lower earnings-price ratios (i.e. higher stock market valuations) are followed by periods of higher equity volatility and vice versa, a finding which squares well with the notion of investor sentiment (see e.g. Brown and Cliff, 2005; Baker and Wurgler, 2007) where overly optimistic investors drive up stock prices to a level not warranted by fundamentals which is eventually followed by a sharp decline in stock prices (and vice versa). Another important predictor is the short-term reversal factor, which has been found to be related to (il)liquidity in Nagel (2012). On a *daily* frequency, the short-term reversal return is positively related to the VIX and the return to supplying liquidity (Nagel, 2012). Table 2 shows that the predictive coefficient of the short-term reversal return is negative in our *monthly* data which implies that high returns to supplying liquidity are negatively related to future equity volatility. Given that we are working at lower frequency, it might be the case that high returns to liquidity provision are quickly exploited on average so that a high short-term reversal return is eventually followed by a decline in volatility.

[Insert Table 3 about here]

Table 3 shows results for the shorter sample period (1983-2010) for all four major asset classes. The results are based on the full set of 38 macro-finance variables. While it is the case that mostly market-specific variables show up as predictors of volatility of a particular market, it is interesting to note that there are some variables which emerge as *common* predictors across asset classes. In particular, these variables are associated with money market stress and funding (il)liquidity (see e.g. Brunnermeier and Pedersen, 2009; Brunnermeier, Nagel, and Pedersen, 2009). The TED spread is among the best forecasting variables for foreign exchange, fixed income as well as equity market volatility, and is positively related to future volatility in all three cases. This finding makes economic sense, since a higher TED spread signals higher illiquidity and perceived counterparty risk, which in turn is known to be a driver of volatility in financial markets. In a similar vein, the default spread (or credit spread) shows predictive power for equity and bond markets. Again, this finding makes sense economically since higher default spreads are associated with higher market leverage and the latter is a direct driver of equity volatility (e.g. Merton, 1974). In addition, since higher perceived default risk typically triggers portfolio rebalancing on behalf of investors, it is not surprising to see that the default spread also moves bond volatility.

There are minor differences in the results for stock market volatility between the long sample and the recent sample. Some differences are to be expected as the explanatory variables for the long sample only comprise a subset of those for the recent sample. For the recent sample the dividend-price ratio (D-P) rather than the earnings-price ratio seems to matter and the leverage effect is less strong. Again, the TED spread and the default spread (DEF) show up as useful predictive variables.

One noteworthy effect in the case of foreign exchange market volatility is the predictive content of the average forward discount (AFD). As documented in Lustig, Roussanov, and Verdelhan (2010), this variable, measuring the average interest rate differential vis à vis the U.S., is associated with time-varying risk premia in FX markets. In fact, this variable is also a primary and robust predictor of FX market volatility. Macroeconomic variables like money growth, inflation, and output fluctuations also matter to some degree which seems comforting given the weak link between these standard exchange rate fundamentals and first moments of currency returns (the so-called "disconnect puzzle", see e.g. Obstfeld and Rogoff, 2000).[14]

Our results show that the term spread (T-S), a well-known bond return predictor (e.g. Campbell and Shiller, 1991), is also among the best predictors for future bond market volatility. Another bond related variable, namely the default spread (DEF) also helps in predicting bond volatility. Finally, the S&P500 turnover (TURN) has a bearing upon bond volatility. While there does not seem to be a close theoretical link between stock market turnover and bond market volatility in the earlier literature, one possibility for such effects may be portfolio rebalancing on behalf of investors. Our results on commodity volatility again suggest that financial conditions seem to matter most, whereas variables proxying for conditions in the real economy do not matter that much in explaining financial volatility. On the other hand, model uncertainty seems to be larger, such that the Bayesian BMA t-ratios are much lower compared to the other asset classes and the economic relationships are less robust in vase of commodity market volatility.

In sum, the results suggest a close economic link between return and volatility predictability in equity, foreign exchange and bond markets. The most successful predictors are those associated with measures of funding market (il)liquidity and heightened credit risk as well as time-varying risk premia. Macro variables like output fluctuations, news about the business cycle or inflation only appear to be of minor importance.

Model Selection Approach. For comparison, Tables 4 and 5 report in-sample results based on classical model selection for the long and recent samples, respectively. Results for the three top-performing model specifications according to the BIC are tabulated.

[Insert Tables 4 and 5 about here]

There is a large degree of overlap between the important explanatory variables identified via the BMA approach and the model selection approach. There is not much difference between the predictive power of the best-performing models when measured by their adjusted R^2s. Again,

[14] The results for the individual exchange rates are similar to the aggregate FX results and can be found in the Online Appendix (see Table IA.3). The main difference is that the explanatory power is smaller for the individual currencies than for the aggregate FX market volatility which makes sense since individual currency pairs are subject to large idiosyncratic movements whereas our aggregate FX market volatility measure averages over different currencies and, hence, provides a smoother time-series of movements in the value of the U.S. dollar.

common predictors of volatility across asset classes are the TED spread and the default spread (DEF). Asset specific measures, known to be useful to predict returns (e.g. AFD in case of foreign exchange, Cochrane-Piazessi (C-P) factor, and term spread (T-S) for bonds) are the most successful volatility predictors.

4.2. Benchmarks with Higher-Order Autoregressive Terms

To provide further evidence on the importance of macro-finance predictors for financial volatility, we also present results for specifications where we allow for more autoregressive terms of volatility. More specifically, we repeat our BMA and model selection exercises for the case where we include up to three AR terms and let the data speak about the usefulness of these additional autoregressive terms relative to our macro-finance predictors. The results in this section are based on the long equity market sample since this sample seems best suited for an analysis like this where a lot of power is needed to discriminate between alternative models. We present results for the shorter samples in the Online Appendix (see Tables IA.2, IA.4, IA.5, IA.7).

[Insert Table 6 and Table 7 about here]

Tables 6 and 7 show results for both the BMA approach and classical model selection, respectively. As can be seen from Table 6, allowing for more AR terms does not change our main conclusions about the macro-finance predictors discussed above. The market excess return (MKT), the short term reversal (STR) factor, and the default spread (DEF) still show up as the most important financial predictors with high posterior inclusion probabilities and large BMA t-ratios (in absolute terms). Regarding the AR terms themselves, we find that they also show up with high posterior inclusion probabilities, but that many financial predictors actually dominate them in terms of posterior inclusion probabilities. Thus, including economic and financial information still seems valuable even after accounting for these higher-order autoregressive terms.

We find very similar results for the classical model selection procedure. The top three models all select the first two lags of realized volatility whereas the third autoregressive lag is selected only in case of model (i) and (ii). However, we also find that all three top models select the market return (MKT), the STR factor, and the default spread (DEF), and that predictive coefficients for all three are highly statistically significant. In addition, the earnings-price ratio (E-P) shows up in models (ii) and (iii). Hence, these results are very much in line with what we find for the benchmark case with only one lag of volatility in Table 4 above.

4.3. Out-of-Sample Analysis

We proceed by investigating the out-of-sample (OOS) predictive power of macro-finance variables for future financial volatility. This exercise seems interesting since it is relevant to know whether market participants could usefully employ macro-finance information to improve their volatility forecasts in a real-time setting. However, it should also be noted that performing OOS tests also has drawbacks since it reduces the power of tests due to relying on shorter sub-samples (see e.g. Inoue and Kilian, 2004). Hence, our in-sample tests are best suited to investigate whether there is a link between macro-finance information and future volatility in the population whereas the out-of-sample tests in this section are more suitable for investigating whether this link can actually be exploited in real time.

Summary statistics for the evaluation of out-of-sample forecasts are reported in Tables 8 and 9. The out-of-sample forecasts are generated recursively with an expanding window.[15] We evaluate the out-of-sample forecast results based upon three types of model selection approaches (best model according to BIC, AIC, and R^2) as well as three variants of forecast combination methods (BMA based on analytical evaluation of the models after trimming the set of predictors, BMA weights from the MC^3 sampling approach, and equal weights of the forecasts by all evaluated models (EW)).[16]

To conduct inference on forecast accuracy, we also report bootstrap p-values to test if the macro-finance augmented models outperform the benchmark in terms of mean square forecast

[15] We also checked the OOS performance based on rolling window schemes, which might be a better choice in case of structural breaks (see e.g. Pesaran and Timmermann, 2007; Pesaran and Pick, 2009, on the tradeoff between bias-reduction and efficiency when choosing estimation windows for forecasting). Our results indicate superior results for expanding window schemes, which suggests that efficiency gains due to less estimation uncertainty outweigh potential bias-reduction in the presence of structural breaks.

