Option-Implied Network Measures of Contagion and Stock Return Predictability

Manuela Pedio

University of Bristol (Bristol) and Baffi
Carefin Research Centre (Milan)
manuela.pedio@bristol.ac.uk

November 2023

Abstract

The Great Financial Crisis of 2008 – 2009 has raised the attention of policy-makers and researchers about the interconnectedness among the volatility of the returns of financial assets as a potential source of risk that extends beyond the usual changes in correlations and include transmission channels that operate through the higher order co-moments of returns. In this paper, we investigate whether a newly developed, forward-looking measure of volatility spillover risk based on option implied volatilities shows any predictive power for stock returns. We also compare the predictive performance of this measure with that of the volatility spillover index proposed by Diebold and Yilmaz (2008, 2012), which is based on realized, backward-looking volatilities instead. While both measures show evidence of in-sample predictive power, only the option-implied measure is able to produce out-of-sample forecasts that outperform a simple historical mean benchmark.

Key words: connectedness, volatility networks, implied volatility, realized volatility, equity return predictability, spillover risk

JEL Codes: G12, G17

1. Introduction

The financial crisis of 2007-2009 has taken under the spotlight the role of financial asset connectedness as a source of systematic risk. Such risk would operate through different channels as network connections inflate the exposures to systematic risk factors and reduce any diversification benefits. Yet, a unified framework for the measurement of systematic network risk has been elusive. Nonetheless, a number of heterogenous approaches have appeared in the literature, often measuring different and not directly comparable quantities. In this paper, we propose a novel, forward-looking volatility spillover index implicit in the network structure of individual stock option-implied volatilities, in contrast to backward-looking realized volatilities employed by the earlier literature following the seminal work by Diebold and Yilmaz (2012).

One of the first attempts to capture connectedness across financial assets has been made by Engle and Kelly (2012) who proposed the equi-correlation approach, based on the average of the pairwise linear correlations across asset returns; Billio, Getmansky, Lo, and Pelizzon (2012) developed a number of statistical measures of connectedness based on principal components analysis and on networks constructed using the notion of Granger causality. Other authors have chosen not to model connectedness explicitly, but they propose to compute the risk measures for individual firms conditional on the system being under distress to account for the risk of potential spillovers. Examples of such approaches are the CoVaR developed by Adrian and Brunnermeir (2016), the systemic expected shortfall (SES) proposed by Acharya, Pedersen, Philippon, and, Richardson (2017), and the SRISK advocated by Brownlees and Engle (2016).

Diebold and Yilmaz (2009, 2012) have exploited the concept of forecast error variance decomposition (henceforth FEVD) applied to a vector autoregressive (VAR) model applied to forecast (stock) realized volatilities to compute a measure of aggregate asset volatility connectedness that they call (volatility) spillover index. More specifically, using the FEVD, Diebold and Yilmaz measure what portion of the forecast error of the historical volatility of a stock (or any other asset) is due to innovations to the volatilities of the other stocks in the system, interpreted as a weighted, directed graph. Consequently, an increase in their index (defined as the ratio between the sum of all the elements of the FEVD matrix excluding those on the main diagonal and the sum of all the elements of the FEVD matrix) signals an increase of the spillover of volatility shocks from one stock (asset) to the others. When the final goal is to capture spillover (or "contagion") risk, this approach has several advantages. As a matter of fact, while asset returns tend to co-move also in tranquil times, their volatilities only move together in times of market turmoil, and this makes volatility spillover indices powerful predictors of crisis regimes and bear states. In addition, the use of the FEVD enables a researcher to capture forms of contagion occurring in complex, non-linear ways that go beyond a simple increase in the contemporaneous correlations among the assets. Indeed, in their framework, an increase in aggregate volatility connectedness may be caused either by an increase of the direct links between the (volatilities of) the asset returns as captured by the lead-lag relationship in the VAR by an increase in the covariances of their innovations, or (as it is most likely during a crisis) by a combination of the two. Of course, disentangling the two drivers of the dynamics of the volatility spillover index may be highly informative.

In this paper, we build on the Diebold and Yilmaz's seminal research and extend it in several important ways. First, we propose to extrapolate the volatility spillover index from the network of option-implied volatilities, in contrast to realized volatilities. The main advantage of using option market information is that option prices are forward looking by their nature and therefore they embed the (risk-neutral) expectations of the investors about future volatility over the remaining life of the option. More precisely, if option markets were efficient, (at-the-money) option implied volatility should be regarded as an unbiased (under the risk-neutral measure) forecast of the future realized volatility of the underlying between time *t* and the maturity of the option. Previous literature (see, e.g., Christensen and Prabhala, 1998; Fleming, 1998; Blair, Poon, and Taylor, 2001) has empirically shown that – despite being a biased forecast of ex-post realized volatility – implied volatility has a larger information content concerning future volatility than past realized volatility. Therefore, in this paper we build on this empirical finding that a network based on implied volatilities may be more informative about future volatility spillovers than a network based on realized volatilities.

To this purpose, we collect the option implied volatilities for options on common stocks traded on regulated U.S. stock markets from the IvyDB database of OptionMetrics for a period spanning from January 2006 to December 2017. We base the construction of the implied volatility network (and consequently of the implied volatility spillover index) on at-the-money (ATM) options with a maturity closest to 60 days and that were traded at least once a week in our sample period.[1] Based on these filters, we select a panel of 70 stocks that were characterized by liquid options and that were included in the S&P 500 index over our sample period.

Second, we assess the predictive content of (time-varying) volatility connectedness – as summarized by the spillover index – for the equity risk premium and for individual stock returns. More precisely, we compare the forecasting power of two version of the spillover index: the one based on realized, backward-looking volatilities (RV) as in Diebold and Yilmaz (2009, 2012) and the one based on implied, forward-looking volatilities (IV) that we propose. To the best of our knowledge, this is the first paper that attempts to investigate whether a measure of (volatility) spillover risk has out-of-sample (henceforth, OOS) predictive power for stocks returns. In this respect, there are a few papers that relates to ours. Allen, Bali, and Tang (2012) develop a measure of systemic risk called CATFIN and find that it predicts future economic downturns as well as the cross-section of equity excess returns. However, CATFIN specifically captures the risk of spillovers from the financial sector to the real economy. Conversely, our analysis is not limited to the financial sector; instead, we investigate the predictive power of changes in the transmission of volatility shocks of a set of stocks representative of the entire S&P 500 index. Piccotti (2017) argues that (time-varying) financial contagion risk is non diversifiable and therefore it should be related to

[1] There are several reasons that motivate the use of ATM options. First, as stated above, when markets are efficient, ATM implied volatility should be an efficient forecast of future realized volatility. Second, ATM options are the most sensitive to changes in volatility. Third, ATM options are typically the most liquid (see, e.g., Baltussen et al., 2018).

the equity risk premium. Despite using a variance decomposition approach similar in spirit to Diebold and Yilmaz, similarly to Allen et al. (2012) and in contrast to us, he only features the financial sector as a source of contagion and does not use option-implied information.

Buraschi and Tebaldi (2017) postulate the existence of a super-critical equilibrium in which the equity risk premium is composed by two terms: one which captures the linear exposure to instantaneous market risk, and a network risk premium proportional to firms' exposures to cascades of firm-specific distress shocks. Billio, Caporin, Panzica, and Pelizzon (2016) have extended the classical, ICAPM, factor-based model to include network effects and exploit spatial econometrics to postulate a framework in which the network structure act as an inflating factor for systematic exposure to common risk factors. However, differently from us, these papers do not feature contagion as a result of network dynamics in the second moment of the (option-implied) distribution of stock returns.

We report a number of interesting results. First, the (changes in) both the RV-based and IV-based spillover indices show in-sample predictive power for the equity risk premium and for the excess returns of more than one-third of the stocks under investigation. Interestingly, while the slope coefficient associated with the IV-based spillover index is positive, signaling that an increase in the index is associated with higher future excess returns (to compensate for higher risk), the opposite is true for the RV-based spillover index. Notably, the results hold true when the two indices are both used in a predictive regression for stock excess returns (and for the aggregate equity risk premium). This confirms that the two indices carry different information content and that one is not able to subsume the other.

However, when the out-of-sample (OOS) predictive accuracy is examined, the IV-based spillover index shows a considerably stronger forecasting power than its RV-based counterparts. More specifically, the RV-based spillover index is not able to outperform a simple, historical mean forecast (which is used as a no-predictability benchmark in our analysis), as it delivers a negative OOS R-square (as defined by Campbell and Thompson, 2008). Conversely, the IV-based spillover index yields a positive OOS R-square of 2.11%, as far as the predictive regression for the equity risk premium is concerned. The results from the individual stock predictive regressions are similar: when the RV-based spillover index is used as a predictor, the average OOS R-square turns out to be negative; conversely, the average OOS R-square for predictive regressions relying on the IV-based spillover index is positive and equal to 0.33%.

Because it is well known that the existence of an appreciable statistical forecasting accuracy does not always generate economic value to an investor willing to exploit predictability, we corroborate the results concerning the equity risk premium implementing two alternative investment strategies: a simple switching strategy by which the investor allocates all her wealth alternatively to stocks or to the riskless bond, depending on the predicted sign of excess stock returns (as proposed by Pesaran and Timmerman, 1995); a mean-variance (MV) strategy applied to asset menu consisting of equity and the riskless bond. These exercises confirm the earlier results: an investor using the forecasts based on the RV spillover index will not be able to consistently outperform an investor who relies on a simple historical mean forecast; on the contrary a MV investor who exploits the forecasts based on the IV spillover index will obtain a utility gain of 3.29% on annualized basis, which may also be interpreted as the maximum fee that could be charged to switch from a MV strategy based on historical mean forecasts to IV spillover index-based ones.

Interestingly, much of the predictive power of the IV-based spillover index is expressed in times of high volatility. Indeed, when we split our sample period in two sub-samples characterized by high and low volatility, none of the two spillover indices displays a positive OOS R-square in the low volatility period; on the contrary, the IV spillover index largely outperforms the benchmark in the high volatility period, as shown by an OOS R-square equal to 6.65%.

Finally, because the literature has emphasized that the VIX index shows some predictive power for the equity risk premium (see, e.g., Banerjee, Doran, and Peterson, 2007), we investigate whether the IV spillover index is just a more complex way to capture the same information already contained in the VIX. We find that the inclusion of (the changes of) the VIX index in the predictive regression for the equity risk premium does not subsume the predictive power of the IV spillover index.

The rest of the paper is organized as follows. In Section 2, we present the methodology employed to construct the implied (realized) volatility index. In Section 3, we describe the data together with the filters that we have applied to construct time-series of option implied volatilities from the panel data available in Optionmetrics. In Section 4, we compare the spillover index based on option implied volatilities with its counterpart obtained from realized volatilities (as in Diebold and Yilmaz, 2009, 2012). In Section 5, we compare the predictive power of implied and realized volatility spillover indices for the equity risk premium and the excess returns of individual stocks both in-sample and OOS. In Section 6, we discuss whether the predictive power displayed by the implied volatility spillover index is subsumed by the VIX index. Section 7 concludes.

2. Methodology

2.1. Measuring volatility connectedness among assets

To construct the realized and implied volatility spillover indices, we rely on the procedure suggested by Diebold and Yilmaz, which starts from the estimation of a standard VAR(p),

$$\widetilde{\boldsymbol{y}}_t = \boldsymbol{\nu} + \sum_{i=1}^{p} \mathbf{A}_i \widetilde{\boldsymbol{y}}_{t-i} + \boldsymbol{u}_t, \tag{1}$$

where $\boldsymbol{\nu}$ is a vector of intercepts, $\widetilde{\boldsymbol{y}}_t$ and $\widetilde{\boldsymbol{y}}_{t-i}$ are vectors collecting the log of the realized (or, as in our application, option-implied) volatilities of K assets at t and t-i, respectively, and $\boldsymbol{u}_t \sim IID\ (0, \boldsymbol{\Sigma}_u)$ is a vector of independently and identically distributed disturbances. For the sake of illustration, in what follows, we consider a zero-mean VAR(1) for the de-meaned variables, $\boldsymbol{y}_t = \widetilde{\boldsymbol{y}}_t - \boldsymbol{\mu}$,

$$\boldsymbol{y}_t = \mathbf{A}_1 \boldsymbol{y}_{\boldsymbol{t-1}} + \boldsymbol{u}_t, \tag{2}$$

where $\boldsymbol{\mu} = (\mathbf{I}_{\mathrm{K}} - \mathbf{A}_{\mathbf{1}})^{-\mathbf{1}} \boldsymbol{\nu}$. This is without loss of generality because any VAR(p) process can always be rewritten as a de-meaned VAR(1) through a companion form transformation.[2] Additionally, we

[2] See Hamilton (1994), Chapter 10, for a complete derivation.

assume that the VAR model in (2) is covariance stationary, which is a necessary and sufficient condition for the process to possess a (infinite) moving average representation, by Wold's representation theorem.

It can be shown that, given a stable VAR(p), the minimum mean square error (MSE) predictor for forecast horizon h at forecast origin t is the conditional expected value:

$$\boldsymbol{E_t}(\boldsymbol{y}_{t+h}) = \boldsymbol{E_t}(\boldsymbol{y}_{t+h}|\Omega_t) = \boldsymbol{E_t}(\boldsymbol{y}_{t+h}|\{\boldsymbol{y}_s|s \leq t\}). \tag{3}$$

This predictor minimizes the MSE of each component of $\boldsymbol{y}_t$, i.e., if we call $\boldsymbol{y}_t^*(h)$ any h-step predictor at origin t, we obtain that

$$\text{MSE}[\boldsymbol{y}_t^*(h)] \geq \text{MSE}[\boldsymbol{E_t}(\boldsymbol{y}_{t+h})]. \tag{4}$$

This can be seen by noting that

$$\text{MSE}[\boldsymbol{y}_t^*(h)] = E\{[\boldsymbol{y}_{t+h} - E_t(\boldsymbol{y}_{t+h}) + E_t(\boldsymbol{y}_{t+h}) - \boldsymbol{y}_t^*(h)] \times [\boldsymbol{y}_{t+h} - E_t(\boldsymbol{y}_{t+h}) + E_t(\boldsymbol{y}_{t+h}) - \boldsymbol{y}_t^*(h)]'\} = \text{MSE}[\boldsymbol{E_t}(\boldsymbol{y}_{t+h})] + E\{[E_t[\boldsymbol{y}_{t+h}] - \boldsymbol{y}_t^*(h)][E_t[\boldsymbol{y}_{t+h}] - \boldsymbol{y}_t^*(h)]'\}, \tag{5}$$

where the fact that $E\{[\boldsymbol{y}_{t+h} - E_t[\boldsymbol{y}_{t+h}]][E_t[\boldsymbol{y}_{t+h}] - \boldsymbol{y}_t^*(h)]'\} = 0$ has been exploited (see Lütkepohl, 2005, for further details).

Therefore, the optimal h-step-ahead prediction for $\boldsymbol{y}_{t+h} = \mathbf{A}_1^h \boldsymbol{y_t} + \sum_{i=0}^{h-1} \mathbf{A}_1^i \, u_{t+h-i}$ is given by its conditional expected value, i.e.,

$$\widehat{\boldsymbol{y}}_{t+h|t} = \boldsymbol{E_t}(\boldsymbol{y}_{t+h}) = \mathbf{A}_1^h \boldsymbol{y_t}, \tag{6}$$

which yields a h-step-ahead forecast error equal to

$$\widehat{\boldsymbol{y}}_{t+h|t} - \boldsymbol{y}_{t+h} = \sum_{i=0}^{h-1} \mathbf{A}_1^i \, u_{t+h-i} = \sum_{i=0}^{h-1} \boldsymbol{\phi}_i \, u_{t+h-i}, \tag{7}$$

where the vectors $\boldsymbol{\phi}_i$ collect the coefficients of the moving average representation of the VAR. Therefore, the forecast error covariance matrix is

$$\boldsymbol{\Sigma}_y(h) = \text{E}\left[\left(\sum_{i=0}^{h-1} \mathbf{A}_1^i \, u_{t+h-i}\right)\left(\sum_{i=0}^{h-1} \mathbf{A}_1^i \, u_{t+h-i}\right)'\right] = \sum_{i=1}^{h-1} \mathbf{A}_1^i \boldsymbol{\Sigma}_u (\mathbf{A}_1^i)'. \tag{8}$$

We shall denote with $\Theta(h)$ the h-step-ahead FEVD matrix; each element $\theta_{j,k}(h)$ of $\Theta(h)$ measures the share of total variability of $\hat{y}_{j,t+h}$ (i.e., the h-step-ahead forecast of the variable *j*), that is due to a shock to the variable y_k. Obviously, the diagonal element $\theta_{j,j}(h)$ is the proportion of total variability of $\hat{y}_{j,t+h}$ due to its own innovation. Formally,

$$\theta_{j,k}(h) = \frac{\sigma_{jj}^{-1} \sum_{i=0}^{h-1} (\boldsymbol{e}_j' \boldsymbol{\phi}_i \boldsymbol{\Sigma}_u \boldsymbol{e}_k)^2}{\sum_{i=0}^{h-1} (\boldsymbol{e}_j' \boldsymbol{\phi}_i \boldsymbol{\Sigma}_u \boldsymbol{\phi}_i' \boldsymbol{e}_j)}, \tag{9}$$

where σ_{jj} is the standard deviation of the error term for the *j*th equation and $\boldsymbol{e}_j$ is a selection vector that lists one as the *j*th element and zeros elsewhere. Notably, while the FEVD relies on the orthogonality of the shocks, the generalized FEVD (GFEVD) in (6), firstly proposed by Pesaran and Shin (1998), uses the original, non-orthogonalized shocks, but appropriately accounts for the correlations among them. This avoids the need to enforce orthogonality through identification schemes, for instance, in the form of a Cholesky factorization, which would heavily depend on the ordering of the variables. Moreover, such schemes would turn out to be unsuitable to our high-dimensional application, as a large number of different and equally plausible orderings would in principle be possible. Importantly, due to the covariance between the original shocks, $\sum_{k=1}^{K} \theta_{j,k}(h) \neq 1$, which would instead be the case in a standard FEVD. Therefore, following Diebold and Yilmaz (2012), we normalize each entry of the GFEVD matrix as

$$\tilde{\theta}_{j,k}(h) = \frac{\theta_{j,k}(h)}{\sum_{k=1}^{K} \theta_{j,k}(h)}. \tag{10}$$

The sum of the non-diagonal elements of the *j*th row of the FEVD matrix is the total contribution of volatility shocks to the rest of the system to the uncertainty (as measured by the forecast error) on the volatility of asset *j*; conversely, the sum of the non-diagonal elements of the *j*th column is the total contribution of a volatility shock to asset *j* to the uncertainty on the volatility in the rest of the system. Overall, an increase (decrease) of the shares of forecast errors that are explained by other variables than the one being predicted denotes an increase (decrease) of the connectedness among the assets, i.e., an increase in the proportion of idiosyncratic (volatility) shocks transmitted to (and received from) the rest of the system. The aggregate volatility connectedness is well captured by the volatility spillover index, computed as

$$SI(h) = \frac{\sum_{\substack{k,j=1 \\ k \neq j}}^{K} \theta_{j,k}(h)}{\sum_{k,j=1}^{K} \theta_{j,k}(h)}. \tag{11}$$

In order to capture the time-varying nature of volatility connectedness, we use a rolling window approach. More precisely, to compute $SI_t(h)$, we estimate the VAR model using the 50 more recent observations of the realized (implied) volatility time series (that we shall describe in Section 3) and the associated FEVD; then we proceed recursively until the end of the sample.

