Bocconi

WORKING PAPER SERIES

WORKING PAPER
N. 190
DECEMBER 2022

The Empirical Performance of Option Implied Volatility Surface-Driven Optimal Portfolios

by Massimo Guidolin and Kai Wang

Università Bocconi
BAFFI CAREFIN
Centre for Applied Research on International Markets, Banking, Finance, and Regulation

The Empirical Performance of Option Implied Volatility Surface-Driven Optimal Portfolios*

Massimo Guidolin

University of Liverpool
Management School and Baffi
CAREFIN, Bocconi University

Kai Wang

CAFD, Central
University of Finance
and Economics

December 2022

Abstract

We apply a two-step strategy to forecast the dynamics of the volatility surface implicit in option prices to all American-style options written on the stocks that have entered the Dow Jones Industrial Average Index between 2004 and 2016. We explore whether the implied volatilities extracted through the two-step approach help improve the out-of-sample performance of minimum-variance portfolios. We find that, by using option-implied volatilities in estimating the covariance matrix, the ex-post volatility of the minimum-variance portfolio is lower when compared with the equal-weighted portfolio and a minimum-variance portfolio simply derived from the historical, sample covariance matrix estimator. Moreover, over most of our 13-year sample, the realized Sharpe, Sortino and information ratios increase when the sample covariance matrix estimator is replaced with its implied counterpart. However, the benefits of using option-implied information are countered by an increase in portfolio turnover that may imply higher (implicit) transaction costs. We also apply shrinkage methods to both the sample covariance estimator and the implied covariance estimator and note that they often lead to significant improvements in portfolio performance.

Keywords: Equity options, Implied volatility surface, Predictability, optimal portfolios

JEL code: G11, G17, G13.

1. Introduction

Equity options are often used by market participants for speculation and risk hedging purposes. However, in the last 15 years both academic researchers and market practitioners

* We would like to thank two anonymous referees for constructive comments and for stimulating us to improve the paper and Runmei Xue for her help with codes and data collection. Massimo Guidolin gratefully acknowledges that this research was funded by MIUR—PRIN Bando 2017—prot. 2017TA7TYC.

have been drawn to option markets because of their forward-looking nature and the rich information about future payoffs and their volatility that they may reveal (see, e.g., Christoffersen, Jacobs, and Chang, 2013). In fact, a substantial amount of research focuses on forecasting the future volatility of the underlying asset by means of implied volatility (henceforth, IV) (see, e.g., Poon and Granger, 2003; Chun et al., 2019). Moreover, an extensive literature has focused on modeling and forecasting the very dynamics of the implied volatility surface (henceforth, IVS).[1] The idea is that if the IVS contains useful predictive signals for underlying quantities, understanding the dynamics of the IVS will unlock the possibility to form accurate and efficient forecasts of volatility. In this paper, we explore whether and how such predictability of the IVS may generate valuable forecasts within a standard portfolio set-up.

This on-going, progressive research program faces two challenges. On the one hand, the empirical evidence on the predictability of implied volatilities remains somewhat mixed. On the other hand, even stronger doubts exist as to whether forecasting the IVS may support profitable trading strategies and correctly inform economic decisions. Dumas, Fleming, and Whaley (1998) investigate whether the dynamics of the S&P 500 index implied surface can be predicted over different time periods and find that their specifications are unstable over time for the purpose of option pricing and hedging. Gonçalves and Guidolin (2006, henceforth GG) propose a modeling approach for the time-series properties of the S&P 500 index option IVS by combining a cross-sectional approach to fitting the IVS similar to Dumas et al.'s with the application of vector autoregression (VAR) models to the time series of estimated cross-sectional coefficients. In practice, in their first step, they fit a linear model to every daily cross section of option contracts IVs and get the time series of the estimated volatility function coefficients; in the second step, they fit a VAR to predict any time variation in the estimated first-step coefficients. GG find that the of the S&P 500 IVS are reliably predictable.[2] However, they also report that these predictability patterns fail to be exploitable in an economically significant way since no abnormal profits can be obtained as

[1] Contrary to the constant volatility assumption of Black and Scholes' (1973) model (henceforth, BS), a number of empirical studies have shown that volatilities implicit in options written on the same underlying asset vary across strike prices and maturities, resulting in a non-flat three-dimensional implied volatility surface (see Cont and Fonseca, 2002; Kim et al., 2010; Carr and Wu, 2016).

[2] Fengler et al. (2003) investigate the index IVS dynamics using principal component analysis while Fengler et al. (2007) focus on semiparametric factor models.

transaction costs increase. On a brighter note, Goyal and Saretto (2009) find that there is both a statistically and economically significant predictable pattern in the dynamics of implied volatility by using information from the cross-section of implied volatilities across various stock options. Based on the two-step approach in GG, Bernales and Guidolin (2014) explore the possibility that the dynamics of the implied volatility surface of individual stocks may be associated with movements in the volatility surface of S&P 500 index options and provide strong evidence of predictability in the cross-section of equity options and of dynamic linkages between the volatility surface of equity and S&P 500 index options. They also report that such predictable dynamic patterns can be exploited to generate abnormal risk-adjusted returns before transaction costs. Yet, Konstantinidi, Skiadopoulos, and Tzagkaraki (2008) consider a number of European and U.S. implied volatility indices and show that both point and interval forecasts obtained from standard models cannot outperform the random walk model in an out-of-sample setting.

The objective of our paper is to assess the economic value of the two-step approach proposed by GG. To this end, we adopt a portfolio selection approach that applies Markowitz's mean-variance (henceforth, MV) framework. The vector of mean returns and the covariance matrix are the two main inputs needed to calculate an optimal MV portfolio. Conventionally, sample mean and sample covariance matrix are estimated from historical data. However, the resulting optimal weights are then known to be extremely sensitive to the estimation errors produced by sample moments; consequently, the out-of-sample (henceforth, OOS) performance have often proved to be disappointing.[3] In our work, rather than improving the sample moments estimated from historical data, we adopt a forward-looking approach to adjust the covariance matrix estimates by applying IV forecasts obtained through the two-step approach of GG.

There are several studies that are relevant but different from our paper. Kostakis, Panigirtzoglou, and Skiadopoulos (2011) extract the implied distribution of the S&P 500 index and maximize investors' utility by considering an investment universe that consists of the S&P 500 index and a riskless asset. They show that the optimal portfolio obtained by

[3] Michaud (1989) shows that imprecisely estimated inputs often lead to an unreasonable portfolio compositions since large weights tend to be assigned to the assets with the highest expected returns; yet such assets are also the most likely to suffer of positive estimation error. His arguments were reinforced by subsequent studies, such as Best and Grauer (1991) and Green and Hollifiled (1992). For a general survey of the literature on portfolio choice, see Brandt (2007).

making use of these implied distributions is preferred to the one derived from the historical distribution since it generates better OOS performance in terms of Sharpe ratios. DeMiguel, Plyakha, Uppal, and Vilkov (2013) investigate which aspects of option-implied information are more effective in improving the OOS performance of portfolios consisting of the S&P500 stock constituents. They extract the implied information through the model-free approach outlined by Bakshi, Kapadia, and Madan (2003) and find that option-implied volatilities can help reduce portfolio volatility while using option-implied correlations do not lead to significant improvements.[4] Recently, Kyriacou, Olmo, and Strittmatter (2021) have explored the role of option-implied information adopting the long-term asset allocation perspective offered in Campbell and Viceira (1999) and show that two state variables, the risk-premium and the market price of risk constructed from option-implied density are unable to outperform their historical counterparts using the Sharpe and Sortino ratios.[5]

Specifically, our investable universe consists of the stocks composing the Dow Jones Industrial Index over a 12-year sample from September 1, 2004 to August 31, 2016. For each constituent stock in the index, we model and forecast the corresponding IVS by applying the two-step approach introduced by GG. We then compute the stock-specific volatility by using a weighted average of the IV forecasts, where the open interest provides the weights. Given the large impact of estimation error incorporated in mean estimates (see Best and Grauer, 1991; Chopra and Ziemba, 1993) on the portfolio weights and the fact that simple Black-Scholes IVs cannot directly (i.e., unless further assumptions are introduced) reveal information on expected returns, we focus our analysis on minimum-variance portfolios, which require solely the estimation of the covariance matrix. In addition, DeMiguel, Plyakha, Uppal, and Vilkov (2013) show that the estimated covariance matrix based on option-

[4] Neumann and Skiadopoulos (2013) analyze the predictable patterns in the higher-order risk neutral moments (RNMs) extracted from the prices of S&P 500 index options using the model free approach of Bakshi, Kapadia, and Madan and show that the one-day-ahead skewness forecasts rather than the volatility forecasts lead to economically significant profits when transaction costs are not considered. Brinkmann, Kempf, and Korn (2015) and Kempf, Korn, and Sabning, (2015) show that option-implied information on higher-order dependencies is useful to portfolio optimization and that the more risk averse investors are, the more they may benefit from an implied estimator of higher-order (co-)moments.

[5] As noted by Birke and Pilz (2009), parametric approaches incur the risk of misspecification and of introducing systematic errors in the estimates of state price. The current consensus is that nonparametric models generally provide more reliable estimates of risk-neutral distributions, which is also consistent with the finding of Taboga (2016) who shows that a semi-nonparametric model allows to estimate the option-implied distributions quite precisely.

implied correlations is highly unstable, and therefore, we choose to use the historical correlation in calculating pair-wise covariances. After constructing the implied covariance matrix estimator, we plug it along with a "grand" (i.e., shrinked towards the cross-sectional average) mean into the optimization and obtain the minimum-variance portfolio. For comparison, we also consider the equal-weighted portfolio and a MV portfolio derived from sample variances and covariances. Moreover, we compute the minimum-variance portfolios in three cases. Firstly, when no constraints are imposed. Secondly, we add a non-negativity restriction in the portfolio selection process. Thirdly, an additional 20% upper bound on each individual stock position is introduced.[6] The performance metrics we adopt are realized ex-post volatility, the realized Sharpe, Sortino, and information ratios, the maximum drawdown, and a measure of implicit, maximum transaction costs relative to the DJ30 index which trade-offs the realized Sharpe ratio with portfolio turnover. Finally, we consider both weekly and daily rebalancing intervals.

We firstly apply the analysis to stocks that have always been the constituents of the Dow Jones Industrial Index over our full 12-year sample. Our results demonstrate that the volatility of the minimum-variance portfolio based on the implied covariance matrix estimator could be reduced when compared to the one derived from the sample covariance matrix estimator. The improvement is more significant after we impose the non-negativity and the 20% upper bounds constraints together. The results concerning the realized Sharpe and Sortino ratios are instead mixed. In the unconstrained case, the implied covariance matrix estimator fails to produce an improvement in the Sharpe ratios when compared with their sample counterparts, while in the remaining two cases, it leads to a higher realized risk-adjusted performance. As far as turnover is concerned, a general increase is observed across all three cases when option-implied volatility forecasts are plugged in the covariance matrix estimates. A higher level of turnover indicates that more rebalancing trades are needed when using the implied covariance estimator in the construction of the minimum-variance portfolio. However, the empirical results concerning the implicit transaction costs fail to

[6] The reason to apply an additional upper bound is that while in practice money managers cannot hold a specific stock in excess of self- and legally-imposed limits, mean-variance optimization is well-known to lead to rather extreme weights. For instance, anticipating some results later discussed in our empirical analysis in Section 5, without any constraints, the maximum weight obtained by minimum-variance portfolios based on the sample covariance predictor is as high as 63% (to Johnson & Johnson), and maximum for the option-implied minimum-variance portfolio is 87% (to Intel).

show a drastic decay of realized Sharpe ratio performance due to a higher turnover. We also apply shrinkage methods to both the sample covariance estimator and the implied covariance estimator and find that they fail to yield systematic improvements in realized performance even though at the margin they often help by reducing turnover and drawdown statistics.

Next, considering that the composition of the Dow Jones Industrial Index had changed five times during the sample period, we decompose the 12-year period into 6 unequally spaced sub-periods based on the adjustment of the components of the index. For all sub-periods, we focus on the analysis under no-short sale constrains because—when compared with the other two cases—this provides the best OOS performance of the minimum variance portfolios. Some results are common across different sub-periods. Firstly, we report again an increase in turnover when we replace the sample covariance matrix estimator with its option-implied predicted counterpart. Secondly, as reported earlier, regardless of the rebalancing frequency, realized volatility is always reduced by using IVS predictions in the covariance matrix estimates. As far as the Sharpe ratio is concerned, the values are negative for certain periods. However, negative Sharpe ratios are uninformative since a higher value does not necessarily indicate better risk-adjusted performance.[7] In such circumstances, we use instead the information ratio to rank portfolio performances. In fact, even in periods of negative realized Sharpe ratios, the corresponding, realized information ratios are positive, relatively high and they tend to generally favor portfolio strategies based on our two-step IVS approach.