[16] For computational reasons, the number of Monte-Carlo draws for the MC^3 algorithm in the out-of-sample exercise is set to 1,000 as opposed to the in-sample results which are obtained with 500,000 draws.

errors. We rely on a model-based wild bootstrap imposing the null of no predictability by macro-finance variables as described in Section 3 and Appendix B.3. These bootstrap p-values ($\#\mathrm{TU}^{bs} < TU$) are computed as the proportion of Theil's U statistics in the artificial bootstrap samples that are smaller than the sample Theil's U. Thus, these p-values are one-sided and test the null of equal predictive performance against the alternative of superior performance of the model including macro-finance predictors against the benchmark.

[Insert Table 8 about here]

OOS Predictability: U.S. Equity Market (Long-term): Table 8 reports out-of-sample forecast evaluation results for the long-term U.S. equity sample for different starting values for the first forecast (i.e. different lengths of the initialization periods). Models augmented by economic variables generally tend to outperform a simple autoregressive benchmark. This is corroborated by the bootstrap p-values and the Giacomini and White (2006) tests which indicate superior forecast performance relative to the benchmark. There is no particular forecast method, however, that appears preferable over the others in terms of mean-squared errors over all sub-periods. Forecast combination approaches, however, tend to do well in that the Mincer-Zarnowitz forecast optimality restrictions are in many cases not rejected.

As indicated by the previous in-sample BMA results which include higher order AR terms (Table 6), many macro-finance predictors receive higher probabilities of inclusion compared to AR(2) or AR(3) terms. These results indicate that a first-order autoregressive model serves as a natural candidate benchmark. In the Online Appendix, we also report additional results where we consider the AR(3) model as an alternative benchmark. Also when higher order autoregressive benchmarks are considered, the macro-finance augmented forecasting models perform fairly well in terms of out-of-sample forecast accuracy: Theil's U is generally smaller than one, the OOS-R^2 are sizable and the statistical inference based on the bootstrap and the Giacomini-White test indicate that forecasts relying on economic variables outperform forecasts by the AR(3) model. While these benchmarks are admittedly simple, and more sophisticated reduced-form time series models – which may capture some features of the data better – could potentially generate better forecasts, we deem it an encouraging success that macro-finance predictors perform generally well in these simple out-of-sample forecast experiments.

[Insert Table 9 about here]

OOS Predictability: Multi-asset Class (Short-term): Out-of-sample evaluation results for the short-term multi asset class sample are reported in Table 9. While these results indicate that the macro-finance augmented models perform rather well vis-à-vis the AR(1) benchmark, the results are less favorable compared to the case of the long-term U.S. equity sample. Moreover, results reported in the Online Appendix (Table IA.10) show less out-of-sample success for macro-finance augmented models when higher order benchmarks are considered. This echoes results by Paye (2012) that macro-finance variables are less powerful predictors for equity volatility in recent data. In the same vein, our results on the differences in predictability of U.S. equity volatility across the two samples indicate that it is generally useful to rely on long-term data in order to investigate the predictive power by macro-finance variables.

Dynamic Out-of-Sample Performance. We investigate the dynamics of out-of-sample predictability for the long-term equity volatility sample in Figure 3. The figure is based on Net-SSE plots similar to Goyal and Welch (2003, 2008). The graphs show the cumulative sum of squared forecast errors of the benchmark model minus the squared errors of a forecast model based on economic variables: Net-SSE$(\tau) = \sum_{t=1}^{\tau}(e_{b,t}^2 - e_{a,t}^2)$, where $e_{b,t}$ is the forecast error of the benchmark (AR(1) or AR(3)), and $e_{a,t}$ is the error of the model of interest. Hence, a positive slope in the figure indicates a superior performance of the macro-finance augmented model relative to the benchmark at a particular point in time.

[Insert Figure 3 about here]

The plot of dynamic out-of-sample performance shows an overall very good performance of the augmented model for U.S. equity volatility for both considered benchmarks. As is well known from the literature on the predictability of returns (Goyal and Welch, 2008; Timmermann, 2008) the OOS forecast performance can be fairly variable over time. For instance, models based on macro-finance predictors performed rather poorly in the late 1990s / early 2000s. Interestingly, however, macro-finance variables provided informative predictive content beyond autoregressive benchmarks over the most recent 2007-2009 financial crisis.

5. Conclusion

This study provides a comprehensive analysis of volatility predictability in financial markets by economic variables. The main goal is to shed more light on the ultimate economic sources of financial volatility. Compared to the previous literature, we extend the analysis in three directions: First, we employ a long sample of stock market volatility and also look at the volatility of three additional asset classes, namely foreign exchange, bonds, and commodities. Second, we allow for a comprehensive set of predictive variables which goes far beyond existing studies in the literature on the economic drivers of volatility. Third, we use comprehensive model selection and forecast combination procedures to assess whether economic variables are useful and robust predictors of financial volatility.

We find that there is meaningful information contained in economic variables that helps in predicting future volatility for all four asset classes under study. Importantly, this predictive content by economic variables goes beyond the information contained in the history of the time series of realized volatility. Our results are also supportive of financial volatility predictability by macroeconomic and financial variables in a realistic out-of-sample setting. The economic variables that are the most robust predictors of volatility are those that have sensible economic interpretations. In particular, variables that proxy for credit risk and funding (il)liquidity consistently show up as common predictors of volatility across several asset classes. Variables capturing time-varying risk premia (such as valuation ratios for equities, or interest rate differentials in foreign exchange) also perform well as predictors of volatility. In contrast to these financial predictors, variables that proxy for macroeconomic conditions, are much less informative about future volatility. Thus, our results suggest that especially channels that emphasize the effects of leverage, credit risk and funding illiquidity as well as time-variation of risk premia are the most promising candidates for understanding the economic drivers of financial volatility.

References

Andersen, T. G., T. Bollerslev, P. Christoffersen, and F. X. Diebold (2006): "Volatility and Correlation Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliot, C. W. Granger, and A. Timmermann, pp. 777–878. Elsevier, Amsterdam.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 529–626.

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Ang, A., and G. Bekaert (2007): "Stock Return Predictability: Is it there?," *Review of Financial Studies*, 20(3), 651–707.

Ang, A., R. Hodrick, Y. Xing, and X. Zhang (2006): "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61, 259–299.

Avramov, D. (2002): "Stock Return Predictability and Model Uncertainty," *Journal of Financial Economics*, 64(3), 423–458.

Baba, N. (2009): "Dynamic Spillover of Money Market Turmoil from FX Swap to Cross-Currency Swap Markets: Evidence from the 2007-2008 Turmoil," *Journal of Fixed Income*, 18, 24–38.

Baker, M., and J. Wurgler (2007): "Investor Sentiment in the Stock Market," *Journal of Economic Perspectives*, 21, 129–151.

Black, F. (1976): "Studies in Stock Price Volatility Changes," *American Statistical Association, Proceedings of the Business and Economic Statistics Section*, pp. 177–181.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., G. Tauchen, and H. Zhou (2009): "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies*, 22(11), 4463–4492.

Brown, G. W., and M. T. Cliff (2005): "Investor Sentiment and Asset Valuation," *Journal of Business*, 78, 405–440.

Brunnermeier, M., S. Nagel, and L. Pedersen (2009): "Carry Trades and Currency Crashes," *NBER Macroeconomics Annual 2008*, 23, 313–347.

Brunnermeier, M., and L. Pedersen (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201–2238.

Cakmakli, C., and D. van Dijk (2010): "Getting the Most Out of Macroeconomic Information for Predicting Stock Returns and Volatility," Tinbergen Institute Discussion Paper 2010-115/4.

Campbell, J. Y., and R. J. Shiller (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1(3), 195–228.

——— (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495–514.

Campbell, J. Y., and S. B. Thompson (2008): "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies*, 21, 1509–1531.