At least two remarks are in order. First, as pointed out by Diebold and Yilmaz (2014), there exists an obvious parallel between a FEVD matrix and the adjacency matrix of a weighted, directed network, in which the shares of the forecast error decomposition assigned to each asset in the graph represent the "distances" between them, considered in pairs.[3] In this respect, the spillover index represents a measure of the overall connectivity of the volatility network. Second, in this framework, an increase in aggregate volatility connectedness may be caused either by an increase of the direct links between the (volatilities of) the assets as captured by the lead-lag relationship in the VAR, by an increase in the covariances of their innovations, or (as it is most likely during a crisis) by a combination of the two.

2.2. Estimation of a large dimensional VAR through LASSO

Because in our application we base the construction of the spillover indices on high-dimensional (N=70) VAR models, we deem the conventional least square estimation inappropriate and resort instead on the least absolute shrinkage and selection operator (LASSO) introduced by Tibshirani (1996). More specifically, the LASSO is a regularization

[3] The adjacency matrix $\mathbf{A}$ of a simple network is filled with zero and one entries; more precisely A_{ij} is equal to one when there exists a link between entity *i* and entity *j* and to zero otherwise. In the case of a weighted network (such as the one described by the FEVD), the entries are weights that denote the strength of the link and not only its existence. In addition, the FEVD represents a directed network, since it does not have to be symmetric; therefore, the strength of the link between i and j may differ from the strength of the link between *j* and *i.*

technique that imposes a $L1$ penalty on the least square objective function, shrinking coefficients towards zero. Because of the nature of the penalty, LASSO shrinks some coefficients towards zero (also setting some of them precisely to zero), thus producing sparse vector autoregressive matrices. This is particularly suitable to our application, as many of the off-diagonal coefficients are likely to be zero in the true, unobserved model.[4]

In practice, the LASSO estimator is obtained by solving

$$\min_{\mathbf{A},\boldsymbol{\nu}} \sum_{t=1}^{T} \left\| \mathbf{y}_t - \boldsymbol{\nu} - \sum_{i=1}^{p} \boldsymbol{\Phi}_i \mathbf{y}_{t-1} \right\|^2 + \lambda \|\boldsymbol{\Phi}\|_1 \tag{12}$$

where $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \dots, \boldsymbol{\Phi}_P)$, $\|\boldsymbol{\Phi}\|_1$ is the L_1-norm of the matrix $\boldsymbol{\Phi}$, λ is a tuning parameter, and the rest of the notation is consistent with that employed in equation (1).

Because the LASSO objective function is not differentiable, the problem has to be solved numerically. In particular, following Friedman et al. (2010), we use a coordinate descent algorithm that consists of partitioning equation (9) into scalar subproblems for each $[\boldsymbol{\Phi}]_{i,j}$ which we solve component-wise, and iterating until convergence. The tuning parameter λ is selected via data-driven cross-validation, i.e., starting from a grid of potential values, we select the one that minimizes the one-step-ahead mean square forecast error (MSFE).[5] While other regularization techniques are also available, our preference for the LASSO is justified by the fact that it has been shown to outperform several conventional subset selection methods (such as, for example, stepwise regressions) . For instance, Hsu et al. (2008) perform a simulation study to evaluate the forecasting performance of a set of different variable selection procedures for

---

[4] In our VAR model, all the volatility series are treated as endogenous, i.e., the implied (realized) volatility of stock j depends on its own lags and on the past realizations of implied (realized) volatility of all the other stocks in the system. However, in normal times we expect that the implied volatility of stock j depends on the lags of only a subset of stocks (e.g., the ones in the same industry). However, it is important to notice that sparsity in the VAR matrix does not imply sparsity in the FEVD matrix as in the generalized FEVD the shocks are not orthogonal.

[5] To estimate the model, we use the R package BigVAR by Nicholson, Matteson, and Bien (2019). Additional details about the solution algorithm and the cross-validation procedure for the choice of λ can be found in Nicholson, Matteson, and Bien (2017).

VAR models (including LASSO); they find that LASSO not only yields the lowest one-step-ahead mean square forecast error, but also has the highest precision in estimating $\boldsymbol{\Sigma}_u$, which is particularly relevant in our application, because the FEVD will depend on its estimate.

3. Data

Our data come from a number of sources. To construct the IV spillover index, we collect option data from the IvyDB database by OptionMetrics, which contains daily, closing bid and offer prices, trading volumes, open interest, strikes, maturities, and the common "greeks" for all the US-listed index and equity options. In addition, OptionMetrics data also include the IVs computed in correspondence to the midpoint of the best closing bid and best closing offer prices of each option.[6] Because options on individual stocks have an American-style exercise feature, option implied volatilities are computed using an algorithm based on the binomial tree model of Cox, Ross, and Rubinstein (1979) to account for the early exercise premium.

We retain only options on common stocks traded on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the National Association of Securities Dealers Automated Quotation (NASDAQ). Furthermore, by applying standard filters used in the literature (see, e.g., Bali and Hovakimian, 2009; Driessen, Maenhout, and Vilkov, 2009; Goyal and Saretto, 2009; Baltussen, Van Bekkum, and Van der Grient, 2018), we exclude options on closed-end funds and real estate investment trusts (REITs) and options with zero open interest or zero trading volume on any given day.[7] We also apply a set of filters to clean the

---

[6] The option price used to compute implied volatilities is an average between the maximum bid and the minimum ask, selected across all the exchanges the contract is traded on. Up to March 4, 2008, option prices used in the implied volatility calculation are end-of-day prices. Since March 5, 2008, OptionMetrics has started capturing the best bid and best ask prices as close to 4 p.m. as possible in the attempt to better synchronize the reported option prices with the closing price of the underlying. The problem of non-synchronous trading between stocks and options due to different closing times of the exchanges has been pointed out by Battalio and Schultz (2006). However, this does not appear to represent a relevant issue for the purposes of our analysis, also because most of our data concern a period posterior to March 2008.

[7] To achieve this goal, we merge the information contained in the Option Price file with the security data from the Security file. We retain only options whose underlying stock has *Issue Type* equal to zero

data from mis-reported prices, outliers, and microstructural biases (see, e.g., Goyal and Saretto, 2009 and Baltussen et al., 2018). Specifically, we discard observations for which the bid-ask spread exceeds 50% of the average between the best bid and the best offer or it is lower than the minimum tick size (which is 0.05 USD for options trading below 3 USD and 0.10 USD in all other cases). We also delete observations with missing or extreme values for the implied volatility (less than 3% or higher than 150%).[8]

Given that estimating a VAR model on a panel of implied volatilities requires constructing regularly spaced time-series, we need to select one observation of the implied volatility for each underlying stock at each date. Because it would be impossible to find a sufficiently large number of stocks with at least one option trading on every day, we settle for a weekly frequency. Additionally, while the OptionMetrics dataset starts in 1996, we restrict our sample to the period January 2006 – December 2017, because before this date the number of options available would be insufficient to support our application. For instance, over the period January 1996 – February 2003, Carr and Wu (2008) are able to find only 35 options on individual stocks with at least 600 days of active trading (which represents approximately one-third of the total number of trading days in their sample).

To select one observation of the IV for each underlying stock and for each week, we use the following rules to ensure that the resulting time-series are as homogenous as possible in terms of time to maturity and moneyness of the options that are used to construct them. First, we retain only put options with an effective time to maturity (that we compute as the difference between the stated maturity and the calendar date) ranging between 7 and 120 days, as short-term options are usually the most liquid and actively traded. We focus on at-the-money (ATM)

(which corresponds to common stocks), a SIC code different from 6720 – 6730 and 6798 (because these codes identify closed-end funds and REITs) and an *Exchange Flag* equal to 1, 2, 4, and 8 (corresponding to NYSE, AMEX, NASDAQ and NASDAQ Small Cap, respectively). Additionally, for each date, we only retain the observations with strictly positive *Volume* and *Open Interest*.

[8] These filters also remove all the observations for which the implied volatility is set to -99.99 by OptionMetrics. These are options with non-standard settlement, options for which the midpoint of the bid-ask price is below the intrinsic value, whose vega is below 0.5, for which the implied volatility calculation fails to converge or the underlying closing price is not available.

options, which are the most sensitive to changes in volatility (i.e., have maximum vega) and therefore we focus our analysis only on options with moneyness ranging between 0.9775 and 1.0225. We compute moneyness as the ratio between the strike and the closing price of the underlying. In this respect, our sample construction choices are close to Goyal and Saretto's (2009), who also build time-series of option implied volatilities. When more than one option with these characteristics is available on a given day, we retain the contract closer to having 60 days left to maturity.[9]

Once we have distilled one observation per day per each underlying stock, we build a Wednesday-to-Wednesday weekly time series. However, when no option with the required characteristics happened to have been traded on a Wednesday for a given underlying, we take the previous day's observation; if no option had been traded on Tuesday too, we use the Thursday's observation. Only residually, we also rely on Monday and Friday observations, but this happens in less than 5% of the records in our sample. If in a given week there is no traded option for an underlying stock, we record that date as a missing value.[10]

To include a stock option series in our analysis, we require that less of 5% of the observations be missing values and that no more than three consecutive missing values to appear in the sample. The remaining missing values that do not cause the exclusion of an IV time series are filled using a one-month rolling mean estimate.[11] Table 1 lists the stocks that satisfy our

---

[9] We choose options close to 60 days to maturity because the average time to maturity of the options left in our sample after filtering is around 60 days.

[10] There is a total of 1,018 missing values in our sample for an average of 15 missing observations for each of the 70 series. Considering that each series includes 617 observations for a total sample of 43,190 observations, missing values represent less than 2.5% of the sample. The maximum number of missing values for a series is 30 in the case of AON Plc. There are three stocks with zero missing values, namely Apple, Costco Wholesale Corp, and Intuitive Surgical. For what concerns the cross-sectional dimension, on any given date in the sample, there are on average two stocks for which the value of the implied volatility is not observed.

[11] We believe that this choice is likely to have a minimal impact on our results. Appendix A shows the series of the implied volatility of AON Plc, which is the series with the maximum number of missing values, under two different assumptions concerning the length of the window for the moving-average used to fill the missing values: 1 month (i.e., 4 weekly observations, which is our chosen length) and 3 months (12 weekly observations). We note that the two series are almost indistinguishable.

requirements and that are therefore included in the analysis. These are mostly S&P 500 stocks, the largest one being Apple, with an average market capitalization of USD 399 billion over the sample period and the smallest is Abercrombie & Fitch, with an average market cap of just USD 3 billion. In Table 1, we also report the industry to which the stocks belong. The selected stocks cover a broad set of different sectors, including technology, energy, consumer discretionary, financials, industrials, and health care.

For the stocks in the list, we also compute realized volatilities. To match the implied volatilities, which represent the expectation at time t of the annualized volatility over the next 60 days (given that we choose options close to 60 days to maturity), we compute the annualized 60-day realized volatilities as

$$RV_{t,t+60} = \sqrt{\frac{365}{60}\sum_{i=1}^{60}\left(\frac{S_{t+i} - S_{t+i-1}}{S_{t+i-1}}\right)^2}, \tag{13}$$

where S_t is the price of the stock at time t, retrieved from Bloomberg. We elect to use Bloomberg prices instead of the stock prices provided by OptionMetrics, because the latter are not adjusted for stock splits. Table 1 shows the average realized and implied volatilities for each stock over the sample period. Average realized volatilities range from 16% for Kimberly-Clark Corp to 45% for Abercrombie & Fitch; the implied volatilities range from 17% to 46%. Notably, the average implied volatilities tend to be higher than the average realized volatilities, consistently with the literature that has documented the existence of a positive spread between realized and implied volatilities (see, e.g., Bali and Hovakimian, 2009).

Bloomberg prices are also used to construct a weekly series of returns for each stock in the sample. Excess returns are obtained by subtracting the one-month Treasury bill rate from the Federal Reserve Economic Data (FRED) repository of the St. Louis Fed. However, as one may object that our data are collected at weekly frequency, we also compute excess returns by subtracting the one-week US-based LIBOR rate (also collected from FRED) as a robustness check. Finally, we obtain the closing values of the S&P 500 Index and the Chicago Board

Options Exchange (CBOE) S&P 500 Volatility Index (the VIX) from the Wharton Research Data Service (WRDS).

4. Implied and realized volatility spillover indices

The first step of our analysis is to construct alternative volatility spillover indices using either implied or realized volatilities. Following the procedure described in Section 2, we recursively estimate a VAR model for the implied (realized) volatilities of the 70 stocks in our sample using a 50-week rolling window (meaning that the first value of the two indices are obtained with reference to December 13, 2006).[12] Besides the rolling window length, there are other two key choices that we need to make, namely the number of lags to be included in the VAR model and the forecast horizon of the FEVD. As far as the choice of the VAR order is concerned, we rely on a standard model selection procedure based on the Bayesian information criterion (BIC). In the case of the VAR estimated on realized volatilities, our specification search shows that a VAR(1) model yields a BIC equal to 7.285, while the BIC is equal to 7.291 and 7.293 for the VAR(2) and VAR(3), respectively. For what concerns implied volatilities, a VAR(1) yields a BIC of 7.282, while the BIC is equal to 7.290 and 7.291 for VAR(2) and VAR(3), respectively. Therefore, we estimate a VAR(1), which minimizes the BIC for both the implied and the realized volatilities In both cases the estimated VAR matrix is rather sparse, as we expected.[13] Notably, the use of the LASSO algorithm allows us to estimate the rolling VAR even if the

[12] In a further robustness check, we also estimate the two indices using a 100-week rolling window. The resulting indices are characterized by dynamics that resemble those obtained using a 50-week rolling window but are smoother and therefore less informative. A comparison of the two indices estimated using alternatively the 50- and 100-week rolling windows is performed in Figure A2 of Appendix A.

[13] Figure A3 of Appendix A shows the sparsity plot of the vector autoregressive matrix of the VAR fitted on the realized volatilities (the one concerning implied volatilities is not reported as it is almost identical), which is available as a standard output from the R package BigVAR. Each of the squares represents one of the 70×70 coefficients. The darker is the colour, the bigger is the coefficient (in absolute value). A white square denotes that the coefficient has been set to zero. The picture shows that there are large coefficients on the main diagonal (as expected, because volatility tends to be highly persistent) but a lot of the coefficients out of the main diagonal has been set to zero.

number of observations available for each recursion (which is 50 periods times 70 stocks, i.e., 3500) is less than the number of the parameters to be estimated (i.e., the 4900 coefficients in the vector-autoregressive matrix).

As far as the forecast horizon of the FEVD is concerned, no optimal selection procedure exists. In fact, as pointed out by Diebold and Yilmaz (2014), different horizons may carry different information. As we increase the forecast horizon, we get close to the unconditional variance decomposition, that is obtained when $H \to \infty$. Conversely, in our application, we are more interested in the short-term volatility spillovers and therefore we believe that $H = 2$ may represent an appropriate horizon. However, to check the robustness of our indices to alternative assumptions, we also experiment with different choices of H. In particular, in Figure 2, we depict the RV (Panel A) and the IV (Panel B) spillover indices computed setting $H = 2$. The dotted lines represent the mean and median values obtained when we compute the indices setting H at all the possible horizons between 2 and 10 weeks. As one can notice, despite being higher in levels, the dotted lines approximately describe the same evolution as the indices based on a 2-week ahead forecast horizon. This is true for both the RV and the IV indices and entails that in our application different horizons convey almost the same information. Therefore, in what follows we focus exclusively on the case $H = 2$.

Both the IV and the RV indices show three peaks. The first peak corresponds to the financial crisis of 2007-2009; the second one straddles the period 2010-2012, approximately corresponding to the European sovereign crisis; the third peak starts in 2013/2014 and ends in 2016/2017 (the exact timing depends on whether we examine the IV or at the RV index). While the first two peaks have an obvious interpretation and also correspond to sharp increases in the VIX, the third peak is harder to explain.[14] However, at least to some extent, the last peak appears to coincide with the tightening of the US monetary policy that started in

[14] Interestingly, similar peaks are also visible in the aggregate SRISK Index computed according to the methodology proposed by Brownlees and Engle (2016). Their updated index can be found at https://vlab.stern.nyu.edu/welcome/srisk.

December 2015 that triggered what has been dubbed the "Taper tantrum" by some market commentators.[15] Interestingly, the IV index started to increase before its RV counterpart did both during the financial crisis and during the last peak, thus corroborating our conjecture that the IV spillover index could be a better real-time predictor of equity returns that its RV counterpart.

To check that the dynamics of our indices do not depend on our specific choices concerning the (underlying) stocks included in the sample, we also randomly exclude ten stocks from our list and compute afresh the spillover indices. The exercise is repeated ten times, and the results are plotted in Figure 2. In particular, the solid line represents our baseline RV (IV) spillover index (based on the entire sample and assuming $H = 2$); the dotted lines represent the minimum and the maximum values of the indices estimated using the randomly selected subsamples of 60 stocks. We observe that the differences between the baseline indices and the dotted lines are minimal and therefore we conclude that stock selection is not the main driver for the observed behavior of the indices.

