Mean-Variance Market Timing the U.S. Stock Market *

Pezzo Luca† Lei Wang‡ Duygu Zirek§

November 11, 2021

Abstract

While recently the after-cost profits of many anomalies are close to zero, investing according to the Mean-Variance (MV) criterion has never been so rewarding. In particular, minimizing over transaction costs restores credibility in the capability of MV strategies to efficiently target risk premia by timing stock risk premia, additionally lowering downside risk and enhancing scalability. More generally, market timing and estimation error are important drivers behind the dynamics of MV profitability in the U.S. stock market over the last century.

JEL-Classification: C61, D23, G11

Keywords: Mean-Variance, Market-timing, Estimation Error, Transaction Costs, Profitability

*We are thankful to Phil Dybvig, Thomas Maurer, Mihail Velikov and Raja Velu for many helpful comments and suggestions.

†University of New Orleans, department of Economics and Finance, +1 314 422 6083, lpezzo@uno.edu.

‡University of New Orleans, department of Economics and Finance, +1 716 589 4183, lwang10@uno.edu.

§Montclair State University, +1 646 713 4585, zirekd@montclair.edu.

1 Introduction

Recent years have witnessed a boost in the overall quality and efficiency of the U.S. stock market.[1] Using the words in Chordia et al. (2014): "The improvements in trading technology and liquidity are dramatic and quite unprecedented". In line with these findings, Chen and Velikov (2020) find that profits from trading in anomalies disappeared. In contrast, we find that investing according to the Mean-Variance (MV) criterion over all the common stocks traded on the NYSE/AMEX and NASDAQ has never been so rewarding. Several of our strategies perform better then the equally and value weighted market portfolios as well as broad ETFs such as the SPDR (mimicking the S&P 500 index) and the IWM (mimicking the Russell 2000 index), earning annualized out-of-sample after-cost Sharpe ratios well above 1. On the contrary, the top-right graph of Figure 4 shows how the Sharpe ratios of the leading 23 anomalies analyzed in Novy-Marx and Velikov (2016) are almost always below 0.5 recently.

We analyze the performance of three standard textbook types of MV strategies – the MV Tangency Portfolio (MVTP), the Global Minimum Variance Portfolio (GMVP), and four frontier portfolios targeting an annualized risk premium of 1% (MVP1), 5% (MVP5), 10% (MVP10) and 15% (MVP15) respectively – and their cost-optimized versions when appropriate.[2] The two main contribution of our study are the findings that: 1) the stabilizing role of cost minimization on portfolio weights can restore the credibility of MV strategies to efficiently target risk premia, 2) market timing and estimation error are two key drivers behind the profitability of MV strategies.

We find two robust patterns in the profitability of MV strategies over the recent past. While the insights from the first one are mostly helpful in refreshing the typical known properties of MV strategies, those from the second ones are new and represents the first main contribution~~s~~ of this paper. We first document how $GMVP$, the simplest among all MV strategies, is very profitable and stable, earning after-cost Sharpe ratios of 1.12 and 1.29 in the absence of price pressure (depending on whether we use the past 6 or 12 months to estimate the covariance matrix). More generally, textbook MV strategies tend to be very

[1] See for example Hendershott et al. (2011), Chordia et al. (2011) and Chordia et al. (2014).

[2] Our cost-optimized strategies are versions of GMVP and the four frontiers portfolio that minimizes over transaction cost inside the MV optimization as we detail in Section 2.1. There is no cost-optimized equivalent to the MVTP strategy since its cost-optimized analog depends on a given level of risk aversion.

profitable only when the impact of stock risk premia estimates is minimal and their main source of profitability comes from timing stock covariances. That is, those strategies cannot handle the estimation error present in the estimates of stock risk premia. What is new here is the finding that the recent performance of $GMVP$ is among the best over the last 100 years. The rest is already well known.[3] *The second observed pattern, which represents the first main result of this paper, is the finding that minimizing over costs within the MV setup is a way to restore credibility in the ability of MV strategies to efficiently target risk premia.* MV strategies minimizing portfolio variance *and* transaction costs subject to a targeted level of risk premium have Sharpe ratios indistinguishable from those of GMVP and dominates their textbook analogs (not minimizing over costs) at targeting higher premia. Once stock risk premia are estimated using at least the past 10 years of data, cost-optimized MV strategies targeting premia of 5,10 and 15% per annum have Sharpe ratios never below 1.2. In contrast, standard analogs have ratios almost always smaller than 1. Furthermore, cost-optimizing strategies are less prone to downside risk and more scalable (i.e. their profits are eroded less from the price impact of trades).

We can explain these findings by analyzing the role played by trading costs in the MV optimization. Olivares-Nadal and DeMiguel (2016) show how trading costs in the MV optimization can be thought as tuning parameters for stabilizing the performance. More generally, explicitly accounting for costs represents an economically sounded way to impose bounds on the MV weights. This is because, as we assume in our paper (and most of the literature),[4] when costs are proportional or proportional and quadratic there exists a no trading region around the optimal allocation absent costs, and it is optimal to trade only when the before-cost risk-return trade-off is enough far away from such optimum. We show this point by illustrating how bounding the weights through cost minimization dominates the adoption of standard MV strategies in the presence of either non-negative weights (no short-selling allowed) or in the additional presence of uniform upper bounds on the maximal positions in any single stock.

The stabilizing role played by cost-minimized weights allows to: i) unlock and profitably exploit the information contained in stock risk premia, ii) attenuate downside risk, and

[3]See for example the literature on volatility timing, e.g. Fleming et al. (2001), and Moreira and Muir (2017), and on the performance of $GMVP$, e.g. Jagannathan and Ma (2003), and Clarke et al. (2006, 2011)

[4]See for example Dybvig and Pezzo (2020) and reference therein.

iii) scale up. First we argue that the stabilizing role of cost-optimized weights allows to market time stock risk premia by smoothing out the extra noise that prevents MV standard strategies to exploit the information contained in the stock premia. This is because we find that cost-optimized strategies that dominate their standard analogs derive a considerable portion of their Sharpe ratios from market timing stock risk premia in contrast to their less profitable analogs. Moreover the most profitable standard strategies load very little on stock premia estimates and mostly derive their profitability from timing stock covariances (while actually suffering from stock premia estimation error when the stock premia estimates are considered as inputs). Second, we find that cost-optimized strategies uniformly have lower exposures to downside risk as measured by the worst observed loss (Maximum Draw Down) and the average time to recover from it. This is because the lower volatility present in the realized returns helps to cap losses. Finally, explicitly modeling for price impact, by means of quadratic transaction costs in our setup, enables cost-optimized strategies to consistently scale up more (and be less sensitive to changes in price impact). This is an important new insight in that it suggests that also large institutional investors can gain efficient market exposures by exploiting the MV criterion.

The fact that cost-optimized strategies targeting higher premia are more profitable than their standard analogs can be rationalized by the Frazzini and Pedersen (2014) "betting against beta" theory. The systematic pattern found in the alphas (higher) and betas (lower) of cost-optimized strategies vis-a-vis those of standard strategies matches the theory's predictions. Therefore, these findings are sustained by an equilibrium model populated by investors with different levels of trading constraints. While the original theory attributes such constraints to different margin requirements in our re-interpretation we attribute them to the presence of trading costs, a different type of trading constraint.

It is instructive to compare the benefits of optimizing over transaction costs across different markets. In the U.S. stock market, where estimation error is quite pervasive, especially in the forecast of risk premia, the predominant role of cost minimization is that of stabilizing the performance. Using a setup very similar to ours, Maurer et al. (2020) analyze the impact of costs in the FX markets. In such markets, as the authors show, estimation error is much less pervasive, primarily because good and simple forecasts are available for the currency risk premia. As a consequence the main benefits of cost minimization come from the direct reduction of execution costs. Also the implications for price impact are different. While explicitly modeling price impact as a trading cost always stabilize and substantially improves the after-cost performance, profits of all our strategies are mostly eroded at a slow, approximately linear, rate. This is in sharp contrast to the exponential profit decay found by Maurer et al. (2020) in the FX markets while conducting the *exact same* type of analysis. What is crucial here is the *number* of assets invested. While in the FX markets no more than 29 currencies are jointly traded, we invests in thousands of stocks. This makes portfolio weights relatively small. Accordingly, rebalancing induces small buy and sell orders generating a modest price impact from trading.[5]

According to our discussion so far a large chuck of the profits from the most profitable strategies (GMVP and the cost-optimized strategies targeting higher returns) comes from market timing at least one of the first two moments of the stock return distribution. We also find that the least profitable strategies tends to suffer the most from estimation error. More generally, *our second main finding is that market timing and estimation error are key (and closely related) drivers behind the dynamics of MV profitability in the U.S. stock market over the last century.* In our setup market timing and estimation error are mechanically related. We measure their impact on the performance of a given MV strategy as the Sharpe ratio differential between the real-time version of it and its unimplementable version where at least one of the first two moments is kept fixed at its in-sample look-ahead biased estimate (instead of being estimated recursively based on the information available at the time of execution). Since the only difference between the two versions of the strategy comes from the time-variation in the estimates for the moments, when the differential is positive(negative), i.e. the real-time performance is higher(lower), we are capturing the impact of market timing (estimation error). We find that market timing and estimation error have substantial impact on the performance of MV strategies, with average Sharpe ratio increments (for the case of market timing) and decrements (for the case of estimation error) between 0.3 and 0.45, as well as 0.26 and 0.66 respectively. Moreover, MV profitability displays a procyclical pattern with magnitudes increasing in expansions and decreasing in recessions over time. Such dynamics can be explained by market timing in expansions and estimation error in recessions. As a consequence, we find substantial positive correlation between the

[5] The same argument holds true even while investing in the smaller universe of $SP500$ stocks as shown in Appendix E.

profitability of MV strategies and their level of market timing (with low values associated to estimation error), particularly for cost-optimized strategies. The average correlation over the last 100 years is between 0.28 and 0.56 for standard strategies and between 0.5 and 0.82 for cost-optimized strategies. Further corroborating the hypothesis of a causal relation between market timing and estimation error and the profitability of our MV strategies, the majority of most profitable strategies have market timing levels belonging to the top tercile, almost always corresponding to positive gaps (our way to detect actual market timing activity). At the same time, most of the least profitable strategies have market timing levels not higher then the median, almost always corresponding to negative gaps (our way to detect estimation error).

The implementation of MV strategies requires estimates for stocks risk premia and covariances. We deal with the well known issues associated with the estimation of the first two moments for the entire cross section of stock returns by imposing a conditional factor structure on the data generating process and exploiting the superior predictability of the second moments at higher frequency. The conditional covariance matrix is estimated recursively using the previous 6 or 12 months of daily returns, adjusting for infrequent trading and missing values. While the vector of risk premia is generated by the two-step noise-filtering missing-at-random procedure developed in Gagliardini et al. (2016) recursively using the previous 5, 10 or 20 years of monthly returns. More details are provided in Section 2.2.