Chauvet, M., Z. Senyuz, and E. Yoldas (2010): "What Does Realized Volatility Tell Us About Macroeconomic Fluctuations," Working Paper.

Christiansen, C., A. Ranaldo, and P. Söderlind (2011): "The Time-Varying Systematic Risk of Carry Trade Strategies," *Journal of Financial and Quantitative Analysis*, 46(4), 1107–1125.

Clark, T. E., and K. D. West (2006): "Using Out-of-sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis," *Journal of Econometrics*, 135, 155–186.

——— (2007): "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models," *Journal of Econometrics*, 138, 291–311.

Cochrane, J. H., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95(1), 138–160.

Cooper, I., and R. Priestley (2009): "Time-Varying Risk Premiums and the Output Gap," *Review of Financial Studies*, 22(7), 2801–2833.

Cremers, K. J. M. (2002): "Stock Return Predictability: A Bayesian Model Selection Perspective," *The Review of Financial Studies*, 15(4), 1223–1249.

Da, Z., and E. Schaumburg (2009): "The Pricing of Volatility Risk across Asset Classes and the Fama-French Factors," Working Paper, Northwestern University.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Diebold, F. X., and K. Yilmaz (2009): "Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets," *Economic Journal*, 119, 158–171.

Diebold, F. X., and K. Yilmaz (2010): “Macroeconomic Volatility and Stock Market Volatility, Worldwide,” in *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, ed. by T. Bollerslev, J. Russell, and M. Watson, pp. 97–116. Oxford University Press, Oxford.

Engle, R. F. (1982): “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 50, 987–1007.

Engle, R. F., E. Ghysels, and B. Sohn (2008): “On the Economic Sources of Stock Market Volatility,” *Working Paper*.

Fama, E. F., and K. R. French (1993): “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, 33, 3–56.

Faust, J., S. Gilchrist, J. H. Wright, and E. Zakrajsek (2011): “Credit Spreads as Predictors of Real-Time Economic Activity: A Bayesian Model-Averaging Approach,” Working paper, John Hopkins University.

Fernandez, C., E. Ley, and M. F. J. Steel (2001): “Benchmark Priors for Bayesian Model Averaging,” *Journal of Econometrics*, 100, 381–427.

Fleming, J., C. Kirby, and B. Ostdiek (1998): “Information and Volatility Linkages in the Stock, Bond, and Money Markets,” *Journal of Financial Economics*, 49, 111–137.

Fong, W. M., and K. H. See (2001): “Modelling the Conditional Volatility of Commodity Index Futures as a Regime Switching Process,” *Journal of Applied Econometrics*, 16, 133–163.

Fornari, F., and A. Mele (2010): “Financial Volatility and Economic Activity,” Working Paper, LSE.

Giacomini, R., and H. White (2006): “Tests of Conditional Predictive Ability,” *Econometrica*, 74, 1545–1578.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993): “On the Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,” *Journal of Finance*, 48, 1779–1801.

Goyal, A., and I. Welch (2003): “Predicting the Equity Premium with Dividend Ratios,” *Management Science*, 49, 639–654.

——— (2008): “A Comprehensive Look at the Empirical Performance of Equity Premium Prediction,” *Review of Financial Studies*, 21, 1455–1508.

Inoue, A., and L. Kilian (2004): “In-sample or Out of Sample Tests of Predictability: Which One Should We Use?,” *Econometric Reviews*, 23(4), 371–402.

Koop, G. (2003): *Bayesian Econometrics*. Wiley, Chichester.

Leamer, E. E. (1978): *Specification Searches: Ad Hoc inference with Non-experimental data.* Wiley.

Ludvigson, S. C., and S. Ng (2009): "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, 22(12), 5027–5067.

Lustig, H., N. Roussanov, and A. Verdelhan (2011): "Common Risk Factors in Currency Markets," *Review of Financial Studies*, 24, 3731–3777, Working Paper UCLA.

Lustig, H., N. L. Roussanov, and A. Verdelhan (2010): "Countercyclical Currency Risk Premia," Working Paper, UCLA.

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

——— (2008): "Understanding Stock Market Volatility – A Business Cycle Perspective," Working Paper, LSE.

Menkhoff, L., L. Sarno, M. Schmeling, and A. Schrimpf (2012): "Carry Trades and Global Foreign Exchange Volatility," *Journal of Finance*, forthcoming.

Merton, R. C. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29, 449–470.

Nagel, S. (2012): "Evaporating Liquidity," *Review of Financial Studies*, forthcoming.

Nelson, D. B. (1991): "Conditional heteroskedacticity in asset returns," *Econometrica*, 59, 347–370.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–708.

Obstfeld, M., and K. Rogoff (2000): "The Six Major Puzzles in International Macroeconomics: Is there a Common Cause?," in *NBER Macroeconomics Annual 2000*, ed. by B. Bernanke, and K. Rogoff. MIT Press, Cambridge.

Pastor, L., and R. Stambaugh (2003): "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111, 642–685.

Patton, A. J., and K. Sheppard (2009): "Optimal Combinations of Realised Volatility Estimators," *International Journal of Forecasting*, 25, 218–238.

Paye, B. S. (2012): "Deja Vol: Predictive Regressions for Aggregate Stock Market Volatility Using Macroeconomic Variables," *Journal of Financial Economics (forthcoming)*.

Pesaran, M., and A. Pick (2009): "Forecast Combination across Estimation Windows," *Journal of Business and Economic Statistics*, 29, 307–318.

Pesaran, M., and A. Timmermann (2007): "Selection of Estimation Window in the Presence of Breaks," *Journal of Econometrics*, 137, 134–161.

Scheinkman, J., and W. Xiong (2003): "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 111, 1183–1219.

Schrimpf, A. (2010): "International Stock Return Predictability Under Model Uncertainty," *Journal of International Money and Finance*, 29, 1256–1282.

Schwert, G. W. (1989): "Why does Stock Market Volatility Change over Time?," *Journal of Finance*, 44, 1115–1153.

Stock, J. H., and M. W. Watson (2004): "An Empirical Comparison of Methods for Forecasting Using Many Predictors," Working paper, Harvard University.

Timmermann, A. (2006): "Forecast Combinations," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. W. J. Granger, and A. Timmermann. Amsterdam: North-Holland.

Timmermann, A. (2008): "Elusive Return Predictability," *International Journal of Forecasting*, 24, 1–18.

Wright, J. H. (2008): "Bayesian Model Averaging and exchange rate forecasts," *Journal of Econometrics*, 146, 329–341.

Zellner, A. (1986): *Bayesian Inference and Decision Techniques: Essays in Honour of Bruno de Finetti*chap. On Assessing Prior Distributions and Bayesian Regression Analysis With g-Prior Distributions, pp. 233–243. North Holland, Amsterdam.

A. Data Description

[Insert Table A.1 about here]

B. Methodological Details

In this Appendix, we provide some additional details on the Bayesian methods that are underlying the results discussed in the main text. We first describe the elicitation of prior distributions in the Bayesian Model Averaging (BMA) setup. We then provide some details on the Markov Chain Monte Carlo Model Composition algorithm (MC^3) which is used for sampling from the set of models $\mathcal{M}_1, ..., \mathcal{M}_{2^\kappa}$.

B.1. Prior Elicitation

For ease of exposition, we denote the dependent variable by Y, which is a $T \times 1$ vector of realized volatility as in Eq. (2). The predictive variables are collected in a matrix Z_j which has dimension $T \times k_j$ depending on the particular model $\mathcal{M}_j$. We are considering a linear regression model with i.i.d. errors which are assumed to be normal with mean zero and variance σ^2. It is common in the BMA setup to work with the strict exogeneity assumption of the regressors such that a closed form expression for the likelihood can be derived (See, e.g. Wright, 2008).[17]

We follow most of the extant BMA literature and choose to work with a natural conjugate prior distribution for the model parameters $p(\beta_j|\mathcal{M}_j)$ and $p(\sigma^2)$. Thus, our prior on the predictive coefficients β_j conditional on σ^2 is taken to be a normal distribution

$$\beta_j|\sigma^2 \sim \mathcal{N}(0, \sigma^2\phi(Z_j'Z_j)^{-1}), \tag{B.1}$$

which is centered around zero, i.e. it is expected a-priori that there is no predictive power by the

[17]Of course, in a time-series setup as the one considered here, strict exogeneity is typically violated. Nevertheless, given that this violation is generally considered to be of minor relevance for the forecasting problem, the literature (e.g. Stock and Watson, 2004; Wright, 2008; Faust, Gilchrist, Wright, and Zakrajsek, 2011) generally assumes strict exogeneity, which provides an elegant theoretical framework for model averaging.

economic variables, and ϕ is a hyperparameter. A higher ϕ means a less informative prior (i.e. a higher prior variance), whereas a lower ϕ (approaching zero) induces more shrinkage towards the non-forecastability case. This prior specification is also known as a so-called g-prior framework and is originally due to Zellner (1986). The prior on the predictive coefficients is proper – an important feature to obtain meaningful Bayes factors for model comparison – but it is relatively uninformative, where the amount of informativeness is controlled by the ϕ hyperparameter. The prior on σ^2 is a standard improper prior, proportional to $1/\sigma^2$.