5. Predictive power of the spillover indices

In this Section, we evaluate the in-sample and OOS predictive power of the IV spillover index for the (excess) returns of the S&P 500 Index and for the individual stocks included in our sample; we shall compare such empirical performances with those offered by the RV counterpart and a standard benchmark, i.e., the historical mean (similar, for instance, to Campbell and Thompson, 2008 and Welch and Goyal 2008). More precisely, we estimate the predictive regression

$$r_{t+1,j} = \alpha_j + \beta_j^{(RV)} \Delta RV_t + \beta_j^{(IV)} \Delta IV_t + \varepsilon_{t+1,j} \tag{14}$$

where $r_{t+1,j}$ is the weekly excess return (over the one-month T-bill) of an individual stock j or of the S&P 500 index, ΔRV_t (ΔIV_t) is the change between time $t-1$ and t of the realized

[15] See e.g., https://www.cnbc.com/id/100829208.

(implied) volatility spillover index, and $\varepsilon_{t+1,j}$ is an i.i.d shock with zero mean and volatility σ_j. As we are mostly interested in comparing the predictive performance of the two spillover indices, we also estimate (11) after either setting $\beta_j^{(RV)} = 0$ or $\beta_j^{(IV)} = 0$, alternatively. In the next subsection, we discuss the in-sample results, while in the following subsections, we examine the OOS predictive performance of the spillover indices.

5.1. In-sample results

Table 2 reports the estimates of the predictive regressions of the excess returns of the S&P 500 on (the changes of) both the spillover indices (model I) and (the changes of) each of the two alternative indices (models II and III), together with the R-squares of the regressions. A first interesting result that we report is that, while equity (excess) returns load negatively on the (changes of) the RV index in the previous period, they load positively on (changes of) the IV index in the previous period. More specifically, S&P 500 excess returns display a slope coefficient of -0.31 on the lagged (changes of) RV index and of 0.46 on the (changes of) IV index. Both coefficients turn out to be statistically significant both in the individual predictive regressions and when the two indices are used in combination to forecast the equity risk premium. The estimates of the coefficients of the two predictors change only slightly when they are used in combination.

Table 3 displays the results concerning the predictive regressions for the individual stock (excess) returns. To save space and foster interpretation, we only report the average of the coefficients across each sector.[16] In particular, we consider the following seven sectors: consumer discretionary (11 stocks), energy (11 stocks), financials (6 stocks), health care (10

---

[16] To save space, in this case, we do not report the results for the regressions where the (changes of) the two indices in the previous period are used simultaneously to predict the stock excess returns. Overall, the results are coherent with what has been discussed for the prediction of the equity risk premium: both spillover indices show in-sample predictive power for the stock excess returns. Indeed, in the case of approximately one-third of the stocks, both coefficients are statistically significant at least at a 5% test size level; this proportion grows to a half when a test size level of 10% is considered. A summary of the results is available in Appendix B (Table B1).

stocks), industrials (11 stocks), materials (8 stocks), and technology (5 stocks). The remaining 8 stocks are aggregated under the category "Other" because we do not have enough observations to compute meaningful averages for the sectors to which they belong. We also report the average value of the slope coefficients across all stocks.

The results obtained for the individual stocks tend to mimic what we already commented in the case of the equity risk premium. Notably, the average of the coefficients obtained from the regressions on the individual stocks ("β coeff. All" in Table 3), which are -0.36 and 0.41 for the RV and the IV, respectively, are not far from the estimates that we obtained for the equity risk premium (i.e., -0.31 and 0.46 for RV and IV spillover indices, respectively). Additionally, the variability of the coefficients across all stocks is quite moderate, especially in the case of RV regressions, where the minimum value for the estimated slope coefficients is -0.78 and the maximum is -0.10, with a standard deviation of 0.15 only. There is more heterogeneity as far as the slope coefficients in the IV regressions are concerned; indeed, they range from 0 to 1.03, with a standard deviation of 0.21. Interestingly, among the sectors, technology stocks are those that imply the smallest (in absolute value) coefficients for both the RV and IV predictive regressions (their averages are -0.23 and 0.34, respectively). Conversely, materials and energy are the sectors with the largest (in absolute value) average estimated coefficients (-0.44 and -0.42, respectively) for what concerns the RV regressions; energy and financial stocks are those implying the largest average coefficients in the IV regressions (0.71 and 0.52, respectively).

Figures 3 and 4 depict the distributions of the estimated slope coefficients and the associated t-statistics obtained from the RV (Figure 3) and IV (Figure 4) regressions. In the case of the RV regressions, we note that the coefficients turn out to be statistically significant for more than a half (namely, 51) of the stocks, when we set a test size of 10%; however, the number decreases to 10 when we impose a more restrictive test size of 1%. For what concerns the IV regressions, the distribution of the estimated coefficients appears to be bimodal, with most of the coefficients ranging between 0.26 and 0.39 and then between 0.65 and 0.77. Similarly to the RV regressions, 50 coefficients are statistically significant at a 10% test size; however, in this case 20 stocks (i.e., one-fourth of the total) display an estimated slope that is also significant at a 1% size.

All in all, we find evidence of in-sample predictive power from both the RV and the IV indices, even when they are used jointly (which means that they are both useful to predict stock returns). The R-squares are generally low as they range from 0.73 to 2.06 percent and from 0.92 to 2.41 percent, for the RV and IV index, respectively; the R-squares for the equity risk premium predictive regressions are 1.44% for RV index and 2.51% for the IV index, respectively. However, these values are comparable with those reported in the literature for equity premium in-sample predictive regressions (usually estimated on monthly data). For instance, in their seminal study, Fama and French (1988) report monthly R-square statistics of approximately 1% for a predictive regression on the dividend price ratio; more recently Campbell and Thompson (2008) report values of the R-square that range between 0.05% and 3.48% from predictive regressions of the (monthly) excess returns of the S&P 500 on a broad set of predictors (e.g., the dividend yield, the term spread, the book-to-market ratio, etc.).

5.2. Out-of-sample predictive performance

Considering that a forecaster is typically more interested in the OOS than in the in-sample predictive performance, in this subsection, we analyze the results obtained when we recursively estimate the predictive regression and recursively use the information available at time t to forecast the excess return of the S&P 500 (or of each of the individual stocks) at time t+1 as

$$\hat{r}_{t+1|t,j}^{(m)} = \hat{\alpha}_j^{(m)} + \hat{\beta}_j^{(m)} x_{t,m}, \tag{15}$$

where $x_{t,m}$ is the change of the RV (IV) index between t and t-1, and $\hat{\alpha}_j^{(m)}$ and $\hat{\beta}_j^{(m)}$ are the ordinary least square (OLS) estimates of the regression coefficients obtained using the data available at time t. In our application, we use data from December 20, 2006 through December 5, 2007 as the initial estimation period and we obtain the forecast for the excess return of stock/index j over the week of December 5 through 12 in 2007 using the changes of the RV

(IV) index over the week Nov. 28 – Dec. 5, 2007. Next, we proceed recursively by adding one observation to our estimation sample in an expanding window fashion, until the end of the OOS period (i.e., December 27, 2017).

To evaluate the OOS predictive performance of the two alternative indices, we use the standard OOS statistics suggested in Campbell and Thompson (2008), Goyal and Welch (2008), and Rapach et al. (2010). For each predictive regression, we compute the mean square forecast error (MSFE) for the security/index j and the predictor $m = RV, IV$ as

$$MSFE_j^{(m)} = \frac{1}{n_2}\sum_{s=1}^{n_2}\left(r_{j,n_1+s} - \hat{r}_{j,n_1+s|n_1+s-1}^{(m)}\right)^2 \qquad (16)$$

where $n_1 = 50$ is the number of observations that are used as the initial in-sample estimation period, $n_2 = T - n_1$ is the number of observations in the OOS period, $\hat{r}_{j,n_1+s|n_1+s-1}^{(m)}$ is the forecast of the excess return of the asset j obtained as in (17), and r_{n_1+s} is the realized return that is actually observed. We also compute the MSFE of a model that assumes constant expected (excess) returns implying that the historical average is the best prediction for future excess returns:

$$MSFE_j^{(bmk)} = \frac{1}{n_2}\sum_{s=1}^{n_2}\left(r_{j,n_1+s} - \bar{r}_{j,n_1+s}\right)^2 \qquad (17)$$

where $\bar{r}_{j,n_1+s} = \frac{1}{n_1+s-1}\sum_{t=1}^{n_1+s-1} r_t$.

As it is typical of the literature, for each predictive model, we report the difference between the square root of the MSFE (RMSFE) of the benchmark model, denoted as $RMSFE_j^{(bmk)}$ and the RMSFE of the predictive model, denoted as $RMSFE_j^{(m)}$.This difference, denoted as ΔRMSFE, is positive when using (the change of) the RV (IV) index as a predictor reduces the forecast error. When the ΔRMSFE is positive, we also test whether the gain in predictive accuracy is statistically significant, i.e., we test $H_0: MSFE_j^{(bmk)} \leq MSFE_j^{(m)}$ against $H_a: MSFE_j^{(bmk)} > MSFE_j^{(m)}$. Because the standard Diebold and Mariano (1995) and West (1996) (DMW) statistics have a non-standard asymptotic distribution when comparing forecasts from nested models, as it is our case because the benchmark corresponds to $\hat{\beta}_j^{(m)} = 0$, we rely on the MSFE-adjusted statistic proposed by Clark and West (2007, henceforth CW). As suggested by CW, we first compute

$$\tilde{d}_{j,n_1+s}^{(m)} = \left(r_{j,n_1+s} - \bar{r}_{j,n_1+s}\right)^2 - \left[\left(r_{j,n_1+s} - \hat{r}_{j,n_1+s|n_1+s-1}^{(m)}\right)^2 - \left(\bar{r}_{j,n_1+s} - \hat{r}_{j,n_1+s|n_1+s-1}^{(m)}\right)^2\right] \quad (18)$$

and then we regress $\tilde{d}_{j,n_1+s}^{(m)}$ on a constant for $s = 1, \ldots n_2$; the MSFE-adjusted statistic is the t-statistic corresponding to the constant.

Additionally, we compute the OOS R-square, firstly proposed by Campbell and Thompson (2008), which measures the proportional reduction in the MSFE of the predictive regression forecasts relative to the historical mean benchmark:

$$R2(OOS)_j^{(m)} = 1 - \frac{MSFE_j^{(m)}}{MSFE_j^{(bmk)}}. \quad (19)$$

A negative value of $R2(OOS)_j^{(m)}$ indicates that the predictive model *m* fails to outperform the historical mean for stock *j* (or the S&P index).

However, as emphasized, for instance, by Campbell and Thompson (2008), the OOS R-square is insufficient to gauge whether the additional amount of return predictability (if any) obtained through the use of the two spillover indices is economically meaningful. For instance, Dal Pra, Guidolin, Pedio, and Vasile (2018) note that best model in terms of statistical predictive accuracy are not necessarily the ones that deliver the maximum economic value. In addition, as observed by Rapach et al. (2010), the OOS R-square neglects the risk borne by an investor over the holding period. For this reason, with reference to the evaluation of the predictive power of the spillover indices for the equity risk premium, we also implement two different allocation strategies based on the alternative forecasting models.

First, following Pesaran and Timmerman (1995), we construct a simple switching strategy, whereby the investor uses the forecasts based on the predictive regressions to allocate all the available wealth alternatively to stocks or risk-free bills, depending on the sign of the forecasted equity risk premium (i.e., when the predicted sign is positive, the investor allocates all her wealth to equity and viceversa). More precisely, the realized wealth at the end of each holding period (which in our case is equal to one-week) is equal to:

$$W_{t+1} = W_t[(1 + r_{t+1})I(ES)_t + (1 + rf_{t+1})(1 - I(ES)_t] \quad (20)$$

where W_t is wealth at the beginning of the period (which we normalize to 1, as typical in the literature), r_{t+1} is the realized excess equity return between t and $t+1$, rf_{t+1} is the rate of the one-month T-bill between t and $t+1$, and $I(ES)_t$ is a dummy variable that equals 1 when the predicted sign is positive and all wealth is invested in the index, and 0 otherwise. This exercise is recursively repeated over the OOS period, so that on every weak the investor selects her optimal portfolio based on all the data available up to that point. The (annualized) average return and Sharpe ratio (SR) achieved using the competing forecasting models are then compared using the same statistics computed when the historical mean forecast is employed.

Second, following Campbell and Thompson (2008), Welch and Goyal (2008), and Rapach, Strauss, and Zhou (2010), we also compute the (annualized) returns, Sharpe ratios and realized utilities obtained by a mean-variance investor who allocates her wealth between stocks and risk-free bills (at a weekly frequency) using the forecasts of the equity risk premium from the alternative predictive models. More precisely, the investor is supposed to maximize

$$\mathrm{U}(W_{\mathrm{t+1}}) = E_t[\mathrm{W}_{t+1}] - \frac{\gamma}{2} Var_t[W_{\mathrm{t+1}}], \quad (21)$$

with an investment horizon equal to one week and a risk aversion coefficient, γ, equal to 3.[17] Terminal wealth depends on realized asset returns and on the selected portfolio weights in standard, linear ways. This allows us to optimize an objective function that reflects total one-

[17] The choice of a risk aversion coefficient of 3 is quite standard in the predictability literature (see, e.g., Campbell and Thompson, 2008 and Welch and Goyal, 2008). However, our main results are robust to different choices of the risk aversion coefficient. For instance, an investor with a risk aversion coefficient of 5 who uses (the changes of) the IV spillover index as a predictor of the equity excess returns would achieve an (annualized) utility gain of 3.64%; similarly, an investor with a risk aversion coefficient of 3 who uses the same predictor would obtain a utility gain of 3.29% per annum, as we shall discuss below.

period portfolio returns. An investor determines the optimal weights to be assigned to the risky asset at time t (to be held fixed until time $t+1$) according to the formula

$$\omega_t{}^* = \frac{1}{\gamma}\frac{\hat{r}_{t+1|t}^{(m)}}{\hat{\sigma}_{t+1|t}^2}, \tag{22}$$

where $\hat{r}_{t+1|t}^{(m)}$ is the equity risk premium forecast based on the predictive model m and $\hat{\sigma}_{t+1|t}^2$ is a historical estimate of the covariance matrix (similar to Campbell and Thompson, 2008, Goyal and Welch, 2008, and Rapach et al., 2010).[18] The allocation to the risk-free asset is then simply equal to $1-\omega_t{}^*$. As a benchmark, we also compute the weights obtained when the historical mean forecast is used in the optimization process, i.e.,

$$\omega_t{}^{*(bmk)} = \frac{1}{\gamma}\frac{\bar{r}_{t+1|t}}{\hat{\sigma}_{t+1|t}^2}. \tag{23}$$

The average, realized utility level from each predictive model is computed as

$$\tilde{\nu}^{(m)} = \tilde{\mu}_P^{(m)} - \frac{1}{2}\gamma\tilde{\sigma}_P^{2\,(m)}, \tag{24}$$

where $\tilde{\mu}_P^{(m)}$ and $\tilde{\sigma}_p^{2\,(m)}$ are the sample mean and variance of the ex-post, realized returns over the OOS period from the optimal portfolio formed by exploiting model (m) to originate the forecasts of the equity risk premium. We also compute $\tilde{\nu}^{(0)}$, the average utility that the investor obtains when she uses the historical mean forecast in the optimization process, as in (23). The difference between $\tilde{\nu}^{(i)}$ and $\tilde{\nu}^{(0)}$ is the utility gain arising from using a predictive model for the equity risk premium and can be interpreted as the risk-free compensation an investor is willing to pay to switch from a strategy based on the historical mean to a strategy based on each of the predictive models proposed. A predictive model generates economic value with respect to the benchmark if the utility gain is positive.

5.2.1 Individual stock predictions

---

[18] Similarly to Campbell and Thompson (2008) and Rapach et al. (2010), we constrain the weight attached to the equity to be positive and we allow a maximum leverage of 50% (i.e., $\omega_t{}^*$ is set to be lower than or equal to 150%).

In Table 4, we report the predictive accuracy statistics for the individual stocks. To foster interpretation, we do not report the results for the 70 individual stocks but focus instead on their averages across sectors. For this reason, the CW statistics for the statistical significance of ΔRMSFE are not reported. However, complete results are displayed in Table 2B of Appendix B. Interestingly, the RV index does not show any OOS predictive power for the excess returns of the individual stocks as it fails to outperform the benchmark. On the contrary, the IV spillover index significantly outperforms the benchmark for many of the individual stocks.[19] On average, when the changes of the IV index are used as predictor for the individual stock returns, the OOS R-squares is 0.33% but this figure increases to 0.76% when we only consider the stocks with a positive OOS R-squares. Most of the predictability seems to come from the energy sector, which deliver average values of the R-square coefficient equal to 1.05%, while technology stocks display a negative average R-square and consumer discretionary stocks have an average R-square close to zero.

In Table 4, we report the percentage of correct sign predictions (computed as an average across sectors). Interestingly, prediction models based alternatively on the RV and IV spillover indices show approximately the same proportion of correct sign predictions, slightly above 50%; this proportion is in turn not dissimilar from the one displayed by the benchmark model. More precisely, on average a predictive model based on the RV spillover index display 51.51% of correct sign predictions, while this percentage is equal to 50.88 for the IV predictive model and to 50.66 for the benchmark model (i.e., the historical mean prediction). Therefore, we conclude that the different forecasting power of the two predictive models (and the fact that the IV predictive model outperforms the benchmark as far as RMSE is concerned) may not come from a higher ability of the IV index to forecast the correct sign of future stock returns; instead, we conjecture that the outperformance of the IV predictive model is linked to a better

[19] In particular, Table B2 in the Appendix B shows that when the changes in the RV spillover index are used to predict stock returns, the Campbell and Thompson R-squares are never positive and statistically significant. Instead, when the changes in the IV spillover index are used as predictors, 45 out of 70 stocks display a positive OOS R-square. The differences in the MSFE between the benchmark and the predictive model are statistically significant (at least at a 10% test size) in 40 cases.

ability to produce accurate forecast during times of market distress, a hypothesis that we shall investigate further in Section 5.4.