Our results are robust to the way we measure costs (with or without TAQ data), which factor model (market versus Fama and French (1993) 3-factors) we use to estimate the conditional stock risk premia μ_t and the conditional covariance matrix of stock returns V_t, as well as to the specific re-balancing frequency (monthly vs. quarterly) and the composition of the stock universe.[6] From our robustness analysis we additionally learn that: 1) the trade-off between bias and estimation error in the covariance matrix estimator is clearly in favor of a reduction of the latter,[7] 2) adopting a parsimonious specification for the return generating process and filtering out part of the noise in the risk premia estimates through

[6]Whether or not we reduce the stock universe to: i) the S&P 500 constituents, ii) the same sub-sample of stocks implicitly selected by the most stringent analyzed rolling-window-length combination of 12 months for estimating V_t and 20 years for estimating μ_t, or iii) the even more stringent sub-sample of stocks required to estimate V and μ unconditionally).

[7]This is because the performance of our strategies are (much) better when we estimate V_t via factor models rather then the Ledoit and Wolf (2017, 2020) shrinkage estimator.

the Gagliardini et al. (2016) approach is the best way, among those analyzed in this paper, to handle stock risk premia,[8] 3) there are economically sizable diversification benefits in investing in all common stocks.[9]

The rest of the paper is structured as follows: section 2 details the setup and the strategies employed in our analysis, including the measurement of market timing and estimation error. Section 3 describes the U.S. stock market investment opportunity set over the last century, highlighting the macro-trends useful for our analysis. Section 4 documents the recent stark profitability of MV strategies, singling out the relevant patterns. Section 5 analyzes the role of cost minimization in enabling MV strategies to efficiently target risk premia. Section 6 shows how market timing and estimation error are important drivers behind the profitability of the analyzed MV strategies over the last century. Section 7 presents the summary of the robustness analysis which is detailed in the Internet Appendix. Section 8 concludes.

2 Setup

The stock universe is taken to be that of all common stocks trading on the NYSE, AMEX and NASDAQ[10] from January 1926 through December 2017.

Let t be the end of a given period, and N_t the number of available stocks to trade in at that time. Define r_t and C_t to be the N_t-vector of realized excess returns and proportional cost respectively. Returns are taken as the Center for Research in Security Prices (CRSP) holding period returns in excess of the Ibbotson and Associates 1-month risk-free rate,[11] while costs are the Chen and Velikov (2020) effective spreads.[12]

The trading dynamics for a generic strategy j works as follow: at any time t the positions of the strategy inherited from $t-1$, $\theta_t^{0,j}$, equals the $t-1$ post-trade positions θ_{t-1}^{j} (which can

[8] The one-factor market model within the Gagliardini et al. (2016) is more efficient than the Fama and French (1993) 3-factor formulation or the estimation of the factor models under the standard Fama and MacBeth (1973) procedure.

[9] Performances are lower if we restrict the stock universe to the $SP500$ stocks.

[10] Stocks with the first two digits of the Center for Research in Security Prices (CRSP) variable "shrcd" equal to 10 and 11.

[11] If the period is a quarter we compute the implied quarterly excess returns by compounding the monthly ones.

[12] A measure of the implied half bid-ask spreads. We are grateful to the authors for sharing their data.

be a vector of 0s if the strategy has just been set up) plus the vector of adjustments due to the shocks associated with the realizations of the returns, $r_t \cdot \theta^j_{t-1}$, where $\cdot$ defines the element-by-element product. Then the time t weight vector θ^j_t is set up given the available information up to t and the following transaction costs are incurred $TC^j_t = \sum_i^{N^j_t} |\theta^j_{t,i} - \theta^{0,j}_{t,i}|C_{t,i}$, where N^j_t is the number of stocks accessible for trading by strategy j. From our definition of trading costs it follows that the cost of trading each single stock i, $C_{t,i}$, is proportional to the size of trade in that stock, $|\theta^j_{t,i} - \theta^{0,j}_{t,i}|$. In $t+1$ the following after-cost excess return realizes $r^{AC,j}_{t+1} = r'_{t+1}\theta^j_t - TC^j_t$. Notice that the returns of our strategies are all obtained out-of-sample and there is no look-ahead bias, i.e., such returns obtains in real-time.

2.1 Strategies

We analyze three standard textbook types of MV strategies – the MV Tangency Portfolio (MVTP), the Global Minimum Variance Portfolio (GMVP), and four frontier portfolios targeting an annualized risk premium of 1% (MVP1), 5% (MVP5), 10% (MVP10) and 15% (MVP15) respectively – and their cost-optimized versions when appropriate[13] – namely, the cost-minimized version of the GMVP (labeled GMVPtc), and that of the four frontier portfolios (labeled MVP1tc,MVP5tc,MVP10tc and MVP15tc respectively). Moreover, we evaluate our MV strategies against several popular U.S. market benchmarks: the equally (EW) and value (VW) weighted portfolios as well as the S&P 500 ETF SPDR, and the Russell 2000 ETF IWM, which are only available in the most recent time periods.

We next describe our strategies in details.

2.1.1 Equally Weighted portfolio (EW)

Is the strategy which weight vector at generic time t is defined as

$$\theta^{EW}_t \equiv \frac{1}{N^{EW}_t}$$

[13]There is no cost-optimized equivalent to the MVTP strategy since its cost-optimized analog depends on a given level of risk aversion.

where N_t^{EW}, with a slight abuse of notation, is both the number and the set of stocks that at time t have a trading cost estimate.

2.1.2 Value Weighted portfolio (VW)

Is the strategy which weight vector at generic time t is defined as

$$\theta_t^{VW} \equiv \frac{MktCap_{t,i}}{\sum_i^{N_t^{VW}} MktCap_{t,i}}$$

where N_t^{VW} is the number of stocks, and subset of N_t^{EW}, for which we have an estimate for the market capitalization.

2.1.3 Global Minimum Portfolio (GMVP) and cost-optimized analog (GMVPtc)

The Global Minimum Variance Portfolio is the strategy which weights minimize the variance of all possible weights that sum up to 1

$$\theta_t^{GMVP} \equiv \arg \min_{\{\theta_t | \mathbf{1}'\theta_t = 1\}} \{\theta_t' V_t \theta_t\} = \frac{V_t^{-1}\mathbf{1}}{\mathbf{1}' V_t^{-1}\mathbf{1}}$$

where $\mathbf{1}$ is the vector of 1s.

A cost-optimized version of GMVP additionally minimizes over transaction costs. That is, the optimal weight vector, θ_t^{GMVPtc}, is the solution of the following problem

$$\begin{aligned} \min_{\{\Delta_t^+ \geq 0, \Delta_t^- \geq 0\}} & \{\theta_t' V_t \theta_t + (\Delta_t^+ + \Delta_t^-)' C_t\} \\ s.t. \quad \theta_t &= \theta_t^0 + \Delta_t^+ - \Delta_t^- \\ 1 &= \mathbf{1}'\theta_t \end{aligned}$$

where Δ_t^+ and Δ_t^- are the non-negative vectors of stock purchases and sales to be added and subtracted to the initial position $\theta_t^0 = \theta_t^{0,GMVPtc}$ inherited from $t-1$ to get the after-trade optimal vector of positions $\theta_t = \theta_t^{GMVPtc}$. Notice that when $C_t = 0$, $\theta_t^{GMVPtc} = \theta_t^{GMVP}$. If $C_t \geq 0$ than $\theta_t^{GMVPtc} \neq \theta_t^{GMVP}$. Moreover, both vectors are uniquely identified as long as V_t is positive-definite.

In practice for both strategies V_t is estimated using the $N_t^{GMVP} \subseteq N_t^{EW}$ number of stocks for which we have estimates for the matrix entries.

2.1.4 Generic Frontier Portfolio (MVPE) and cost-optimized analog (MVPEtc)

A frontier portfolio is defined as a vector of weights, summing up to 1, minimizing the portfolio variance given a target risk premium level E

$$\theta_t^{MVPE} \equiv \arg \min_{\{\theta_t | \mathbf{1}'\theta_t = 1, \mu_t'\theta_t = E\}} \{\theta_t' V_t \theta_t\} = \frac{V_t^{-1}}{D_t}[\mathbf{1}(B_t - EA_t)\mu_t(EC_t - A_t)]$$

where $A_t = \mathbf{1}'V_t^{-1}\mu_t, B_t = \mu_t'V_t^{-1}\mu_t, C_t = \mathbf{1}'V_t^{-1}\mathbf{1}$ and $D_t = B_tC_t - A_t^2$.

Similarly to the case of GMVP, a cost-optimized version of MVPE additionally minimizes over transaction costs. That is, the optimal weight vector, θ_t^{MVPEtc}, is the solution of the following problem

$$\min_{\{\Delta_t^+ \geq 0, \Delta_t^- \geq 0\}} \{\theta_t' V_t \theta_t + (\Delta_t^+ + \Delta_t^-)' C_t\} \quad s.t.$$
$$\begin{aligned} \theta_t &= \theta_t^0 + \Delta_t^+ - \Delta_t^- \\ E &= \theta_t' \mu_t \\ 1 &= \mathbf{1}'\theta_t. \end{aligned}$$

When there are no costs MVPE and MVPEtc are identical and both strategy weight vectors are uniquely identified as along as V_t is positive-definite.

In practice for both strategies V_t and μ_t are estimated using the $N_t^{MVP} \subseteq N_t^{GMVP}$ number of stocks for which we have estimates for both the matrix and the vector entries.

2.1.5 Mean-Variance Tangency Portfolio

Finally the Mean-Variance Tangency Portfolio is the proportion of (risky) stocks present in the optimal portfolio of a Mean-Variance investor with coefficient of risk aversion λ investing in the (risky) stocks and a risk-free asset. That is

$$\theta_t^{MVTP} \equiv \frac{V_t^{-1}\mu_t}{\mathbf{1}'V_t^{-1}\mu_t} = \frac{1}{\mathbf{1}'(\frac{\lambda}{2}V_t^{-1}\mu_t)}(\frac{\lambda}{2}V_t^{-1}\mu_t) = \frac{1}{\mathbf{1}'(\frac{\lambda}{2}V_t^{-1}\mu_t)} \arg\max_{\{\theta_t\}}\{\theta_t'\mu_t - \frac{\lambda}{2}\theta_t' V_t \theta_t\}$$

notice how θ_t^{MVTP} is not a function of λ.[14] In practice V_t and μ_t are estimated using the $N_t^{MVP} \subseteq N_t^{GMVP}$ number of stocks for which we have estimates for both the matrix and the vector entries.

2.2 Estimation of parameters

In practice MV strategies require estimates for the vector of risk premia and the covariance matrix of returns. We deal with the well know issues associated with the estimation of the first two moments for the entire cross section of stock returns by imposing a conditional factor structure on the data generating process and exploiting the superior predictability of the second moments at higher frequency. Let f_t be a k-th vector of traded factor returns and assume the data generating process for the vector of excess stock return to be

$$r_t = a + B' f_t + e_t \tag{1}$$

where e_t represents an independent mean-zero disturbance vector with diagonal covariance V_e and each column of B contains the k stock-specific factor loadings.

Using daily returns over the previous 6 or 12 months, we recursively estimate (1) and define the monthly covariance matrix of returns at time t as $V_t = \frac{12}{252}(B_t' V_{f,t} B_t + V_{e,t})$ where $B_t, V_{f,t}$ and $V_{e,t}$ are estimates for B, V_e and the covariance matrix for the factors V_f over the past 6 to 12 months. We deal with the unbalanced nature of the panel of returns and the illiquidity of small stocks by estimating (1) separately for each stock. The former issue is dealt with by only retaining stocks that have at least 80% of the observations over the previous 6 or 12 months. The latter problem is tackled by estimating the i-th stock loading B_i (i-th column of B) as the average of the slopes of three different variations of (1) where f_t is in turn regressed on $r_{i,t-1}, r_{i,t}$, and $r_{i,t+1}$.