The likely reason for the predominantly superior out-of-sample statistical (and arguably, economic value) out-performance of our two-stage modelling approach lies with its ability to capture the (joint, simultaneous) predictability in stock and option returns induced by models of asset pricing with learning dynamics in which simple rules of Bayesian updating of beliefs induce (co-)persistence in asset returns (see, e.g., Beber and Brandt, 2006, Ederington and Lee, 1996, Engle, 2001, Guidolin and Timmermann, 2003, Timmermann, 2001). In particular, a learning process followed by option market participants could provide

[7] For example, consider two portfolios that both generate negative mean returns over a certain period. Compared with portfolio B, portfolio A has a lower mean return and a higher volatility that result in a higher (less negative) Sharpe ratio. However, this does not imply that portfolio A achieves a better performance than portfolio B.

an explanation for the existence of a precisely estimable dynamic relationship between the IVS of equity options and the IVS of market index options, see, e.g., Bernales and Gudolin (2015). Although the mapping between our two-stage approach and the optimizing behaviour of a representative investor who learns the process of the underlying asset fundamentals is not straightforward, our results suggest the presence of a strong Markov structure in the IVS. In fact, such a structure is so persistent to be exploitable in a recursive portfolio construction exercise.

The rest of this paper is organized as follows. Section 2 describes the data. Section 3 outlines the specification of the model. Section 4 lists the OOS statistical measures of prediction accuracy that we shall use. Section 5 analyzes the economic value of the two-step GG's methodology by exploring the OOS performance of different minimum-variance portfolios. Section 6 performs a few important robustness checks, concerning the number of stock option series investigated and the impact of transaction costs. Section 7 concludes.

2. The Data

The investment universe we consider consists of the components of the Dow Jones (DJ) Industrial Average during the period September 1, 2004 - August 31, 2016, for a total of 2517 trading days. All options are traded on the Chicago Board Options Exchange (CBOE). Dow Jones 30 index options are European-style while individual stock options are American-style[8]. Options data are downloaded from OptionMetrics, which contains the expiration dates, highest (lowest) closing bid (ask), strike prices, trading volume, and open interest. Data for stock-specific information are obtained from the Center for Research in Security Prices (CRSP), including spot stock prices, daily returns, dividend payment dates, and dividend cash amounts. As for the risk-free rate, we use the LIBOR as the proxy for maturities shorter than one year. For maturities longer than one year, the discount curve is bootstrapped from the Eurodollar future contracts. Furthermore, to match the maturity of each option contract, we complement the discount curve offered by the data by simple linear interpolation.

---

[8] We select the stocks included in the DJ30 index because of computational reasons: in some parts of the paper we need to contrast results for the set of the index components with the index itself and the burden involved by larger indices would be heavy (prohibitive in the case of the S&P 500 also because mean-variance asset allocation techniques suffer from instability and excessive weight concentration over large-scale cross-sections).

To compute the implied volatility of the Dow Jones 30 index options, we follow OptionMetrics and "invert" BS model by equating the theoretical price to the mid-point quote price of the option. The spot price of the DJ 30 index is adjusted for the discrete dividend payments in ways similar to GG (2006). To compute the implied volatility of the individual stock options, a binomial tree model, following Cox, Ross, and Rubinstein (1979), is used.[9]

The composition of the Dow Jones 30 index has changed five times over our sample period. Table 1 presents the index constitution over time. Due to scarcity of relevant, traded options data, we discard three stocks, namely Arconic Inc., General Motors Co. and Travelers Cos. Inc. The resulting total number of stocks is 36, out of which 24 were constituents of DJ 30 index over the entire sample period. In the following, we use the 24 stocks always in the index for the baseline tests and then expand a portion of such tests to the additional 12 stock option series that belonged to the DJ30 just for a sub-period, only with reference to the relevant sub-period of inclusion in the DJ30, of course.

[Table 1 about here]

We follow the literature, e.g., GG (2006), and apply the following filters to our raw data. First, we remove options for which the open interest is zero. Second, we exclude the options with maturities shorter than six trading days, because short-term options contain little information regarding the time value dimension of the IVS. Third, we discard the contracts with absolute moneyness exceeding 20% because their prices are usually rather noisy, especially in the case of individual equity options.[10] Fourth, we dismiss options that violate basic no-arbitrage bounds, such as upper and lower bounds for call and put prices and call-put parity relationships. Fifth, we remove options with a mid-point price lower than $0.30 to mitigate the impact of the price discreteness bias. Finally, we exclude options with non-standard settlement agreements.

After data cleaning, we get a total number of observations that equals 714,188 data points with a daily maximum of 377 and a daily minimum of 138 observations for DJ 30 index

---

[9] In practice, we obtain data on implied volatilities from OptionMetrics directly. Citing their manual, "The CRR model usually requires a very large number of sub-periods to achieve good results (typically, N > 1000), and this often results in a large computational requirement. The IvyDB proprietary pricing algorithm uses advanced techniques to achieve convergence in a fraction of the processing time required by the standard CRR model (…)" that however always allows a depth of at least 1,000 steps (N).

[10] The absolute moneyness is defined as $|(K/S) - 1|$, where K and S are the strike price and underlying stock or the equity index price, respectively.

options. In the case of the individual stock options, in total we have more than 3 million observations. Specifically, JPMorgan Chase displays the highest number of traded contracts, with 153379 observations in total. On a daily basis, Intl Business Machines Corp presents the highest daily number of traded contracts, with a maximum of approximately 190 active contracts.[11] Table 2 reports the summary statistics for the filtered implied volatilities of DJ 30 index options and for the whole dataset of equity stock options, respectively. We also sub-divide the data into 15 groups by time to maturity and moneyness. Specifically, in the time-to-maturity dimension, we define a contract to be *short-term* if the trading days to expiration (DTE) are less than 60 days, *medium-term* if DTE are between 60 and 180 days, and *long-term* if DTE days exceed 180. In the moneyness dimension, we define an option to be *deeply in the money* or *deeply out of the money* (DITM/DOTM) if the absolute moneyness is larger than 10%, *in the money* or *out of the money* (ITM/OTM) if the absolute moneyness is between 2% and 10% and *at the money* (ATM) with absolute moneyness equal to or less than 2%.

[Table 2 about here]

The average IV we record for equity options is ranging from 20% to 40%, while that of Dow Jones 30 index options is between 15% and 30%, indicating that the individual stocks are more volatile than the index. Out of the entire dataset for the individual stock options, approximately 65% is represented by OTM and DOTM options and 80% by short- and medium-term options. Such dominance is also found for DJ 30 index options. This is consistent with industry practice because option trading is mostly executed for short-term hedging or speculative purposes; moreover, in the case of American-style options, ITM options tend to be often exercised early and as such finding them in the data is rarer vs. ATM and OTM options. Our filtered data also display evidence on the shape of the IVS as a function of time to maturity and moneyness. For example, in Table 2b we note that for put options, IVs display an asymmetric smile for all maturities, while in the case of call options, a smile pattern is observed only for short and medium-term contracts, as long-term call options are best characterized by downward sloping skews as one moves from DITM to DOTM. In the time-to-maturity dimension, in both the case of DJ 30 and individual stock options, the ATM

[11] The lowest daily number is two contracts which clearly insufficient to estimate a regression. Therefore we impose a requirement of a minimum number of daily observations (traded contracts) equal to 15 when estimating (1) below. When in the case of an individual equity option series, the daily number of contracts turned out to be smaller than 15, we used the estimated betas from the previous day to replace the missing values of the estimated OLS coefficients.

IVs are increasing in DTE while a relatively flat pattern is observed in the case of ITM contracts. Additionally, the IVs of OTM options are generally higher than those of ATM options and the level of the IVs are higher for puts than for calls.[12]

3. A Two-Step Approach to Model the Dynamics of the Implied Volatility Surface

To predict the dynamics of the IVS, we adopt the two-step approach in GG (2006). First, we fit a linear model to every daily cross section of option contracts IVs and obtain the time series of the estimated volatility function coefficients. Second, we estimate a VAR to capture and predict any time variation in the estimated first-step coefficients. Specifically, *for each underlying stock in our sample*, we estimate the following cross-sectional model:[13]

$$ln\sigma_i = \beta_0 + \beta_1 M_i + \beta_2 M_i^2 + \beta_3 \tau_i + \beta_4 (M_i \times \tau_i) + \varepsilon_i \qquad (1)$$

i =1, ..., N, where *N* is the number of option contracts available in each daily cross section. M_i is moneyness defined as $\ln(K_i/S)$, where K_i and S are the strike and underlying spot prices, respectively; τ_i is the time to maturity (measured as a fraction of year, i.e., $\tau_i = DTE_i/252$) and ε_i is a random disturbance capturing measurement error. We take the logarithm of the implied volatility as the dependent variable to avoid fitting negative values.

For each underlying stock on each day, we estimate the coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)'$ by OLS. In particular, β_0 is the intercept such that under the classical BS assumptions, β_0 would be the log of the constant IV while $\beta_1, \beta_2, \beta_3, \beta_4$ should be all identically zero (or at least, in a sample, they should not be statistically significant).

[Figure 1 about here]

We plot in Figure 1 the time series of the daily estimates of the components of $\boldsymbol{\beta}$ for the DJ 30 index and their average across the individual stocks included in the index. The estimated coefficients are remarkably unstable over the sample period. For instance, the estimated $\hat{\beta}_0$ peaks during the financial crisis and then reverts back to lower values, closer to its historical mean. Moreover, because the estimated intercept proxies for the log value of IV under BS, the higher estimated intercept value for individual stock options compared with the DJ 30 index options shown in Figure 1 is consistent with investors agreeing to pay higher

---

[12] Because investors prefer to buy downside protection, they are willing to pay a higher premium for protective put options, thus resulting in a higher level of OTM put IVs.

[13] To prevent the notation to become cluttered we shall drop a further pedix (or apex) that refers to the individual stock under analysis (or to the DJ30).

prices for the individual stock components of the index vs. the index itself, which probably reflect less than unit correlations among individual components. The plots for $\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$, and $\hat{\beta}_4$ in the case of individual equities are remarkably less volatile vs. the ones for the DJ30 index but this is likely to mechanically derive from the fact that in Figure 1 only averages across component stocks are plotted. In fact, in unreported plots concerning the individual stocks, we find that most of the time, the estimated coefficients are more volatile than the ones inferred from the index, which is compatible with the IVS of individual stocks being more predictable than the DJ30 implied volatilities.

Following GG, we next fit a VAR model to the vector time series of daily cross-sectional coefficient estimates. Our previous analysis has shown that the beta estimates vary significantly across periods. In unreported results, we also document that beta estimates display a strong correlation with their own lags as well as with the lags of the estimates of the other estimated betas. However, both the autocorrelation coefficients and the partial autocorrelation coefficients of the beta estimates decrease sharply with the coefficient order, in a manner which is generally consistent with a low-order VAR modelling strategy. Therefore, in the following, we develop our empirical analysis with reference to the case of a VAR(1) estimated on the vector time series of estimated OLS coefficients from the first step:

$$\widehat{\boldsymbol{\beta}}_{t+1} = \boldsymbol{A}_0 + \boldsymbol{A}_1 \widehat{\boldsymbol{\beta}}_t + \boldsymbol{\epsilon}_{t+1}. \tag{2}$$

In (2), $\widehat{\boldsymbol{\beta}}_t$ contains the IVS coefficient estimates from the regression estimated at time t and $\boldsymbol{A}_0$, $\boldsymbol{A}_1$ are the matrices of intercepts and autoregressive coefficients, respectively. $\boldsymbol{\epsilon}_{t+1}$ is an IID $(\mathbf{0}, \boldsymbol{\Omega})$ error term. We employ a 252-day rolling window to estimate this VAR model for each individual stock and the DJ 30 index.[14] As a result, over the entire sample period, we obtain a total of 2,266 OLS estimates for $\boldsymbol{A}_0$ and $\boldsymbol{A}_1$, denoted as $\{\widetilde{\boldsymbol{A}}_{0,t}\}$ and $\{\widetilde{\boldsymbol{A}}_{1,t}\}$.

Next, we use the following model to forecast the one-day ahead beta coefficients:

$$\widetilde{\boldsymbol{\beta}}_{t+1} = \widetilde{\boldsymbol{A}}_{0,t} + \widetilde{\boldsymbol{A}}_{1,t} \widehat{\boldsymbol{\beta}}_t \tag{3}$$

Finally, we forecast the one-day ahead IVS at time t for each underlying stock using the newly predicted beta coefficients,

[14] To explore the VAR model with the highest forecasting power, we have performed preliminary comparisons of two alternative recursive estimation designs: a rolling window scheme vs. an expanding window one. The findings show that a rolling window VAR has better forecasting performance vs. the expanding one. Complete results are available from the Authors upon request.

$$E\left(ln\hat{\sigma}_{i,t+1}\right) = \tilde{\beta}_{0,t+1} + \tilde{\beta}_{1,t+1}M_i + \tilde{\beta}_{2,t+1}M_i^2 + \tilde{\beta}_{3,t+1}\tau_i + \tilde{\beta}_{4,t+1}(M_i \times \tau_i), \qquad (4)$$

where M_i is the moneyness and τ_i is the time-to-maturity, respectively.

4. Statistical Performance of Out-of-Sample Forecasts

In this Section, we assess the OOS performance of the forecasts derived from the two-step approach. We follow the literature, e.g., GG (2006) and Christoffersen and Jacobs (2004), and establish the random walk as a benchmark for comparison purposes. This benchmark is applied to both the IVS coefficients and to the IVs derived from exponentiating the predictions in (4).[15] Hence, to assess the OOS performance of the two-step approach (henceforth called the VAR(1) model) and of the random walk, we compute the following three summary measures of predictive accuracy:

1. The root mean squared prediction error (RMSE), i.e., the square root of the average squared deviations of the forecasts of the estimated beta coefficients (IVs) from the beta coefficients (the BS IVs) on the next trading day.
2. The mean absolute prediction error (MAE), i.e., is the average of the absolute difference between the forecasts of estimated beta coefficients (IVs) and the beta coefficients (BS IVs) on the next trading day.
3. The mean correct prediction (MCP) of the direction of change in either estimated betas or in IVs, i.e., the average frequency for which the change in beta coefficients (IVs) predicted by the model has the same sign as the realized change in betas (BS IVs).