5.2.2 Equity risk premium predictions

Although the evidence in favor of individual stock (excess) return forecastability is undeniable albeit not overwhelming, in Table 5 we document the robust predictive ability of the IV spillover index for the market-wide, aggregate equity risk premium (as proxied by the excess returns on the S&P 500 index) both in terms of statistic accuracy and of economic value. Albeit the recursively estimated slope coefficients from the RV and IV predictive regressions (plotted in Figure 5, Panels A and B, respectively) are always statistically significant (with the exception of a short period before the outburst of the financial crisis), only the IV spillover-based predictive regression manages to consistently outperform the benchmark.[20]

In particular, the OOS R-square of the IV predictive regression is 2.11%, more than twice the average obtained on individual stocks. Following Goyal and Welch (2008), in Figure 6, we plot the cumulative difference in the squared forecast errors (CDSFE) for the historical average vis-à-vis the RV (Panel A) and the IV (Panel B) forecasts. A visual inspection of the plots can reveal whether the predictive regressions have a lower MSFE than the historical mean in any given period by simply taking a segment that joins the beginning and the end of the period of interest: if the curve is higher (lower) at the end of the segment relative to the beginning, then the predictive regression has a lower (higher) MSFE than the historical average during that period. Panel A shows that the RV spillover index outperforms the historical mean only during a short period at the end of 2008 (in correspondence to the outburst of the financial crisis). The CDSFE sharply declines in the middle of 2010 and remains flat since then, meaning that from 2011 to the end of the OOS period, the historical mean and the RV predictive regression display equal predictive power.

---

[20] Also in these OOS regressions, the two slope coefficients display opposite signs. While the IV coefficient peaks at the beginning of the sample but then remains flat around 0.5-0.6 for the rest of the OOS period, the RV slope shows more variability but stabilizes around -0.35 after 2011.

Panel B allows us to identify at least two periods in which the IV spillover index has considerably outperformed the historical average. The first one is from the beginning of the OOS period (December 2007) to the end of 2009 (corresponding to the financial crisis); then, after a short period in which the historical mean has outperformed the IV predictive model (from the end of 2009 to end of 2010), the CDSFE raises again from the beginning through end of 2011. On the contrary, from 2011 to the end of the OOS period there is only episodic evidence of predictability with the CDSFE turning completely flat after 2014. This is not surprising for at least two reasons. First, it has been shown (see, e.g., Rapach et. al, 2010) that OOS stock return predictability is mostly a recessionary phenomenon, irrespective of the choice of the predictor. Therefore, it is unsurprising that both predictors seem to work during the latest recession that started in December 2007 and ended in June 2009, according to the National Bureau of Economic Research (NBER) dating system. More interestingly, the predictive performance of the IV spillover index appears to be associated with periods of turmoil and contagion in financial markets. This is not unexpected, as the one we are proposing is a forward-looking measure of the proportion of uncertainty around future volatility that is propagated in the system; as such, it may timely track increases in the risk premium required by the investors. We shall return to this point in subsection 5.4, where we investigate the link between our forecasts and the aggregate level of risk aversion, as measured by the VIX Index.

In addition, the IV predictive regression outperforms both the RV predictive regression and the historical mean benchmark in terms of economic value. In particular, when a simple switching strategy of the type described in Section 5.2 is implemented, an investor relying on the IV spillover-based forecast would achieve an average (annualized) return equal to 8.25%, which is significantly larger than the average return earned by an investor using the RV-based forecasts (5.9%) or than the one relying on the historical mean forecast (5.5%). This is despite the fact that both the historical mean and the RV-based forecast are more accurate as far as the sign predictions are concerned (they both display a percentage of correct sign predictions equal to 52%, in contrast to the slightly lower 51.62% of the IV-based forecasts). In addition, in order to consider also the risk borne by an investor, we compute the SRs derived from the alternative predictive models, and we find that an investor using the IV spillover-driven predictions would achieve an annualized SR equal to 0.79. Notably, an investor relying on the RV forecasts would realize a SR of 0.59, which is only slightly higher than the one obtained using the historical mean forecast (which is equal to 0.55). This happens even though the RV spillover forecasts were found to yield considerably higher average returns than those implied by the historical mean. This implies that using the RV-based forecasts entails taking on more risk than using the historical mean forecast.

The results concerning the stronger predictive power of the IV spillover index also hold when a mean-variance investor is considered. In this case, the investor using the IV-based forecasts would achieve a (massively) superior risk-adjusted performance than an investor using the RV-based forecasts (obtaining a SR of 0.71 vs. 0.31) and a slightly better performance compared to an investor who relies on the historical mean forecast (the latter achieves an annualized SR of 0.70). Moreover, a MV investor basing her portfolio choices on the IV spillover index forecast would obtain a higher average realized utility than an investor who employs historical mean or RV-based forecast. In particular, she would face an (annualized) utility gain of 3.29%, which can be interpreted as the (annualized) fee that she ought to be willing to pay to have access to a predictive model that filters information from the network of option implied volatilities.

Finally, we also report the predictive accuracy and the economic value of a forecast that is produced by regressing the excess returns on the changes in both the spillover indices simultaneously employed as predictors. Notably, while the RV index does not display a superior predictive power compared to the benchmark, when used in combination with the IV index it helps to produce an increase in the forecasting performance. Indeed, the predictive model based on both the indices, although it fails to display a positive OOS R-square, outperforms all the other models (including the benchmark) in terms of the economic value that it is able to generate. For instance, when switching strategies are considered, an investor exploiting the predictive model that contains both spillover indices would achieve an annualized SR of 0.90, which is considerably higher than 0.71 (which is the SR obtained using the IV-based forecasts). As far as the MV strategies are considered, an investor would be willing to pay 3.60% per year in order to get access to the forecasts based on both spillover indices, which is higher than the 3.29% scored by the IV-based forecasts.

These results are not surprising. Indeed, the IV- and RV-based forecasts based on the IV and RV indices are weakly correlated (their full-sample correlation is close to zero, namely -0.005). As noted by Rapach et al. (2010), when two predictors produce forecasts that are weakly (or un-) correlated, using both the predictors may help to stabilize the forecast.[21]

5.3. Long-horizon return predictability

Because actual investors would be probably interested in prediction and investment horizons longer than one week, in this subsection, we discuss the results obtained when we set the forecasting horizon to 4, 12, and 24 weeks corresponding to 1 month, a quarter, and a semester, respectively. More precisely, we estimate the (direct) predictive regression

$$r_{t+1:t+H,j} = \alpha_j^{(m,h)} + \beta_j^{(m,h)} x_{t,m} + \varepsilon_{t+1:t+H,j}^{(m,h)}, \tag{25}$$

in which $r_{t+1:t+H,j} = \sum_{i=1}^{H} r_{t+i,j}$ is the cumulative H-period excess return on stock j (or on the S&P 500 index), $x_{t,m}$ is the change of the RV (IV) index between t and t-H, and $\hat{\alpha}_j^{(m,h)}$ and $\hat{\beta}_j^{(m,h)}$ are the ordinary least square (OLS) estimates of the regression coefficients obtained using the data available at time t with the same expanding window procedure described in subsection 5.2. The forecast for the H-period excess return on asset j is then given by

[21] It is worthwhile to notice that we could also use a weighting scheme to combine the forecasts coming from the two alternative models. For instance, as recently discussed by Tsiakas, Li, and Zhang (2020), because combining two negatively correlated assets in a portfolio produces a large diversification gain, a forecast combination of two negatively correlated forecasts yields a variance of the forecast error that is lower than the average variance of the forecast errors of the individual models. However, we do not pursue this analysis as the main goal of this paper is to show the predictive power of the IV spillover index, not to find the best predictive model for the market equity risk premium.

$$\hat{r}_{t+1:t+H|t,j} = \hat{\alpha}_j^{(m,h)} + \hat{\beta}_j^{(m,h)} x_{t,m}. \quad (26)$$

To assess the forecasting performance of the alternative models with respect to the benchmark, we use the same statistics proposed in subsection 5.2. However, because we use overlapping returns on the left-hand side of equation (21), the resulting H-step-ahead forecast errors will be autocorrelated by construction. Therefore, when we test for equal predictive accuracy with respect to the benchmarks, we use autocorrelation consistent standard errors to compute the t-statistics.[22]

Table 6 has a similar structure to Table 4, but each panel reports the results for a different forecast horizon: 1 month (Panel A), 3 months (Panel B), and 6 months (Panel C). In the first row of each panel, we report the results concerning the equity risk premium. Instead, in the case of the individual stocks, we report the average values across different industries (similarly to Table 4). While the results for a one-month horizon closely resemble those obtained for the one-step-ahead predictions that we have discussed in subsection 5.2, when 3- and 6-month horizons are analyzed, we find considerable long-term predictability vis-à-vis the historical mean for both predictors and not only for the IV spillover index. In particular, when we examine the predictability of the equity risk premium, we obtain values of the R-square that are equal to 4.75% and 8.40% (for RV and IV predictive regressions, respectively) for $H = 12$ and to 12.58% and 15.59% for $H = 24$.

However, we shall refrain from interpreting Table 6 as a stark evidence in favor of long horizon return predictability. Indeed, recent literature (see, for instance, Boudoukh et al., 2019), strikes a cautionary note on the interpretation of the results on long-term predictability, as the corrections that have been proposed to address the autocorrelation issue coming from the use of overlapping returns may prove insufficient to avoid systematic under-rejection of the

---

[22] More precisely, we follow the procedure proposed by Clark and West (2007): we compute $\tilde{d}_{j,n_1+s}^{(m,H)}$ as in (18) and we regress it on a constant. However, we use the Newey-West standard errors to compute the t-statistic associated to the constant. We reject the null hypothesis of equal predictive ability when the statistic is greater than +1.282 (for a one-sided 10% sized test) or +1.645 (for a one-sided 5% size test). We only report the significance of the results concerning the equity risk premium.

null hypothesis of equal predictive ability. In any event, we can safely state that the proportion of correct sign prediction increases with the forecast horizon for both spillover-driven predictors, reaching values in excess of 65%. For instanc,, when we forecast the equity premium at a six-month horizon, the proportion of correct sign predictions are 66.93% in the case of the RV index and 63.81% for the IV index.

5.4. Forecasting during low and high volatility regimes

As discussed in Section 5.2 when commenting Figure 6, in light of the literature we suspect that the IV spillover index may carry a strong predictive relationship with equity returns especially during bear markets. To test this hypothesis, we compute the RMSFE for the IV and RV spillover index-based and the historical mean forecasts during times of high and low market volatility. To assign the observations to these two regimes, we use the S&P 500 implied volatility index, namely the VIX, whose increase is often regarded as an indicator of distress in the market. In particular, we adopt the following method: at any time *t,* we compare the current value of the VIX index with its 1-year moving average (52 weekly observations); if this value is 20% above the moving average, we classify the time *t* observation as belonging to a high-volatility regime; otherwise, we classify the observation as belonging to a low-volatility period.

In Table 7, we evaluate the forecasting power of the two spillover indices (and of a combination of the two) for the equity risk premium under the two regimes defined above. As we conjectured, while in the low-volatility regime none of the predictive models is able to outperform the historical mean (all the OOS R-squares are negative), in the high-volatility regime the IV-based forecasts display a positive (and rather large) OOS R-square (6.65%). In order to assess the economic value of the alternative forecasting models, we also report the SRs of the switching and the MV investing strategies for the alternative predictive models (and for the historical mean benchmark). Interestingly, despite it is not able to outperform the benchmark in terms of statistical predictive accuracy, the IV-based predictive model implies a better risk-adjusted performance compared to the historical mean in both the regimes. However, in the low-volatility regime, an investor using the IV spillover predictions would only slightly outperform an investor using the historical mean forecasts (achieving a SR of 1.61 vs. 1.49 when a simple switching strategy is considered and of 1.57 vs. 1.53 when a MV strategy is applied). Conversely, when the high-volatility period is considered, a MV investor using the IV-based forecasts would achieve a SR of -0.84, which is far less dreadful vs. the -2.30 obtained by an investor who relies on the historical mean and hence fails to exploit the predictability patterns we have uncovered.

Notably, while the RV spillover index-based model strongly underperforms the IV-based one in terms of economic value in the high-volatility regime, the difference is weaker during the low-volatility regime. On the contrary, during the low-volatility regime a MV investor using the RV-based forecasts would slightly outperform an investor using the IV-based forecasts in terms of risk-adjusted performance. An (unreported) analysis of the MV weights shows that employing a predictive model based on the IV spillover index allows an investor to massively switch the allocation towards the risk-free bond in a more timely manner than the alternative predictive model based on the RV spillover index during times of distress (and especially between the end of 2008 and the beginning of 2009). This seems to confirm our intuition that the IV spillover index is able to predict market distress much better than the RV index does.

6. Spillover effects vs. the predictive power of the VIX

Considering that (the changes in) the IV spillover index strongly outperforms the RV spillover index, especially in times of high market volatility, one may wonder whether this differential predictive power may come from the fact that we are capturing spillover effects or we are just featuring the same information that is embedded in option-implied volatilities. Indeed, previous literature (see, e.g., Banerjee et al., 2007) has shown that the VIX index displays some predictive power for equity returns. In addition, Ang et al. (2006) argue that aggregate volatility risk, as proxied by innovations to the VIX index is priced in the cross section of stock returns. Although their claim is not a predictive one in a formal sense and there is a clear logical difference between implied volatilities and building an aggregate network spillover index based on IVs, such findings are of course consistent with the existence of a direct predictive power for changes in implied volatilities for stock returns. Therefore, in this subsection, we estimate a one week ahead predictive regression of aggregate excess returns on the (changes) in the VIX index and we assess its forecasting performance in-sample and OOS, in comparison with the results based on the IV spillover index. We also estimate a predictive regression that contains both the VIX and the IV spillover indices to assess whether the slope coefficient of the IV spillover index remains statistically significant also when the (changes in the) VIX is included in the regression.

The results of this analysis are reported in Table 8. Panel A shows that the loading of (the changes of) the VIX index is small but statistically significant when the VIX is used alone in the predictive regression. However, when the IV spillover index is included in the predictive regression, the slope coefficient of the VIX index turns out not to be statistically different from zero. Panel B shows the predictive accuracy of the VIX index (alone and in combination with the IV spillover index) when it is used to recursively forecast the equity risk premium OOS. The panel also reports results on the economic value that is generated when such forecasts are used to form a mean-variance portfolio or, alternatively, to implement simple switching strategies as described in Section 5. Interestingly, none of the two predictive regressions outperforms the historical mean benchmark as far as the forecasting accuracy is concerned. In fact, the OOS R-square is equal to -1.96% when the changes in the VIX index are used as the sole predictor and to -0.33% when the VIX and the IV spillover index are used in combination. Interestingly, this means that the addition of the VIX index to an IV-based predictive regression deteriorates its forecasting power even though, at least in our sample, the VIX by itself possess no OOS predictive accuracy.

As far as the economic value is concerned, a predictive regression including the VIX does not outperform the historical mean benchmark as it is evident from the fact that the expected utility gain is negative. Although the expected utility gain from the predictive regressions based on both the VIX and the IV spillover index is positive, it is equal to 0.23% only in annualized terms, which is lower compared to 3.29% that was achieved in Table 5 when the IV spillover index was used as the sole predictor. Finally, unreported results show that the correlation between the IV- and the VIX-based forecasts is rather moderate (0.30). Therefore, we conclude that the IV index is not just a different (and more convoluted) way to capture aggregate volatility risk but captures instead different, richer information useful to predict market excess returns.

7. Robustness checks

In this section, we test the robustness of our results concerning the predictive accuracy of the alternative spillover indices to a set of different assumptions. In Panel A of Table 9, we report the forecasting accuracy statistics that we obtain when the excess returns are computed by subtracting the 1-week US LIBOR instead of the 1-month T-bill rate. Notably, the results are very similar to those already commented in Section 5. For instance, the OOS R-square for a predictive regression of the equity risk premium on (the changes of) the IV spillover index is 2.08% when the 1-week US LIBOR is used, which is really close to the 2.11% that we reported in Table 5. Similarly, the OOS R-square for a predictive regression of the equity risk premium on (the change of) the RV index is -4.80%, while it was -4.74% in Table 5.[23]

In Panel B, we report the results that we obtain when we compute the RV (IV) index as the average of the indices resulting from a sub-sample of 60 randomly picked stocks (we repeat the experiment ten times). The results are in line with those already reported and commented in Section 5. For instance, the OOS R-square for a predictive regression of the equity risk premium on the (changes in the) IV and RV spillover indices are 2.01% and -5.12%, respectively. All the main results concerning the individual stocks previously shown in Table 4 are confirmed as well. For instance, the energy sector is still the one showing the largest amount of predictability, while the technology sector is the one showing the lowest strength of predictability (as measured by the OOS R-square). Therefore, we conclude that our results are not driven by a specific choice of the stocks included in the panel.

[23] This is not surprising considering that the correlation between the two proxies of the risk-free rate (namely, the 1-week US LIBOR and the 1-month T-bill rate) is 0.85.

Finally, in Panel C of Table 9, we show the statistics that we obtain when we compute the RV (IV) index as the average of the indices resulting from nine alternative GFEVD where the forecasting horizon H varies from 2 to 10. Although the overall message does not change and we continue to find evidence of predictive power for the IV spillover index at least as far as the equity risk premium is concerned, the OOS R-squares are lower in this case, and they turn negative for most of the sectors when we try to forecast individual stock returns. In particular, the OOS R-square corresponding to the equity risk premium predictive regression drops to 0.93% (from 2.10% when only H = 2 was considered). Moreover, as far as the individual stocks are concerned, only the energy sector displays a positive average R-square of 0.11%. This confirms our intuition that a short-horizon FEVD is more informative for the sake of our application. Also in this case, the RV spillover index fails to display forecasting power for either the equity risk premium or the individual stock excess returns as it always yields negative R-squares.

Considering that our analysis was carried out at a weekly frequency, and therefore that the spillover index series can be noisy, we have also experimented with a number of transformations of the two indices to investigate whether eliminating the noise can improve their predictive ability. For instance, we regress the weekly returns on a moving average of the changes of the indices over the previous month; in addition, we also compute the changes as the difference between the value of each index at time t and a moving average of the values assumed by the index over the previous month. However, none of these transformations improves the predictive ability of the RV or of the IV spillover indices; conversely, they lead to a loss of important information and actually decrease the OOS forecasting power of the two indices.