Given these assumptions, the expression for the marginal likelihood takes the following form

$$p(D|\mathcal{M}_j) \propto (1+\phi)^{-k_j/2} S_j^{-T}, \tag{B.2}$$

where $S_j^2 = Y'Y - Y'Z_j(Z_j'Z_j)^{-1}Z_j'\frac{\phi}{1+\phi}$. The expression in (B.2) is important since it enters Eq. (4) and thus plays an essential role for the computation of posterior model probabilities $p(\mathcal{M}_j|D)$. Given the likelihood and the prior, the posterior mean of the predictive coefficients takes the form

$$\beta_j|D = \frac{\phi}{1+\phi}(Z_j'Z_j)^{-1}Z_j'Y. \tag{B.3}$$

In this BMA setup there are two modeling choices which require input by the researcher. First, the hyperparameter ϕ must be selected, which controls the degree of informativeness of the prior on the predictive coefficients. We select the ϕ hyperparameter according to the simulation-based recommendations in Fernandez, Ley, and Steel (2001). The second choice is that we assign equal prior probability on the models, i.e. we take $1/2^\kappa$ as the prior model probability $p(M_j)$. This implies a prior probability of inclusion for each predictive variable of $\pi = 1/2$ as in Faust, Gilchrist, Wright, and Zakrajsek (2011).

B.2. MC^3 Algorithm

The MC^3 algorithm is a Markov Chain Monte Carlo method of sampling from the distribution of models and has similarities with a Metropolis-Hastings algorithm. For each run r of the algorithm, a candidate model M^* is drawn from the model space $\mathcal{M}_1, ..., \mathcal{M}_{2^\kappa}$ which can either be accepted – if it improves on the model drawn in the previous draw $M^{(r-1)}$ – otherwise it is rejected. If the drawn model is rejected then the chain remains at the previous model $M^{(r-1)}$. The acceptance probability $\Xi(M^{(r-1)}, M^*)$ is expressed as

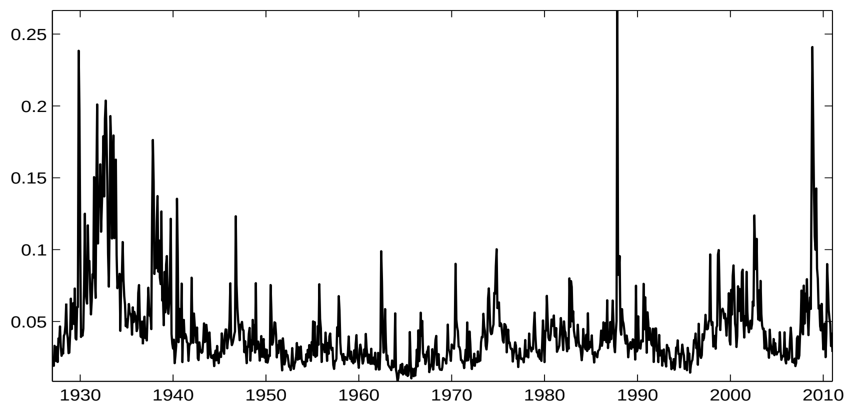
$$\Xi(M^{(r-1)}, M^*) = \min \left\{ \frac{p(D|M^*)p(M^*)}{p(D|M^{(r-1)})p(M^{(r-1)})}; 1 \right\}, \tag{B.4}$$

and depends on a comparison of the marginal likelihoods of the drawn model vis-a-vis the previous model of the chain as well as a comparison of the model priors (which are equal in our case). If the number of Monte Carlo draws is large (in our case 500,000) the fraction of draws for the different models converges to the posterior model probability. In order to ensure that the starting value of the chain does not affect the results a burn-in period of 50,000 draws is used.

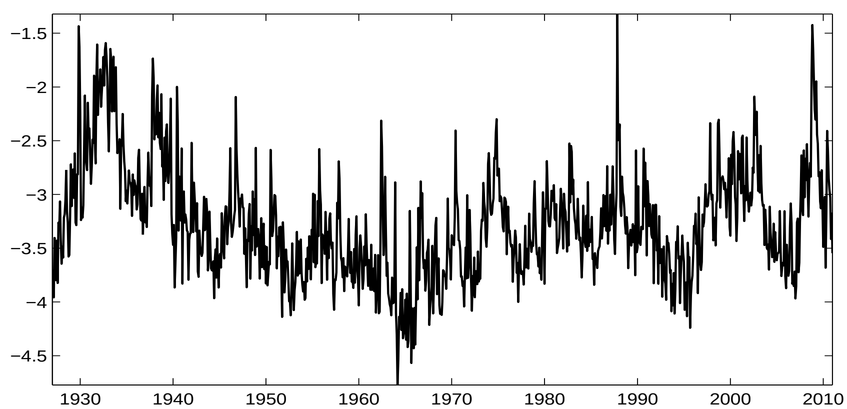
B.3. Bootstrap Procedure for Out-of-Sample Evaluation

The bootstrap procedure is a model-based wild bootstrap (imposing the null of no predictability by macro-finance variables) and is a variant of the approach considered in Clark and West (2006). The wild bootstrap ensures accurate inference in the presence of conditional heteroskedasticity. In each bootstrap iteration the following steps are performed: (i) A series of i.i.d. standard normal innovations η_t is drawn. (ii) AR(1) models are fitted for both the dependent variables $RV_{i;t}$ as well as each of the κ macro-finance variables in z_t and the residuals ($\hat{\epsilon}_t$, $\hat{\nu}_t$) are saved. (iii) Artificial bootstrap series $RV_{i;t}^{bs}$ and z_t^{bs} are constructed based on the estimated AR(1) parameters and the innovations $\hat{\epsilon}_t\eta_t$, $\hat{\nu}_t\eta_t$. The starting observations of the bootstrap series $RV_{i;0}^{bs}$ and z_0^{bs} are drawn randomly from the actual series. (iv) The artificial bootstrap data are used to generate recursive forecasts based on models relying on the bootstrapped explanatory macro-finance variables as well as the benchmark AR(1). The corresponding Theil's U statistics $\text{TU}^b s$ are computed. (v) We compute bootstrap p-values as the fraction of times that Theil's U in the bootstrap samples is below the one observed in-sample. Hence, these p-values are one-sided and test the null of equal predictive performance against the alternative of superior performance of the model including macro-finance predictors vis-a-vis the benchmark. The number of bootstrap iterations is set to 1,000.

Figure 1. U.S. equity market volatility over the long-run (1926-2010)



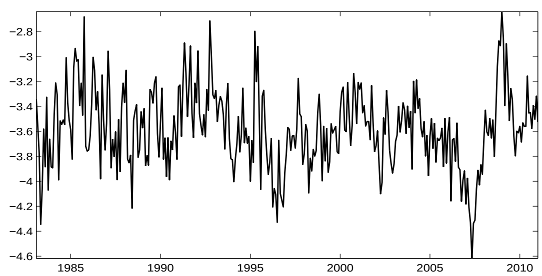
(a) Realized Volatility U.S. Stock Market



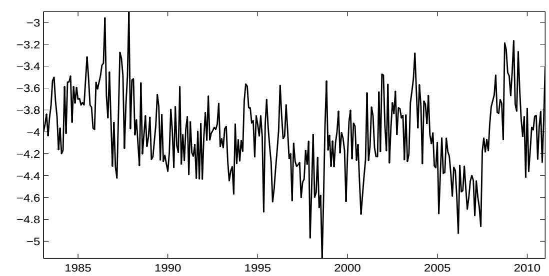
(b) (Log) Realized Volatility U.S. Stock Market

Notes: This figure shows realized aggregate U.S. stock market volatility over the long-term sample period 12/1926-12/2010. Panel A shows the level of realized equity market volatility while Panel B shows log realized volatility.