4.1 Out-of-Sample Performance of the Forecasts of the Beta Coefficients

We compute measured of OOS predictive accuracy for each individual stock option series as well as for DJ30 index options. In panel A of Table 3, we find that the VAR(1) model outperforms the RW benchmark under all measures for all but the intercept beta (i.e., $\hat{\beta}_0$) coefficient estimates. This pattern holds not only for DJ30 Index, but also for the whole sample of individual stock options. For instance, in the case of $\hat{\beta}_1$ for the DJ30 Index options, the average OOS RMSE (MAE) returned by the VAR(1) is 0.1746 (0.1234), while it increases to 0.2614 (0.1655) in the case of the RW model. Similar patterns emerge in the case of

[15] Harvey and Whaley (1992) comment that "(…) while the random walk model might appear naïve, discussions with practitioners reveal that this model is widely used in trading index options."

$\hat{\beta}_2$, $\hat{\beta}_3$, and $\hat{\beta}_4$. Moreover, both the OOS RMSE and MAE yield lower values DJ30 options vs. those obtained for the whole sample of individual stock options across all beta coefficient estimates. Among the five estimated betas, $\hat{\beta}_3$ achieves the lowest OOS RMSE and MAE values, indicating that the superiority of the VAR(1) model is more prominent in forecasting the shape of the IVS in the time-to-maturity dimension vs. other dimensions. When moving to the MCP, it is important to emphasize that the values of the frequency of mean correct predictions exceed 50% for all beta estimates based on the VAR(1) model. In a no-predictability environment, we should expect values for MCP of about 50% but the results in Panel A of Table 3 suggest that we can forecast the beta coefficient estimates using a simple VAR(1) model.

[Table 3 about here]

To formally assess the statistical significance of the differences in OOS performance of the VAR(1) model compared to the random walk benchmark, we use the equal predictive ability test proposed by Diebold and Mariano (1995). As a loss function to construct the test statistic, we consider two types of performance indicators: the difference in squared forecast errors and the difference in absolute forecast errors. A Newey-West (1987) heteroskedasticity and autocorrelation consistent (HAC) variance estimator is used to calculate the Diebold and Mariano test. Panel B of Table 3 reports the values of the statistic and associated significance levels for all beta coefficient estimates. In the case of individual stock options, we reject the null hypothesis of equal forecast accuracy of the VAR(1) model compared to the benchmark random walk model for all but the intercept beta coefficient. More precisely, with the only exception of $\hat{\beta}_0$, we observe that the test statistics across beta coefficients are negative and significant, indicating that the OOS superior performance of the VAR(1) model is statistically significant. The result applies also to DJ30 index options, but statistical significance only obtains in the case of $\hat{\beta}_2$ and $\hat{\beta}_3$.

4.2 Out-of-Sample Performance of Implied Volatility Forecasts

After obtaining the beta coefficient estimates through the second-stage VAR(1) model, we compute the forecasts of IVs using equation (4). To examine the OOS performance of the IV forecasts, we perform afresh the analyses presented in Table 3 using forecasts of IVs. Panel A of Table 4 reports the results for the DJ30 Index as well as for the whole set of individual stock option series. Considering in this case the average of the performances totaled over the individual stock options, the average OOS RMSE and MAE for annualized IVs are 4.23% and 2.29% for the VAR(1) model, while they massively increase to 25.91% and 13.13% for the benchmark RW. Similar results are obtained in the case of the DJ30 Index. The best-performing model according to these measures is once more the VAR(1). When moving to MCP, we observe that the MCP yielded by the two-step model is 53.61% on average for individual stocks (while this is not defined in the case of RW which predicts no change), suggesting that the best performance is provided again by the VAR(1) model. Similar results are obtained for the DJ30 Index.

[Table 4 about here]

Panel B of Table 4 shows the results of equal predictive accuracy tests for the VAR(1) model against the benchmark RW, again based on Diebold and Mariano's test (1995). In this case, the test is applied to the one-day-ahead forecasts reported in Panel A of Table 4. As a loss function, we employ the differences between the squared (absolute) forecast errors from the VAR(1) model and the squared (absolute) forecast errors from the benchmark random walk model. A Newey-West (1987) HAC variance estimator is used to compute the Diebold and Mariano test statistic. With one exception, results in Panel B of the Table 4 show that the null hypothesis of equal predictive accuracy for the VAR(1) model and the benchmark model is strongly rejected. The negative and significant test statistic suggests that the out-of-sample superior performance of the VAR(1) is statistically significant, which also confirms that time variation in the IVS is statistically important.

Panel B of Table 4 also shows that using two alternative tests of equal predictive accuracy, also the MCP scores of the two-step methodology significantly outperforms (with p-values of 0.000) a 50% benchmark MCP corresponding to random sign prediction. The two tests of equal predictive accuracy are a standard DM test and Pesaran and Timmermann's (1992) test (PT). In the case of the DM test, we first randomly simulate correct prediction values (0/1) over our out-of-sample forecast periods such that the benchmark MCP is 50%. Then, we define the loss function as the differences between the correct prediction values from the VAR(1) model and the correct prediction values from the benchmark. We calculate the DM test statistic for each simulation and report the average DM test statistics for 1000 simulations. In the case of the PT test, we follow Pesaran and Timmermann (1992) and

Skabar (2013) and calculate the success ratio as well as the sign independence ratio based on the correct prediction values from our VAR(1) model.[16] Then, we test whether the success ratio from our VAR(1) model is statistically higher than a 50% benchmark by calculating the PT test statistic, which has been shown follows an asymptotic standard normal distribution.

5. Economic Value

The results in the previous section suggest that the IVS is highly predictable in a statistical sense. Following GG, we examine in this section the economic value of this predictability when applied to portfolio selection. Specifically, we compute a range of minimum-variance portfolios that exploit the alternative IVS predictions and compare their OOS performances.

5.1 Portfolio Construction

We explain first the portfolio methods used in our analysis, including the equal-weighting approach, the minimum-variance portfolio based on the historical covariance matrix estimator, and the minimum-variance portfolio based on the forecasted implied covariance matrix estimator. We include the first two types of portfolios in our analysis for a comparison purposes because they do not depend on any option-implied information. Besides, we also compare the realized, OOS performance of different strategies with the returns of Dow Jones 30 index, which represents a passive investment strategy.

The equal-weighted portfolio assigns the same weight to each stock in the investment universe. If an equal-weighted portfolio were able to yield a performance comparable (or even superior) to more complicated strategy, investors would not have any incentives to select a more sophisticated strategy. Moreover, DeMiguel, Garlappi, Nogales and Uppal (2009) find that, for portfolios built from large-, medium- and small-cap S&P indices, an equal-weighted portfolio with monthly rebalancing outperforms a range of mean-variance strategies across performance metrics such as mean return, the Sharpe ratio and realized Jensen's alpha. Therefore, we also entertain the performance of an equally-weighted portfolio.

[16] The success ratio is defined as the fraction of days in an out-of-sample test period for which the sign of the change is predicted correctly while the sign independence ratio is defined as the expectation of the correct predictions.

Because our main goal is to examine the contribution of the revisions of the forecasts of the covariance matrix driven by our two-stage methodology, it is desirable to reduce the impact of other estimation error as much as possible. This applies in particular to the errors potentially generated by expected return predictions. As Jagannathan and Ma (2003) state, "The estimation error in the sample mean is so large nothing much is lost in ignoring the mean altogether when no further information about the population mean is available". Therefore, in order to identify the contribution of the predicted IVs in improving portfolio selection performance, we set the mean return on all assets to be equal. Specifically, we assume the grand mean return across all individual stock and the DJ30 option series to be the expected return for every underlying stock (and index). As a result, this is equivalent to simply find the optimal portfolio that minimizes the variance of portfolio returns. Thus, the optimal minimum-variance weights can be computed as

$$\boldsymbol{w}_{t+1|t}^{min} = \frac{\widehat{\Sigma}_{t+1|t}^{-1}\mathbf{1}}{\mathbf{1}'\widehat{\Sigma}_{t+1|t}^{-1}\mathbf{1}}, \tag{5}$$

where $\mathbf{1}$ is a vector of ones and $\widehat{\Sigma}_{t+1|t}$ is the covariance matrix predicted at time t for time t + 1. We focus on two types of covariance matrix predictors in our analysis: the historical sample covariance estimator and the implied covariance estimator. The former is estimated simply using sample moment estimates, while the latter is constructed by using our forecasts of implied volatilities derived in Section 4. Specifically, on the main diagonal of the implied covariance matrix, we predict the implied variance of each stock by using the weighted average of the volatility forecasts obtained from the above two-step approach, where the open interest provides the weights to be applied to different strike prices and maturities. In the case of the covariance forecasts off the main diagonal of the implied covariance matrix, we follow Mostowfi and Stier (2013) and resort to one-year daily historical correlation estimates to obtain the so-called "quasi-implied covariance".[17] However, as is well documented in the literature that uses sample moments to estimate the covariance matrix, this may create considerable estimation errors that may flaw the realized OOS performance of mean-variance portfolios (see Michaud, 1989). Therefore, we also consider the *shrinkage*

[17] There is no consensus on the best method to extract the implied correlation from option prices and DeMiguel, Plyakha, Uppal, and Vilkov (2013) demonstrate that the benefits from using option-implied correlation seem limited in portfolio selection as their recorded improvement of the out-of-sample performance is not always statistically significant.

methods proposed by Ledoit and Wolf (2003a, 2003b) to estimate the covariance matrix.[18]

We rebalance our portfolios at either a daily or at a weekly frequency. In the case of daily rebalancing, we find a set of optimal weights every day and hold that portfolio for one day; in the case of weekly rebalancing, in correspondence to every Wednesday (or Thursday or Tuesday, in case a Wednesday does not happen to be a trading day), we form a portfolio that is then held for a week, similarly to DeMiguel, Plyakha, Uppal, and Vilkov (2013). Next, by multiplying the daily or weekly returns by the number of rebalancing times in one year, we compute the OOS annualized return.

We use a 252-day rolling window horizon to compute the mean return and historical covariance matrix in the case of methods based on sample moments. At time t, we determine a set of optimal weights $\boldsymbol{w}_t$ for the minimum-variance portfolio, and the out-of-sample return is computed as $r_{t+\Delta t} = \boldsymbol{w}'_t \boldsymbol{r}^{sample}_{t+\Delta t}$, where Δt equals to 1 for daily rebalancing and 5 for weekly rebalancing. To formally assess the OOS performance of the minimum-variance portfolios obtained under alternative specifications of the covariance matrix predictors, we consider five widely used criteria to measure realized performance: volatility, the Sharpe ratio, trading turnover, the Sortino ratio, and the maximum drawdown return.[19]

5.2 Out-of-Sample Realized Portfolio Performance

Firstly, we apply the portfolio optimization analysis on the 24 stocks that have always been the constituents of the Dow Jones 30 index over our sample period. Over this period, the cumulative weight of the 24 stocks is roughly 70% of the whole index on average. The results

[18] The shrinkage method aims at pulling the most extreme coefficients towards "central" (e.g., grand mean) values, thereby systematically reducing estimation error where it matters the most. Specifically, the *shrinkage* estimate of covariance matrix is $\widehat{\boldsymbol{\Sigma}}_{Shrinkage} = \delta \boldsymbol{F} + (1-\delta)\boldsymbol{S}$, where $\boldsymbol{S}$ denotes the sample covariance estimator and $\boldsymbol{F}$ is the shrinkage target. Ledoit and Wolf (2003b) suggest a constant correlation model to construct $\boldsymbol{F}$, i.e., using the average of all the sample correlations as the common constant correlation when constructing the shrinkage target covariances. δ is the *shrinkage constant*, which measures the weight that is given to the shrinkage target and is determined by minimizing the expected distance between the shrinkage estimator and the true covariance matrix (see Appendices A and B of Ledoit and Wolf, 2003b).

[19] The Sortino ratio is a variation of the Sharpe ratio that differentiates downside volatility from total overall volatility by using a portfolio's standard deviation of negative returns instead of the total standard deviation of returns. The maximum drawdown is the maximum observed loss from a peak to a trough of a portfolio, before a new peak is attained, over a given observation period, expressed in percentage terms.

are summarized in Table 5. We consider three cases: panel A shows the results for the unconstrained case; panel B reports the results when a no short-sale constraint is imposed; panel C reports the results after an additional upper bound of 20% is added on the weights to prevent that excessively unbalanced portfolios may emerge.[20] We also apply the shrinkage method of Ledoit and Wolf (2003a, 2003b) to both the sample covariance estimator and the implied covariance predictor, to yield two additional strategies. Finally, we display the results under daily rebalancing in Table 5a and weekly rebalancing in Table 5b.