7. Conclusions

In this paper, we have analyzed whether two alternative indices that measure the average strength of volatility spillovers in a network of stocks constructed following the methodology introduced by Diebold and Yilmaz (2009, 2012), are able to predict equity excess returns. In order to construct the network, we rely alternatively on realized volatility (as suggested by Diebold and Yilmaz) and on implied volatilities, extracted from a set of liquid, nearly at-the-money options traded on the CBOE. Indeed, as option implied volatilities are forward-looking by their nature, our fundamental hypothesis is that they can be used to capture early signals of an increase in the spillovers of volatilities and therefore an increase in systemic risk. In the measure in which systemic risk is at least partially a source of undiversifiable, systemic risk, it is legitimate to expect that forward-looking network spillover indices may contain yet untapped forecasting power.

Our intuition is confirmed by the empirical results that we have reported throughout. First, we note that the loading of excess returns on the IV spillover index is significant both when the index is used as an individual predictor and when it used in combination with the RV index. In addition, the sign of this coefficient is positive, as expected: indeed, an increase in volatility spillovers is linked to a higher systemic (hence, systematic, to some extent) risk and therefore to a higher required risk premium. Second, we report a significant predictive power of (the changes of) the IV spillover index both in-sample and (especially) OOS compared to a no-predictability benchmark (where returns are assumed to be constant and equal to the historical mean). Notably, the stronger predictability also generates significant economic value when different trading strategies are considered (i.e., a simple switching strategy where the investor decides to allocate all her wealth to equity when the forecasted return is positive and a mean-variance asset allocation strategies that considers equities and a risk-free bond as the asset menu). In particular, a mean-variance investor will achieve a utility gain of more than 3% on annualized basis from using a predictive model based on the IV spillover index. In contrast, despite showing a statistically significant coefficient and some evidence of in-sample predictive power, the predictive model based on RV spillovers is never able to outperform the benchmark in OOS experiments or to deliver a positive utility gain.

Interestingly, the predictive power of the IV index is particularly evident in times of high market volatility. In periods of low volatility, the IV spillover index does not outperform the historical mean benchmark, at least in terms of statistical accuracy. However, when markets are highly volatile (as signaled by spikes in the VIX index) the IV spillover index-based forecasts strongly outperform both the benchmark and the RV-based predictions both in terms of statistical accuracy and of economic value. In particular, it is worthwhile to note that an investor using the IV-based forecasts would have reduced her allocation to equity more quickly during 2008-2009 compared to an investor using the historical mean or the forecast based on the RV spillover index. Overall, an investor using RV-based forecasts would benefit in normal times, at the cost of suffering massive losses during the crisis periods. This seems to confirm our intuition that the IV spillover index can be used to perform early detection of situations of market distress. It is also worthwhile to notice that using changes in VIX index to predict equity returns would fail to lead to the same results of the IV index in terms of improvements in the forecasting ability.

These findings lead an important question open: given their dynamic, accurate relationship with aggregate excess returns, is the presence of volatility spillovers a systematic risk-factor priced in the cross-section of stocks? If this were the case, we would expect that stocks with different sensitivities to changes to the IV (RV) index should have different expected returns. A systematic investigation of the implications of our results to asset pricing represents an interesting venue for future research.

REFERENCES

Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106(7), 1705-41.

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *Review of Financial Studies*, 30(1), 2-47.

Allen, L., Bali, T. G., & Tang, Y. (2012). Does systemic risk in the financial sector predict future economic downturns?. *Review of Financial Studies*, 25(10), 3000-3036.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61(1), 259-299.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, 52(5), 2003-2049.

Bali, T. G., & Hovakimian, A. (2009). Volatility spreads and expected stock returns. *Management Science*, 55(11), 1797-1812.

Baltussen, G., Van Bekkum, S., & Van der Grient, B. (2018). Unknown Unknowns: Uncertainty About Risk and Stock Returns. *Journal of Financial and Quantitative Analysis*, 53(4), 1615-1651.

Banerjee, P.S., Doran, J.S., Peterson, D.R. (2007). Implied volatility and future portfolio returns. *Journal of Banking and Finance*, 31, 3183–3199.

Battalio, R., & Schultz, P. (2006). Options and the bubble. *Journal of Finance*, 61(5), 2071-2102.

Billio, M., Caporin, M., Panzica, R., & Pelizzon, L. (2016). The impact of network connectivity on factor exposures, asset pricing and portfolio diversification. Working Paper.

Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of financial economics*, 104(3), 535-559.

Blair, B., Poon, S.H., & Taylor, S. (2000). Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High-Frequency Index Returns. *Journal of Econometrics*, 105, 5-26.

Boudoukh, J., Israel, R., & Richardson, M. (2019). Long-Horizon Predictability: A Cautionary Tale. *Financial Analysts Journal*, 75(1), 17-30.

Brownlees, C., & Engle, R. F. (2016). SRISK: A conditional capital shortfall measure of systemic risk. *Review of Financial Studies*, 30(1), 48-79.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average?. *Review of Financial Studies*, 21(4), 1509-1531.

Carr, P., & Wu, L. (2008). Variance risk premiums. *Review of Financial Studies*, 22(3), 1311-1341.

Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. *Journal of financial economics*, 50(2), 125-150.

Clark,T.E., & West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics,* 138(1), 291–311.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3), 229-263.

Dal Pra, G., Guidolin, M., Pedio, M., & Vasile, F. (2018). Regime shifts in excess stock return predictability: an out-of-sample portfolio analysis. *The Journal of Portfolio Management*, *44*(3), 10-24.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 20(1), 134-144.

Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, 119(534), 158-171.

Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, 28(1), 57-66.

Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1), 119-134.

Driessen, J., Maenhout, P. J., & Vilkov, G. (2009). The price of correlation risk: Evidence from equity options. *Journal of Finance*, 64(3), 1377-1406.

Engle, R., & Kelly, B. (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics*, 30(2), 212-228.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of financial economics*, 22(1), 3-25.

Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of empirical finance*, 5(4), 317-345.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1), 1.

Goyal, A., & Saretto, A. (2009). Cross-section of option returns and volatility. *Journal of Financial Economics*, 94(2), 310-326.

Hsu, N. J., Hung, H. L., & Chang, Y. M. (2008). Subset selection for vector autoregressive processes using lasso. *Computational Statistics & Data Analysis*, 52(7), 3645-3657.

Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press. Princeton, New Jersey.

Lütkepohl, H. (2005). New introduction to multiple time series analysis. Springer Science & Business Media.

Nicholson, W. B., Matteson, D. S., & Bien, J. (2017). VARX-L: Structured regularization for large vector autoregressions with exogenous variables. *International Journal of Forecasting*, 33(3), 627-651.

Nicholson,W., Matteson,D., & Bien, J. (2019). BigVAR: Dimension Reduction Methods for Multivariate Time Series. R package version 1.0.4. https://CRAN.R-project.org/package=BigVAR.

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1), 17-29.

Pesaran, M.H., Timmermann, A. (1995). Predictability of stock returns: robustness and economic significance. *Journal of Finance*, 50, 1201-1228.

Piccotti, L. R. (2017). Financial contagion risk and the stochastic discount factor. *Journal of Banking & Finance*, 77, 230-248.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, 23(2), 821-862.

Buraschi, A., & Tebaldi, C. (2017). Asset Pricing Implications of Systemic Risk in Network Economies. Working Paper.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso*. Journal of the Royal Statistical Society: Series B*, 58(1), 267-288.

Tsiakas, I., Li, J., & Zhang, H. (2020). Equity premium prediction and the state of the economy. *Journal of Empirical Finance*, 58, 75-95.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4), 1455-1508.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica: Journal of the Econometric Society*, 1067-1084.

Table 1

List of stocks

This table contains the list of all the stocks included in the analysis and the industry to which they belong. It also reports their average market cap (in millions of USD), and their mean realized and implied volatilities (expressed in percentage points) over the sample period January 2006 – December 2017.

Ticker	Name	Sector	Avrg. Mkt Cap (mln)	Mean RV	Mean IV
DISH	Dish Network Corp	Communication	18,370.36	32.99	39.17
ANF	Abercrombie & Fitch	Consumer Discretionary	3,411.05	45.30	46.46
AZO	Autozone Inc	Consumer Discretionary	13,859.06	21.89	24.85
BBBY	Best Buy Co Inc	Consumer Discretionary	10,675.54	29.29	30.72
CCL	Carnival Corp	Consumer Discretionary	32,017.40	29.60	31.28
KSS	Kohls Corp	Consumer Discretionary	13,075.98	31.38	32.86
NKE	Nike Inc	Consumer Discretionary	55,077.01	24.18	25.44
ROST	Ross Stores Inc	Consumer Discretionary	12,691.40	25.98	29.38
SWK	Stanley Black & Decker Inc	Consumer Discretionary	11,041.04	26.26	26.79
TOL	Toll Brothers Inc	Consumer Discretionary	4,628.63	35.84	38.68
VFC	VF Corp	Consumer Discretionary	16,900.24	25.25	26.72
WHR	Whirlpool Corp	Consumer Discretionary	8,654.96	34.83	34.93
CL	Colgate-Palmolive Corp	Consumer Staples	48,487.43	16.69	18.56
COST	Costco Wholesale Corp	Consumer Staples	43,466.27	20.31	21.75
KMB	Kimberly-Clark Corp	Consumer Staples	33,798.53	15.74	17.27
APA	Apache Corp	Energy	29,518.49	34.10	34.31
APC	Anadarko Petroleum Corp	Energy	32,853.87	37.17	38.29
BHGE	Baker Hughes a Co	Energy	22,703.04	37.27	36.93
CNQ	Canadian Natural Resources Ltd	Energy	35,517.55	38.24	37.99
CVX	Chevron Corp	Energy	190,215.78	22.65	22.97
EOG	EOG Resources Inc	Energy	34,398.97	34.79	35.67
HES	Hess Corp	Energy	20,691.63	37.38	37.62
NOV	National Oilwell Varco Inc	Energy	22,019.34	39.61	39.61
OXY	Occidental Petroleum Corp	Energy	61,644.42	30.02	30.85
RDC	Rowan Cos Plc	Energy	3,303.47	42.10	43.40
SLB	Schlumberger Ltd	Energy	98,291.87	31.39	32.19
AON	AON Plc	Financials	19,257.02	20.71	23.13
AXP	American Express Co	Financials	64,263.06	29.02	29.65
COF	Capital One Financials Corp	Financials	29,988.63	36.50	36.19
GS	Goldman Sachs Inc	Financials	77,555.44	31.02	30.88
LM	Legg Mason Inc	Financials	5,455.50	37.19	36.86
PNC	PNC Financial Services Group Inc	Financials	35,309.92	30.64	29.24
AMGN	Amgen Inc	Health Care	80,273.10	24.21	26.15
BAX	Baxter International Inc	Health Care	32,730.18	19.93	22.03
CELG	Celgene Corp	Health Care	49,424.31	32.18	33.72
CI	Cigna Corp	Health Care	20,043.98	30.44	32.27
DHR	Danaher Corp	Health Care	39,654.32	20.63	22.53
ESRX	Express Scripts Co	Health Care	34,830.85	27.18	29.58
GILD	Gilead Sciences Inc	Health Care	72,978.24	28.63	30.60
ISRG	Intuitive Surgical Inc	Health Care	16,215.98	35.92	38.27
LH	Laboratory Corp of America	Health Care	9,652.85	19.62	22.33
MCK	McKesson Corp	Health Care	26,400.59	24.52	24.92
BA	Boeing Co	Industrials	72,243.62	24.99	26.40
CAT	Caterpillar Inc	Industrials	50,490.74	28.56	29.80
CMI	Cummins Inc	Industrials	17,881.93	35.69	35.33
FDX	FedEx Corp	Industrials	34,953.76	25.92	27.17
GD	General Dynamics Corp	Industrials	34,366.31	21.36	22.54
GWW	WW Grainger Inc	Industrials	11,359.98	23.73	25.24
HON	Honeywell International Inc	Industrials	56,493.67	22.48	23.93
LLL	L3 Technologies Inc	Industrials	9,887.53	21.16	23.11
PCAR	Paccar Inc	Industrials	18,125.18	30.57	32.24
RTN	Raytheon Co	Industrials	26,823.65	19.65	21.01
UNP	Union Pacific Corp	Industrials	57,516.07	25.99	27.13
PH	Parker-Hannifin Corp	Machinery	13,693.50	27.46	29.12
APD	Air Products & Chemicals Inc	Materials	21,920.80	23.33	24.44
CCJ	Cameco Corp	Materials	8,796.96	39.43	39.06
MLM	Martin Marietta Materials	Materials	6,408.09	31.91	34.29
MMM	3M Co	Materials	75,117.83	19.05	20.45
MON	Monsanto Co	Materials	46,268.01	27.49	29.68
NUE	Nucor Corp	Materials	15,307.93	34.03	34.90
PX	Praxair Inc	Materials	30,204.44	21.36	22.82
SCCO	Southern Copper Corp	Materials	25,527.68	39.12	39.27
RL	Ralph Loren Corp	Retail Discretionary	10,099.89	32.70	33.26
TPR	Tapestry Inc	Retail Discretionary	12,860.11	34.76	36.10
AAPL	Apple Inc	Technology	398,732.91	29.58	32.81
CERN	Cerner Corp	Technology	12,588.18	28.20	31.01
CTSH	Cognizant Technology Solution	Technology	22,533.15	32.53	33.48
IBM	International Business Machine Corp	Technology	170,723.40	19.62	21.08
INTU	Intuit Inc	Technology	18,296.69	24.92	26.70
ETR	Entergy Corp	Utilities	14,618.84	19.41	20.82

Table 2

Predictive regressions for the equity premium

This table reports the results of the estimation of the predictive regression

$$r_{t+1} = \alpha + \beta_{RV}\Delta\text{RVSI}_t + \beta_{IV}\Delta\text{IVSI}_t + \varepsilon_{t+1},$$

where r_{t+1} is the weekly excess return (over the one month T-bill) of the S&P 500 index, and ΔRVSI_t (ΔIVSI_t) is the change between time $t-1$ and t of the realized (implied) volatility spillover index. The sample period is December 2006 – December 2017. Models (II) and (III) are restricted versions of the baseline predictive regression above in which the coefficients β_{IV} and β_{RV} are alternatively set to be equal to zero. The R-square coefficients are reported in percentages (e.g., i.e., 1.00, means 1.00%).

	(I)	(II)	(III)
Intercept	0.0014	0.0013	0.0014
(t-stat)	(1.4674)	(1.4029)	(1.4663)
β coeff. RV	-0.3116	-0.2959	
(t-stat)	(-3.0871)	(-2.8971)	
β coeff. IV	0.4602		0.4463
(t-stat)	(3.9845)		(3.8390)
R-square (%)	4.11	1.44	2.51

Table 3

Predictive regressions for the excess returns of individual stocks

This table reports the results of the estimation of a set of predictive regressions of the type

$$r_{t+1,j} = \alpha_j^{(m)} + \beta_j^{(m)} x_{t,m} + \varepsilon_{t+1,j}^{(m)},$$

where $r_{t+1,j}$ is the weekly excess return (over the 1-month T-bill) of an individual stock j, and $x_{t,m}$, with $m = \Delta RVSI, \Delta IVSI$, is the change between time $t-1$ and t of the realized (implied) volatility spillover index. The sample period is December 2006 – December 2017. We report the mean, median, minimum, and the maximum values of the slope coefficients (for $x_{t,m} = RV_t$ and $x_{t,m} = IV_t$, respectively) and the R-square coefficients across industry sectors. The R-square coefficients are reported in percentages (e.g., 1.00, means 1.00%). For each of the two sets of predictive regressions, we also report the number of significant slope coefficients at the 10-, 5-, and 1-percent test size levels.

	RV Regression					IV Regression				
	Mean	Median	Std Dev.	Min	Max	Mean	Median	Std Dev.	Min	Max
β coeff. - All	-0.36	-0.36	0.15	-0.78	-0.10	0.44	0.40	0.21	0.00	1.03
β coeff. -Consumer Discretionary	-0.40	-0.39	0.22	-0.78	-0.13	0.35	0.37	0.11	0.15	0.49
β coeff. -Energy	-0.42	-0.40	0.13	-0.67	-0.25	0.71	0.66	0.18	0.35	1.03
β coeff. - Financials	-0.33	-0.36	0.14	-0.48	-0.12	0.52	0.51	0.22	0.31	0.74
β coeff. -Health Care	-0.31	-0.31	0.12	-0.49	-0.11	0.36	0.36	0.16	0.17	0.74
β coeff. - Industrials	-0.34	-0.33	0.08	-0.46	-0.19	0.40	0.41	0.11	0.22	0.60
β coeff. - Materials	-0.44	-0.41	0.16	-0.69	-0.22	0.48	0.36	0.22	0.27	0.78
β coeff. - Technology	-0.23	-0.22	0.10	-0.38	-0.10	0.34	0.35	0.19	0.04	0.53
β coeff. - Others	-0.35	-0.32	0.14	-0.54	-0.18	0.35	0.30	0.21	0.00	0.70
R square (%)	0.73	0.62	0.44	0.05	2.06	0.92	0.86	0.59	0.00	2.41
N. of significant β coeff. (α=10%)	51					50				
N. of significant β coeff. (α=5%)	32					44				
N. of significant β coeff. (α=1%)	10					20				

Table 4

Out-of-sample forecast evaluation for the excess returns of individual stocks

This table reports statistics on forecast errors for out-of-sample (OOS) recursively estimated predictive regressions of individual stock excess returns on (changes of) the two alternative spillover indices. We report the difference in root mean square forecast error (RMSFE) between each of the two predictive models and the historical mean benchmark, $\hat{r}_{t+1,j|t} = \sum_{i=1}^{n_2} r_{i,j}$. A positive value of ΔRMSFE means that the predictive model has a lower RMSFE than the benchmark. We also report Campbell and Thompson's (2008) OOS R-square (OOS R2) and the percentage of correct sign predictions for the two alternative predictive models and for the benchmark. The R-squares and the proportion of correct sign predictions are expressed as percentages, e.g., 1.00, means 1.00%.