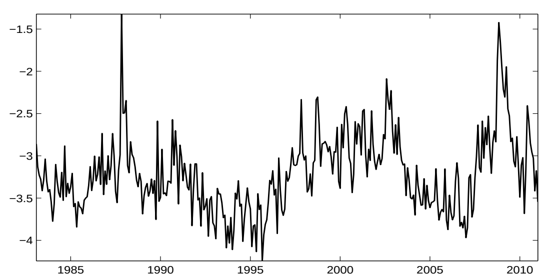
Figure 2. Volatility - Other Asset Classes (Short-term Sample)



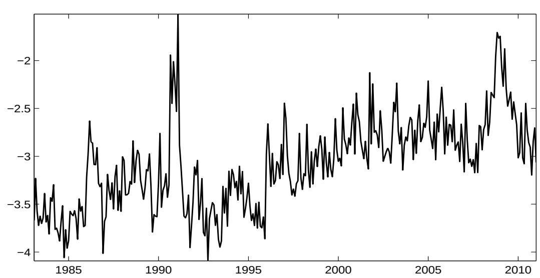
(a) FX-Aggregate



(b) Bonds



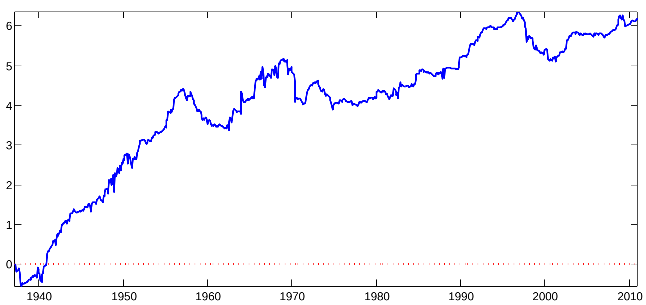
(c) Stocks



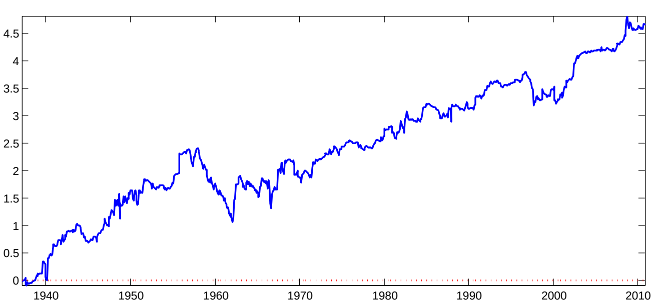
(d) Commodities

Notes: This figure shows (log) realized volatility for several asset classes over the short-term sample period from 01/1983-12/2010. Panel A shows an aggregate measure of foreign exchange market volatility (plots for major rates are shown in the internet appendix), Panel B shows log realized volatility for U.S. bonds, Panel C shows log realized volatility for the U.S. equity market for a shorter sample period than Figure 1, Panel D shows the evolution of commodity market volatility calculated from daily returns on the S&P GSCI index.

Figure 3. Equity Volatility (Long-Term Sample): Time-Variation of Out-of-Sample Performance



(a) AR(1)



(b) AR(3)

Notes: The figure shows the time-variation of the out-of-sample performance of forecasts based on the MC^3-BMA approach. Net-SSE is the cumulated difference of squared forecast errors of the benchmark model and the model of interest (BMA): Net-SSE(τ) = $\sum_{t=1}^{\tau}(e_{b,t}^2 - e_{a,t}^2)$, where $e_{b,t}$ is the forecast error of the benchmark, and $e_{a,t}$ is the error of the model of interest. An increase of the slope represents a better forecast performance of the forecast model at the particular point in time. Panel A considers an AR(1) model for realized volatility as the benchmark while Panel B considers an AR(3) benchmark.

Table 1. Summary Statistics: Realized Volatility

Panel A.	Mean	Std.	Skew.	Kurt.	JB p-val.	AC(1)	AC(2)	AC(3)
Stocks (1926-2010)	-3.27	0.52	0.72	3.80	0.00	0.78	0.70	0.64
EURUSD	-3.58	0.31	0.07	3.36	0.38	0.53	0.42	0.32
JPYUSD	-3.58	0.34	0.20	3.56	0.04	0.42	0.34	0.26
GBPUSD	-3.67	0.36	0.40	3.25	0.01	0.61	0.56	0.46
CHFUSD	-3.49	0.29	0.15	3.52	0.09	0.45	0.38	0.27
FX-Aggr.	-4.06	0.39	-0.11	3.36	0.33	0.66	0.58	0.51
Bonds	-4.01	0.36	0.00	3.11	0.94	0.58	0.49	0.49
Stocks (1983-2010)	-3.22	0.45	0.81	4.44	0.00	0.71	0.63	0.58
Commod.	-3.08	0.47	0.35	3.03	0.03	0.77	0.73	0.68

Panel B.	**Correlations**							
	EURUSD	JPYUSD	GBPUSD	CHFUSD	FX-Aggr.	Bonds	Stocks	Commod.
EURUSD	1.00							
JPYUSD	0.39	1.00						
GBPUSD	0.73	0.23	1.00					
CHFUSD	0.88	0.47	0.69	1.00				
FX-Aggregate	0.81	0.33	0.82	0.73	1.00			
Bonds	0.33	0.14	0.33	0.33	0.34	1.00		
Stocks	0.23	0.38	0.14	0.17	0.25	0.32	1.00	
Commod.	0.02	0.27	0.02	0.04	0.03	0.04	0.48	1.00

Notes: The table shows summary statistics of realized volatility for stock markets, foreign exchange (FX), bonds, and commodity markets. The realized volatility series are defined as the log of the square root of the realized variance. The reported statistics in Panel A include the mean, standard deviation (Std.), Skewness (Skew.), Kurtosis (Kurt.), the p-value from the Jarque-Bera test for normality (JB p-val.) as well as first (AC(1)), second (AC(2)), and third order (AC(3)) autocorrelation coefficients. Panel B reports the correlations between the different volatility series. The sample period is from 01/1983-12/2010 with the exception of the equity market volatility where also statistics based on the long-term sample from 12/1926-12/2010 are reported.

Table 2. Predictive Regressions for Equity Market Volatility (1926-2010), BMA

Composite Model			Post.	Post.			*Top 5 Models*				
No.	Variable	$\pi\|D$	Mean	STD	t-ratio		(i)	(ii)	(iii)	(iv)	(v)
1	RV(t-1)	**1.00**	0.617	0.027	22.95		1	1	1	1	1
2	DEF	**1.00**	0.117	0.014	8.51		1	1	1	1	1
3	MKT	**1.00**	-0.051	0.011	-4.53		1	1	1	1	1
4	E-P	**0.97**	-0.037	0.012	-2.99		1	1	1	1	1
5	STR	**0.89**	-0.030	0.015	-2.05		0	1	1	1	1
6	T-B	0.07	-0.001	0.005	-0.21		0	1	0	0	0
7	LTR	0.05	-0.001	0.003	-0.17		0	0	1	0	0
8	T-S	0.05	-0.001	0.004	-0.16		0	0	0	1	0
R^2_a	0.645					R^2	0.645	0.641	0.646	0.646	0.645
R^2_b	0.606					$\bar{R}^2$	0.643	0.640	0.644	0.643	0.643

Notes: This table reports in-sample predictability results for U.S. equity market volatility (long-term sample) obtained from a Bayesian Model Averaging approach with an MC^3 algorithm. The results are obtained with a set of predictors which contains the lagged dependent variable RV(t-1); results based on a benchmark allowing for higher order AR terms are reported in Table 6. The table displays the results for the 8 best predictors, as ranked according to the posterior probability of inclusion $\pi|D$ (sorted in descending order). The table reports posterior means, standard deviation and BMA t-ratios of the best predictors (reflecting model uncertainty). Inclusion of the specific variable in the Top 5 models (according to the posterior model probability) is indicated by 1. R^2_a denotes a pseudo-R^2 based on the composite Bayesian model, R^2_b shows the R^2 of the benchmark model. Unadjusted (adjusted) R^2 ($\bar{R}^2$) are reported for the five best model specifications. Table A.1 contains a description of the abbreviations for the different predictive variables. The sample period is 12/1926-12/2010.