Using monthly returns over the previous 5,10 or 20 years we recursively estimate (1) and define the monthly vector of stock risk premia at time t as $\mu_t = b_t' \lambda_t$. Specifically, we follow the noise-filtering missing-at-random two-pass approach of Gagliardini et al. (2016) in its simpler time-invariant risk premia formulation as in Berrada and Coupy (2015). b_t is the first pass estimate for B in (1) for those stocks with enough non-missing observations as

[14] A cost-optimized version would instead depend on λ. This is why $MVTP$ has no cost-optimized analog.

established by a statistical filter. λ_t equals the conditional mean of the factors' returns plus a miss-pricing adjustment. Such adjustment is estimated in the second pass by regressing a_t, the intercept from (1) for the retained stocks, on b_t via weighted least squares. The exact steps followed in our estimation are described in Section 2.1 of Berrada and Coupy (2015). The advantage of this method over the standard Fama and MacBeth (1973) approach lies in the noise reduction in the premia achieved through the first pass filter and the second pass adjustment condensing the information contained in the whole cross-section.

Throughout our main analysis we parsimoniously specify f_t as the market factor (time series of the CRSP value-weighted portfolio excess returns). In the online appendix we show how our results are robust to the Fama and French (1993) 3-factor specification. We confirm the superiority of the Gagliardini et al. (2016) risk-premia estimates over the standard Fama and MacBeth (1973) approach (finding that the 1-factor specification under the Gagliardini et al. (2016) is indeed the most efficient approach). We also find that estimating a large covariance matrix via a factor model rather than via the nonlinear Ledoit and Wolf (2017, 2020) Shrinkage estimator induces a substantial improvement in the out-of-sample profitability (thus validating the claim that a reduction of estimation error over bias is what matters for large scale covariance estimators). Finally, in line with the literature[15] we confirm in an unreported robustness check (available upon request) that estimating the covariance matrix at the monthly frequency is suboptimal from a performance standpoint.

2.3 Market timing and estimation error

In this paper we primarily measure the ability of a given MV strategy to time the market or alternatively to suffer from the impact of estimation error by comparing its out-of-sample real-time performance with that under a scenario where the first and/or second moments of the stock return distribution are kept fixed at their in-sample look-ahead biased estimates.

Specifically, define $\{t\}_{t=1}^{T}$ as the reference sample. Let $\theta_t^j(\hat{\mu}_t, \hat{V}_t)$ be the weight vector of strategy j at time $t \in \{t\}_{t=1}^{T}$ for given conditional estimates for the stock risk premia, $\hat{\mu}_t$, and their covariance matrix, $\hat{V}_t$. Define $\hat{\mu}$ and $\hat{V}$ as the fixed in-sample look-ahead biased estimates. Then if the out-of-sample performance from $\{r_{t+1}^j(\hat{\mu}_t, \hat{V}_t)\}_{t=0}^{T-1}$ is different

[15]See Fleming et al. (2001) (and references therein) as well as Moreira and Muir (2017).

than that from $\{r^j_{t+1}(\hat{\mu}, \hat{V})\}_{t=0}^{T-1}$ such difference is entirely due to the time variation in the estimates for the risk premia and/or their covariance matrix. In particular, if the gap is positive, i.e. the real-time performance is higher, this is because the strategy is benefiting from the time-variation in the weights coming from the conditional moments. That is, the strategy is displaying market timing abilities. On the other hand, if the gap is negative, i.e. the real-time performance is negative, this is because the extra variation in the weights coming from the conditional estimates $\hat{\mu}_t$ and $\hat{V}_t$ is compromising the overall performance, thus representing estimation error.

3 The U.S. stock market investment opportunity set

Two quantities an investor always needs to consider while forming a portfolio strategy are the number of available stocks and the cost of trading those stocks. The number of available stocks are proportional to the achievable degree of diversification but also to the amount of estimation error a strategy (which requires stock-level estimates) might suffer from. While trading costs can be thought of proxies for market frictions (e.g. liquidity, or simply anything that prevents trades to be those available in a frictionless world).

Figure 1 represents the U.S. stock market investment opportunity set over the monthly period January 1926 - December 2017. The left y-axis reports in blue the time series of the Chen and Velikov (2020) cross-sectional median effective spreads (implied half bid-ask spreads). While the right y-axis displays in orange the time series of the total number of common stocks available from the CRSP database as defined in Section 2. Combining the structural brakes in these time series we can divide roughly the last 100 years into four sub-samples: 1926-44 as a sample with high costs and low number of stocks, 1945-72 as a sample where costs turn to low, 1973-02 as a sample characterized by the spikes in the number of available stocks (due to the inclusion of the NASDAQ stocks in CRSP) but also by very high trading costs, and 2003-17 as a sample characterized by the all-times-low costs level.

We argue that these two quantities (costs and number of stocks) roughly pick-up the long-term cyclicality in the U.S. stock market investment opportunity set. 1926-44 is a recessionary period (including the 1929-39 Great Depression, followed by the WWII) defined by an exceptionally high documented level of market volatility (see Schwert (1990)). The post-war sample 1945-72 is an expansionary period as documented by the steady growth of the stock market relative to the broader economy. Such growth is captured by the green time series which reports the ratio of the stock market capitalization to the U.S. Gross Domestic Product (GDP) as in Ludvigson et al. (2020). 1973-02 is again mostly recessionary as indicated by the 74-94 two-decade-long sluggish market growth (green line) occupying more than two-third of the sub-sample. Finally, as documented by several recent studies, after 2002 an exceptional boost in market quality takes place. We observe a simultaneous dramatic drop in transaction costs and increment in volume traded (captured by the dot-dashed magenta line and measured as the cross-sectional median of the monthly median daily number of trades as reported by CRSP) and a steady decreasing pattern for the number of available stocks (orange line) culminating in an average market growth (green line) at record heights.[16]

3.1 The U.S. Mean-Variance investment opportunity set

It has been known since the seminal work of Markowitz (1952) that the Capital Allocation Line (CAL) describes the mean-variance frontier in the presence of a risk-free asset and a generic number of risky assets with risk premia μ and variance-covariance V. Therefore, the slope of the CAL, the highest achievable Sharpe rario $SR = \sqrt{\mu V^{-1} \mu}$, is a sufficient statistic to characterize the MV investment opportunity set.

Figure 2 plots the time series of *ex-ante* MV Sharpe ratios obtainable in real time at every date t by suitably estimating μ_t and V_t only using information up to t (blue lines, left y-axes). The top graph plots such time series when the vector of risk premia μ_t and the covariance matrix of returns V_t are estimated over a rolling-window covering the previous 60 months. Similarly the center and bottom graphs plot the time-series for rolling windows

[16]The beginning of 2003 corresponds to the inclusion of the NYSE autoquote trading system. This event can be regarded as an exogenous shock capturing the sharp increment in algorithmic trading and ultimately resulting in a liquidity boost (see Hendershott et al. (2011)). Moreover, as highlighted by the black dashed vertical line, in August 2005 the Security Exchange Commission introduced the NMS Regulation, which opened the gates for competition among market venues. Angel et al. (2015) and Chordia et al. (2011), among others, argue that these changes brought an unprecedented increment in market quality and market efficiency. Finally, we can see how the increment in competition is consistent with the observed downward patter in the number of available stocks: an unreported analysis (available upon request) shows how 55% of the variation in the number of stocks after 2002 is explained by stocks being dropped out of the NYSE/AMEX and NASDAQ (adding the explanatory power of stocks merged brings the explained variation up to 80% while still keeping the expiatory power of the number of stock dropped highly statistically significant).

of 120 and 240 months respectively. Underneath each sub-sample date, we report the *ex-post* MV Sharpe ratio, computed by using the fixed in-sample estimates for μ and V. The time series of the U.S. market growth relative to the economy are also super-imposed in the same graphs (green lines, right y-axes). While the solid lines represent the already discussed ratio between the U.S. market cap. and the GDP, the dot-dashed series displays the ratio between the U.S. market cap and the U.S. total personal consumption expenditures. Finally, the horizontal dashed green lines enable to quickly locate periods where market growth is faster/slower that that of the economy (namely, when the solid and dot-dashed green time series are above/below it).

We can see how MV investment opportunities are: i) volatile: especially when the moments (i.e. μ and V) of the ex-ante Sharpe ratios are estimated at the higher frequencies (shorter rolling windows), ii) pro-cyclical: the ex-ante Sharpe-ratios are highly correlated with the stock market growth (as reported in the graphs' legends, approximately 47% at the 60-month rolling window, 60% at the 120-month rolling window, and 83% at the 240-month rolling window), iii) increasing over time: the ex-ante Sharpe ratios display a marked long-term increasing trend, while the in-sample estimates gets higher as time goes by and jump at an all-time record high of 0.85 starting in 2003.

In summary, roughly over the past century, when the U.S. stock market goes up, trading frictions (as measured by median bid-ask spreads) go down and the MV investment opportunities go up and vice-versa. Moreover, these long-term trends are well captured by our sub-samples. In Section 6.2 we further show how the documented ex-ante real-time pro-cyclical and increasing pattern displayed by the MV opportunities is systematically reflected in the ex-post out-of-sample performance of our MV strategies.

4 Recent stark mean-variance profitability

According to the discussed characteristics of the investment opportunity set, investing following the MV criterion never seemed so appealing as recently. MV strategies are designed to maximize the Sharpe ratio in an i.i.d. world and: i) we observed a secular long-term increasing trend in the ex-ante MV Sharpe ratio (matched by an ex-post Sharpe ratio after 2002 almost 2.5 times bigger than ever before), ii) the market is perceived as more efficient (see Chordia et al. (2011)) and, iii) returns' auto-correlation at the analyzed (monthly and quarterly) frequencies are (notoriously) low. We find indeed that, in sharp contrast to investing in anomalies, the real-time profitability of MV strategies has never been so high as recently.

Figure 3 displays the out-of-sample after-cost annualized Sharpe ratios of the equally (EW) and valued weighted (VW) strategies as they compare to those of our MV strategies: namely the Mean Variance Tangency Portfolio (MVTP), the Global Minimum Variance Portfolio (GMVP), and four frontier strategies targeting an annualized risk premium of 1% (MVP1), 5% (MVP5), 10% (MVP10) and 15% (MVP15) respectively. The reported Sharpe ratio of a given MV strategy is the average across six strategies only differing in the estimates for μ_t and V_t.[17] Solid lines report the ratios of standard textbook strategies while dotted lines (only available for the GMVP and the four frontier portfolios) those belonging to transaction-cost-optimized analogs. Finally, the different colors refers to the four different sub-samples introduced in Section 3. Notice how the after-cost performances of the MV strategies post 2002 (displayed in green) have never been so high, and how cost optimized strategies (in dotted green) dominates the textbook analogs (in solid green) at targeting higher returns with average Sharpe ratios above 1. One of the main result of this paper is the finding that minimizing over cost within the MV setup is a way to restore credibility in the ability of MV strategies to efficiently target risk premia.