	RV			IV			Benchmark
	ΔRMSFE	OOS R2	Correct sign	ΔRMSFE	OOS R2	Correct sign	Correct sign
Avrg. All	-0.0006	-2.94	51.51	0.0001	0.3283	50.88	50.66
Avrg. Energy	-0.0007	-2.70	49.99	0.0003	1.0534	51.38	48.80
Avrg. Consumer Discret.	-0.0007	-3.48	52.55	0.0000	0.0548	51.10	50.79
Avrg. Financials	-0.0010	-4.04	50.60	0.0001	0.2645	49.65	49.46
Avrg. Health Care	-0.0004	-1.97	51.85	0.0000	0.0930	50.78	51.54
Avrg. Industrials	-0.0003	-1.40	52.31	0.0001	0.3597	50.49	50.72
Avrg. Technology	-0.0014	-7.07	53.52	0.0000	-0.2793	53.22	54.51
Avrg. Materials	-0.0006	-3.08	50.45	0.0001	0.2169	50.93	50.90
Avrg. Others	-0.0004	-2.30	51.10	0.0001	0.4976	49.98	50.14

Table 5

Out-of-sample forecast evaluation for the aggregate equity risk premium

This table reports statistics concerning the forecast errors from out-of-sample (OOS) recursively estimated predictive regressions for the S&P 500 excess returns on (changes in) two spillover indices (either specified together or used alternatively). We report the root mean square forecast error (RMSFE) and the difference between the RMSFE of each of the predictive models and the historical mean benchmark, $\hat{r}_{t+1,j|t} = \sum_{i=1}^{n_2} r_{i,j}$. A positive value of ΔRMSFE means that the predictive model has a lower RMSFE than the benchmark. We compute Clark and West's (2007) MSFE-adjusted statistic to assess whether a positive ΔRMSFE is statistically significant. A rejection of the null that $\Delta\text{RMSFE} = 0$ at a 10-percent size is denoted by *, while a rejection at a 5-percent size is denoted by **. We also report Campbell and Thompson's (2008) OOS R-square (OOS R2) and the percentage of correct sign predictions for the alternative predictive models. In addition, we report the statistics concerning the economic value of the alternative forecasting models. In particular, we report the average annualized returns and Sharpe ratios (SR) of a switching strategy (Pesaran and Timmermann, 1995) in which the investor takes a long position in the equity at any time a positive return is forecasted, while she invests in the risk-free bond otherwise. Finally, we report the average annualized returns, Sharpe ratios (SR), and average realized utility for a mean-variance asset allocation strategy ($\gamma = 3$). The average utility gain represents the (annualized) fee that the investor would be willing to pay to access the spillover index-based models relative to the historical average benchmark forecast. The R-squares, the proportion of correct sign predictions, the annualized returns, and the annualized realized utility (and utility gains) are all expressed as percentages, e.g., i.e., 1.00, means 1.00%.

	Predictive Accuracy			Switching Strategies			Mean Variance Asset Allocation			
Predictive variable	RMSFE	ΔRMSFE	OOS R2	Correct sign	Ann. return	Ann. SR	Ann. return	Ann. SR	Ann. realized utility	Ann. utility gain
RV	0.0241	-0.0006	-4.74	52.19	7.21	0.59	3.78	0.31	1.87	-1.19
IV	0.0233	0.00025**	2.11	51.62	8.25	0.79	8.26	0.71	6.35	3.29
RV + IV	0.0239	-0.0004	-3.04	50.48	10.00	0.90	9.35	0.68	6.66	3.60
Benchmark (Hist- mean)	0.0235	-	-	52.19	5.71	0.55	3.35	0.70	3.06	-

Table 6

Long-horizon return predictability

This table reports statistics on forecast errors for out-of-sample (OOS) recursively estimated long-horizon predictive regressions

$$r_{t+1:t+H,j} = \alpha_j^{(m,h)} + \beta_j^{(m,h)} x_{t,m} + \varepsilon_{t+1,j}^{(m,h)},$$

where $r_{t+1:t+H,j} = \sum_{i=1}^{H} r_{t+i}$, H is the forecast horizon and j indicates either the S&P 500 or one of the individual stocks in the sample; $x_{t,m}$ is the vector of the change in each of the two alternative spillover indices ($m = \Delta RV, \Delta IV$). H is set to 4, 12, and 24 weeks (covered by Panels A, B, and C, respectively) corresponding to 1, 3, and 6 months. For each H, we report the difference in the root mean square forecast error (RMSFE) between each predictive model and the historical mean, the OOS R-square, and the percentage of correct sign predictions, as in Table 2. In the case of the S&P 500, we compute Clark and West's (2007) MSFE-adjusted statistic to assess whether a positive ΔRMSFE is statistically significant. A rejection of the null that ΔRMSFE $= 0$ at a 10-percent size is denoted by *, while a rejection at a 5-percent size is denoted by **. The OOS R-square and the proportion of correct sign predictions are expressed as percentages, e.g., 1.00, means 1.00%.

	RV			IV			Benchmark
	ΔRMSFE	OOS R2	Correct sign	ΔRMSFE	OOS R2	Correct sign	Correct sign
Panel A - H = 4 (1-month ahead)							
S&P 500	-0.0006	-2.81	59.77	0.0006**	2.47	54.98	55.75
Avrg. All	-0.0009	-2.07	53.85	0.0006	1.49	52.71	51.63
Avrg. Energy	-0.0012	-2.18	49.74	0.0004	0.82	50.73	48.36
Avrg. Consumer Discret.	-0.0011	-3.03	55.49	0.0006	1.34	53.88	52.68
Avrg. Financials	-0.0005	-0.88	50.73	0.0010	2.09	48.66	47.54
Avrg. Health Care	-0.0010	-2.19	56.07	0.0006	1.54	54.56	54.35
Avrg. Industrials	-0.0004	-0.95	54.95	0.0006	1.61	53.13	52.37
Avrg. Technology	-0.0011	-3.11	57.62	0.0007	1.94	57.32	56.67
Avrg. Materials	-0.0004	-0.91	53.07	0.0006	1.43	52.71	51.70
Avrg. Others	-0.0011	-3.35	53.71	0.0006	1.68	51.10	50.12
Panel B - H = 12 (3 months ahead)							
S&P 500	0.0019**	4.75	66.93	0.0035**	8.40	63.81	60.89
Avrg. All	0.0027	3.28	57.11	0.0050	6.36	55.51	53.31
Avrg. Energy	0.0056	5.78	52.62	0.0066	6.86	52.99	48.34
Avrg. Consumer Discret.	0.0003	-0.07	57.15	0.0035	4.48	55.15	53.75
Avrg. Financials	0.0033	3.64	53.50	0.0061	6.63	50.13	48.99
Avrg. Health Care	0.0011	2.24	60.74	0.0040	5.52	59.22	57.82
Avrg. Industrials	0.0033	4.34	59.46	0.0048	6.64	56.60	54.12
Avrg. Technology	0.0029	3.95	64.59	0.0053	7.67	63.42	61.75
Avrg. Materials	0.0054	6.94	56.25	0.0062	8.14	55.40	52.99
Avrg. Others	0.0004	-0.05	54.35	0.0040	6.10	52.50	51.09
Panel C - H = 24 (6 months ahead)							
S&P 500	0.0084**	12.58	72.71	0.0106**	15.59	71.12	66.53
Avrg. All	0.0113	9.52	58.68	0.0139	11.93	57.60	54.13
Avrg. Energy	0.0151	10.64	52.28	0.0165	11.69	51.20	46.25
Avrg. Consumer Discret.	0.0067	4.88	60.00	0.0094	7.61	58.91	56.83
Avrg. Financials	0.0104	7.32	51.76	0.0168	11.72	49.93	46.58
Avrg. Health Care	0.0104	10.15	63.09	0.0129	12.04	61.81	58.82
Avrg. Industrials	0.0120	10.76	60.34	0.0153	13.85	59.02	53.57
Avrg. Technology	0.0124	12.18	67.17	0.0134	13.14	67.13	64.06
Avrg. Materials	0.0160	13.56	59.99	0.0170	14.53	59.29	57.74
Avrg. Others	0.0075	7.80	56.47	0.0112	12.20	55.50	51.97

Table 7

Out-of-sample forecast evaluation under different volatility regimes

This table reports the OOS R-square for three alternative predictive models (based on changes in the RV spillover index, on changes of the IV spillover index, and on both the predictive variables) over two regimes, namely, a high-volatility and a low-volatility regime. Each observation is classified as belonging to a (low-) high-volatility period if the level of the VIX is (below) above the 120% of the value of the moving average of the VIX over the previous year. The annualized Sharpe ratios (SR) for the switching and the mean-variance (MV) investment strategies based on the three alternative predictive models and on the historical mean are also reported. The R-squares are expressed in percentages, i.e., 1.35 means 1.35%.

	Low Volatility			**High Volatility**		
Predictive variable	OOS R2	SR (Switch. Strategy)	SR (MV)	OOS R2	SR (Switch. Strategy)	SR (MV)
RV	-1.35	1.10	1.62	-7.86	-1.94	-2.14
IV	-2.88	1.61	1.57	6.65	-0.65	-0.84
RV + IV	-3.54	1.36	1.60	-2.60	-0.70	-0.95
Benchmark (Hist. Mean)	-	1.49	1.53	-	-2.00	-2.30

Table 8

VIX vs. IV spillover index-based predictive regressions

The table reports the in-sample and OOS results of a predictive regression for the aggregate equity risk premium based on the VIX spillover index (I) and on both the VIX and the IV spillover index (II). In particular, panel A shows the results of the estimation of the two predictive regressions and the in-sample R-square. Panel B reports the OOS R-square, the percentage of correct sign prediction, the annualized Sharpe ratio of a switching investment strategy, the annualized Sharpe ratio of a mean-variance asset allocation strategy, and the annualized utility gain from a mean-variance strategy that exploits predictability vs. a strategy that relies on the historical mean forecast. The R-squares, the proportion of correct sign predictions, and the annualized utility gain are all expressed in percentages, i.e., 1.00 means 1.00%.

Panel A - In sample		
	(I)	(II)
Intercept	0.0014	0.0014
(t-stat)	(1.4102)	(1.4631)
β coeff. VIX	0.0007	0.0004
(t-stat)	(2.2975)	(1.3183)
β coeff. IV		0.40280
(t-stat)		(3.2252)
R-square	0.91	2.80
Panel B - OOS		
ROOS R2	-1.96	-0.33
Sign Prediction	47.62	48.57
Ann. SR switch. strategy	0.16	0.45
Ann. SR MV strategy	-0.01	0.43
Ann. Utility gain	-5.79	0.23

Table 9

Robustness to alternative assumptions

This table reports the same out-of-sample forecasting accuracy measures as in Table 2 but under different assumptions concerning: (i) the risk-free rate used to calculate the excess returns; (ii) the stocks included in the VAR model from which the two spillover indices are computed; (iii) the forecast horizon of the forecast error variance decomposition (FEVD) from which the two spillover indices are computed. In particular, in panel A, we use the 1-week USD based LIBOR instead of 1-month T-bill to compute excess returns. The results in Panel B are based on average RV and IV spillover indices where the average is computed across a set of spillover indices obtained using random subsamples of 60 stocks (out of 70). The results in Panel C are based on average RV and IV spillover indices when the average is computed across a set of spillover indices obtained using alternative forecast horizons (namely, $h = 1,2, \dots ,10$).

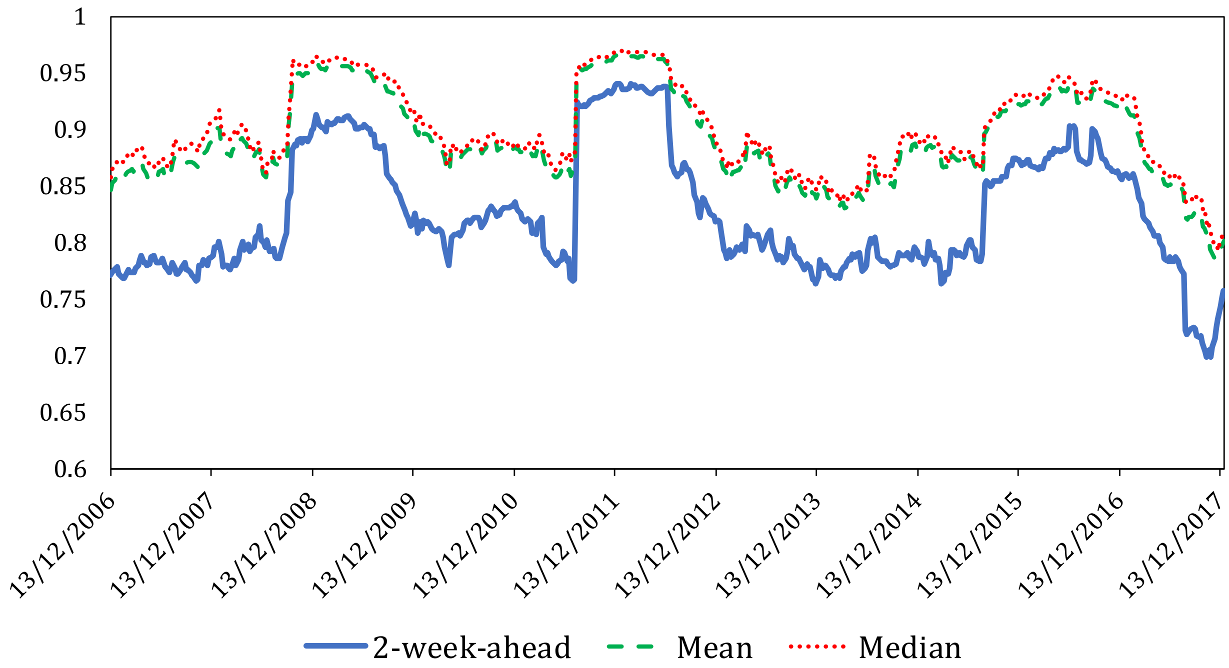
	RV			IV			Benchmark
	ΔRMSFE	OOS R2	Correct sign	ΔRMSFE	OOS R2	Correct sign	Correct sign
Panel A - Changing assumption on risk-free rate							
S&P 500	-0.0006	-4.80	51.05	0.0002	2.08	49.90	52.38
Avrg. All	-0.0006	-2.97	51.25	0.0001	0.32	50.78	50.50
Avrg. Energy	-0.0007	-2.72	49.54	0.0003	1.05	51.55	48.80
Avrg. Consumer Discret.	-0.0007	-3.52	52.36	0.0000	0.05	51.27	51.08
Avrg. Financials	-0.0010	-4.05	49.90	0.0001	0.26	49.46	48.48
Avrg. Health Care	-0.0004	-1.99	51.62	0.0000	0.08	50.70	51.73
Avrg. Industrials	-0.0003	-1.42	51.93	0.0001	0.35	50.04	50.53
Avrg. Technology	-0.0014	-7.13	53.33	0.0000	-0.28	52.46	54.40
Avrg. Materials	-0.0007	-3.12	50.26	0.0001	0.21	50.76	50.55
Avrg. Others	-0.0005	-2.32	51.33	0.0001	0.49	50.12	49.52
Panel B - Changing assumption on selected stocks							
S&P 500	-0.0006	-5.12	51.24	0.0002	2.01	51.62	52.38
Avrg. All	-0.0007	-3.04	51.11	0.0001	0.31	50.93	50.50
Avrg. Energy	-0.0008	-3.06	49.54	0.0002	0.85	51.52	48.80
Avrg. Consumer Discret.	-0.0007	-3.18	51.46	0.0000	0.08	51.38	51.08
Avrg. Financials	-0.0010	-4.10	49.33	0.0001	0.34	49.46	48.48
Avrg. Health Care	-0.0004	-2.17	51.90	0.0000	0.10	50.44	51.73
Avrg. Industrials	-0.0003	-1.40	52.05	0.0001	0.42	50.86	50.53
Avrg. Technology	-0.0014	-7.48	53.30	0.0000	-0.18	53.03	54.40
Avrg. Materials	-0.0007	-3.26	50.43	0.0001	0.19	51.07	50.55
Avrg. Others	-0.0005	-2.39	51.12	0.0000	0.43	49.86	49.52
Panel C - Changing assumption on FEVD forecasting horizon							
S&P 500	-0.0003	-2.81	52.00	0.0001	0.93	51.24	52.38
Avrg. All	-0.0003	-1.57	51.00	0.0000	-0.24	50.11	50.50
Avrg. Energy	-0.0005	-2.03	49.02	0.0000	0.11	49.42	48.80
Avrg. Consumer Discret.	-0.0002	-1.02	52.26	-0.0001	-0.22	50.15	51.08
Avrg. Financials	-0.0007	-2.64	49.71	0.0000	-0.20	49.84	48.48
Avrg. Health Care	-0.0003	-1.62	50.78	0.0000	-0.22	49.89	51.73
Avrg. Industrials	-0.0001	-0.48	51.64	-0.0001	-0.43	50.25	50.53
Avrg. Technology	-0.0006	-3.37	53.52	-0.0001	-0.49	52.19	54.40
Avrg. Materials	-0.0003	-1.48	50.55	0.0000	-0.26	50.88	50.55
Avrg. Others	-0.0002	-1.30	51.21	-0.0001	-0.37	49.26	49.52

Figure 1

Volatility spillover indices

This figure plots the realized (Panel A) and implied (Panel B) volatility spillover indices over the period December 2006 – December 2017, estimated using 50-week rolling windows. The solid line corresponds to the index computed from a 2-week-ahead forecast error variance decomposition. We also report the mean and the median of the indices obtained by experimenting over all the possible forecast horizons used in the variance decomposition, between 2- and 10-week-ahead.