[Table 5 about here]

In all cases, both the equal-weighted and the minimum-variance portfolio strategies achieve better realized performance compared to passively holding the Dow Jones 30 index in terms of out-of-realized annualized mean return, Sharpe ratio, Sortino ratio, and maximum drawdown. However, although an equally-weighted portfolio provides impressive OOS mean returns, its volatility is almost always significantly higher than that of minimum-variance strategies. For example, in the case of daily rebalancing, when no constraints are imposed, the volatility of the minimum-variance portfolio obtained from the option-implied covariance matrix forecast is 13.56%, while the realized volatility of the equally weighted portfolio is 19.65%. When a non-negativity constraint is introduced, the difference in realized volatilities between the equally weighted and the other minimum-variance portfolios is also evident. This result is consistent with our expectations because equally-weighted portfolios are not explicitly targeting the minimization of risk. We plot in Figure 2 the realized volatility of the one-year rolling window returns on the Dow Jones 30 index vs. the realized returns on the minimum-variance portfolios based on a sample covariance estimator and the implied variance predictor in the unconstrained case. From the plot we can see that the minimum-variance strategies achieve a remarkably lower volatility especially during bear market phases.

[Figure 2 about here]

[20] The application of option-implied information to the portfolio problem fails per se to mitigate the weight concentration problem typical of minimum-variance optimization. Instead of optimizing the risk-reward ratio, the objective of constructing minimum-variance portfolios is minimizing the risk only. Therefore this frequently leads to a solution with pronounced concentration in low volatility stocks. If we group the stocks in our sample by different levels of volatilities, that is, we consider $IV < 20\%$ as low volatility, $20\% < IV < 25\%$ as medium volatility and $IV > 25\%$ as high volatility, we find that on average stocks with higher volatilities receive much lower weights compared to stocks with low volatilities.

When we focus on the performance of different minimum-variance portfolios, under a volatility metrics, in all panels based on daily rebalancing, using the option-implied information lowers realized OOS volatility by 0.51%, 0.49%, and 31.91% respectively in comparison to the minimum-variance portfolio derived from a sample covariance matrix estimator. A similar pattern is found under weekly rebalancing. Besides, the highest Sharpe ratio (0.87) is obtained for minimum-variance portfolios based on the implied covariance matrix estimator when non-negativity constraints are imposed. More specifically, under daily rebalancing, the Sharpe ratio of the minimum-variance portfolio built from implied covariance matrix estimator increases by 16% vs. the portfolio based on the sample covariance matrix estimator, when short-sale restrictions bind. When moving to realized portfolio turnover, its value is high for the minimum-variance portfolios derived from the option-implied covariance. A similar improvement is obtained with respect to the Sortino ratio (1.13) and also in this case it is coupling no short sale constraints with shrinking that delivers the highest reward-to risk ratios. Although turnover is often higher under IVS-based strategies vs. equally-weighted portfolio, the maximum drawdown is the least when IVS information is used, especially when non-negative weights constraints are imposed. This partially expected, as implied volatility incorporates the market expectation of the future volatility of the returns on the underlying asset, it should be frequently updated over time thus increasing realized turnover.

The shrinkage methodology we apply does not yield a significant impact on realized portfolio performance. The shrinkage constant we obtain ranges between 42.8% and 43.7% for the historical sample covariance matrix estimator and between 0.53% and 0.59% in the case of the option-implied covariance matrix predictor, which explains the particularly pronounced lack of performance improvement in the latter case. One explanation may be that shrinkage methods work well when the number of stocks is relatively large compared to the number of observations, which is not our case.[21] However, even the modest improvements always go in the direction of increasing the realized reward-to-risk ratios when shrinkage is applied. We further compute the tracking error vs. the performance of the

[21] Additionally, in the original paper by Ledoit and Wolf (2003b), they apply shrinkage to identifying the minimum-tracking error portfolio. As Ledoit and Wolf point out, the tracking-error efficient portfolios are not mean-variance efficient and therefore different from the strategies of our interest. In fact, with the only exception of one panel of Table 5b (when the ShrIVS strategy is the best), in our tests, 1/N is the strategy that minimizes the average tracking error.

DJ30 for each case. Interestingly, the tracking error of the equal-weighted portfolio is lower than that of minimum-variance strategies, which shows that outperforming the benchmark index comes from departures from the index at appropriate times irrespectively of the correlation with the DJ30 (importantly, a non-zero mean tracking error does not imply a low correlation, but a low correlation is sufficient for a large tracking error to emerge). Especially in the case of daily rebalancing, the mean tracking error of the minimum-variance portfolio based on the sample covariance estimator is ranging from 45bps to 100 bps, and using variances implied from the IVS further increases the tracking error slightly. However, adding weight constraints helps to reduce the tracking error of the minimum-variance strategies. Moreover, in Table 5b, weekly rebalancing leads to a mean tracking error of minimum-variance portfolios that are not substantially different from the equally-weighted strategy.

Moving to an analysis of the impact of different portfolio constraints, we find that, after imposing non-negativity restrictions, the improvement in the realized volatility of the minimum-variance strategy based on both the option-implied and the sample, historical covariance matrices climb up. Imposing non-negativity restrictions on the weights helps to reduce the sampling error and hence it is expected to lead to a better out of sample performance. Another noticeable finding triggered by adding a non-negativity constraint is the improvement in the realized turnover ratio: for instance in Table 5a, a non-negative portfolio weight reduced the average, recorded turnover to less than half under daily rebalancing. Additionally imposing an upper bound on the weights, further cuts the average turnover ratio, as one would anticipate.

One may wonder how stable over time can these portfolio results be. To deal with this issue, we exploit one natural circumstance. As already mentioned, over our sample period, the components of Dow Jones 30 index have experienced five changes. For instance, on February 18th, 2008, Honeywell International Inc. and Altria Group Inc. were deleted from the index, while one day later, on February 19th, 2008, Bank of America and Chevron Corp, were added to it. Although there are several reasons why companies are removed and included to the Dow Jones index, we exploit these changes to sub-divide our sample in six unequally spaced sub-periods and perform our economic value analysis afresh.[22]

[22] The six periods are: September 1, 2004 - February 18, 2008 (period 1); February 19, 2008 - September 21, 2008 (period 2); September 22, 2008 - June 7, 2009 (period 3); June 8, 2009 - September 23, 2012 (period 4); September 24, 2012 - September 22, 2013 (period 5); after

We report the overall performances for each of the sub-samples in Table 6. Because Table 5 has shown that the minimum-variance portfolios achieve their best realized performance only after the non-negativity restrictions are imposed, we now focus on the case with no short-sale constrains enforced in each sub-period. In the case of Table 6a, across all six sub-samples, we fail to observe a homogeneous result by which minimum variance strategies outperform the equally-weighted naïve one. For instance, in period 5 (2012-2013), the equally-weighted portfolio outperforms all minimum-variance portfolios in terms of out-of-sample mean, Sharpe, and Sortino ratios.[23] However, for most of the sub-samples and especially during the bear market regimes, minimum-variance strategies generate better performance and among them option-implied ones appear to be superior. For instance, in the sixth and final (post 2013) sub-sample, under daily rebalancing, the minimum variance strategies based on two-step IVS predictions lead to an annualized Sharpe ratio of 2.16 (2.14 when shrinking is applied) vs. 1.57 (0.99) when the historical sample covariance matrix is used; in fact, historical moments are outperformed by an equally-weighted portfolio (the realized Sharpe ratio is 2.05). Moreover, even in sub-periods in which IVS-driven strategies fail to outperform (such as period 5), there are realized performance indicators (e.g., realized volatility) for which the two-step approach preserve its competitiveness.

[Table 6 about here]

Moreover, in line with our preceding findings, the turnover increases when we build portfolios based on the option-implied covariance matrix. This applies to both daily and weekly rebalancing. However, realized volatility is lowered by a different extent by including the option-implied information in covariance matrix estimates. During bear regimes, generally featuring high risks, for example, the third sub-period (2008), when compared with the equally-weighted portfolio, the minimum-variance strategy based on the implied covariance estimator yields a reduction in volatility almost equal to 50% (49.8%) vs- simpler

---

September 23, 2013 (period 6).

[23] Plyakha, Uppal and Vilkov (2012) explain that a likely source of outperformance in total return produced by equal-weighted portfolio is the overweighting of small stocks and stocks characterized by high idiosyncratic volatility. During times of extremely accommodative monetary policy, such stocks achieve quite an impressive performance. Compared with minimum-variance portfolios, the equally-weighted strategies would participate to market upturns to a stronger extent. However, this explanation is unlikely to apply to our results as the DJ 30 index is by construction only composed of (very) large companies, the so-called “blue chip” stocks.

strategies.[24]

When in Table 6b, we switch to a weekly rebalancing frequency, the results are qualitatively homogeneous with respect to Table 6a, in the sense that although the realized performance indicators yield different statistics, their rankings across different strategies and the relative performance of the two-step IVS models are similar to the case of daily rebalancing. The only notable difference is that the only performance index with respect to which the DJ30 index manages to outperform the remaining strategies in all sub-samples is the max drawdown return, which reflects the cap-weighted nature of indices.

Although it remains difficult to speculate on the reasons for why in a few specific sub-samples our two-step approach may turn out to be less competitive in relative terms, anchoring our commentary to the likely source of attractive performance—the power to capture recursive parameter learning within a rational asset pricing framework—may provide some insights. Paradoxically, when the parameter of a SDF-based option pricing model were to be known and leaning fails to occur recursively over time, we would expect the IVS-based method to be relatively the least accurate. Moreover, when learning were to occur discontinuously or in jumps and not in the recursive fashion typical of dynamic economic models, we speculate that our two-step approach could be less competitive. In fact, this is for instance confirmed by Table 6, in which the sub-period in which the two-step IVS-driven methodology turns out to perform the worst in correspondence to the outbreak of the Great Financial Crisis in 2008, when in the face of a drastically different environments, investors' beliefs may have jumped and learning occurred in quick "leaps".[25]

6. Robustness Checks

6.1 Including all individual stock option series

A natural concern is that the main thrust of our results may be limited to only the individual stock options persistently included in the DJ30 index. Even though to expand our empirical

[24] The realized Sharpe ratios for all the portfolios in sub-periods 2 and 3 (2008-2009) are negative. Negative Sharpe ratios are of course typical of bear regimes and period 3 represents exactly the peak of the financial crisis. Such negative values do not provide meaningful information since even the risk-free asset outperforms the risky investments and negative risk premia are incompatible with the equilibrium (or even no arbitrage).

[25] It is however less compelling to speculate that the sub-sample 2012-2013 may correspond to a sub-sample or no or vanishing learning (see Guidolin and Timmermann, 2003, for an exact definition) that would cause the two-step approach to lose pseudo OOS predictive power in our backtest.

tests to a much longer list of stocks (e.g., all the constituents of the S&P 500), different national markets, or even alternative asset classes (e.g., options on currencies in multi-currency portfolios), would be desirable, for reasons of computational feasibility, we have at first tested the robustness of our results by including all the options series written on underlying stocks that have ever belonged to the DJ30 index, also for just portions of our overall sample period. This way the total number of underlying stocks covered by our paper has increased from an original count of 24 to 36.

The new Table 7 presents the results of the same economic value tests as in Table 5b, when the sample of individual stock option series is expanded to 36. Interestingly, especially when nonnegative weight constraints are imposed, the results are tilted even more in favor of IVS-based portfolio strategies than when only the 24 stocks that always belonged to the DJ30 were investigated, in the sense that the distance between IVS-based strategies under shrinking and the equally-weighted and historical sample based ones grows in favor of the former. However, although they remain rather attractive in absolute terms, the reported realized reward-to-risk performance results in Table 7 are worse than in Table 5b. For instance, the maximum annualized Sharpe ratio obtained under the two-step IVS strategies declines from 0.86 in Table 5b to 0.72 in Table 7; the maximum annualized Sortino ratio declines from 1.12 in Table 5b to 0.90 in Table 7. Yet, because our paper intended to establish that a two-step IVS implied minimum-variance strategy may be competitive vs. a number of well-established benchmarks, our conclusions appear to be robust to extending the breadth of the test stocks, as also the benchmarks struggle when the extension is applied.

6.2 Accounting for transaction costs

One potential issues with the economic value results in Tables 5-7 is that even though it provides higher realized risk-adjusted performances, the two-step IVS driven strategies tend to lead to higher realized mean turnover statistics. As a result, we follow DeMiguel, Plyakha, Uppal, and Vilkov (2013) and calculate the "equivalent transaction cost", that measures how large the transaction costs would need to be to equate any particular performance metric (e.g., the Sharpe ratio) of a given strategy with that for the benchmark strategy (in our case, the DJ30 index).[26] Therefore, this implies statistic measures the resilience of the risk-adjusted performance of a strategy to imposing transaction costs. In our tests, we set the unit transaction cost to equal 10 bps (see French, 2008).

Results are reported in Table 8 in which compile four panels, according to whether the rebalancing is performed daily or weekly and whether only the original sample of 24 stocks permanently in the DJ30 over our sample or the extended cross-section of 36 stock option series are used. Several interesting results emerge. First, the equivalent transaction cost for most strategies, especially when constraints are imposed, tends to exceed 10 bps, suggesting that they are generally better than the DJ30 index even after controlling for the fact that differently from the index they imply a higher turnover. Second, because an equally weighted strategy implies much less trading than less passive strategies, unless constraints are imposed, 1/N leads to the highest equivalent transaction costs. Third, when constraints are imposed and all the 36 stock option series are employed, it is a tight contest between shrunk historical sample moments and two-step IVS strategies as to which model implies the highest resilience to transaction costs. It is in fact reassuring that such competitive implicit transaction costs (in the order 21-25 bps under nonnegative weights and 48-64 bps when also upper bounds on the weights are imposed) emerge when all available stock option data are used.