The recent performance of MV strategies is particularly impressive even when taken into perspective. While investing in the entire market, either equally or through a value weighted scheme or via broad ETFs (such as the S&P 500 SPDR or the Russell 2000 IWM), generates an after cost Sharpe ratio between 0.6 and 0.75,[18] the average Sharpe ratios of our most profitable MV strategies are in the 1.05-1.25 neighborhood. Therefore gaining market exposure via the MV criterion is much more efficient. Also, as we show in Section 5.4, especially with respect to cost-optmized strategies, these outcomes are likely obtainable not just by marginal price-taker investors. Indeed cost-optimize strategies appear to scale up nicely even in the presence of substantial price impact and the same is true for $GMVP$

[17] For each given MV strategy we compute six strategies only differing in the length of the rolling window used to estimate the risk premia (employing either the past 60, 120 or 240 months of data) and the covariance matrix (employing either the past 6 or 12 months of data).

[18] Where EW and VW produce a Sharpe ratio of 0.62 and 0.73, and the S&P 500 SPDR and the Russell 2000 IWM a ratio of 0.70 and 0.63 respectively.

if we estimate V_t using the past year of daily data. Moreover this recent profitability is in sharp contrast with that from investing in anomalies. Figure 4 compares the performance of our strategies with those of the leading 23 anomalies analyzed in Novy-Marx and Velikov (2016). While before 2003 the average profitability of anomalies was similar and perhaps higher than those of MV strategies, after 2002 it is only slightly above zero on average.

4.1 Main features of recent mean-variance profitability

Two main features characterize the recent stark profitability of mean-variance strategies: $GMVP$ is the simplest way to profitably gain market exposure (given an estimate for the covariance matrix we have a closed form solution for the optimal weights) and cost-optimized strategies are valid alternatives, especially when it comes to target risk premia. Decomposing the strategies' Sharpe ratios reveals that the main source of profitability of $GMVP$ (expectedly) comes from its variance reduction while the documented superior Sharpe ratios of cost-optimized strategies targeting higher returns come both from lower volatilities and higher average excess returns.

Figure 5 decomposes the post-2002 average performance shown in Figure 3 (solid green for the standard strategies and dotted-green for the cost-optimized MV strategies). Each graph only differs in the rolling-window lengths used to estimate the conditional risk premia μ_t and the covariance matrix V_t for our set of analyzed MV strategies. EW and VW are reported each time for convenience but they do not depend on such estimates. Each row correspond to a different rolling window for μ_t, going from 60 months (or 5 years) in the top graphs to 240 months (or 20 years) in the bottom ones. Similarly, each column fixes the window for V_t to 6 months (left column) or 12 months (right column). Every point estimate (ex-post after cost annualized Sharpe ratio) is denoted with a green bold circle. Sharpe ratios of given strategies that are statistically better at the 10% level than: the EW, the VW, both the SPDR and the IWM ETFs,[19] and their standard analogs (if they refer to cost-optimized strategies) are marked with a star, a diamond, a square and a circle respectively. Markers for standard strategies are reported in black while for cost-optimized analogs in red.

$GMVP$ has the highest most stable performance with a annualized after-cost out-of-

[19] Which mimic the exposure to the $S\&P$ 500 and the Russell's 2000 indices.

sample Sharpe ratio of 1.12 and 1.29 when the covariance V_t is estimated using the last 6 months and year of data respectively. Moreover its performance is higher (at least at the 10% level) than that of EW (0.61) and VW (0.73) as well as those of the market ETFs, $SPDR$ (0.70) and IWM (0.63). Even if the finding of the recent performance of $GMVP$ being among the best over the last 100 years is new, the efficient out-of-sample performances of this strategy and the fact that a major portion of it comes from the ability to time stock covariances (which we confirm in Section 5.2) are not (see Clarke et al. (2006, 2011), Fleming et al. (2001), Jagannathan and Ma (2003) and Moreira and Muir (2017)). What is also new, and documented later, is the fact that $GMVP$ can actually scale up.

One of the main results of this paper is the finding that cost minimization is a way to restore the credibility of MV strategies to efficiently target risk premia. When the stock risk premia μ_t are reliably estimated (with a rolling window of at least 10 years), Figure 5 shows how cost-optimized strategies targeting higher return (namely $MVP5tc$,$MVP10tc$ and $MVP15tc$) are valid alternative to $GMVP$. With respect to the 10% statistical level their performances are indistinguishable from that of $GMVP$, with Sharpe ratios which are always significantly grater than those of EW, VW, and the market ETFs $SPDR$ and IWM (with point estimates above 1.2). Moreover, their performances is always economically better than those of their standard analogs (with a minimum out-performance of 0.14 for the strategies targeting an annualized premium of 5% under the rolling window configuration 10 years - 1 year for μ_t and V_t respectively), with the statistical significance being more pronounced when V_t is estimated over the past 6 months as indicated by the red circle around the markers of $MVP5tc$,$MVP10tc$ and $MVP15tc$. In Section 5 we show how cost-optimized strategies are also less prone to downside risk (measured as the worse suffered loss and the average time to recover from it) and scale up better then standard strategies (including $GMVP$).

We conclude this section by decomposing the Sharpe ratios of our analyzed strategies. The leftmost top two plots of Figure 6 show the average excess returns for the standard and the cost-optimized strategies (the Sharpe ratio numerators) while the leftmost bottom two plots report the respective volatilities (the Sharpe ratio denominators). As expected, the main source of profitability of $GMVP$ comes from its ability to shrink the portfolio variance. Both versions of $GMVP$, whether we estimate V_t using the past 6 or 12 months, have a volatility of roughly 5%, one of the lowest among all strategies.[20] The volatilities of cost-optimized strategies when we estimate the stock risk premia with at least 10 years of data are uniformly below 15% and less spread out that those from the analog standard strategies. By the same token, the average excess return of cost-optimized strategies display a more consistent upward pattern. The higher degree of consistency in the volatility and average return patterns of cost-optimized strategies is in line with the stabilizing role played by the cost-optimizing MV weights which we further discuss in Section 5. Therefore we conclude that the documented superior Sharpe ratios of cost-optimized strategies targeting higher returns come both from lower volatilities and higher average excess returns. This insight is indeed confirmed by the rightmost graphs of Figure 6 plotting the difference in average returns (top) and volatilities (bottom) between the cost-optimized and the standard MV strategies. Cost-optimized strategies have average returns at least 1.36% higher (mostly between 2 and 4% higher), while the volatility of $MVP5$, $MVP10$ and $MVP15$ when stock risk premia are estimated using at least 10 years of data are almost always lower (i.e. the volatility differential is negative).

5 Cost-minimization to efficiently target risk premia

The stabilizing role played by cost-minimized weights allows to: i) unlock and profitably exploit the information contained in stock risk premia, ii) attenuate downside risk, and iii) scale up. In particular, the fact that cost-optimized strategies targeting higher premia are more profitable than their standard analogs can be rationalized by the Frazzini and Pedersen (2014) "betting against beta" theory.

5.1 Cost-minimization as a good way to stabilize the performance

Standard MV weights are simple and available in close form. However small variation in their estimates often cause large swings in the weights making the performance of such strategies unstable (see for example DeMiguel et al. (2009)). Olivares-Nadal and DeMiguel (2016) show how trading costs in the MV optimization can be thought as tuning parameters for

[20]The higher Sharpe ratio when V_t is estimated annually, 1.29 versus 1.12 when V_t uses the past 6 months of data, therefore mainly come from the higher average return of approximately 9.33% versus 8.70%.

stabilizing the performance. In line with this interpretation we find that the performances of cost-optimized strategies are much less impacted by estimation error.

More generally, explicitly accounting for trading costs represents an economically sounded way to impose bounds on such MV weights. This is because in the presence of costs it is not always optimal to trade. When costs are proportional as in our case (or proportional and quadratic when we consider price impact), there exists a no trading region around the portfolio weights that would have been optimal in the absence of costs. So that at any given time t if our initial position (inherited from time $t-1$ after returns in t have realized) in some stocks is too close to the optimal one we do not trade in those stocks because the marginal benefits are lower than the marginal costs. As a result weights become less sensitive to changes in μ_t or V_t and we adjust positions to the optimal before cost risk-return trade-off only when we are far enough away from it. For a comprehensive analysis of the topic we refer the interested reader to Dybvig and Pezzo (2020) and references therein.

Minimizing costs within the MV optimization is not the only way to bound the weights. The common practice is actually to impose no-short sale constraints and/or upper bounds on the maximum exposure on given stocks. However, most of the times these constraints are not economically justified and even labeled as "wrong" (see Jagannathan and Ma (2003)). Table 1 reports the Sharpe ratio differentials between the cost-optimized strategies and the ratios of the standard strategies: in the presence of no-short selling constraints (top panel), in the additional presence of a uniform upper bound of 1% (central panel), and 2% (bottom panel). Notice how imposing ad-hoc constraints vis-a-vis cost-optimized constraints is almost always sub-optimal (generating positive differentials). The drag in the performance is particularly marked for standard strategies targeting specific risk-premia. These results show how bounding the MV weights by minimizing over costs is a superior strategy economically justified by the aim of preserving the before-cost risk-return trade-off.

5.2 Cost-minimization to unlock stock premia information

We argue that the stabilizing role of cost-optimized weights allows to market time stock risk premia by smoothing out the extra noise that prevents MV standard strategies to exploit the information contained in the stock premia.

We first show that, once stock premia are estimated using at least the previous 10 years of data, unlike their less profitable standard analogs, the profitable cost-optimized strategies targeting higher returns derive a substantial portion of their profitability from timing stock risk premia (more than twice that from targeting stock covariances and substantially more as compared to the stock premia timing taking place in the analog standard strategies if any). In the process we also confirm the literature insight that the most profitable standard strategies derive most of their profitability from timing the stock covariances while suffering from estimation error contained in stock premia when their estimates are considered as inputs (as in Clarke et al. (2006, 2011), Fleming et al. (2001), Jagannathan and Ma (2003) and Moreira and Muir (2017)). Indeed we find an inverse relationship between their profitability and ability to time covariances and the level of premium they target.

Figure 7 and 8 plot performance ratios to evaluate the market timing abilities of our strategies when positive. The performance ratio of a given strategy is the ratio of the Sharpe ratio gap between our analyzed real-time strategy and its version where at least one of the first two moments of the returns distribution has been fixed to the in-sample look-ahead biased estimate divided by Sharpe ratio of the real-time version of the strategy. For a given rolling window length for V_t, six months in Figure 7 and 1 year for Figure 8, the leftmost(rightmost) graphs plot the ratios belonging to standard(cost-optimized) strategies. In particular, in the top graphs the vector of stock premia are artificially held fixed in the unimplementable version of our strategies, in the middle graphs is the covariance matrix which is kept fixed, and in the bottom plots both moments are kept fixed. Therefore the top graphs capture the abilities of our real-time analyzed strategies to time the stock risk premia, the middle graphs the abilities to time the stock covariances and the bottom graphs the abilities to jointly time the premia and the covariances. Finally the starred markers signal gap statistically different from zero at the 10% level.