Table 3. Predictability of Volatility for Several Asset Classes (1983-2010)

Panel A.	*Composite Model*			Post.	Post.			*Top 5 Models*				
Aggregate-FX	No.	Variable	$\pi \mid D$	Mean	STD	t-ratio		(i)	(ii)	(iii)	(iv)	(v)
	1	RV(t-1)	**1.00**	0.426	0.062	6.92		1	1	1	1	1
	2	AFD	**0.99**	0.073	0.023	3.16		1	1	1	1	1
	3	M1A	**0.93**	0.072	0.032	2.25		1	1	1	1	1
	4	INFA	**0.84**	0.052	0.032	1.63		1	1	1	1	0
	5	TED	**0.77**	0.044	0.031	1.40		1	1	1	1	1
	6	IPM	**0.69**	-0.043	0.044	-0.98		1	1	0	1	1
	7	H-S	**0.58**	-0.028	0.029	-0.96		0	0	1	1	0
	8	INFM	0.34	-0.012	0.021	-0.59		0	1	0	1	0
	R_a^2	0.545					R^2	0.541	0.532	0.540	0.539	0.547
	R_b^2	0.433					$\bar{R}^2$	0.531	0.523	0.530	0.529	0.536
Panel B.	*Composite Model*			Post.	Post.			*Top 5 Models*				
Bonds	No.	Variable	$\pi \mid D$	Mean	STD	t-ratio		(i)	(ii)	(iii)	(iv)	(v)
	1	RV(t-1)	**1.00**	0.413	0.071	5.80		1	1	1	1	1
	2	TURN	**0.95**	-0.049	0.022	-2.23		1	1	1	1	1
	3	DEF	**0.86**	0.060	0.034	1.75		1	1	1	1	1
	4	T-S	**0.73**	0.045	0.035	1.28		0	1	1	1	0
	5	TED	0.29	0.012	0.023	0.52		0	1	0	0	0
	6	M1A	0.23	0.011	0.025	0.44		0	0	0	0	1
	7	C-P	0.21	0.009	0.021	0.42		1	0	0	0	0
	8	CAP	0.17	-0.004	0.017	-0.24		0	0	0	1	0
	R_a^2	0.433					R^2	0.422	0.419	0.429	0.428	0.427
	R_b^2	0.335					$\bar{R}^2$	0.415	0.412	0.420	0.419	0.418
Panel C.	*Composite Model*			Post.	Post.			*Top 5 Models*				
Stocks	No.	Variable	$\pi \mid D$	Mean	STD	t-ratio		(i)	(ii)	(iii)	(iv)	(v)
	1	RV(t-1)	**1.00**	0.450	0.064	7.08		1	1	1	1	1
	2	D-P	**0.96**	-0.119	0.041	-2.92		1	1	1	1	1
	3	TED	**0.92**	0.070	0.032	2.17		1	1	1	1	1
	4	DEF	**0.92**	0.095	0.041	2.34		1	1	1	1	1
	5	MKT	**0.69**	-0.042	0.033	-1.25		1	0	1	1	1
	6	E-P	0.33	-0.022	0.039	-0.58		1	0	0	0	0
	7	MSCI	0.26	-0.012	0.026	-0.48		0	1	0	0	0
	8	AFD	0.22	0.009	0.019	0.45		0	0	1	0	0
	R_a^2	0.591					R^2	0.579	0.586	0.576	0.584	0.583
	R_b^2	0.507					$\bar{R}^2$	0.573	0.578	0.570	0.576	0.576
Panel D.	*Composite Model*			Post.	Post.			*Top 5 Models*				
Commod.	No.	Variable	$\pi \mid D$	Mean	STD	t-ratio		(i)	(ii)	(iii)	(iv)	(v)
	1	RV(t-1)	**1.00**	0.596	0.059	10.18		1	1	1	1	1
	2	TURN	**0.67**	-0.028	0.025	-1.14		1	1	0	1	0
	3	C-P	**0.60**	-0.036	0.035	-1.00		1	0	1	1	1
	4	T-B	**0.55**	-0.037	0.040	-0.92		1	0	1	0	1
	5	D-P	0.49	-0.031	0.037	-0.83		0	1	0	1	0
	6	DIFF	0.44	-0.023	0.031	-0.73		1	0	1	0	1
	7	H-S	0.43	-0.021	0.029	-0.73		0	1	0	0	0
	8	M1M	0.41	-0.016	0.023	-0.70		0	0	1	0	0
	R_a^2	0.659					R^2	0.656	0.650	0.648	0.648	0.648
	R_b^2	0.597					$\bar{R}^2$	0.650	0.645	0.643	0.643	0.643

Notes: This table reports in-sample predictability results for aggregate FX volatility (Panel A.), U.S. bond market volatility (Panel B.), equity market volatility (short-term sample, Panel C.) and commodity market volatility (Panel D.) obtained from a Bayesian Model Averaging approach with a MC^3 algorithm. The results are obtained with a set of predictors which contains the lagged dependent variable RV(t-1); results based on a benchmark allowing for higher order AR terms are reported in the Online Appendix. The table displays the results for the best 8 predictors, as ranked according to the posterior probability of inclusion $\pi|D$ (sorted in descending order). The table reports posterior means, standard deviation and BMA t-ratios of the best predictors (reflecting model uncertainty). Inclusion of the specific variable in the Top 5 models (according to the posterior model probability) is indicated by 1. R_a^2 denotes a pseudo-R^2 based on the composite Bayesian model, R_b^2 shows the R^2 of the benchmark model (AR(1)). Unadjusted and (adjusted) R^2 ($\bar{R}^2$) are reported for the best 5 model specifications. Table A.1 contains a description of the abbreviations for the different predictive variables. The sample period is 01/1983-12/2010.

Table 4. Equity Market Volatility (1926-2010), Classical Model Selection

	Top 3 Models		
	(i)	(ii)	(iii)
RV(t-1)	**0.62**	**0.61**	**0.61**
	22.73	22.07	22.12
MKT	**-0.05**	**-0.06**	**-0.05**
	-4.04	-4.17	-4.11
STR	**-0.03**		**-0.04**
	-3.66		-3.75
DEF	**0.12**	**0.11**	**0.12**
	9.53	9.44	9.53
E-P	**-0.04**	**-0.04**	**-0.04**
	-3.71	-4.00	-3.37
T-B			-0.01
			-1.51
R^2	0.645	0.641	0.646
$\bar{R}^2$	0.643	0.640	0.644
BIC	-2.285	-2.280	-2.279

Notes: The table shows results of in-sample predictive regressions for U.S. equity market volatility (long-term sample) based on a classical model selection approach. The results are obtained with a set of predictors which contains the lagged dependent variable RV(t-1), results based on a benchmark allowing for higher order AR terms are reported in Table 7. Predictive regressions results for the three top-performing models (based upon the BIC) are reported. Significant coefficients (at the 5% level based on HAC standard errors) are bold-printed and the corresponding classical t-statistics are reported below. Unadjusted (adjusted) R^2 are reported for the 3 best model specifications. Table A.1 contains a description of the abbreviations for the different predictive variables. The sample period is 12/1926-12/2010.