6.3 Realized information ratios

In Table 6, especially when weekly rebalancing had been assumed, featured two sub-samples (2008 and 2008-2009, i.e., before and after Lehman's demise but both marked by the Great Financial Crisis) returning systematically negative Sharpe and Sortino ratios. This is puzzling because in a portfolio allocation perspective, if one had known of such sub-samples, she would have avoided stocks and the DJ30 index altogether. Yet, ex-ante this would not have been known for sure and it remains to be seen whether the risky strategies—those based on

---

[26] To estimate the equivalent transaction cost, first, for each level of transaction cost, we compute the time series of returns for a given strategy/model and the DJ30 index benchmark when adequate transaction costs are imputed:

$$\tilde{r}_{t+\Delta t}^{strategy} = \tilde{r}_{t+\Delta t}^{strategy} - \sum_{j=1}^{N} \left| w_{j,t+\Delta t}^{strategy} - w_{j,t}^{strategy} \right| \times TC^{strategy},$$

where N is the number of stocks and . $TC^{strategy}$ is the unknown of interest. Then, we compute the performance metrics using these returns. Finally, we numerically search for the level of transaction costs that makes the performance metric the same for the strategy being evaluated and the appropriate benchmark strategy.

IVS implicit information in particular—could still yield positive, relative risk-adjusted performances over these sub-samples. To test this hypothesis, in Table 9, with reference to the same stock option series covered by Table 6b, we have also computed and reported the information ratio (IR), defined as the estimated intercept of a regression of the returns on a given portfolio strategy on DJ30 index returns divided by the standard deviation of the residuals of this very regression (also known as the tracking error). The IR is a well-known measure of the consistency of a portfolio manager's level of skill and ability to generate excess returns relative to a benchmark.[27]

Table 9 presents afresh a selection of the performance measures in Table 6b but adds a rightmost column that shows the annualized IR. Visibly, even in the sub-samples (2008-2009) in which the Sharpe and Sortino ratios had been negative, the IR turns out to be positive and rather sizeable, an indication that at least in relative terms the performance of most of the portfolio strategies under consideration was positive if not really attractive, as IRs close to 1 are reported also in the most problematic sub-samples. In fact, also during 2008-2009, the two-step IVS strategy returns the highest IR (0.99) while in this very period, the equally weighted strategy is the only the leads to a negative IR. Even though the performance of the IVS-based strategies fails to dominate in all sub-samples, Table 9 provides a picture similar to Table 6b but avoids reporting negative risk-adjusted performance indices.

7. Conclusion

Gonçalves and Guidolin (2006) have explored the predictability of the dynamics in the S&P 500 index IVS by proposing a two-step approach. In this paper, our goal has been to evaluate the economic value of their two-step approach in the context of portfolio selection. The investable universe we consider consists of all stocks included in the Dow Jones 30 index over a sample period September 1, 2004 - August 31, 2016.

In line with the original results in GG, we find that the two-stage approach outperforms the benchmark random walk model and provides accurate forecasts in terms of the estimated beta coefficients in the IVS regression as well as of the resulting IVs. In our

[27] By construction, the IR of the benchmark index is not defined or it is conventionally thought of as zero, as the average excess performance at the numerator is zero. Typically, annualized IRs exceeding 0.5 are considered strong and IRs of 1 or more are rather exceptional. We thank an anonymous referee for bringing to our attention to the potential usefulness of IRs in our discussion of realized measures of economic value.

portfolio optimization design, we set all individual stock mean returns to equal the grand mean across all stocks and focus on the prediction of the covariance matrix. We estimate the stock-specific implied volatility as the weighted average of the IVs forecasts obtained through the two-step approach, using the open interest as weights. Next, we compute the implied covariances as the products between the corresponding historical correlations and the implied stock-specific volatilities. We then form the minimum-variance portfolios based on the implied covariance matrix estimator. Our analyses reveal that over the entire sample period, minimum-variance strategies outperform the Dow Jones 30 index both in terms of realized volatility and of Sharpe ratios, indicating that a minimum-variance portfolio is preferable to a passive, value-weighted investment strategy. Secondly, when compared with the equal-weighted portfolio, the minimum-variance strategies always achieve much lower volatility, though they fail to dominate the equal-weighted portfolio in terms of Sharpe ratio and portfolio turnover. Finally, the minimum-variance strategies based on option-implied IVS forecasts generally (e.g., in most sub-samples) lead to lower realized portfolio volatilities vs. MV strategies built using sample moments and also to higher realized Sharpe and Sortino ratios, with the exception of the 2008-2009 financial crisis sub-sample, when however all realized Sharpe ratios turn negative. However, using option-implied second moments in portfolio construction may increase portfolio turnover and therefore expose to high(er) transaction costs. Yet, imposing non-negativity constraints massively reduces transaction costs without spoiling the realized performance in terms of both volatility and Sharpe ratios.

The benefits of applying the IVS-based forecasts in building the covariance matrix for portfolio selection could be further explored in several ways. First, future research could focus on a larger investable universe by taking a wider equity index into consideration, e.g., the S&P 500 and its optionable constituents, or by considering traded options written on other asset classes, e.g., currencies. Second, given the existence of a variety of alternative approaches to forecast the IVS (see, e,g., Chalamandaris and Tsekrekos, 2011; Chen, Zhou, and Li, 2016; Cao, Chen, and Hull, 2020), it would also be interesting to compare their differential economic value by employing their forecasts in the portfolio selection problem that we have used in our research design. Third, even though the two-step method outperforms all benchmarks over the complete back-testing sample, further analysis has uncovered some interesting time variation in its relative, realized economic value performance that may carry some structure. As discussed earlier, it may be interesting to explicitly link our reduced-form approach to some structural asset pricing model with learning (as in Bernales and Guidolin, 2015) to try and also understand the sources of such instability, also with a view of additionally improving its overall portfolio performance.

References

Bakshi, G., N., Kapadia, and Madan, D. (2003). Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. *Review of Financial Studies*, 16, 101-143.

Beber, A., and Brandt, M.W., 2006. The Effect of Macroeconomic News on Beliefs and Preferences: Evidence from the Options Market. *Journal of Monetary Economics*, 53, 1997–2039.

Bernales, A., and Guidolin, M. (2014). Can We Forecast the Implied Volatility Surface Dynamics of Equity Options? Predictability and Economic Value Tests. *Journal of Banking and Finance,* 46, 326-342.

Bernales, A., and Guidolin, M. (2015). Learning to Smile: Can Rational Learning Explain Predictable Dynamics in the Implied Volatility Surface? *Journal of Financial Markets*, 26, 1-37.

Best M. J., and Grauer, R. R. (1991). On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *Review of Financial Studies*, 4, 315-342.

Black, F., and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637-654.

Birke, M., and Pilz, K. F. (2009). Nonparametric Option Pricing with No-Arbitrage Constraints. *Journal of Financial Econometrics*, 7, 53-76.

Brandt, M., W. (2007). Portfolio Choice Problems, in Y. Ait-Sahalia, and L. P. Hansen (eds.), *Handbook of Financial Econometrics*, St. Louis, MO: Elsevier.

Brinkmann, F., Kempf, A., and Korn, O. (2015). Forward-Looking Measures of Higher-Order Dependencies with an Application to Portfolio Selection. *Working paper*, University of Cologne.

Campbell, J. Y., and Viceira, L. M. (1999). Consumption and Portfolio Decisions when Expected Returns Are Time Varying. *Quarterly Journal of Economics*, 114, 433-495.

Cao, J., Chen, J., and Hull, J. (2020). A Neural Network Approach to Understanding Implied Volatility Movements. *Quantitative Finance*, 20, 1405-1413.

Carr, P., and Wu, L. (2016). Analyzing Volatility Risk and Risk Premium in Option Contracts: A New Theory. *Journal of Financial Economics*, 120, 1-20.

Chalamandaris, G., and Tsekrekos, A. E. (2011). How Important is the Term Structure in Implied Volatility Surface Modeling? Evidence from Foreign Exchange Options. *Journal of International Money and Finance*, 30, 623-640.

Chen, S., Zhou, Z., and Li, S. (2016). An Efficient Estimate and Forecast of the Implied Volatility Surface: A Nonlinear Kalman Filter Approach. *Economic Modelling*, 58, 655-664.

Chopra, V. K., and Ziemba, W. T. (1993). The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice. *Journal of Portfolio Management*, 19, 6-11.

Christoffersen, P., and Jacobs, K. (2004). The Importance of the Loss Function in Option Valuation. *Journal of Financial Economics*, 72, 291-318.

Christoffersen, P., Jacobs, K., and Chang, B., Y. (2013). Forecasting with Option-Implied Information. *Handbook of Economic Forecasting* 2, 581-656.

Chun, D., Cho, H., & Ryu, D. (2019). Forecasting the KOSPI200 spot volatility using various volatility measures. *Physica A: Statistical Mechanics and its Applications,* 514, 156-166.

Cont, R., and da Fonseca, J. (2002). Dynamics of Implied Volatility Surfaces. *Quantitative Finance,* 2, 45-60.

Cox, J., Ross, S., and Rubinstein, M. (1979). Option Pricing: A Simplified Approach. *Journal of Financial Economics*, 7, 229-263.

David, A., and Veronesi, P. (2002). Option Prices with Uncertain Fundamentals: Theory and Evidence on the Dynamics of Implied Volatilities*. Working Paper*.

DeMiguel, V., Garlappi, L., Nogales, F. J., Uppal, R. (2009). A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. *Management Science*, 55, 798-812.

DeMiguel, V., Plyakha, Y., Uppal, R., and Vilkov, G. (2013). Improving Portfolio Selection Using Option-implied Volatility and Skewness. *Journal of Financial and Quantitative Analysis,* 48, 1813-1845.

Diebold, F. X., and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13, 253-63.

Dumas, B., Fleming, J., and Whaley, R. (1998). Implied Volatility Functions: Empirical Tests. *Journal of Finance*, 53, 2059-2106.

Ederington, L., ad Lee, J.H. (1996). The Creation and Resolution of Market Uncertainty: the Impact of Information Releases on Implied Volatility. *Journal of Financial and Quantitative Analysis*, 31, 513–539.

Engle, R.F., 2001. GARCH 101: the Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, 15, 157–168.

Fengler, M. R., Hardle, W. K., and Villa, C. (2003). The Dynamics of Implied Volatilities: a Common Principal Components Approach. *Review of Derivatives Research*, 6, 179-202.

Fengler, M. R., Hardle, W. K., and Mammen, E. (2007). A Semiparametric Factor Model for Implied Volatility Surface Dynamics. *Journal of Financial Econometrics*, 5, 189-218.

French, K. R. (2008). Presidential address: the cost of active investing. *Journal of Finance*, 63, 1537-1573.

Gonçalves, S., and Guidolin, M. (2006). Predictable Dynamics in the S&P500 Index Options Implied Volatility Surface. *Journal of Business,* 79, 1591-1635.

Goyal, A., and Saretto, A. (2009). Option Returns and the Cross-Sectional Predictability of Implied Volatility. *Journal of Financial Economics*, 95, 310-326.

Guidolin, M., and Timmerman, A. (2003). Option Prices under Bayesian Learning: Implied Volatility Dynamics and Predictive Densities. *Journal of Economic Dynamics and Control*, 27, 717-769.

Harvey, C., and Whaley, R. (1992). Market Volatility Prediction and the Efficiency of the S&P100 Index Options Market. *Journal of Financial Economics*, 31, 43-73.

Jagannathan, R., and Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance*, 58, 1651-1683.

Kempf, A., O., Korn, and Sabning, S. (2015). Portfolio Optimization Using Forward-looking Information. *Review of Finance,* 19, 467-490.

Kim, M. J., Lee, S. Y., Hwang, D. I., Kim, S. Y., & Ko, I. K. (2010). Dynamics of implied volatility surfaces from random matrix theory. *Physica A: Statistical Mechanics and its Applications,* 389, 2762-2769.

Konstantinidi, E., Skiadopoulos, G., and Tzagkaraki, E. (2008). Can the Evolution of Implied Volatility be Forecasted? Evidence from European and US Implied Volatility Indices. *Journal of Banking & Finance*, 32, 2401-11.

Kostakis, A., Panigirtzoglou, N., and Skiadopoulos, G. (2011). Market Timing with Option-Implied Distributions: A Forward-Looking Approach. *Management Science*, 57, 1231-1249.

Kyriacou, M., Olmo, J., and Strittmatter, M. (2021). Optimal Portfolio Allocation Using Option-Implied Information. *Journal of Futures Markets*, 41, 266-285.

Ledoit, O., and Wolf, M. (2003a). Improved Estimation of the Covariance Matrix of Stock Returns With an Application to Portfolio Selection. *Journal of Empirical Finance*, 10, 603-621.

Ledoit, O., and Wolf, M. (2003b). Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management*, 30, 110-119.

Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is Optimized Optimal?. *Financial Analysts Journal*, 45, 31-42.

Mostowfi, M., and Stier, C. (2013). Minimum-Variance Portfolios Based on Covariance Matrices Using Implied Volatilities: Evidence from the German Market. *Journal of Portfolio Management,* 39, 84-92.

Neumann, M., and Skiadopoulos, G. (2013). Predictable Dynamics in Higher-Order Risk-Neutral Moments: Evidence from the S&P 500 Options. *Journal of Financial and Quantitative Analysis*, 48, 947-977.