Focusing on the blue and red lines in the rightmost graphs of the two figures, representing the parametrizations of cost-optimized strategies which use at least 10 years of data to estimate μ_t, we notice that: the gaps are more positive in the top graphs than in the middle ones, and the patterns in the bottom graphs mostly resemble the ones from the top graphs. The first fact tell us that most of the market timing abilities of cost-optimized strategies come from timing the stock premia. For $MVP10tc$ and $MVP15tc$ the impact of premia timing is always significant and at least twice as much as that from volatility. Moreover the ratio for $MVP10tc$ is at least 0.31 and that of $MVP15tc$ at least 0.54, that is, stock primia timing explain roughly at least one-third and half of the total performance of our real-time strategies. The fact that the patterns in the bottom graphs (measuring the impact of jointly timing premia and covariances) mostly resemble the ones from the top graphs (measuring the exclusive impact of premia timing), with performance ratios which are always statistically significant for $MVP10tc$ and $MVP15tc$, confirms that the premia timing is the predominant type of timing. The performance ratios in the bottom graphs are actually roughly the same as those in the top graphs implying that roughly the totality of the joint timing comes from timing the risk premia.

Focusing on the standard strategies (leftmost graphs in the two figures) tell us a radical different story. The impact of premia timing for standard strategies targeting higher returns when significant is mostly less then half that for the cost-optimized analogs. When V_t is estimated using the previous six months of data the only statistically positive gap is found for $MVP15$ when μ_t is estimated using the previous 20 years of data in the amount of 0.25 versus a ratio of 0.58 for $MVP15tc$. When V_t is estimated annually $MVP10$ has a ratio of 0.14 against a ratio of 0.31 for $MVP10tc$ and $MVP15$ has a ratio of 0.34 vs a ratio of 0.54 for $MVP15tc$. Moreover as reported in Figure 5 standard strategies targeting higher returns are less profitable than their cost-optimized analogs.

Finally, turning to the most profitable standard strategies, $GMVP$ and $MVP1$, notice how a large portion of their profitability comes from timing the stock covariances (at least 60% of the baseline performance for $GMVP$ mechanically accounting for all the joint impact, and at least one-third for $MVP1$ when μ_t is estimated over the previous 20 years). More generally, looking at the patterns of covariance timing among the standard strategies and considering $GMVP$ as a special case of a strategy with no target, we can see a predominant inverse relation between the timing abilities and the level of targeted premium (the average downward patter present in the leftmost middle graphs in the two figures). This suggests that the more important is the role played by the stock premia in the MV optimization the less impact come from timing the stock covariances. This decreasing pattern is roughly matched by the real-time performances of standard strategies reported in Figure 5. Therefore the bigger the role played by the stock premia the less the role played by covariance timing and the less efficient the performance. Last but not least, it is also noteworthy to see how the most profitable standard strategies except $GVGP$, namely $MVP1$ and $MVP5$, are the ones that suffer the most from risk premia estimation error (see the negative values in the top left graphs of both Figure 7 and 8. Therefore, the most profitable standard strategies derive most of their profitability from timing the stock covariances while suffering from estimation error contained in stock premia when their estimates are considered as inputs.

In summary, these results suggest that the stabilizing role of cost-optimized weights smooths out the extra noise present in the stock risk premia estimates preventing standard strategies to profitably timing them.

One caveat in this analysis is represented by the different stock universe composition under the conditional versus unconditional estimates of the analyzed moments over which we construct our gaps. In order to compute unconditional moments we require stocks that have been present in the market throughout the sample,[21] while for computing conditional premia, and conditional covariances in particular, we require a (much) shorter horizon. This means that we have more stocks at disposal when we conditionally estimates the first two returns moments. Therefore we might be picking up the joint effect of market timing and the implied stocks availability. To avoid spurious inference in Section C of the Online Appendix we compute Sharpe ratio differentials between our baseline strategies restricted to the stock universe required to compute μ_t and V_t unconditionally over the entire sample and those introduced in this subsection where we keep some of the moments fixed. We find similar results.[22]

5.3 Cost-minimization to reduce downside risk

The stabilizing role of cost minimization also helps reducing the downside risk of a MV strategy as measured via its Maximum Draw Drown (MDD) or the number of years required to recover from the worst loss. The MDD of a strategy is the maximum observed loss in value from a peak to a trough before a new peak is obtained, and it is a popular downside risk measure in the financial industry.

[21] More precisely we require stocks for which we have at least 80% of the data.

[22] We relegate such analysis to Section 7 of the Online Appendix since eliminating stocks that are not present over the entire sample make the strategies not implementable in real time.

Figure 10 makes clear how the the stabilizing role of cost-optimized weights helps reducing MDD. We compare the realized after-cost returns of $MVP15$, referred to as "Std.", and that of its cost-optimized analog $MVP15tc$, referred to as "TC", over the sample 2003-2017 when V_t and μ_t are estimated over a rolling window length of 0.5 and 10 years respectively. MDD is computed as the minimum between the running minimum of the cumulative realized return from a given date and the end of the sample ("Running mi." in the two central graphs) and the cumulative realized return itself ("Cum. ret." in the two central graphs). The rightmost graphs plot the difference over which to take the minimum for the case of $MVP15$ (top graph) and its analog $MVP15tc$ (bottom graph). The leftmost biggest graph plots the time series of realized returns for $MVP15$, in dashed red, and $MVP15tc$, in solid blue (the cumulative sum of which yields the cumulative returns). Notice how the realized returns of $MVP15tc$ are less volatile, with the blue line almost always inside the red one. This is due to the presence of the no-trading region bounding the weights of $MVP15tc$. This element of stability is crucial especially around the financial crises (end of 2008) where the two biggest negative realizations of $MVP15$ are capped (the first) and neutralized (the second) in $MVP15tc$. This results in a running minimum, the trough, of 0.49 for $MVP15tc$ as opposed to a running minimum of -0.15 for $MVP15$ as can be seen by the middle bottom and top graph respectively. Because MDDs are computed as the difference between the reported troughs and the highest peaks before such troughs, we have an MDD of $-0.15 - 0.73 = -0.88$ for $MVP15$ and $0.49 - 0.87 = -0.38$ for $MVP15tc$.

The top graph of Figure 11 reports the differential in the absolute value of MDDs between the standard and the cost-optimized strategies. Notice how the MDDs of all cost-optimized strategies are smaller (negative differentials), with differentials generally increasing in the premium level to be targeted. Almost all MDDs of cost-optimized strategies (except $MVP1tc$ when μ_t uses a rolling window estimation of 5 year, which only experience losses approximately 1% lower than its analog $MVP1$) are at least approximately 8% lower. That is, cost minimization almost always allows strategies to suffer losses in their worst case that are lower by 8% in levels.

A simple comparison of two strategy MDDs does not take into account the effect of potentially different average returns. If a strategy has a lower MDD but a higher average return it might still take less than another strategy with a higher MDD but a lower average return to recover. Hence comparing the average time (in years) a strategy takes to recover from the worst experienced loss might be a better metric. Such metric can be readily computed as the ratio of the absolute MDD over the average excess return of our strategies. This is exactly what the bottom graph of Figure 11 reports. We notice a similar pattern to that displayed in the top graph, making an even stronger case for the role of cost-optimize strategies in reducing downside risk. Cost-optimized strategies take less to recover from their worst losses. As for the absolute MDDs differentials, $MVP1$ is the closest to $MVP1tc$ when μ_t is estimated using a rolling window of 5 years. However, even in this case where the severity of losses among cost-optimized and standard strategies are almost the same, $MVP1tc$ take approximatelly one year less to recover. All other cost-optimized strategies (including $MVP1tc$ under the other analyzed rolling window length combinations for μ_t and V_t) take at least approximately 2 years less to recover with $MVP15tc$, when μ_t is estimated over the past 5 years, taking at least approximately 12 years less.

5.4 Cost-minimization to scale up

Our results so far applies to Mean-Variance investors small enough not to distort prices. This assumption is implied by the fact that we do not model any price impact on the trades we execute. However, in practice as soon as the value of our positions become big enough the price impact of our trades is no more negligible. Therefore an important question, especially for big institutional investors, concerns the scalability of our strategies. The answer to this question depends on the ability of a given strategy to handle price impact.

We find that explicitly modeling price impact as a form of (quadratic) transaction cost makes a difference. Our cost-optimized strategies consistently scale up more and are less sensitive to changes in price impact than our standard strategies. In particular, our results suggest that cost-optimized strategies are profitable options to gain exposure on the market even for large institutional investors.

We model price impact of strategy j at time t as the quadratic cost

$$\frac{1}{2}\pi(\theta_t^j - \theta_t^{0,j})'(\theta_t^j - \theta_t^{0,j}).$$

The parameter π captures the average cross-sectional level of price impact in the market.

Specifically, we should think of π to be inversely related to the liquidity and depth of the market and positively related to the portfolio size of the MV investor. Under our specification if we increment(decrement) the position in stock i with respect to strategy j by $\theta^j_{t,i} - \theta^{0,j}_{t,i}$, the price of stock i increases(decreases) by $\frac{1}{2}\pi(\theta^j_{t,i} - \theta^{0,j}_{t,i})$ cents per dollar traded. Equivalently, we suffer a reduction of $\frac{1}{2}\pi(\theta^j_{t,i} - \theta^{0,j}_{t,i})^2$ in the return of stock i. Quadratic costs are a popular way to model price impact (see for example Dybvig and Pezzo (2020) and discussion therein) and nicely fit our quadratic mean-variance framework.[23]

We could be more general and allow for a stock-specific and time-varying price impact, however estimating these parameters is very hard in practice. For this reason in this paper we follow Maurer et al. (2020) and limit ourselves to a sensitivity analysis where we let π varies over a fixed interval $[0, \overline{\pi}]$. To get sensible estimates for $\overline{\pi}$ we linearize the square-root price impact function derived in Figure 2 of Frazzini et al. (2015) and make use of the size of the trades, $\{\Delta^j_{t,i}\}$, of our analyzed strategies j in the absence of price impact.[24] Frazzini et al. (2015) compute their estimates using the 2000 biggest (by market cap) and most traded (in

[23]While our specification for the price impact vector $\frac{1}{2}\pi(\theta^j_t - \theta^{0,j}_t)$ is linear in the size of the trades $\theta^j_t - \theta^{0,j}_t$, another popular alternative assumes price impact to growth at a square root rate with the size of the trades as in Frazzini et al. (2015). In such specification, the price impact vector is $\frac{1}{2}\pi\sqrt{|\theta^j_t - \theta^{0,j}_t|}$. Under such specification we lose the computational scalability of quadratic programs, therefore executing our algorithms for thousands of stocks become computationally unfeasible.

[24]Frazzini et al. (2015) provide conservative estimates for the price impacts of many stocks in the U.S. market (and internationally) from over a trillion dollars of live trading proprietary data from a large institutional money manager in the monthly period 1998-2013. In particular, they plot the average price impact function (in basis points), which in our setup translates to $\frac{1}{2}\pi\Delta$ where $\Delta = |\bar{\theta} - \bar{\theta}^0|$ represents the absolute average size of our strategies' trades, as a function of the percentage daily traded volume, which in our setup translates to $\frac{AUM\Delta}{\bar{V}} \times 100$ where AUM is the dollar value of the assets under management and $\bar{V}$ the average dollar volume traded.