Table 5. Predictability of Volatility for Several Asset Classes (1983-2010), Classical Model Selection

A. FX-Aggr.	(i)	(ii)	(iii)	***B. Bonds***	(i)	(ii)	(iii)	***C. Stocks***	(i)	(ii)	(iii)	***C. Commod.***	(i)	(ii)	(iii)
RV(t-1)	**0.45**	**0.50**	**0.46**	RV(t-1)	**0.42**	**0.42**	**0.45**	RV(t-1)	**0.55**	**0.61**	**0.59**	RV(t-1)	**0.59**	**0.57**	**0.66**
	7.91	9.34	7.94		7.92	8.58	9.24		10.97	15.19	13.23		11.31	10.74	14.26
AFD	**0.07**	**0.09**	**0.07**	DEF	**0.07**	**0.09**	**0.06**	DEF	**0.05**			C-P	**-0.06**	**-0.07**	
	3.58	4.72	3.68		4.85	6.58	3.43		2.82				-4.32	-4.28	
TED	**0.07**	**0.08**	**0.07**	T-S	**0.06**			MKT	**-0.06**	**-0.06**	**-0.07**	T-B	**-0.07**	**-0.07**	
	5.76	5.68	5.75		3.21				-3.27	-3.17	-3.23		-3.47	-3.65	
M1A	**0.05**		**0.06**	M1A			**0.05**	E-P	**-0.07**	**-0.06**	**-0.07**	D-P			**-0.07**
	2.71		3.00				2.78		-3.87	-4.30	-4.49				-4.09
IPM	**-0.06**	**-0.06**						IPM		**-0.05**	**-0.05**	DIFF	**-0.05**	**-0.06**	
	-4.36	-4.37								-2.17	-2.09		-2.35	-2.40	
				C-P		**0.05**						H-S			**-0.07**
						3.33									-4.25
				TURN	**-0.05**	**-0.05**	**-0.05**	TURN		**-0.04**		TURN	**-0.04**		**-0.04**
					-3.22	-3.29	-3.41			-2.47			-2.60		-2.29
R^2	0.521	0.508	0.516	R^2	0.422	0.419	0.416	R^2	0.550	0.556	0.548	R^2	0.650	0.641	0.641
$\bar{R}^2$	0.513	0.502	0.509	$\bar{R}^2$	0.415	0.412	0.409	$\bar{R}^2$	0.544	0.549	0.543	$\bar{R}^2$	0.645	0.637	0.637
BIC	-2.480	-2.471	-2.471	BIC	-2.508	-2.503	-2.498	BIC	-2.285	-2.283	-2.282	BIC	-2.460	-2.454	-2.453

Notes: The table shows results of in-sample predictive regressions for aggregate FX volatility, U.S. bond market volatility, equity market volatility (short-term sample) and commodity market volatility obtained from a classical model selection approach. The benchmark is an AR(1) model, results allowing for higher order AR terms can be found in the internet appendix. Predictive regressions results for the three top-performing models (based upon the BIC) are reported. Significant coefficients (at the 5% level based on HAC standard errors) are bold-printed and the corresponding classical t-statistics are reported below. R^2 and adjusted R^2 ($\bar{R}^2$) are reported for the 3 best model specifications. Table A.1 contains a description of the abbreviations for the different predictive variables. The sample period is 01/1983-12/2010.

Table 6. Equity Market Volatility (1926-2010), BMA Results: Higher-order AR terms

Composite Model			Post.	Post.			*Top 5 Models*				
No.	Variable	$\pi \vert D$	Mean	STD	t-ratio		(i)	(ii)	(iii)	(iv)	(v)
1	RV(t-1)	**1.00**	0.456	0.035	13.08		1	1	1	1	1
2	MKT	**1.00**	-0.065	0.011	-5.91		1	1	1	1	1
3	DEF	**1.00**	0.077	0.015	5.20		1	1	1	1	1
4	RV(t-2)	**1.00**	0.178	0.039	4.53		1	1	1	1	1
5	STR	**0.99**	-0.041	0.011	-3.74		1	1	1	1	1
6	RV(t-3)	**0.93**	0.102	0.043	2.36		1	0	1	1	1
7	E-P	0.48	-0.012	0.015	-0.83		1	1	0	1	1
8	LTR	0.06	-0.001	0.004	-0.18		0	0	1	0	1
R^2_a	0.665					R^2	0.663	0.665	0.661	0.664	0.666
R^2_b	0.632					$\bar{R}^2$	0.661	0.663	0.659	0.661	0.663

Notes: This table reports in-sample predictability results for U.S. equity market volatility (long-term sample) obtained from a Bayesian Model Averaging approach with an MC^3 algorithm. Additional AR terms (RV(t-2) and RV(t-3)) are included as predictors in the model search besides one autoregressive lag (RV(t-1)). The table displays the results for the best 8 predictors, as ranked according to the posterior probability of inclusion $\pi|D$ (sorted in descending order). The table reports posterior means, standard deviation and BMA t-ratios of the best predictors (reflecting model uncertainty). Inclusion of the specific variable in the Top 5 models (according to the posterior model probability) is indicated by 1. R^2_a denotes a pseudo-R^2 based on the composite Bayesian model, R^2_b shows the R^2 of the benchmark model. Unadjusted (adjusted) R^2 are reported for the 5 best model specifications. Table A.1 contains a description of the abbreviations for the different predictive variables. The sample period is 12/1926-12/2010.

Table 7. Equity Market Volatility (1926-2010), Classical Model Selection: Higher-order AR terms

	Top 3 Models		
	(i)	(ii)	(iii)
RV(t-1)	**0.46**	**0.45**	**0.47**
	12.57	12.50	14.14
RV(t-2)	**0.18**	**0.17**	**0.22**
	4.52	4.42	6.32
RV(t-3)	**0.11**	**0.10**	
	3.82	3.47	
MKT	**-0.06**	**-0.07**	**-0.06**
	-5.02	-5.11	-4.93
STR	**-0.04**	**-0.04**	**-0.04**
	-4.54	-4.29	-4.12
DEF	**0.07**	**0.08**	**0.09**
	6.08	6.43	7.36
E-P		**-0.03**	**-0.03**
		-2.57	-3.01
R^2	0.663	0.665	0.661
$\bar{R}^2$	0.661	0.663	0.659
BIC	-2.330	-2.330	-2.326

Notes: The table shows results of in-sample predictive regressions for U.S. equity market volatility (long-term sample) based on a classical model selection approach. Additional AR terms (RV(t-2) and RV(t-3)) are included as predictors besides one autoregressive term of the dependent variable (RV(t-1)). Predictive regressions results for the three top-performing models (based upon the BIC) are reported. Significant coefficients (at the 5% level based on HAC standard errors) are bold-printed and the corresponding classical t-statistics are reported below. Unadjusted (adjusted) R^2 are reported for the 3 best model specifications. Table A.1 contains a description of the abbreviations for the different predictive variables. The sample period is 12/1926-12/2010.

Table 8. Out-of-Sample Forecast Evaluation: Equity Volatility (Long-term sample)

	BIC	AIC	R^2	BMA	MC^3	EW
Start: 01/1937						
Theil's U	0.970	0.969	0.970	0.965	0.966	0.962
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.00	0.00	0.00	0.00	0.00	0.00
R^2_{OOS}	0.059	0.062	0.059	0.069	0.068	0.074
GW stat.	2.44	2.26	2.10	3.13	3.01	3.52
MZ GLS p-val.	0.08	0.00	0.00	0.16	0.13	0.35
Start: 01/1957						
Theil's U	0.990	0.988	0.987	0.985	0.985	0.979
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.00	0.00	0.00	0.00	0.00	0.00
R^2_{OOS}	0.019	0.023	0.027	0.031	0.029	0.041
GW stat.	0.65	0.74	0.84	1.16	1.08	1.70
MZ GLS p-val.	0.23	0.01	0.00	0.39	0.33	0.68
Start: 01/1977						
Theil's U	0.973	0.973	0.970	0.973	0.973	0.973
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.00	0.00	0.00	0.00	0.00	0.00
R^2_{OOS}	0.053	0.054	0.058	0.054	0.053	0.054
GW stat.	1.88	1.87	2.02	2.00	1.98	2.05
MZ GLS p-val.	0.05	0.00	0.00	0.06	0.06	0.05

Notes: The table shows the results from the evaluation of out-of-sample forecasts based on various forecasting approaches: i) forecasts based on the model with the lowest Schwarz criterion at the forecast date (BIC), ii) forecast based on the model with the lowest Akaike criterion at the forecast date (AIC), iii) forecast from the model with highest adjusted R^2, iv) forecast from a BMA approach with analytical evaluation of posterior model probabilities, v) BMA forecasts based on the MC^3 sampling algorithm and vi) an equally weighted forecast of all evaluated models (EW). OOS results for different start dates of the forecasting scheme are provided. The reported statistics include Theil's U which is the ratio of the RMSE of the model of interest and the RMSE of the benchmark model (TU), the out-of-sample R^2 of Campbell and Thompson (2008). $\#TU^{bs}<TU$ denotes the bootstrap p-value for testing equal predictive performance of the macro-finance augmented model and the AR(1) benchmark against the alternative of superior performance of the model including macro-finance predictors. The bootstrap procedure follows a model-based wild bootstrap methodology as described in section B.3 of the appendix. MZ GLS denotes the GLS version of the Mincer-Zarnowitz statistic and GW stat. denotes the test statistic by Giacomini and White (2006).