Newey, W., and West, K. (1987). A Simple Positive Semi-definite, Heteroskedastic and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55, 703-8.

Pesaran, M. H., and Timmermann, A. (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business & Economic Statistics*, 10, 461-465.

Plyakha, Y., Uppal, R., and Vilkov, G. (2012). Why Do Equal-Weighted Portfolios Outperform Value-Weighted Portfolios? Working paper, EDHEC.

Poon, S. H., and Granger, C. W. J. (2003). Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature,* 26, 478-539.

Skabar, A. (2013). Direction-of-Change Financial Time Series Forecasting Using a Similarity-Based Classification Model. *Journal of Forecasting*, 32, 409-422.

Taboga, M. (2016). Option-implied Probability Distributions: How reliable? How jagged?. *International Review of Economics & Finance*, 45, 453-469.

Timmermann, A. (2001). Structural Breaks, Incomplete Information, and Stock Prices. *Journal of Business and Economic Statistics*, 19, 299–314

Table 1: Index Constituents between September 1, 2004 and August 31, 2016

This table reports the composition of the Dow Jones Industrial Average index between September 1, 2004 to August 31, 2016. The columns designated as “Start Date” and “End Date” denote the date when the stock has been included in and them excluded from the Dow Jones 30 index. When the end date corresponds to the end of the sample, this indicates that for the goals of our paper, at that date the stock was still included in the index.

Company Name	Start Date	End Date
GENERAL ELECTRIC CO	09-01-2004	08-31-2016
EXXON MOBIL CORP	09-01-2004	08-31-2016
COCA-COLA CO	09-01-2004	08-31-2016
INTL BUSINESS MACHINES CORP	09-01-2004	08-31-2016
PROCTER & GAMBLE CO	09-01-2004	08-31-2016
DU PONT (E I) DE NEMOURS	09-01-2004	08-31-2016
UNITED TECHNOLOGIES CORP	09-01-2004	08-31-2016
3M CO	09-01-2004	08-31-2016
MERCK & CO	09-01-2004	08-31-2016
AMERICAN EXPRESS CO	09-01-2004	08-31-2016
MCDONALD'S CORP	09-01-2004	08-31-2016
BOEING CO	09-01-2004	08-31-2016
CATERPILLAR INC	09-01-2004	08-31-2016
DISNEY (WALT) CO	09-01-2004	08-31-2016
JOHNSON & JOHNSON	09-01-2004	08-31-2016
WAL-MART STORES INC	09-01-2004	08-31-2016
HOME DEPOT INC	09-01-2004	08-31-2016
INTEL CORP	09-01-2004	08-31-2016
MICROSOFT CORP	09-01-2004	08-31-2016
AT&T INC	09-01-2004	08-31-2016
JPMORGAN CHASE&CO	09-01-2004	08-31-2016
VERISON COMMUNICATIONS INC	09-01-2004	08-31-2016
PFIZER INC	09-01-2004	08-31-2016
CHEVRON CORP	09-01-2004	08-31-2016
CISCO SYSTEMS INC	06-08-2009	08-31-2016
UNITED HEALTH GROUP INC	09-24-2012	08-31-2016
NIKE INC	09-23-2013	08-31-2016
GOLDMAN SACHS GROUP INC	09-23-2013	08-31-2016
VISA INC	09-23-2013	08-31-2016
HONEYWELL INTERNATIONAL INC	09-01-2004	02-18-2008
ALTRIA GROUP INC	09-01-2004	02-18-2008
AMERICAN INTERNATIONAL GROUP	09-01-2004	09-21-2008
CITIGROUP INC	09-01-2004	06-07-2009
MONDELEZ INTERNATIONAL INC	09-22-2008	09-23-2012
BANK OF AMERICA CORP	02-19-2008	09-22-2013
HP INC	09-01-2004	09-22-2013

Table 2a: Summary Statistics for Implied Volatilities by Maturity and Moneyness, Dow Jones 30 Index Options

This table reports the summary statistics for implied volatilities on Dow Jones 30 index options. The sample period is September 1, 2004-August 31, 2016. The data set is divided into different groups by moneyness and time-to-maturity (see the main text for the relevant group definitions).

		Short-Term		**Medium-Term**		**Long-Term**		Total	%
		calls	puts	calls	puts	calls	puts		
DITM	Obs.	15073	12309	20689	24267	5642	6677	84657	11.85
	Avg. Volume	25	18	10	8	1	2	11	
	Avg. IV	0.3366	0.3003	0.2575	0.2028	0.2326	0.1963	0.2420	
	Std Dev. IV	0.1289	0.1189	0.0700	0.0763	0.0518	0.0656	0.0955	
ITM	Obs.	33031	32739	27164	32211	11433	11229	147807	20.70
	Avg. Volume	56	57	14	11	3	3	24	
	Avg. IV	0.2158	0.1817	0.2065	0.1713	0.2088	0.1950	0.1962	
	Std Dev. IV	0.0909	0.0917	0.0695	0.0753	0.0551	0.0618	0.0759	
ATM	Obs.	22833	20803	17092	20316	18790	17810	117644	16.47
	Avg. Volume	303	328	72	75	18	20	136	
	Avg. IV	0.1589	0.1657	0.1703	0.1630	0.1941	0.1816	0.1740	
	Std Dev. IV	0.0787	0.0871	0.0668	0.0683	0.0642	0.0536	0.0693	
OTM	Obs.	41021	41203	31447	33650	19590	19178	186089	26.06
	Avg. Volume	112	99	48	41	13	9	54	
	Avg. IV	0.1592	0.2184	0.1599	0.2078	0.1779	0.2075	0.1909	
	Std Dev. IV	0.0784	0.0840	0.0680	0.0703	0.0555	0.0558	0.0732	
DOTM	Obs.	30309	46512	33886	35916	17192	14176	177991	24.92
	Avg. Volume	21	25	12	22	7	6	15	
	Avg. IV	0.2444	0.3298	0.1514	0.2611	0.1763	0.2324	0.2413	
	Std Dev. IV	0.0989	0.1071	0.0687	0.0711	0.0569	0.0525	0.0932	
Total	Obs.	295833		276638		141717		714188	100
	Avg. Volume	108		32		7		45	
	Avg. IV	0.2248		0.2003		0.2009		0.2080	
	Std Dev. IV	0.1132		0.0603		0.0788		0.0856	
	%	41.42		38.73		19.84		100	

Table 2b: Summary Statistics for Implied Volatilities Classified by Maturity and Moneyness, All Individual Stock Options

This table reports the summary statistics of implied volatilities on the whole dataset of individual stock options. The sample period is September 1, 2004-August 31, 2016. The data set is divided into different groups by moneyness and time-to-maturity (see the main text for the relevant group definitions).

		Short-Term		**Medium-Term**		**Long-Term**		Total	%
		calls	puts	calls	puts	calls	puts		
DITM	Obs.	180130	123168	35762	56528	951	1770	398309	12.65
	Avg. Volume	279	126	76	27	28	9	91	
	Avg. IV	0.3620	0.3040	0.2519	0.2275	0.2685	0.2572	0.2732	
	Std Dev. IV	0.1225	0.1136	0.0832	0.0817	0.0522	0.0445	0.1029	
ITM	Obs.	137087	139867	2101	6968	966	251	287240	9.12
	Avg. Volume	681	280	125	78	36	4	201	
	Avg. IV	0.2457	0.2132	0.2478	0.2120	0.2473	0.2443	0.2435	
	Std Dev. IV	0.0756	0.0717	0.0656	0.0681	0.0548	0.0407	0.0881	
ATM	Obs.	132817	92877	54792	74778	34766	36548	426578	13.54
	Avg. Volume	1767	1473	462	516	352	300	812	
	Avg. IV	0.2121	0.2069	0.2353	0.2372	0.2428	0.2520	0.2317	
	Std Dev. IV	0.0676	0.0838	0.0833	0.0721	0.0620	0.0677	0.0783	
OTM	Obs.	165391	235317	238015	262591	135251	142723	1179288	37.44
	Avg. Volume	1879	1116	536	329	331	222	735	
	Avg. IV	0.2485	0.2708	0.2312	0.2572	0.2385	0.2418	0.2543	
	Std Dev. IV	0.1043	0.1053	0.0737	0.0758	0.0625	0.0638	0.0811	
DOTM	Obs.	25532	66375	173865	268243	151160	173260	858435	27.25
	Avg. Volume	1439	885	384	245	269	199	570	
	Avg. IV	0.3750	0.4128	0.2437	0.2946	0.2322	0.2752	0.2931	
	Std Dev. IV	0.1289	0.1452	0.0771	0.0814	0.0612	0.0656	0.0932	
Total	Obs.	1298561		1173643		677646		3149850	100
	Avg. Volume	992		278		175		482	
	Avg. IV	0.2851		0.2460		0.2409		0.2513	
	Std Dev. IV	0.1019		0.0762		0.0575		0.0735	
	%	41.23		37.26		21.51		100	

Table 3: Out-of-Sample Average Prediction Errors and Equal Forecast Accuracy Tests Applied on the Estimated Beta Coefficients

Panel A reports the average out-of-sample prediction error measures for the predicted beta coefficients from the IVS regression model. The performance metrics are the Root Mean Squared Prediction Error (RMSE), the Mean Absolute Prediction Error (MAE), and the Mean Correct Prediction (MCP) of the Direction of the Change. Both the VAR(1) model (VAR) and the random walk model (RW) are considered. Panel B shows the results of equal forecast accuracy tests concerning the VAR(1) against the random walk model. The equal forecast accuracy tests are based on Diebold and Mariano (1995) and p-values below 0.05/0.1 (associated to boldfaced test statistics) indicate the possibility to reject the null of equal predictive accuracy (in favor of the VAR model when the statistic is negative).

Panel A: Prediction Error Measures						
		$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
DJ30 Index						
RMSE	VAR	0.0713	0.1746	0.5148	0.0116	0.0692
	RW	0.0593	0.2614	0.8297	0.0307	0.1222
MAE	VAR	0.0549	0.1234	0.2385	0.0080	0.0514
	RW	0.0416	0.1655	0.3974	0.0205	0.0835
MCP	VAR	0.6898	0.7542	0.7952	0.8504	0.7970
	RW	NA	NA	NA	NA	NA
All Sample of Options on Individual Stocks						
RMSE	VAR	0.1357	0.6292	1.2315	0.0220	0.5226
	RW	0.0490	0.8915	1.9245	0.0369	0.7223
MAE	VAR	0.0668	0.4177	0.7986	0.0123	0.3171
	RW	0.0293	0.5721	1.2089	0.0223	0.4414
MCP	VAR	0.7394	0.7909	0.8070	0.8482	0.7888
	RW	NA	NA	NA	NA	NA
Panel B: Forecast Accuracy Tests						
DJ30 Index						
RMSE	DM statistic	0.5632	-1.7630	-1.4837	**-2.5589**	**-2.1342**
	P-value	0.5733	0.0780	0.1380	0.0106	0.0329
MAE	DM statistic	**2.1093**	**-3.5641**	**-3.0898**	**-6.3654**	**-5.3828**
	P-value	0.0350	0.0004	0.0020	0.0000	0.0000
All Sample of Options on Individual Stocks						
RMSE	DM statistic	**6.4034**	**-8.2712**	**-13.0645**	**-4.0827**	**-2.3676**
	P-value	0.0000	0.0000	0.0000	0.0000	0.0180
MAE	DM statistic	**5.8615**	**-8.0625**	**-12.5638**	**-14.2955**	**-8.0244**
	P-value	0.0000	0.0000	0.0000	0.0000	0.0000

Table 4: Out-of-Sample Average Prediction Errors and Equal Forecast Accuracy Tests Applied on Implied Volatilities

Panel A reports the average out-of-sample prediction error measures for the forecasts of implied volatilities. The performance metrics are the Root Mean Squared Prediction Error (RMSE), the Mean Absolute Prediction Error (MAE), and the Mean Correct Prediction (MCP) of the Direction of the Change. Both the VAR(1) model (VAR) and the random walk model (RW) are considered. Panel B shows the forecast accuracy tests for the VAR(1) against the random walk model in the case of RMSE and MAE and against a 50% random sign prediction benchmark in the case of MCP. The forecast accuracy tests are based on Diebold and Mariano (1995) in the case of the squared and absolute value loss and in the column labeled "MCP-DM". The column "MCP-DM" tests equal mean correct prediction vs. a 50% random benchmark applying Pesaran and Timmermann's (1992) test.

Panel A: Prediction Error Measures					
		RMSE	MAE	MCP	
DJ30 Index	VAR	0.0212	0.0139	0.5822	
	RW	0.0243	0.0122	NA	
All Sample of Individual Stocks	VAR	0.0423	0.0229	0.5361	
	RW	0.2591	0.1313	NA	
Panel B: Forecast Accuracy Tests					
		RMSE	MAE	MCP-DM	MCP-PT
DJ30 Index	DM statistic	**-2.1628**	1.7271	**6.1502**	**11.0982**
	P-value	0.0306	0.0841	0.0000	0.0000
All Sample of Individual Stocks	DM statistic	**-8.2523**	**-6.1999**	**6.1444**	**64.3432**
	P-value	0.0000	0.0000	0.0000	0.0000

Table 5a: Out-of-Sample Performance for Different Portfolio Strategies under Daily Rebalancing

This table reports the average values of the out-of-sample performance metrics for different portfolio strategies covered by our analysis. The performance metrics consist of the average realized returns, the realized volatility, the Sharpe ratio, the average turnover, the mean tracking error, the Sortino ratio, and the maximum drawdown. Panel A summarizes the results for the unconstrained, baseline case; Panel B summarizes the results when a short-sale constrain is imposed; Panel C summarizes the results when an additional 20% upper bound on the individual stock weight is added. We denote the equally-weighted portfolio strategy as "1/N", the minimum-variance portfolio based on the historical sample covariance estimator as "H", and the corresponding shrinkage estimators as "ShrH" and "ShrI". The realized performance measure across alternative strategies is boldfaced in each of the panels.