Concentrating on the sub-sample 2003-2017 for a given strategy j and AUM we construct the time series of median daily treading volume for each common stock and we use it to get the best linear approximation that fits our quadratic cost function. We do so by recovering the slope, $\hat{\beta}_{AUM,j}$, of a line going through the origin obtained by vectorizing the $T \times N$ matrix of median daily trading volume and regressing it on the square-root price impact function from Figure 2 of Frazzini et al. (2015). We then obtain the per-dollar linear price impact estimate for stock i at time t as $\hat{\beta}_{AUM,j}\frac{AUM\Delta^j_{t,i}}{V_{t,i}}100$. The time-stock specific price impact parameter $\pi^{AUM,j}_{t,i}$ is then recovered by equating such estimate to our definition of per-dollar price impact $\frac{1}{2}\pi^{AUM,j}_{t,i}\Delta_{t,i}$.

Finally to get $\overline{\pi}$: we first compute at each time t for a given strategy j and AUM, the cross-sectional 90th quantile of the distribution of $\pi^{AUM,j}_{t,i}$. Second, for a given strategy j and AUM, we compute the time-series median of such quantiles. Third, for a given AUM we compute the median of such values across strategies. Fourth, we take the mean across AUM ranging from 1 million through 10000 billion USD.

terms of dollar volume) stocks. In the same spirit, we compute $\overline{\pi}$ retaining the intersection of the 2000 biggest and 2000 most traded stocks (for a total of approximately 1800). We obtain a monotonically increasing upper-bound $\overline{\pi}$ as a function of portfolio sizes ranging from 1 million to 10000 billions USD. In particular for sizes bigger than 10 billions $\overline{\pi}$ caps at a value of 3.42. We therefore present our results for the range $\pi \in [0, 3.4]$

Figure 12 plots the annualized out-of-sample after-cost Sharpe ratios of our standard strategies as a function of the price impact of trades captured by the parameter π. Starred markers highlight ratios higher than those of EW and VW at the 10% level. Every plot is different in that the vector of conditional risk premia μ_t and the covariance matrix V_t are estimated over different combinations of rolling window lengths (expressed in years and reported as titles of each graph).

Notice first how the equally and value weighted portfolios, EW and VW, are insensitive to price impact. Such strategies are not function of μ_t and V_t, therefore their ratios are the same in every graph (black horizontal lines). Focusing on a given graph we see how their ratios remain virtually constant as a function of π. Hence, EW and VW appear to be very scalable.

Looking at standard MV strategies tell us a different story. Their Sharpe ratios are strictly decreasing in the level of price impact and this effect is particularly pronounced when we estimate V_t over the previous six months (see all colored lines). As for the case of $\pi = 0$, $GMVP$ (purple line) remains one among the most profitable standard strategies. While its performance quickly deteriorates when V_t is estimated using the previous 6 months of data (its Sharpe ratio ceases to be statistically better than those of EW and VW for $\pi > 0.4$), it is interesting to notice how it can scale up when we estimate V_t using the previous year of data. Therefore, while standard strategies in general suffer from the adverse effect of price pressure, $GMVP$ can actually scale up if properly parametrized.

Similarly to Figure 12, Figure 13 plots the performance of our cost-optimized strategies as a function of π. Notice now how all the patterns are flatter and no strategy turn unprofitable.[25] More importantly ,results are more robust: there is now not much difference with respect to how we estimate the covariance matrix (whether we use the past 6 or 12

[25]In contrast, for the case of standard strategies the Sharpe ratios of $MVP10$ and $MVP15$ are all negative for $\pi > 0.5$.

months), and given risk premia are estimated over at least 10 years of data all Sharpe ratios remain well above 1 (see the bottom 4 graphs). In particular, the Sharpe ratios of strategies targeting higher premia, i.e. $MVP5tc$ through $MVP15tc$, are always statistically better than EW and VW. Hence, $MVP5tc$ through $MVP15tc$ appear to be profitable strategies to gain exposure on the U.S. stock market even for large institutional investors.

Finally, Figure 14 plots the difference in Sharpe ratios between our cost-optimized and standard MV strategies as a function of π. In general cost-optimized strategies scale up substantially better then their analogs, and they do more so as the severity of price impact increases. Except for $GMVP$ and $MVP1$ that when V_t is estimated annually for low values of π are actually less profitable but not statistically different from the cost-optimied ones, cost-optimized strategies have Sharpe ratios of at least 0.14 higher. Moreover the displayed Sharpe ratio differentials are mostly monotonically increasing in π no matter the length of the rolling window (6 versus 12 months) used to estimated V_t.

5.5 Rationalizing the higher efficiency in targeting risk premia

The fact that cost-optimized strategies targeting higher return are more profitable than their standard (less constrained) analogs can be rationalized by the Frazzini and Pedersen (2014) "betting against beta" theory.

The theory original aim is to explain why empirically the capital market line is found much flatter than it should be under the CAPM. In an economy where different MV investors have different trading constraints (in the form of margin requirements) more constrained investors will optimally choose portfolios with higher betas than otherwise optimal to make up for their impossibility to lever. In particular, this results in an inverse relation between market alphas and betas, with higher alphas and lower betas the more investors are constrained. Also efficient portfolios have maximal Sharpe ratios in correspondence of betas less then one, decreasing(increasing) for higher(lower) betas.

Figure 9 plots the relationship between alphas and betas of our cost-optimized (labeled as TC) and standard (labeled as *Std*) MV strategies under the CAPM (leftmost graphs) and Fama and French (1993) 3 factor representation (rightmost graphs). As usual by now, starred markers refer to figure statistically significant at least at the 10% level.

Under both representations cost-optimized strategies have higher alphas (top graphs) and lower betas (bottom graph). Moreover, the differences in alphas(betas) are monotonically increasing(decreasing) in the level of targeted risk premium, and almost always statistically significant for higher levels of targeted premia once the stock risk premia are estimated over at least the previous 10 years (see all but the magenta and yellow lines). Therefore, constraining the weights by minimizing over costs produces higher alphas and lower betas. Moreover, such pattern is more pronounced the more demanding (i.e. the higher) is the constraint on the minimum risk premium to achieve.

The middle graphs in Figure 9 plots the market betas of our cost-optimized strategies, already shown to be smaller that those of their standard analogs. Starred markets report the two-sided test against the alternative that betas are different from one. As before once we focus on stock risk premia estimates that use at least the past 10 years of data (all but the magenta and yellow lines), all betas are statistically smaller then one. Once again, this pattern is line with the "betting against beta" theory since as reported in Figure 5 such strategies have high Sharpe ratios (almost always above 1.2).

In summary, since cost-minimization is a form of trading constraint imposed on the MV weights, we can use the Frazzini and Pedersen (2014) "betting against beta" theory to justify our results.

6 Market timing and estimation error: drivers of mean-variance profitability

In this section we argue that the market timing abilities of our MV strategies, and the estimation error negatively influencing their performance, are important factors behind the profitability of mean-variance strategies over the last 100 years in the U.S. stock market.

Market timing and estimation error have substantial impact on the performance of MV strategies, with average Sharpe ratio increments (for the case of market timing) and decrements (for the case of estimation error) between 0.3 and 0.45, as well as 0.26 and 0.66 respectively. MV profitability displays a pro-cyclical pattern with magnitudes increasing in expansions and decreasing in recessions over time. Such dynamics can be explained by market timing in expansions and estimation error in recessions. We find substantial positive correlation between the profitability of MV strategies and their level of market timing (with low values associated to estimation error), particularly for cost-optimized strategies. The average correlation over the last 100 years is between 0.28 and 0.56 for standard strategies and between 0.5 and 0.82 for cost-optimized strategies. The majority of most profitable strategies have market timing levels belonging to the top tercile, almost always corresponding to positive gaps (our way to detect actual market timing activity). At the same time, most of the least profitable strategies have market timing levels not higher then the median, almost always corresponding to negative gaps (our way to detect estimation error).

6.1 Magnitudes of market timing and estimation error

In this subsection we quantify the impact of market timing and estimation error by looking at the average *magnitudes* across strategies and $\mu_t - V_t$ rolling-window-length configurations. We find that market timing induces average annualized after-cost out-of-sample Sharpe ratios increments between 0.3 and 0.45 during expansions, while estimation error causes average Sharpe ratios reductions between 0.26 and 0.66 throughout the business cycle. Therefore we conclude that market timing and estimation error have substantial impact on the performance of MV strategies over the last 100 years.

Table 2 displays the dynamics of the average magnitudes of market timing and estimation error for our MV strategies over the last 100 years. Market timing is defined as the average of the differentials, when positive at least at the 10% level, between the Sharpe ratio of our strategies in the baseline setup versus their analogous version where: the vector of stock risk premia is kept fixed at the in-sample average (column label "$\bar{\mu}$"), the covariance matrix of the stocks is kept fixed at the in-sample average (column label "$\bar{V}$"), both moments of the stock return distributions are kept fixed at the in-sample average (column label "$\bar{\mu}$ & $\bar{V}$"). Such differentials are illustrated in column 2 through column 4 and their average across types of market timing (from risk premia, covariances or both respectively) is reported in column 5. Similarly estimation error is defined as the average of the differentials when negative at least at the 10% level. Such differentials are reported in column 6 through column 8 and their average across types of estimation error (from risk premia, covariances or both respectively) is illustrated in column 9. Specifically, a generic numerical entry in the table corresponds to the average of the differentials over all the relevant strategies and across their baseline $\mu_t - V_t$ rolling-window-length configurations. For instance the value 0.2 in the top-left of PANEL A is the average of all the significantly positive differentials between our standard mean-variance strategies and their fixed risk premia analogs under the six different $\mu_t - V_t$ rolling-window-length configurations for the baseline strategies (namely 5-0.5,10-0.5,20-0.5,5-1,10-1,20-1 years).

Market timing induces average annualized after-cost out-of-sample Sharpe ratios increments of 0.3 and 0.45 in the two expansionary sub-sample: 45-72 and 03-17 respectively.[26] Cost-optimized strategies are on average better at market timing with average magnitudes always higher in the aggregate and by type of market timing. In particular, the overall average market timing of cost optimized(standard) strategies is 0.35(0.2) in the sample 45-72 and 0.51(0.38) in the most recent sub-sample. Moreover, the only instance in which a higher Sharpe ratio differential for standard strategies is spotted is in the sub-sample 03-17 with respect to the time-varying management of the stock covariances (column "$\bar{V}$"): PANEL A and B show an increment of 0.51, or 0.53 if we exclude $MVTP$ from the average, against an improvement of 0.33 reported in PANEL C. A finding induced by the documented superior ability of $GMVP$ in market timing stock covariances recently.

If market timing is an important source of profit, estimation error is a well-known serious problem for MV strategies. In contrast to market timing, we detect estimation error throughout the business cycle, i.e. in all four sub-sample 26-44,45-72,73-02 and 03-17.[27] It ranges between Sharpe ratio reductions of 0.26 through 0.66 and it is comparable in magnitude between standard and cost-optimized strategies.[28]

[26]Recall from Section 3 that these sub-samples are characterized by low(er) trading frictions and high(er) market growth.

[27]Recall from Section 3 that while the sub-samples 45-72 and 03-17 are (mostly) expansionary, the sub-samples 26-44 and 73-02 are (mostly) recessionary in that they are associated with high(er) levels of trading frictions and low(er) market growth.