Table 9. Out-of-Sample Forecast Evaluation: Several Asset Classes (Short-term sample)

	BIC	AIC	R^2	BMA	MC^3	EW
FX-Aggregate:						
Theil's U	0.948	0.943	0.950	0.960	0.967	0.942
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.00	0.00	0.00	0.00	0.00	0.00
R^2_{OOS}	0.102	0.112	0.097	0.079	0.065	0.112
GW stat.	1.64	1.72	1.49	1.45	0.86	2.01
MZ GLS p-val.	0.00	0.00	0.00	0.00	0.03	0.00
Bonds:						
Theil's U	0.969	0.959	0.969	0.961	0.972	0.954
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.00	0.00	0.00	0.00	0.00	0.00
R^2_{OOS}	0.061	0.081	0.061	0.076	0.054	0.090
GW stat.	1.14	1.31	1.00	1.63	1.10	1.76
MZ GLS p-val.	0.06	0.07	0.09	0.03	0.06	0.17
Stocks:						
Theil's U	0.998	0.967	0.960	0.976	1.010	0.972
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.06	0.00	0.00	0.00	0.49	0.00
R^2_{OOS}	0.004	0.066	0.078	0.048	-0.020	0.056
GW stat.	0.07	0.91	1.01	0.90	-0.23	1.03
MZ GLS p-val.	0.00	0.00	0.00	0.02	0.00	0.01
Commodities:						
Theil's U	0.971	0.951	0.944	0.942	0.971	0.957
$\#\widehat{TU}_{bs} < \widehat{TU}$	0.01	0.00	0.00	0.00	0.00	0.00
R^2_{OOS}	0.058	0.096	0.108	0.113	0.058	0.085
GW stat.	1.60	2.49	2.49	3.34	1.72	2.68
MZ GLS p-val.	0.00	0.00	0.00	0.01	0.00	0.00

Notes: The table shows the results from the evaluation of out-of-sample forecasts based on various forecasting approaches: i) forecasts based on the model with the lowest Schwarz criterion at the forecast date (BIC), ii) forecast based on the model with the lowest Akaike criterion at the forecast date (AIC), iii) forecast from the model with highest adjusted R^2, iv) forecast from a BMA approach with analytical evaluation of posterior model probabilities, v) BMA forecasts based on the MC^3 sampling algorithm and vi) an equally weighted forecast of all evaluated models (EW). Results for different start dates of the forecasting scheme are provided: The forecasts start in 02/1993 after an initialization period of 10 years. The reported statistics include Theil's U which is the ratio of the RMSE of the model of interest and the RMSE of the benchmark model (TU), the out-of-sample R^2 of Campbell and Thompson (2008). $\#TU^{bs}<TU$ denotes the bootstrap p-value for testing equal predictive performance of the macro-finance augmented model and the AR(1) benchmark against the alternative of superior performance of the model including macro-finance predictors. The bootstrap procedure follows a model-based wild bootstrap methodology as described in section B.3 of the appendix. MZ GLS denotes the GLS version of the Mincer-Zarnowitz statistic and GW stat denotes the test statistic by Giacomini and White (2006).

Table A.1. Overview of Predictive Variables

No.	Variable	Abbrev.	Mean	Std.	Skew.	Kurt.	AC(1)
A. Equity Market Variables and Risk Factors							
1	Dividend Price Ratio (Log) (*, †)	D-P	-3.76	0.39	-0.02	1.88	0.99
2	Earnings Price Ratio (Log) (*, †)	E-P	-3.02	0.43	-1.31	6.49	0.98
3	US Market Excess Return (†)	MKT	0.59	4.57	-0.91	5.77	0.10
4	Size Factor (†)	SMB	0.12	3.23	0.81	11.44	-0.03
5	Value Factor (†)	HML	0.35	3.15	0.05	5.54	0.14
6	Short Term Reversal Factor (†)	STR	0.37	3.44	0.17	8.34	-0.02
7	S&P500 Turnover	TURN	0.01	0.16	-0.07	3.38	-0.51
8	Return MSCI World	MSCI	0.73	4.26	-1.20	6.44	0.13
B. Interest Rates, Spreads and Bond Market Factors							
9	T-Bill Rate (Level) (*,†)	T-B	4.56	2.52	-0.02	2.37	1.00
10	Rel. T-Bill Rate (†)	RTB	-0.18	0.86	-0.30	2.85	0.95
11	Long Term Bond Return (*,†)	LTR	0.81	2.97	0.20	4.78	0.02
12	Rel. Bond Rate (†)	RBR	-0.18	0.63	-0.36	4.49	0.87
13	Term Spread (*,†)	T-S	2.33	1.25	-0.25	1.95	0.96
14	Cochrane Piazzesi Factor	C-P	1.22	1.56	0.41	3.34	0.90
C. FX Variables and Risk Factors							
15	Dollar Risk Factor	DOL	0.12	2.19	-0.34	4.02	0.12
16	Carry Trade Factor	C-T	0.05	2.58	-0.69	4.38	0.18
17	Average Forward Discount	AFD	0.18	0.19	0.87	7.83	0.75
D. Liquidity and Credit Risk Variables							
18	Default Spread (*,†)	DEF	0.11	0.43	2.48	12.30	0.94
19	FX Average Bid-ask Spread	BAS	1.43	5.00	1.92	7.46	0.88
20	Pastor-Stambaugh Liquidity Factor	PS	-0.28	6.83	-1.76	10.49	0.00
21	TED Spread	TED	0.07	0.00	1.78	8.67	0.81
E. Macroeconomic Variables							
22	Inflation Rate, Monthly (*,†)	INFM	0.24	0.31	-1.38	11.31	0.47
23	Inflation Rate, YoY	INFA	2.91	1.26	-0.48	4.41	0.95
24	Industrial Production Growth, Monthly	IPM	0.20	0.66	-1.32	10.18	0.23
25	Industrial Production Growth, YoY	IPA	2.24	4.35	-1.60	7.45	0.98
26	Housing Starts	H-S	-2.20	24.90	-0.04	4.52	0.79
27	M1 Growth, Monthly	M1M	0.40	0.79	1.51	13.79	0.18
28	M1 Growth, YoY	M1A	4.81	4.98	0.29	2.31	0.98
29	Orders, Monthly	ORDM	0.11	1.78	0.00	3.15	-0.19
30	Orders, YoY	ORDA	1.20	6.93	-1.51	8.49	0.93
31	Return CRB Spot	CRB	0.25	2.74	-1.76	17.62	0.25
32	Capacity Utilization	CAP	0.02	0.66	-1.07	8.95	0.25
33	Employment Growth	EMPL	0.11	0.19	-0.37	7.40	0.65
34	Consumer Sentiment	SENT	0.01	4.70	0.07	5.66	0.00
35	Consumer Confidence	CONF	0.02	8.25	-0.29	9.94	0.07
36	Diffusion Index	DIFF	8.68	16.91	-0.64	3.57	0.83
37	Chicago PM Business Barometer	PMBB	55.15	7.33	-0.37	3.37	0.88
38	ISM PMI	PMI	52.08	5.35	-0.39	3.77	0.93

Notes: The table shows the summary statistics for the macro-finance predictive variables. The reported statistics include the mean, standard deviation (Std.), Skewness (Skew.), Kurtosis (Kurt.), as well as the first order autocorrelation coefficient (AC(1)). An asterisk (*) denotes that the variable is also part of the Goyal and Welch (2008) dataset, † denotes that the variable is included in set of predictors in case of the long-term U.S. equity sample from 12/1926-12/2010. The sample period over which the summary statistics for the predictors are computed is from 01/1983-12/2010.