Panel A: RV Spillover Index



Panel B: IV Spillover Index

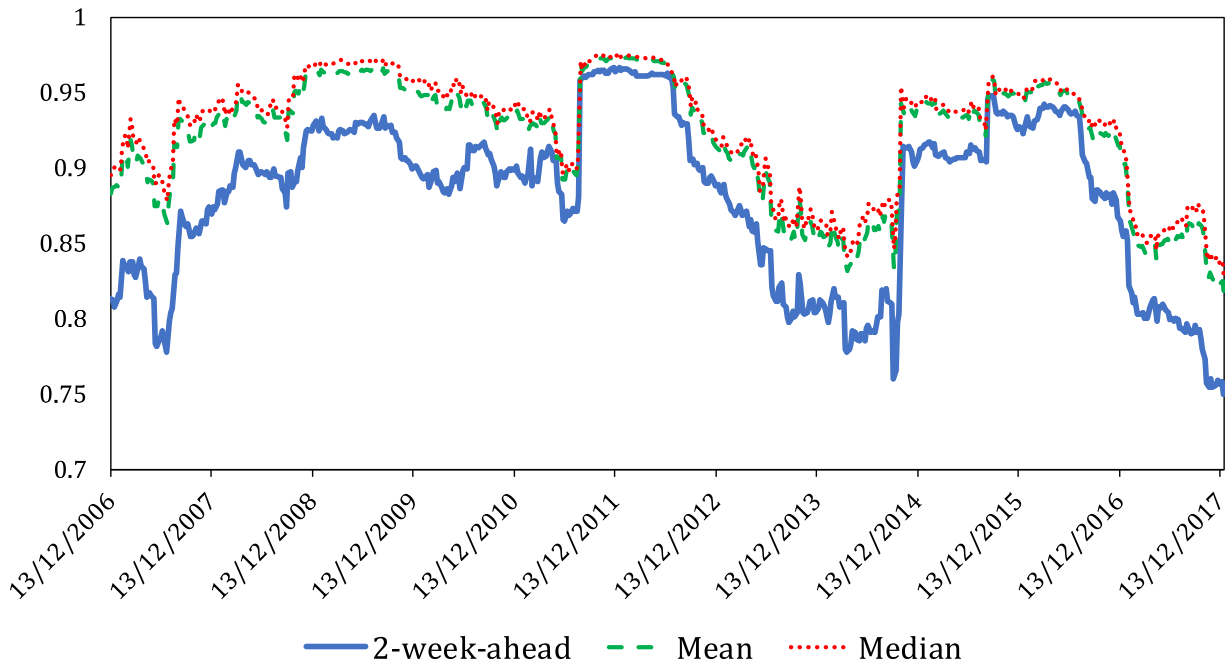
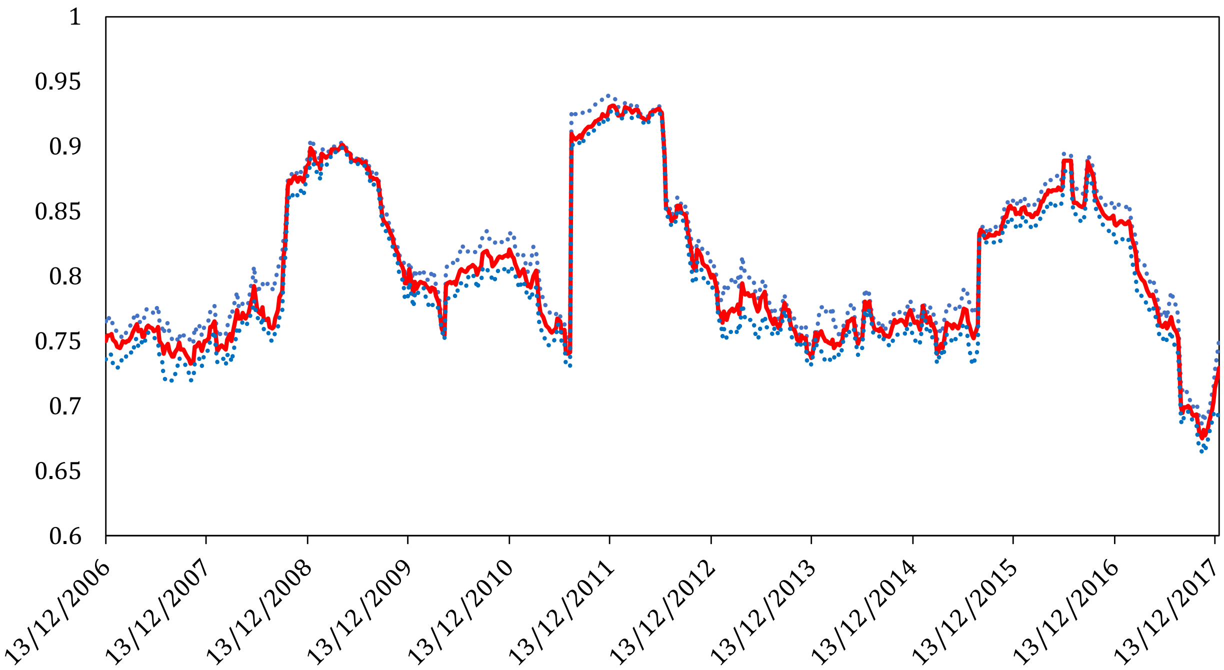


Figure 2

Volatility spillover indices based on a sub-sample of stocks

This figure plots the mean (solid line), the minimum, and the maximum (dotted lines) values of realized (Panel A) and implied (Panel B) volatility spillover indices estimated using random subsamples of 60 stocks (out of 70). They refer to the sample period December 2006 – December 2017 and are recursively estimated from 2-week-ahead forecast error variance decompositions using a 50-week rolling window.

Panel A: RV Spillover Index



Panel B: IV Spillover Index

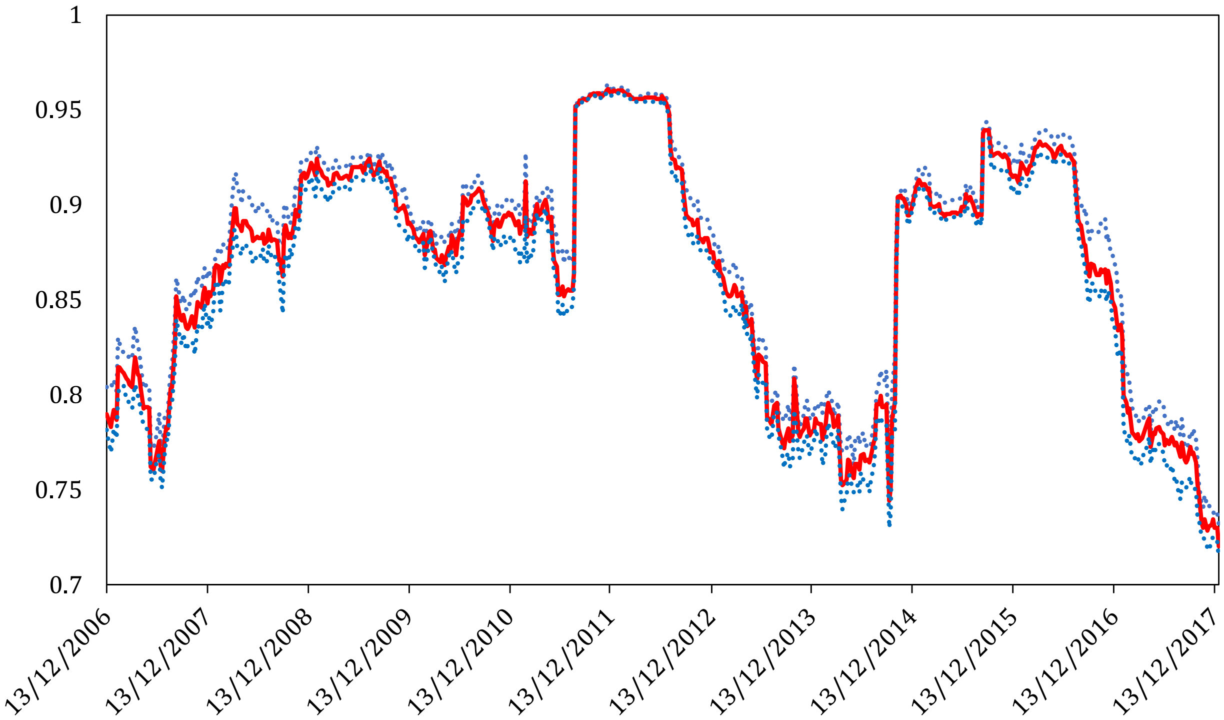


Figure 3

In-sample estimated beta coefficient on the realized volatility spillover index

Panel A displays the distribution of the beta coefficients obtained by regressing each of the 70 stocks that compose the sample on changes in the realized volatility spillover index over the period December 2006 – December 2017. Panel B displays the distribution of the associated t-statistics.

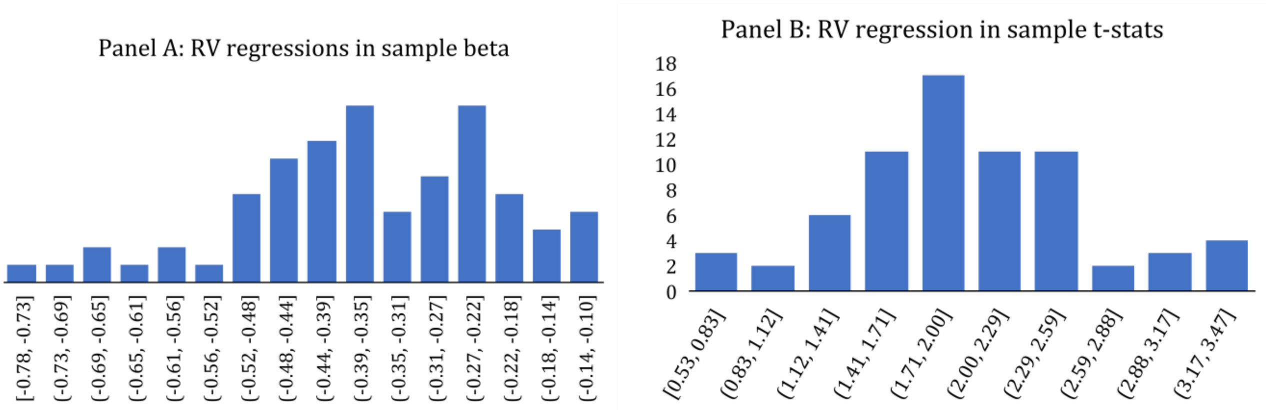


Figure 4

In-sample estimated beta coefficient on the implied volatility spillover index

Panel A displays the distribution of the beta coefficients obtained by regressing each of the 70 stocks that compose the sample on changes in the implied volatility spillover index over the period December 2006 – December 2017. Panel B displays the distribution of the associated t-statistics.

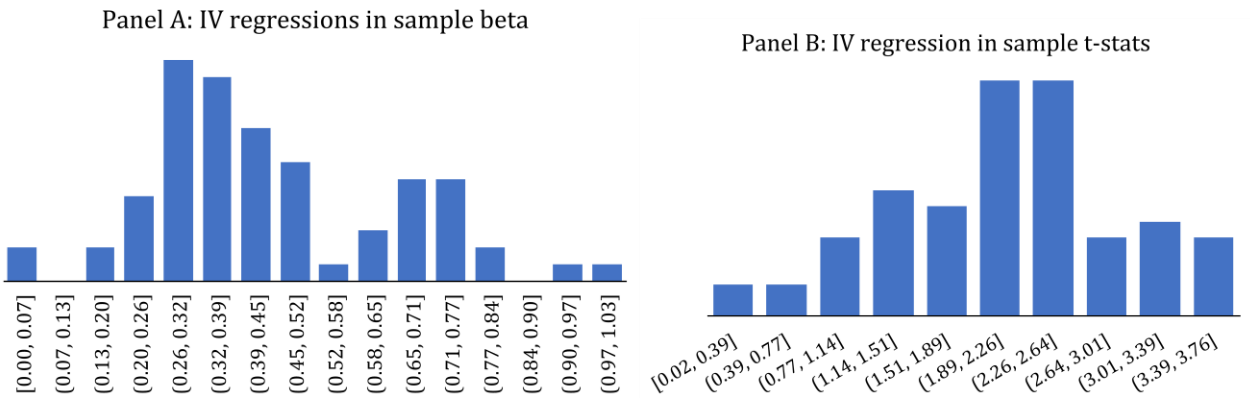
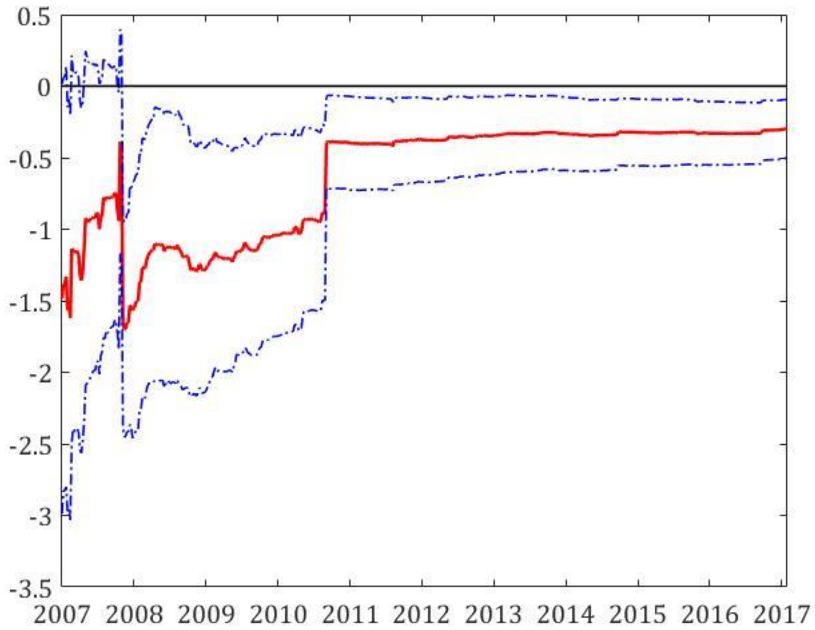


Figure 5

S&P 500 out-of-sample recursive beta coefficients on RV and IV spillover indices

Panel A plots the recursively estimated out-of-sample beta from a regression of the S&P 500 excess returns on the (changes in) the realized volatility index over the period December 2007 – December 2017. Panel B plots the recursively estimated out-of-sample beta from a regression of the S&P 500 excess returns on the (changes of) the realized volatility index over the period December 2007 – December 2017. The dotted lines represent ± 2 standard error confidence bands.

Panel A – S&P 500 Recursive Beta on Realized Volatility



Panel B – S&P 500 Recursive Beta on Implied Volatility

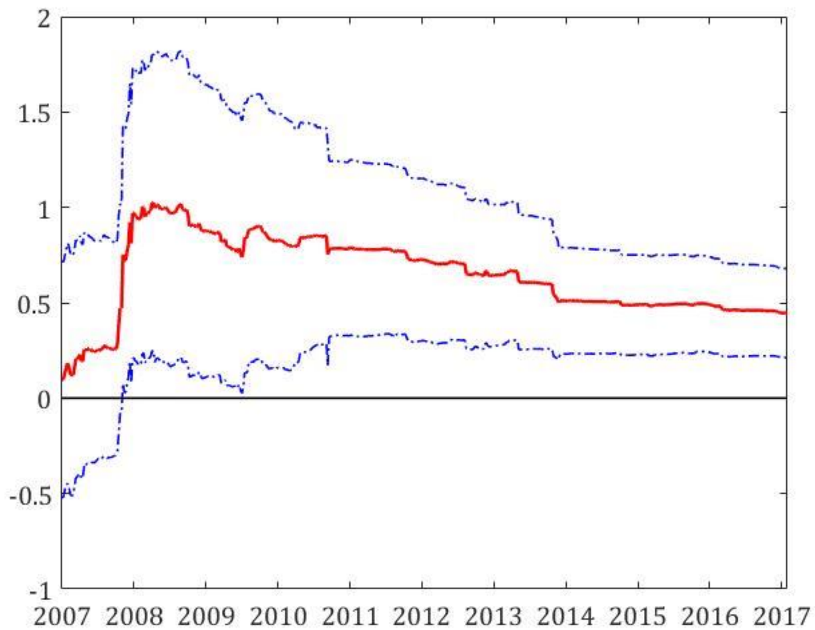
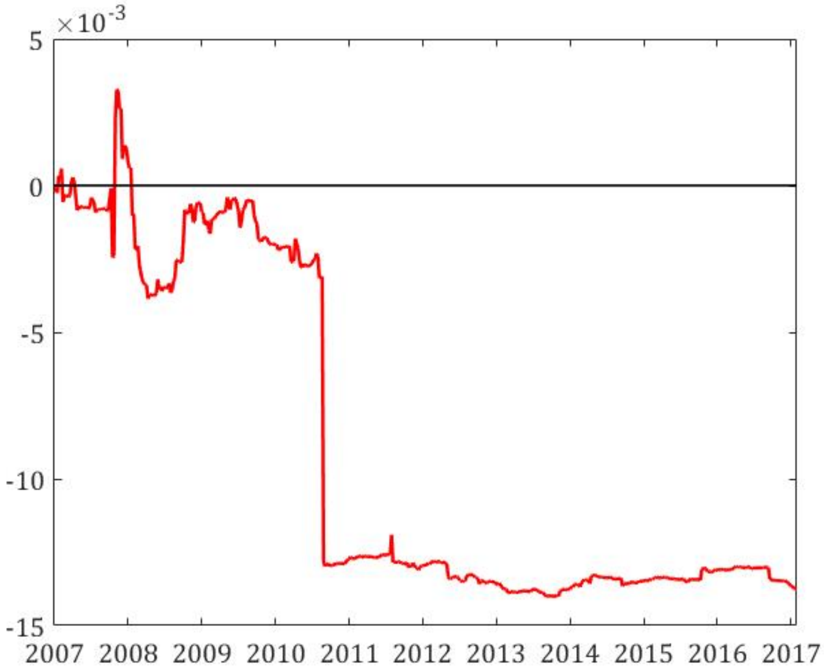


Figure 6

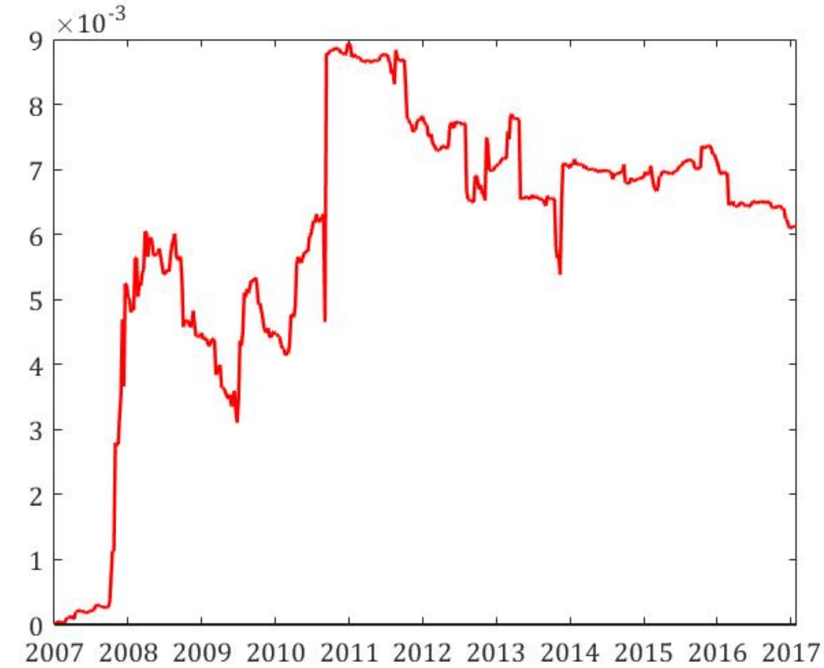
Cumulative squared forecast error differences vs. historical mean predictions

Panel A plots the cumulative squared forecast error for the historical mean benchmark minus the cumulative squared forecast error for the RV-based predictive regression for the S&P 500 over the period December 2007 – December 2017. Panel B plots the cumulative squared forecast error for the historical mean benchmark minus the cumulative squared forecast error for the IV-based predictive regression for the S&P 500 over the period December 2007 – December 2017. An increase in the cumulative squared forecast error signals that the RV (IV) spillover index predictive regression outperforms the historical average and viceversa.