[28]Interestingly, if anything estimation error appears slightly bigger for cost-optimized strategies. This is because using the average within all standard and within all cost-optimized strategies as the statistic to compare this two class of strategies is misleading. In fact the purpose of this subsection is not such comparison rather the focus is on the stand-alone average magnitudes. The reason why the average estimation error is misleading in this context is because it only takes into account the magnitude but not the frequency. For example, if we compare the average estimation error coming from the stock risk premia in the recent sub-sample between standard strategies, -0.49 (column 6, labeled "$\bar{\mu}$", in PANEL A and B), and cost-optimized ones, -0.55 (column 6, labeled "$\bar{\mu}$", in PANEL C), we notice how the latter is bigger (even if slightly). If we take a look at the performance ratios (which are the Sharpe ratio differentials divided by the real-time

6.2 Pro-cyclical increasing profitability

Next, we look at the dynamics of the MV profitability. We conclude that MV profitability displays a pro-cyclical pattern with magnitudes increasing in expansions and decreasing in recessions over time. Such ex-post patterns is correctly anticipated by the ex-ante dynamics of the MV investment opportunities as we discussed in Section 3.1.

Figure 3 not only shows the recent stark profitability of MV strategies during the mostly expansionary sub-sample 2003-2017 (green lines), it also shows how the second best overall performance is achieved during the other mostly expansionary sub-period in our overall sample, the sub-sample 1945-1972 (magenta lines). Analogously if we look at the most recessionary periods in our sample, i.e. the sub-sample 1972-2003 (red lines) and sub-sample 1926-1944 (blue lines), we notice how their profitability is the lowest over the entire sample (with that of 1972-2003 being the lowest over the past 100 years).

6.3 Pro-cyclical increasing market timing and counter-cyclical estimation error

As for the case of profitability, we find for the market timing abilities displayed by our analyzed strategies a pro-cyclical pattern with magnitudes in expansions increasing over time. Moreover, estimation error is found counter-cyclical with magnitudes in recessions decreasing over time. Therefore market timing and estimation error together can explain the dynamics of MV profitability

Table 3, in a fashion very similar to Table 2, displays the dynamics of market timing and estimation error for our MV strategies over the last 100 years. The only difference is that figures are sums of Sharpe ratio differentials (statistically significant at the 10% level) rather than averages. This enable us to focus on the *total amount* of detected market timing and estimation error rather than the average magnitude.

Sharpe ratio of the respective strategies) in the top graphs of Figure 7 and 8, we notice how the -0.55 corresponds to the unique instance of estimation error detected at the 10% level when V_t in the baseline strategy is estimated using the past 6 months (the only negative starred marker in the top right graph of Figure 7). In contrast, -0.49 is the average across several statistically negative differentials corresponding to standard strategies estimated when V_t in the baseline strategy uses the past 6 or 12 months (the negative starred markers in the top left graphs in Figure 7 and 8 respectively).

Notice how, no matter the type of strategy (standard versus cost-optimized), market timing is virtually absent in the recessionary sub-samples 26-44 and 73-02 and is generally present in the remaining two expansionary samples 45-72 and 03-17. In particular, evidence of market timing of any type (from risk premia, covariances or both) are present in the most recent sub-sample and are much higher in magnitudes than those present, either exclusively from the stock risk premia or the covariances, in the 45-72 sub-period. Hence the pattern of market timing in our strategies is pro-cyclical with magnitudes in expansions increasing over time.

Analogously inspecting the sum of the gaps induced by estimation error reveils much bigger impacts (i.e. more negative gaps) from the recessionary periods with magnitudes in recessions decreasing over time. Hence the patter of estimation error, no matter the type of strategy (standard versus cost-optimized), is counter-cyclical with magnitudes in recessionary sub-samples decreasing over time.

6.4 Correlation analysis

The results of this section so far are suggestive of a positive correlation between the profitability of mean-variance strategies and the level of market timing, measured via the differentials between the baseline real-time versions of our strategies and their unimplementable versions where at least one of the first two moments of the return distribution is fixed (with high, or positive, values capturing actual market timing and low, or negative, values actual estimation error).

Table 4 and 5 report such correlation for our standard and cost-optimized strategies respectively over the 4 sub-samples, 26-44,45-72,73-02 and 03-17, covering the last 100 years. In particular, for a given sub-sample each row fixes the length of the rolling window used to estimate the conditional stock risk premia in the baseline setup. Column 4 and 5 look at the differential where the unimplementable strategy versions have fixed in-sample stock risk premia $\bar{\mu}$, thus isolating the time-variation impact of the stock risk-premia estimates. In column 4 the length of the rolling window for the conditional covariance matrix of returns is fixed at 6 months, while in column 5 it is fixed at 1 year. Similarly column 6 and 7 look at the differentials where the unimplementable strategy versions have fixed in sample covariance matrix $\bar{V}$, and column 8 and 9 at the differentials where the unimplementable strategy versions have both moment fixed, i.e. $\bar{\mu}$ and $\bar{V}$.

We find substantial positive correlation between the profitability of MV strategies and their level of market timing , particularly for cost-optimized strategies. The average correlation over the last 100 years is between 0.28 and 0.56 for standard strategies and between 0.5 and 0.82 for cost-optimized strategies. The higher correlation for cost-optimized strategies mostly come from the 1973-2002 sample, and the recent sample with respect to the market timing of stock risk premia. In particular, the inabilities of standard strategies to time risk premia, which we discussed in Section 5, causes the correlation to turn negative.

Table 6 and 7 analyze the relationship between the level of market timing and the Sharpe ratios of the most and least profitable standard and cost-optimized strategies respectively. Given a specific sub-sample, the upper(lower) part of the panel reports the baseline most(least) profitable strategy (column 5/9 when the rolling window length for the conditional covariance is 6 months/1 year) in correspondence of each rolling window length for the vector of conditional stock risk premia (different rows). Column 6 and 10 display the percentile of the market timing level distribution associated with the differentials computed when the unimplementable strategies have their vector of stock risk premia kept fixed at their in-sample level $\bar{\mu}$. Similarly column 7 and 11 (column 8 and 12) report the differentials with respect to unimplementable strategies which conditional covariance matrix (and risk premia vector) is (are) kept fixed at $\bar{V}$ (& $\bar{\mu}$). Finally, whenever the reported percentile is associated with a positive(negative) differential we mark the figure with a $^{+}(^{-})$.

The majority of most profitable strategies have market timing levels belonging to the top tercile, almost always (half the time in the sample 1926-1944) corresponding to positive gaps (our way to detect actual market timing activity). In particular, 71%(53%) of the gaps in the sub-sample 03-17 are in the top tercile for standard(cost-optimized) strategies, of which 83%(88%) are positive. Similarly, the proportions for the sub-sample 73-02 are 59%(80%), of which 80%(100%) are positive. The proportions for the sub-sample 45-72 are 50%(75%), of which 100%(100%). The proportions for the sub-sample 26-44 are 100%(100%), of which 55%(50%) are positive.

At the same time (except for standard strategies recently, see motivation below), most of the least profitable strategies have market timing levels not higher then the median, almost always corresponding to negative differentials (our way to detect estimation error).

In particular, 33%(61%) of the differentials in the sub-sample 03-17 are not higher then the median for standard(cost-optimized) strategies, of which 83%(82%) are negative. Similarly, the proportions for the sub-sample 73-02 are 78%(61%), of which 100%(100%) are negative. The proportions for the sub-sample 45-72 are 94%(83%), of which 76%(73%) are negative. The proportions for the sub-sample 26-44 are 75%(83%), of which 90%(100%) are negative.

The least performing standard strategies in the recent 2003-2017 sample causing the exception in the trend found above – i.e. $MVTP$ when μ_t is estimated at least over the past 10 years and V_t over the past 6 months, as well as $MVP1$ when μ_t uses the past 10 years and V_t the past year, and $GMVP$ when V_t uses the past 1 year – have Sharpe ratios never lower than 0.63, which is very similar to those obtainable by investing in the broad market ETFs (i.e. 0.69 for SPX and 0.62 for IWM). The fact that those strategies have positive differentials with market timing percentiles between 50 and 86 is perfectly in line with the positive correlation between market timing and MV profitability.

7 Robustness

We check the robustness of our findings in several ways. Specifically, we check whether or not they are robust to the way we measure costs (with or without TAQ data), how we estimate the risk premia (with a market model or a 3-factor model a la Fama and French (1993) and the standard Fama and MacBeth (1973) versus the noise filtering approach of Gagliardini et al. (2016)) and the covariance matrix (with factor models or via the Ledoit and Wolf (2017, 2020) shrinkage estimator), as well as to the specific re-balancing frequency (monthly vs. quarterly) and the composition of the stock universe (whether or not we reduce the stock universe to: i) the S&P 500 constituents, ii) the same sub-sample of stocks implicitly selected by the most stringent analyzed rolling-window-length combination of 12 months for V_t and 20 years for μ_t, or iii) the even more stringent sub-sample of stocks required to estimate V and μ in-sample).

Our results are found in general robust. We constructively learn that: 1) the trade-off between bias and estimation error in the covariance matrix estimator is clearly in favor of a reduction of the latter,[29] 2) adopting a parsimonious specification for the return generating

[29]Because the performance of our strategies are much better when we estimate V_t via factor models rather

process and filtering out part of the noise in the risk premia estimates through the Gagliardini et al. (2016) approach is the best way, among those analyzed in this paper, to handle stock risk premia, [30] 3) there are economically sizable diversification benefits in investing in all common stocks.[31]

In particular, in Section B we confirm that: i) the profitability of MV strategies has never been as high as recently, ii) $GMVP$ remains among the most profitable textbook strategies for small investors (in the absence of price impact), and iii) the cost-optimized strategies remain the best way to efficiently target (higher) risk premia once stock risk premia are estimated using at least 10 years of data.

In Section 5.2 we have argued that the stabilizing role of cost-optimized weights allows to market time stock risk premia by smoothing out the extra noise that prevents MV standard strategies to exploit the information contained in the stock premia. In Section C we confirm the robustness of our claim to the evaluated alternative scenarios with the exception of the scenario where our strategies are implemented quarterly. In such scenario our cost-optimized strategies still outperform but no market timing is detected. We ascribe this lack of findings to the small sample size. We also notice how allowing for a richer model for the risk premia and the covariance structure (specifically the Fama and French (1993) 3-factor model versus the single-factor market model) allows our cost-optimized strategies to additionally engage in covariance timing, while preserving their ability to time risk premia (coming from the stabilizing role played by the cost minimization).

In Section D we confirm how cost minimization is also effective in reducing downside risk as measured via the Maximum Draw Down (MDD), or the average time needed to recover from the worst incurred loss, in virtually all the analyzed alternative scenarios.

In the main text we showed how cost-optimized mean variance strategies are very profitable options to gain exposure on the market even for large institutional investors. Due to the computationally intensive nature of the analysis, in Section E we extend this analysis and confirm such insight to a sub-sample of alternative scenarios. Specifically we focus on

then Shrinkage estimators.

[30]The one-factor market model within the Gagliardini et al. (2016) is more efficient than the Fama and French (1993) 3-factor formulation or the single market factor model under the standard Fama and MacBeth (1973) procedure.