Daily Rebalancing							
Strategy	Mean	Volatility	Annualized Sharpe Ratio	Turnover	Tracking Error	Sortino Ratio	Max Drawdown
Panel A: Unconstrained							
1/N	**0.1324**	0.1965	0.6738	**0.3921**	**0.0018**	0.8591	-14.1870
Historical	0.1095	0.1363	**0.8033**	0.6111	0.0098	1.0512	-5.8443
IVS-based	0.0863	0.1356	0.6365	0.8132	0.0098	0.7942	-11.6614
ShrH	0.1060	**0.1345**	0.7876	0.5267	0.0105	**1.0554**	**-5.0876**
ShrIVS	0.0862	0.1355	0.6363	0.8098	0.0098	0.7943	-11.6367
Panel B: Nonnegative weights							
1/N	**0.1324**	0.1965	0.6738	0.3921	**0.0018**	0.8591	-14.1870
Historical	0.1063	0.1420	0.7486	0.1814	0.0063	0.9734	-7.2374
IVS-based	0.1228	0.1413	0.8689	0.3159	0.0062	1.1288	-7.3025
ShrH	0.1041	0.1424	0.7313	**0.1596**	0.0064	0.9534	**-7.2136**
ShrIVS	0.1227	**0.1412**	**0.8690**	0.3158	0.0062	**1.1290**	-7.2941
Panel C: Nonnegative weights and 20% upper limit							
1/N	0.1324	0.1965	0.6738	0.3921	**0.0018**	0.8591	-14.1870
Historical	**0.1489**	0.2178	**0.6837**	0.3160	0.0045	0.8804	-11.4651
IVS-based	0.0948	**0.1483**	0.6389	0.2787	0.0053	0.8532	-10.6449
ShrH	0.1468	0.2183	0.6724	0.3070	0.0045	0.8666	-11.0704
ShrIVS	0.0948	0.1483	0.6392	**0.2786**	0.0053	**0.8835**	**-10.6405**
DJ30 Index	0.0552	0.1936	0.2850	NA	NA	0.3481	-19.3548

Table 5b: Out-of-Sample Performance for Different Portfolio Strategies under Weekly Rebalancing

This table reports the average values of the out-of-sample performance metrics for different portfolio strategies covered by our analysis. The performance metrics consist of the average realized returns, the realized volatility, the Sharpe ratio, the average turnover, the mean tracking error, the Sortino ratio, and the maximum drawdown. Panel A summarizes the results for the unconstrained, baseline case; Panel B summarizes the results when a short-sale constrain is imposed; Panel C summarizes the results when an additional 20% upper bound on the individual stock weight is added. We denote the equally-weighted portfolio strategy as "1/N", the minimum-variance portfolio based on the historical sample covariance estimator as "H", and the corresponding shrinkage estimators as "ShrH" and "ShrI". The realized performance measure across alternative strategies is boldfaced in each of the panels.

Weekly Rebalancing							
Strategy	Mean	Volatility	Annualized Sharpe Ratio	Turnover	Tracking Error	Sortino Ratio	Max Drawdown
Panel A: Unconstrained							
1/N	0.1304	0.1966	0.6633	**0.3922**	**0.0221**	0.8458	-14.1808
Historical	0.1075	0.1409	**0.7629**	0.6449	0.0236	1.0066	-5.9776
IVS-based	0.1013	0.1395	0.7257	0.9216	0.0235	0.9065	-9.4706
ShrH	0.1039	**0.1379**	0.7532	0.5455	0.0241	**1.0191**	**-5.1725**
ShrIVS	**0.1012**	0.1394	0.7256	0.9176	0.0235	0.9067	-9.4384
Panel B: Nonnegative weights							
1/N	0.1304	0.1966	0.6633	0.3922	0.0228	0.8458	-14.1808
Historical	0.1074	0.1429	0.7516	0.2014	0.0227	0.9911	**-6.5636**
IVS-based	0.1219	0.1427	0.8546	0.3801	0.0226	1.1151	-7.1128
ShrH	**0.1036**	0.1426	0.7259	**0.1765**	0.0227	0.9606	-6.9246
ShrIVS	0.1220	**0.1426**	**0.8551**	0.3800	**0.0226**	**1.1159**	-7.1057
Panel C: Nonnegative weights and 20% upper limit							
1/N	0.1304	0.1966	0.6633	0.3922	**0.0221**	0.8458	-14.1808
Historical	0.1485	0.2180	**0.6809**	0.3296	0.0228	0.8756	-11.5098
IVS-based	**0.0956**	0.1486	0.6435	0.3195	0.0225	0.8553	-10.7972
ShrH	0.1452	0.2182	0.6656	**0.3170**	0.0229	0.8578	-10.9676
ShrIVS	0.0956	**0.1486**	0.6438	0.3194	0.0225	**0.8856**	**-10.7870**
DJ30 Index	0.0552	0.1936	0.2850	NA	NA	0.3481	-19.3548

Table 6a: Out-of-Sample Performance for Different Sub-Samples under Short-Sale Constraints and Daily Rebalancing

We denote the equally-weighted portfolio strategy as "1/N", the minimum-variance portfolio based on the historical sample covariance estimator as "H", and the corresponding shrinkage estimators as "ShrH" and "ShrI".

Daily Rebalancing							
Strategy	Mean	Volatility	Annualized Sharpe Ratio	Turnover	Tracking Error	Sortino Ratio	Max Drawdown
Period 1 (2004-2007)							
1/N	0.1085	0.1288	0.8423	0.3299	0.0101	1.1015	-4.1404
Historical	0.1104	0.1032	1.0697	0.2836	0.0089	1.4877	-2.8391
IVS-based	0.1086	0.1029	1.0554	0.3577	0.0089	1.4926	-2.6731
ShrH	**0.1203**	**0.1007**	**1.1951**	**0.2329**	0.0088	**1.7121**	-2.6833
ShrIVS	0.1090	0.1027	1.0618	0.3564	**0.0088**	1.5023	**-2.6514**
DJ30 Index	0.0707	0.1378	0.5131	NA	NA	0.6721	-6.0638
Period 2 (2008)							
1/N	-0.2092	**0.2115**	-0.9893	0.3829	**0.0148**	-1.6451	-8.4141
Historical	-0.1592	0.3151	-0.5051	0.1916	0.0212	-0.8830	-9.1822
IVS-based	-0.3241	0.2950	-1.0986	0.2836	0.0198	-1.9295	-11.8964
ShrH	**-0.0973**	0.3051	**-0.3190**	**0.1544**	0.0206	**-0.5291**	**-7.1650**
ShrIVS	-0.3207	0.2943	-1.0897	0.2816	0.0198	-1.9048	-11.7807
DJ30 Index	-0.3400	0.2453	-1.3863	NA	NA	-1.9638	-9.2363
Period 3 (2008-2009)							
1/N	-0.7831	0.5921	-1.3225	0.4232	0.0372	-2.2479	-24.8809
Historical	-0.4474	**0.3546**	-1.2617	0.3992	**0.0228**	-2.1332	-11.5656
IVS-based	**-0.3736**	0.3954	**-0.9449**	0.3484	0.0253	**-1.6697**	**-10.9037**
ShrH	-0.4528	0.3575	-1.2665	0.3830	0.0230	-2.1299	-12.3507
ShrIVS	-0.3767	0.3954	-0.9528	**0.3482**	0.0253	-1.6835	-10.9679
DJ30 Index	-0.9798	0.5239	-1.8702	NA	NA	-3.0195	-22.1274
Period 4 (2009-2012)							
1/N	0.1621	0.1732	0.9363	0.4370	0.0129	1.2250	-5.6985
Historical	0.1693	0.1610	1.0520	0.3731	0.0125	1.4469	-5.9052
IVS-based	0.1737	0.1551	1.1201	0.4448	0.0121	1.4962	-4.6240
ShrH	**0.1844**	0.1732	1.0650	**0.3322**	0.0131	1.4142	-5.6375
ShrIVS	0.1743	**0.1550**	**1.1217**	0.4422	**0.0121**	**1.4991**	**-4.6224**
DJ30 Index	0.1150	0.1685	0.6823	NA	NA	0.8815	-6.1432
Period 5 (2012-2013)							
1/N	**0.2181**	0.1139	**1.9139**	0.4429	0.0107	**2.6239**	-1.6871
Historical	0.1464	0.1392	1.0520	0.3423	0.0113	1.4527	-2.3559
IVS-based	0.2155	0.1255	1.7166	0.4534	0.0105	2.5151	-1.6025
ShrH	0.1425	0.1413	1.0087	**0.2849**	0.0116	1.4125	-2.5704
ShrIVS	0.2153	0.1255	1.7147	0.4527	**0.0105**	2.5134	**-1.6006**
DJ30 Index	0.1366	**0.1133**	1.2054	NA	NA	1.6036	-2.0096
Period 6 (2013-2016)							
1/N	0.1952	0.0953	2.0479	0.5325	0.0088	3.0177	-1.5556
Historical	0.1518	0.0965	1.5730	0.4860	0.0088	2.3843	-1.5215
IVS-based	**0.2019**	0.0935	**2.1593**	0.6010	0.0087	**3.2226**	**-1.2699**
ShrH	0.0965	0.0977	0.9880	**0.3435**	0.0090	1.6179	-2.1874
ShrIVS	0.1995	**0.0934**	2.1356	0.5969	**0.0087**	3.1959	-1.2920
DJ30 Index	0.1158	0.0993	1.1663	NA	NA	1.5121	-2.3113

Table 6b: Out-of-Sample Performance for Different Sub-Samples under Short-Sale Constraints and Weekly Rebalancing

We denote the equally-weighted portfolio strategy as “1/N”, the minimum-variance portfolio based on the historical sample covariance estimator as “H”, and the corresponding shrinkage estimators as “ShrH” and “ShrI”.

Weekly Rebalancing							
Strategy	Mean	Volatility	Annualized Sharpe Ratio	Turnover	Tracking Error	Sortino Ratio	Max Drawdown
Period 1 (2004-2007)							
1/N	0.1076	0.1208	0.8902	**0.3299**	**0.0024**	1.1920	-14.2344
Historical	0.1331	0.1073	1.2412	0.4313	0.0129	1.7342	-9.1283
IVS-based	0.1227	0.1033	1.1883	0.4967	0.0104	1.5574	-8.8642
ShrH	**0.1346**	0.1068	**1.2608**	0.3364	0.0136	**1.7840**	-9.3286
ShrIVS	0.1226	**0.1032**	1.1886	0.4938	0.0104	1.5563	-8.8502
DJ30 Index	0.0707	0.1378	0.5131	NA	NA	0.6721	**-6.0638**
Period 2 (2008)							
1/N	-0.1593	**0.1777**	-0.8966	0.3823	**0.0326**	-1.5651	-26.8293
Historical	-0.1015	0.2560	-0.3963	0.2644	0.0397	-0.6819	-28.0868
IVS-based	-0.2922	0.2535	-1.1528	0.3827	0.0407	-2.0232	-39.9359
ShrH	**-0.0430**	0.2406	**-0.1788**	**0.2249**	0.0374	**-0.2939**	-21.7453
ShrIVS	-0.2865	0.2518	-1.1379	0.3789	0.0405	-2.0007	-39.4799
DJ30 Index	-0.3400	0.2453	-1.3863	NA	NA	-1.9638	**-9.2363**
Period 3 (2008-2009)							
1/N	-0.9834	0.4956	-1.9841	**0.4229**	0.0876	-3.0068	-70.6487
Historical	-0.5287	0.2826	-1.8707	0.4997	0.0648	**-2.2217**	-41.7377
IVS-based	**-0.5201**	0.2838	**-1.8326**	0.4540	0.0652	-2.2603	-42.9625
ShrH	-0.5501	**0.2801**	-1.9640	0.4801	**0.0641**	-2.3398	-44.2868
ShrIVS	-0.5217	0.2836	-1.8391	0.4528	0.0652	-2.2676	-43.0740
DJ30 Index	-0.9798	0.5239	-1.8702	NA	NA	-3.0195	**-22.1274**
Period 4 (2009-2012)							
1/N	0.1621	0.1679	0.9659	0.4368	0.0327	1.3774	-19.0172
Historical	0.1811	0.1508	1.2013	0.4835	0.0309	1.6681	-18.9544
IVS-based	0.1786	**0.1414**	1.2637	0.5521	0.0297	1.8231	-14.4839
ShrH	**0.1924**	0.1628	1.1812	**0.4275**	0.0320	1.7665	-18.7930
ShrIVS	0.1795	0.1417	**1.2667**	0.5471	**0.0297**	**1.8283**	-14.4929
DJ30 Index	0.1150	0.1685	0.6823	NA	NA	0.8815	**-6.1432**
Period 5 (2012-2013)							
1/N	**0.2136**	0.1069	**2.0550**	0.4429	0.0219	**2.8536**	-5.2186
Historical	0.1184	0.1216	0.9736	0.4822	0.0215	1.2453	-8.2362
IVS-based	0.1888	0.1066	1.7715	0.5901	0.0219	2.2185	-5.3334
ShrH	0.1226	0.1244	0.9851	**0.3951**	0.0216	1.3490	-8.7609
ShrIVS	0.1889	**0.1065**	1.7728	0.5880	**0.0215**	2.2199	-5.3423
DJ30 Index	0.1366	0.1133	1.2054	NA	NA	1.6036	**-2.0096**
Period 6 (2013-2016)							
1/N	0.1950	**0.0906**	2.1533	0.5324	0.0202	3.6371	-4.2356
Historical	0.1649	0.0992	1.6628	0.6209	0.0204	3.2658	-4.1247
IVS-based	**0.2331**	0.0931	**2.5025**	0.7272	**0.0199**	**5.5917**	-2.9795
ShrH	0.0992	0.0993	0.9987	**0.4614**	0.0207	1.8012	-7.6442
ShrIVS	0.2306	0.0932	2.4758	0.7216	0.0199	5.5327	-3.0361
DJ30 Index	0.1158	0.0993	1.1663	NA	NA	1.5121	**-2.3113**

Table 7: Out-of-Sample Performance for Different Portfolio Strategies under Weekly Rebalancing and an Expanded Set of Individual Equity Options

The table replicates panel B of Table 5 when the sample of individual equity option series is expanded to include all stocks ever included in the DJ30. Panel A summarizes the results for the unconstrained, baseline case; Panel B summarizes the results when a short-sale constrain is imposed; Panel C summarizes the results when an additional 20% upper bound on the individual stock weight is added. We denote the equally-weighted portfolio strategy as "1/N", the minimum-variance portfolio based on the historical sample covariance estimator as "H", and the corresponding shrinkage estimators as "ShrH" and "ShrI". The realized performance measure across alternative strategies is boldfaced in each of the panels.