Panel A



Panel B



Appendix A

Figure A1

Implied volatility series of AON Plc under different treatments of the missing values

This figure plots the series of the implied volatility of AON Plc when two different assumptions concerning the treatment of missing values. The dotted line represents the series of the implied volatility when a 1-month moving average is used to replace missing values; the solid line represents the series of the implied volatility when a 3-month moving average is used to replace missing values.

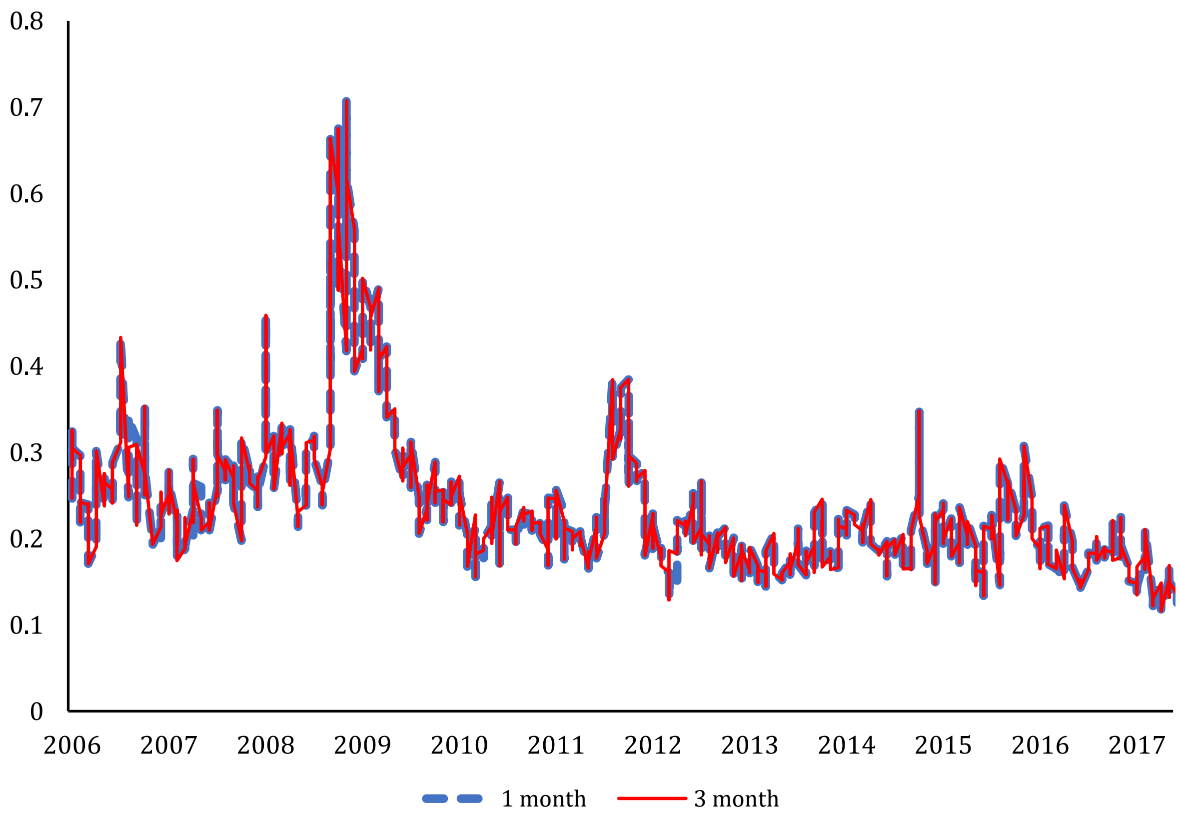
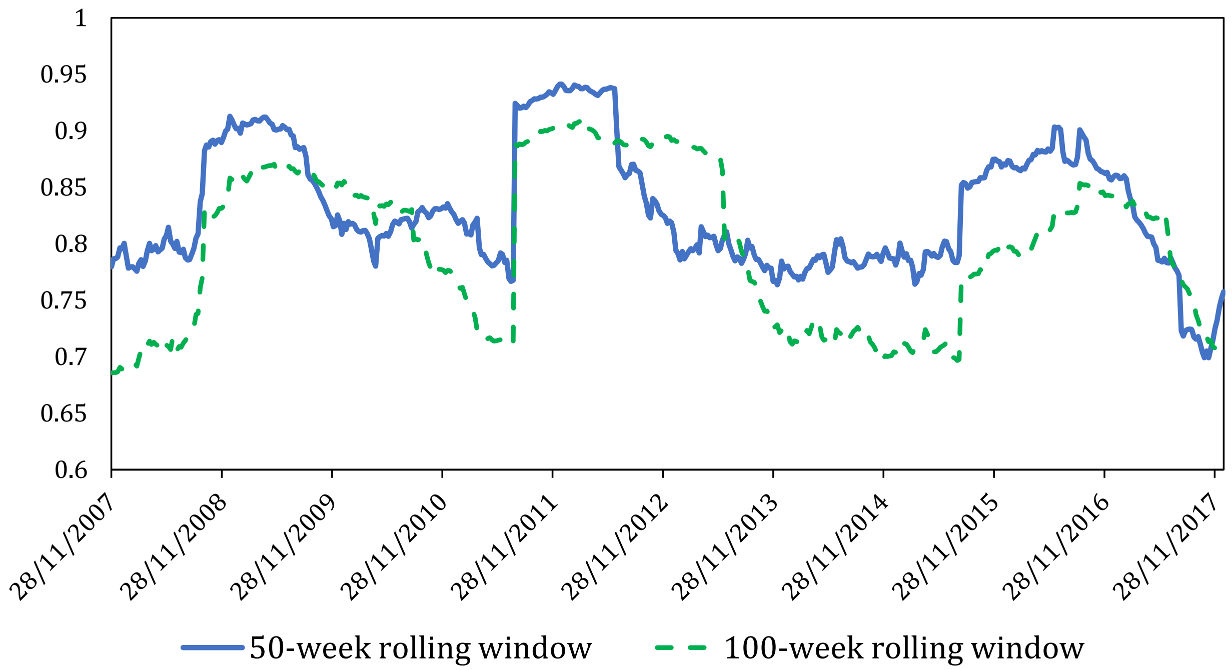


Figure A2

Volatility spillover indices (50- vs. 100-week rolling window)

This figure plots the realized (Panel A) and implied (Panel B) volatility spillover indices over the period December 2007 – December 2017, estimated alternatively using a 50-week rolling window (solid line) and a 100-week rolling window (dotted line).

Panel A: RV Spillover Index



Panel B: IV Spillover Index

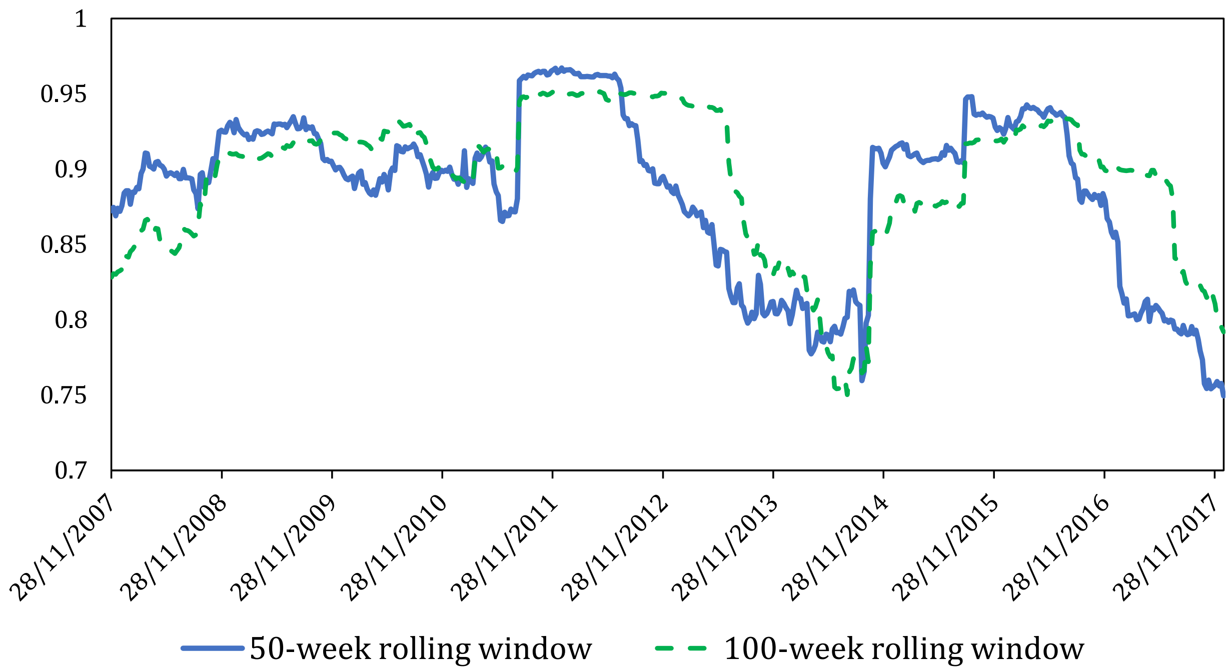
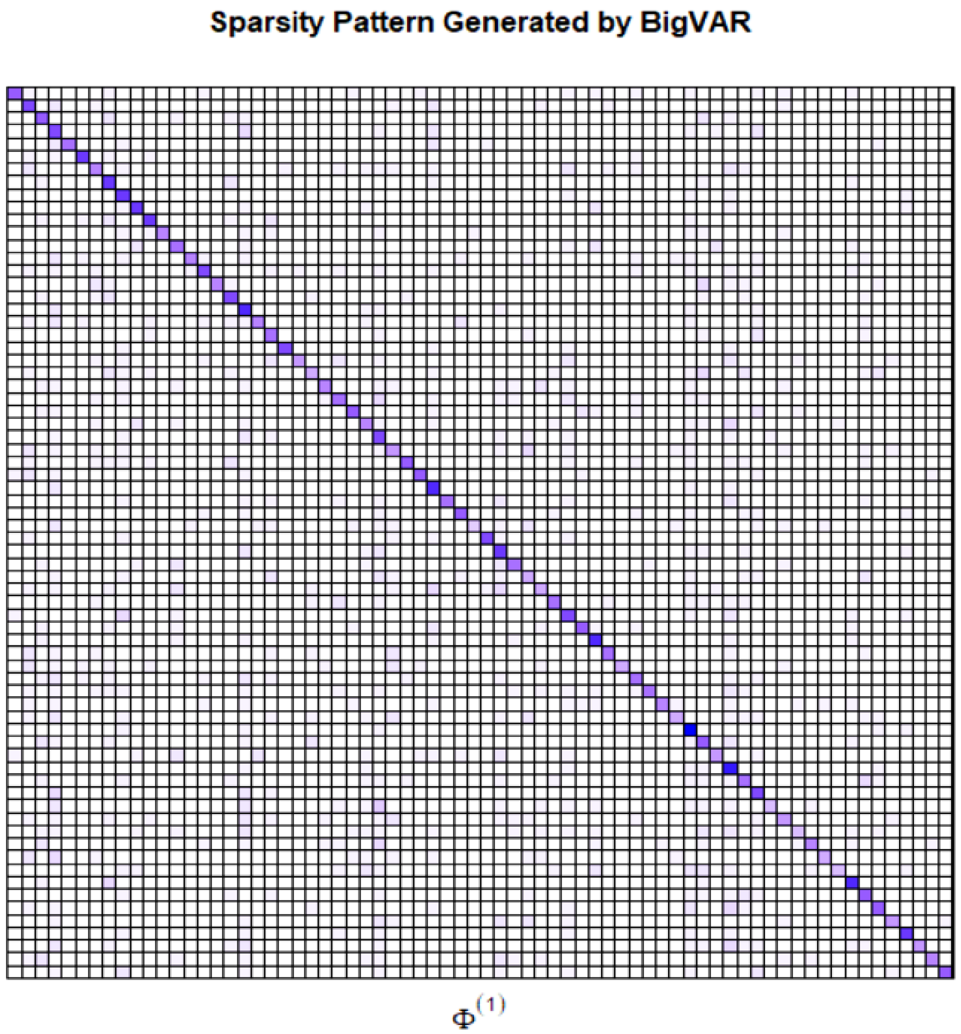


Figure A3

Sparsity plot of the autoregressive matrix of a VAR(1) fitted on realized volatilities

The picture shows the sparsity plot of the autoregressive matrix of a VAR(1) fitted on realized volatilities. Each of the squares represents one of the 70×70 coefficients. The darker is the colour, the larger is the estimated coefficient (in absolute value). A white square denotes that the coefficient has been set to zero.



Appendix B

Table B1

This table reports the results of the estimation of the predictive regression

$$r_{t+1} = \alpha + \beta_{RV}\Delta\mathrm{RVSI}_t + \beta_{IV}\Delta\mathrm{IVSI}_t + \varepsilon_{t+1},$$

where r_{t+1} is the weekly excess return (over the one month T-bill) of an individual stock j, and $\Delta\mathrm{RVSI}_t$ ($\Delta\mathrm{IVSI}_t$) is the change between time $t-1$ and t of the realized (implied) volatility spillover index. The sample period is December 2006 – December 2017. We report the mean values of β_{RV}, β_{IV} and R-square across the sectors. The R-square coefficients are reported in percentages (e.g., i.e., 1.00, means 1.00%). We also report the number of significant β_{RV} and β_{IV} coefficients at 5-percent test size levels.

	RV + IV Regression		
	β coeff. RV	β coeff. IV	R-square (%)
All	-0.78	0.53	1.58
Consumer Discretionary	-0.91	0.50	1.44
Energy	-0.90	0.77	1.34
Financials	-0.70	0.61	2.16
Health Care	-0.67	0.53	1.62
Industrials	-0.68	0.40	1.59
Materials	-0.90	0.56	1.44
Technology	-0.63	0.40	1.73
Others	-0.76	0.42	1.61
N. of significant β_{RV} coeff. (α=5%)		53	
N. of significant β_{IV} coeff. (α=5%)		40	

Table B2

This table reports the Campbell and Thompson (2008) OOS R-squared coefficients for the out-of-sample (OOS) recursively estimated predictive regressions of individual stock excess returns on (changes of) the two alternative spillover indices. The R-squares are expressed as percentages, e.g., i.e., 1.00, means 1.00%. The stars refer to the Clark and West's (2007) MSFE-adjusted statistic; ** (*) denotes that the difference in the mean square forecast error between the historical mean benchmark and the predictive model based on the RV (IV) spillover index is statistically significant for a size of the test of 5% (10%).

Stock ticker	R2 OOS RV Spillover Index	R2 OOS IV Spillover Index	Stock ticker	R2 OOS RV Spillover Index	R2 OOS IV Spillover Index
AAPL	-3.83	0.70**	GWW	-1.31	-0.52
AMGN	-3.79	-0.56	HES	-1.58	1.44**
ANF	-2.08	-0.43	HON	-2.03	0.71*
AON	-1.15	0.57*	IBM	-6.41	-1.54
APA	-2.59	0.73**	INTU	-2.81	-0.58
APC	-4.75	1.11**	ISRG	-2.83	-0.36
APD	-1.43	0.02	KMB	-0.93	0.44
AXP	-5.70	1.02**	KSS	-2.05	-0.33
AZO	-3.54	1.61**	LH	-2.41	-0.46
BA	-1.86	-0.25	LLL	1.29	0.55**
BAX	0.11	-0.03	LM	-5.07	0.91**
BBBY	-1.56	0.27*	MCK	-2.37	0.26
BHGE	-2.68	-0.07	MLM	-2.75	-0.38
CAT	-2.06	0.58**	MMM	-4.41	-0.28
CCJ	-2.68	0.35**	MON	-2.30	0.20*
CCL	-1.01	-0.42	NKE	-4.70	0.58**
CELG	-0.54	1.72**	NOV	-4.15	0.30**
CERN	-7.37	0.42**	NUE	-2.70	1.17**
CI	-1.45	-0.45	OXY	-2.20	1.38**
CL	-3.02	0.85**	PCAR	-1.37	0.71**
CMI	-3.08	0.48**	PH	-0.74	1.73**
CNQ	-1.82	0.42**	PNC	-1.90	-0.58
COF	-6.59	0.40**	PX	-4.49	0.27*
COST	-2.27	1.09**	RDC	-4.44	1.65**
CTSH	-14.96	-0.39	RL	-3.96	-0.60
CVX	-1.69	2.13**	ROST	-4.28	-0.29
DHR	-3.26	0.38	RTN	-0.35	0.27*
DISH	-0.43	-0.53	SCCO	-3.87	0.38**
EOG	-2.78	0.95**	SLB	-1.04	1.54**
ESRX	-2.59	0.47**	SWK	-4.80	0.40*
ETR	-2.64	1.69**	TOL	-5.71	0.04
FDX	-2.27	0.41**	TPR	-4.38	-0.68
GD	-0.99	0.53*	UNP	-1.35	0.50*
GILD	-0.59	-0.04	VFC	-6.29	-0.72
GS	-3.84	-0.73	WHR	-2.26	-0.13