Strategy	Mean	Volatility	Annualized Sharpe Ratio	Turnover	Tracking Error	Sortino Ratio	Max Drawdown
Panel A: Unconstrained							
1/N	0.0712	0.1862	0.3821	**0.4088**	0.0320	0.4275	-45.8008
Historical	0.1322	0.1739	0.7600	5.5083	0.0313	1.0781	-27.3370
IVS-based	**0.1427**	0.1781	0.8012	3.4742	0.0316	1.1425	-21.1355
ShrH	0.1243	**0.1408**	**0.8829**	0.8513	**0.0277**	**1.2295**	**-17.6752**
ShrIVS	0.1351	0.1699	0.7950	2.9501	0.0307	1.1088	-21.5735
Panel B: Nonnegative weights							
1/N	0.0712	0.1862	0.3821	0.4088	0.0320	0.4275	-45.8008
Historical	0.1043	0.1516	0.6878	0.4703	0.0289	0.8350	-32.2913
IVS-based	0.1048	0.1460	0.7182	0.5405	**0.0282**	0.8583	-32.3434
ShrH	0.1051	0.1553	0.6764	**0.3927**	0.0293	0.8561	**-31.5794**
ShrIVS	**0.1051**	**0.1459**	**0.7202**	0.5367	0.0282	**0.8620**	-32.2438
Panel C: Nonnegative weights and 20% upper limit							
1/N	0.0712	0.1862	0.3821	0.4088	0.0320	0.4275	-45.8008
Historical	0.1104	0.1494	0.7393	0.4736	0.0284	0.8981	-32.3006
IVS-based	**0.1108**	**0.1472**	**0.7523**	0.5339	0.0281	**0.8979**	**-30.1183**
ShrH	0.1052	0.1504	0.6998	**0.4076**	0.0286	0.8765	-33.3620
ShrIVS	0.1105	0.1472	0.7508	0.5305	**0.0281**	0.8970	-30.1947
DJ30 Index	0.0552	0.1936	0.2850	NA	NA	0.3481	-19.3548

Table 8: Out-of-Sample Equivalent Transaction Costs of Portfolio Strategies under Weekly Rebalancing and an Expanded Set of Individual Equity Options

The table computes the equivalent transaction cost—a measures of how large such transaction costs would need to be to equate the realized Sharpe ratio of a given strategy with that of benchmark—following DeMiguel, Plyakha, Uppal, and Vilkov (2013). Panel A summarizes the results for the unconstrained, baseline case; Panel B summarizes the results when a short-sale constrain is imposed; Panel C summarizes the results when an additional 20% upper bound on the individual stock weight is added. We denote the equally-weighted portfolio strategy as "1/N", the minimum-variance portfolio based on the historical sample covariance estimator as "H", and the corresponding shrinkage estimators as "ShrH" and "ShrI". The highest equivalent transaction cost measure across alternative strategies is boldfaced in each of the panels.

	1/N	Historical	IVS-based	ShrH	ShrIVS
Panel A: Original sample, daily rebalancing					
Unconstrained	**0.0019**	0.0010	0.0006	0.0011	0.0006
Nonnegative weights	0.0019	0.0032	0.0020	**0.0036**	0.0020
Nonnegative weights and 20% upper limit	0.0019	0.0027	0.0020	**0.0027**	0.0020
Panel B: Original sample, weekly rebalancing					
Unconstrained	**0.0019**	0.0009	0.0006	0.0010	0.0006
Nonnegative weights	0.0019	0.0029	0.0017	**0.0032**	0.0017
Nonnegative weights and 20% upper limit	0.0019	0.0025	0.0017	**0.0026**	0.0017
Panel C: Expanded individual stocks sample, daily rebalancing					
Unconstrained	**0.0013**	0.0002	0.0003	0.0009	0.0003
Nonnegative weights	0.0013	0.0016	0.0014	0.0020	**0.0021**
Nonnegative weights and 20% upper limit	0.0013	0.0016	0.0014	0.0019	**0.0025**
Panel C: Expanded individual stocks sample, daily rebalancing					
Unconstrained	**0.0046**	0.0008	0.0010	0.0033	0.0011
Nonnegative weights	0.0046	0.0052	0.0045	0.0063	**0.0064**
Nonnegative weights and 20% upper limit	0.0046	0.0054	0.0048	**0.0061**	0.0048

Table 9: Out-of-Sample Realized Information Ratios for Different Sub-Samples under Short-Sale Constraints and Weekly Rebalancing

We denote the equally-weighted portfolio strategy as "1/N", the minimum-variance portfolio based on the historical sample covariance estimator as "H", and the corresponding shrinkage estimators as "ShrH" and "ShrI".

Strategy	Mean	Volatility	Annualized Sharpe Ratio	Annualized Information Ratio
Period 1 (2004-2007)				
1/N	0.1076	0.1208	0.8902	**2.1918**
Historical	0.1331	0.1073	1.2412	0.6808
IVS-based	0.1227	0.1033	1.1883	0.7055
ShrH	**0.1346**	0.1068	**1.2608**	0.6622
ShrIVS	0.1226	**0.1032**	1.1886	0.7023
DJ30 Index	0.0707	0.1378	0.5131	—
Period 2 (2008)				
1/N	-0.1593	**0.1777**	-0.8966	0.7798
Historical	-0.1015	0.2560	-0.3963	0.8473
IVS-based	-0.2922	0.2535	-1.1528	0.1654
ShrH	**-0.0430**	0.2406	**-0.1788**	**1.1191**
ShrIVS	-0.2865	0.2518	-1.1379	0.1861
DJ30 Index	-0.3400	0.2453	-1.3863	—
Period 3 (2008-2009)				
1/N	-0.9834	0.4956	-1.9841	-0.0058
Historical	-0.5287	0.2826	-1.8707	0.9804
IVS-based	**-0.5201**	0.2838	**-1.8326**	**0.9932**
ShrH	-0.5501	**0.2801**	-1.9640	0.9437
ShrIVS	-0.5217	0.2836	-1.8391	0.9899
DJ30 Index	-0.9798	0.5239	-1.8702	—
Period 4 (2009-2012)				
1/N	0.1621	0.1679	0.9659	0.2033
Historical	0.1811	0.1508	1.2013	0.3014
IVS-based	0.1786	**0.1414**	1.2637	0.3019
ShrH	**0.1924**	0.1628	1.1812	0.3403
ShrIVS	0.1795	0.1417	**1.2667**	**0.3056**
DJ30 Index	0.1150	0.1685	0.6823	—
Period 5 (2012-2013)				
1/N	**0.2136**	0.1069	**2.0550**	**0.4956**
Historical	0.1184	0.1216	0.9736	-0.1194
IVS-based	0.1888	0.1066	1.7715	0.3358
ShrH	0.1226	0.1244	0.9851	-0.0915
ShrIVS	0.1889	**0.1065**	1.7728	0.3366
DJ30 Index	0.1366	0.1133	1.2054	—
Period 6 (2013-2016)				
1/N	0.1950	**0.0906**	2.1533	0.5533
Historical	0.1649	0.0992	1.6628	0.3390
IVS-based	**0.2331**	0.0931	**2.5025**	**0.8279**
ShrH	0.0992	0.0993	0.9987	-0.1131
ShrIVS	0.2306	0.0932	2.4758	0.8108
DJ30 Index	0.1158	0.0993	1.1663	—

Figure 1: Time Series of the OLS Beta Coefficients Estimates for the Cross-Sectional Model

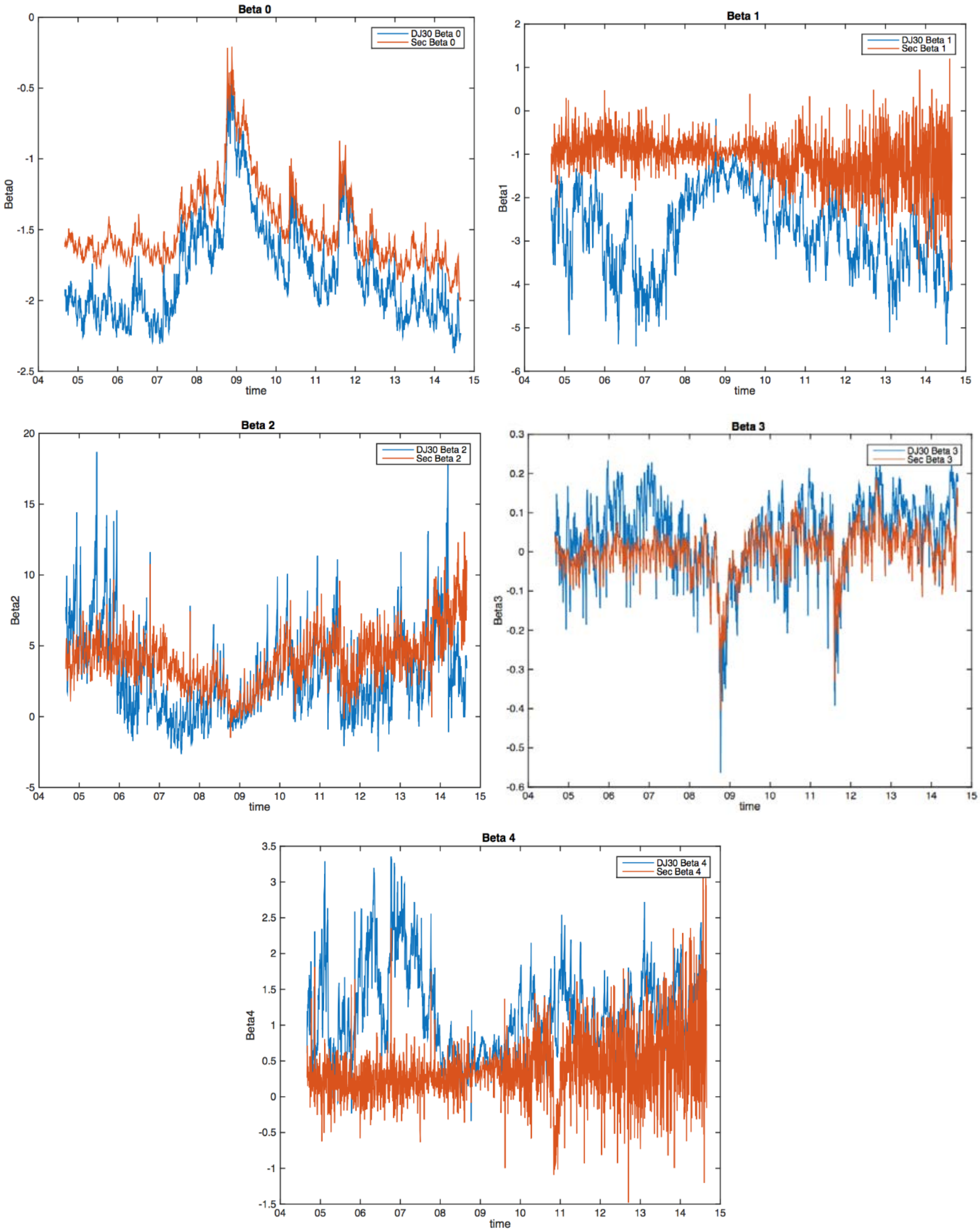

Figure 2: One-Year Rolling Window Volatility

This figure plots the one-year rolling window realized volatility for Dow Jones 30 index and for unconstrained minimum-variance portfolios derived from implied covariance matrix estimator and sample covariance matrix estimator. We denote the minimum-variance portfolio built from the sample covariance matrix estimator as "H" and that built from the implied covariance estimator as "I".