[31]Performances are lower if we restrict the stock universe to the $SP500$ stocks.

the scenarios where: costs are estimated without using TAQ data, stocks are restricted to be $SP500$ constituents, and our strategies are rebalanced quarterly. As in the baseline scenario, we find a substantial performance gap between cost-optimized strategies and standard analogs, which is monotonically increasing in the price pressure severity (or equivalently in the dollar size of the strategies).

Finally in Section F we confirm the main finding that market timing and estimation error have non-negligible impacts on the mean-variance performance and share very similar dynamics over approximately the last 100 years.

All the details are relegated to the Online Appendices.

8 Conclusions

We have analyzed the performance of large-scale Mean-Variance (MV) strategies, invested in all common stocks on the NYSE/AMEX and NASDAQ, over approximately the last century.

In sharp contrast with investing in anomalies, we documented how profitable is to gain exposure on the market via the MV criterion recently. Of particular interest is the insight that the stabilizing role of cost minimization allows to restore the credibility of the capability of MV strategies to efficiently target risk premia. As a matter of fact cost minimization allows to: i) profitably target risk premia (unlike their standard analogs), ii) reduce downside risk (measured by the worst observed loss or the average time to recover from it) and, iii) enhance the scalability of MV strategies (by directly modeling stock-specific price impact as a quadratic trading cost) making such strategies appealing options for large investors. Comparing our results with those found in Maurer et al. (2020) for the FX market highlights the different role played by cost minimization in MV strategies across markets. In the stock market where estimation error is a first order concern, cost minimization enhances the performance mostly by stabilizing the weights. In the FX markets where estimation error in the mean returns is much less problematic, cost minimization mostly enhances the performance via direct execution cost reduction.

Moreover, we argued that market timing and estimation error are two important factors behind the profitability of our analyzed strategies over the last 100 years. They account for substantial portions of the actual out-of-sample after-cost profitability and can explain its long term cyclical dynamics.

Finally, our results are found robust to several alternative scenarios involving different: ways of estimating costs, model parameters, stock universe and rebalancing frequency.

References

Abdi, Farshid, and Angelo Ranaldo, 2017, A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices, *The Review of Financial Studies* 30, 4437–4480.

Angel, J. James, Lawrence E. Harris, and Chester S. Spatt, 2015, Equity Trading in the 21st Century: An Update, *Quarterly Journal of Finance* 5.

Berrada, T. N., and S. Coupy, 2015, Outperforming Naive Diversification Using Stock Level Information., Working paper.

Chen, Andrew Y., and Mihail Velikov, 2020, Accounting for the Anomaly Zoo: a Trading Cost Perspective, Working paper.

Chordia, T. R., R. Roll, and A. Subrahmanyam, 2011, Recent trends in trading activity and market quality, *Journal of Financial Economics* 101, 243–263.

Chordia, T. R., A. Subrahmanyam, and Q. Tong, 2014, Have Capital Market Anomalies Attenuated in the Recent Era of High Liquidity and Trading Activity?, *Journal of Accounting and Economics* 58, 41–58.

Clarke, Roger, Harindra de Silva, and Steven Thorley, 2006, Minimum-Variance Portfolios in the U.S. Equity Market, *Journal of Portfolio Management* 33, 10–24.

Clarke, Roger, Harindra de Silva, and Steven Thorley, 2011, Minimum-Variance Portfolio Composition, *Journal of Portfolio Management* 37, 31–45.

Corwin, Shane A., and Paul Schultz, 2012, A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices, *Journal of Finance* 67, 719–759.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009, Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.

Dybvig, Phillip H., and Luca Pezzo, 2020, Mean-Variance Porfolio Rebalancing with Transaction Costs, Working paper, Washington University and University of New Orleans.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607–636.

Fleming, Jeff, Chris Kirby, and Barbara Ostdiek, 2001, The Economic Value of Volatility Timing, *Journal of Finance* 56, 329–352.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2015, Trading Costs of Asset Pricing Anomalies, Working paper, AQR Capital Management, LLC and Yale University.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

Gagliardini, Patrick, Elisa Ossola, and Oliver Scaillet, 2016, Time-Varying Risk Premium in Large Cross-Sectional Equity Data Sets, *Econometrica* 84, 985–1046.

Hasbrouck, Joel, 2009, Trading costs and returns for US equities: Estimating effectivecosts from daily data, *The Journal of Finance* 64, 1445–1477.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does Algorithmic Trading Improve Liquidity?, *Journal of Finance* 66, 1–33.

Jagannathan, R., and T. Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1683.

Kyle, Albert, and Anna Obizhaev, 2016, Market microstructure invariance: Empirical hypotheses, *Econometrica* 84, 1345–1404.

Ledoit, Olivier, and Michael Wolf, 2008, Robust Performance Hypothesis Testing with the Sharpe Ratio, *Journal of Empirical Finance* 15, 850–859.

Ledoit, Olivier, and Michael Wolf, 2017, Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks, *Review of Financial Studies* 30, 4349–4388.

Ledoit, Olivier, and Michael Wolf, 2020, Analytical Nonlinear Shrinkage of Large-Dimensional Covariance Matrices, *Annals of Statistics* 48, 3043–3065.

Ludvigson, Sydeny C., Dan Greenwald, and Martin Lettau, 2020, How the Wealth Was Won: Factors Shares as Market Fundamentals, Working paper.

Markowitz, Harry, 1952, Portfolio Selection, *Journal of Finance* 7, 77–91.

Maurer, Thomas A., Luca Pezzo, and Mark P. Taylor, 2020, Importance of Transaction Costs for Asset Allocations in FX Markets, Working paper, Washington University.

Moreira, Alan, and Tyler Muir, 2017, Volatility Managed Portfolios, *Journal of Finance* 72, 1611–1644.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Newey, Whitney K., and Kenneth D. West, 1994, Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies* 61, 631–653.

Novy-Marx, Robert, and Mihail Velikov, 2016, A Taxonomy of Anomalies and Their Trading Costs, *Review of Financial Studies* 29, 104–147.

Olivares-Nadal, Alba V., and Victor DeMiguel, 2016, A Robust Perspective on Transaction Costs in Portfolio Optimization, *Operation Research* 66, 733–739.

Schwert, William G., 1990, Indexes of u.s. stock prices from 1802 to 1987, *Journal of Business* 63, 399–426.

The U.S. stock market investment opportunity set

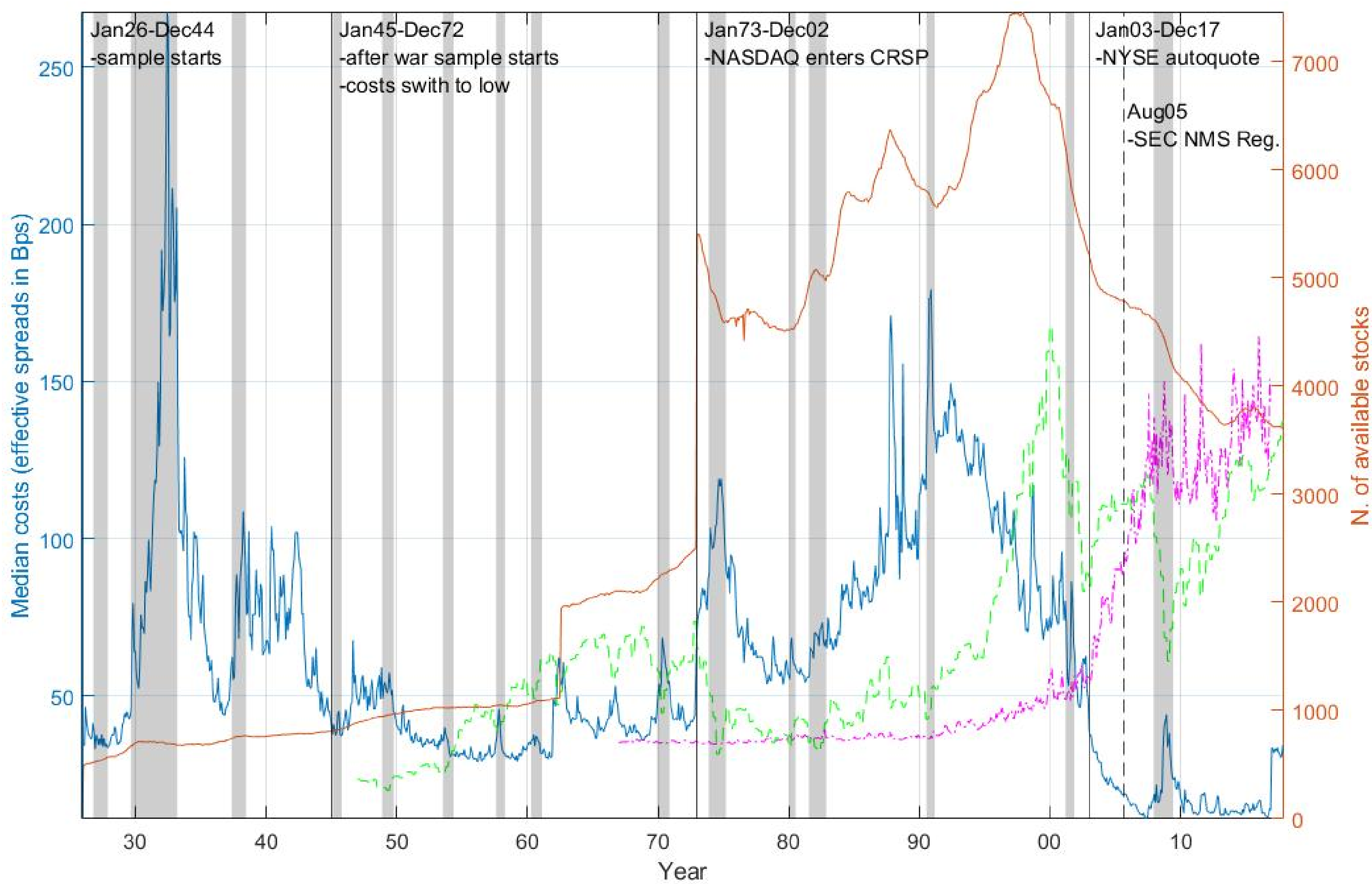

Figure 1: Two quantities that any investor need to consider while forming a portfolio strategy, the number of available stocks and their trading costs, can define the U.S. stock market investment opportunity set highlighting its cyclicality. The figure plots the time series of the Chen and Velikov (2020) cross-sectional median effective spread (implied half bid-ask spread, blue line) and the total number of CRSP common stocks (with "shrcd" first two digits equal to 10 and 11 and corresponding to the common stocks in the NYSE/AMEX and NASDAQ) over the monthly sample period January 1926 - December 2017. The dashed green line represents the ratio of the total U.S. market capitalization over the U.S. Gross Domestic Product as in Ludvigson et al. (2020), while the dot-dashed magenta line displays the cross sectional median of the monthly median daily number of trades as reported in CRSP. The structural brakes in the trading costs and the number of stocks define four subsample: the (mostly) recessionary periods 1926-44 and 1973-02 and the (mostly) expansionary period 1945-72 and 2003-17. The vertical dashed black line displays the date, August 2005, the SEC NMS Regulation (opening the gate to competition among market venues) became effective.

The U.S. Mean-Variance investment opportunity set

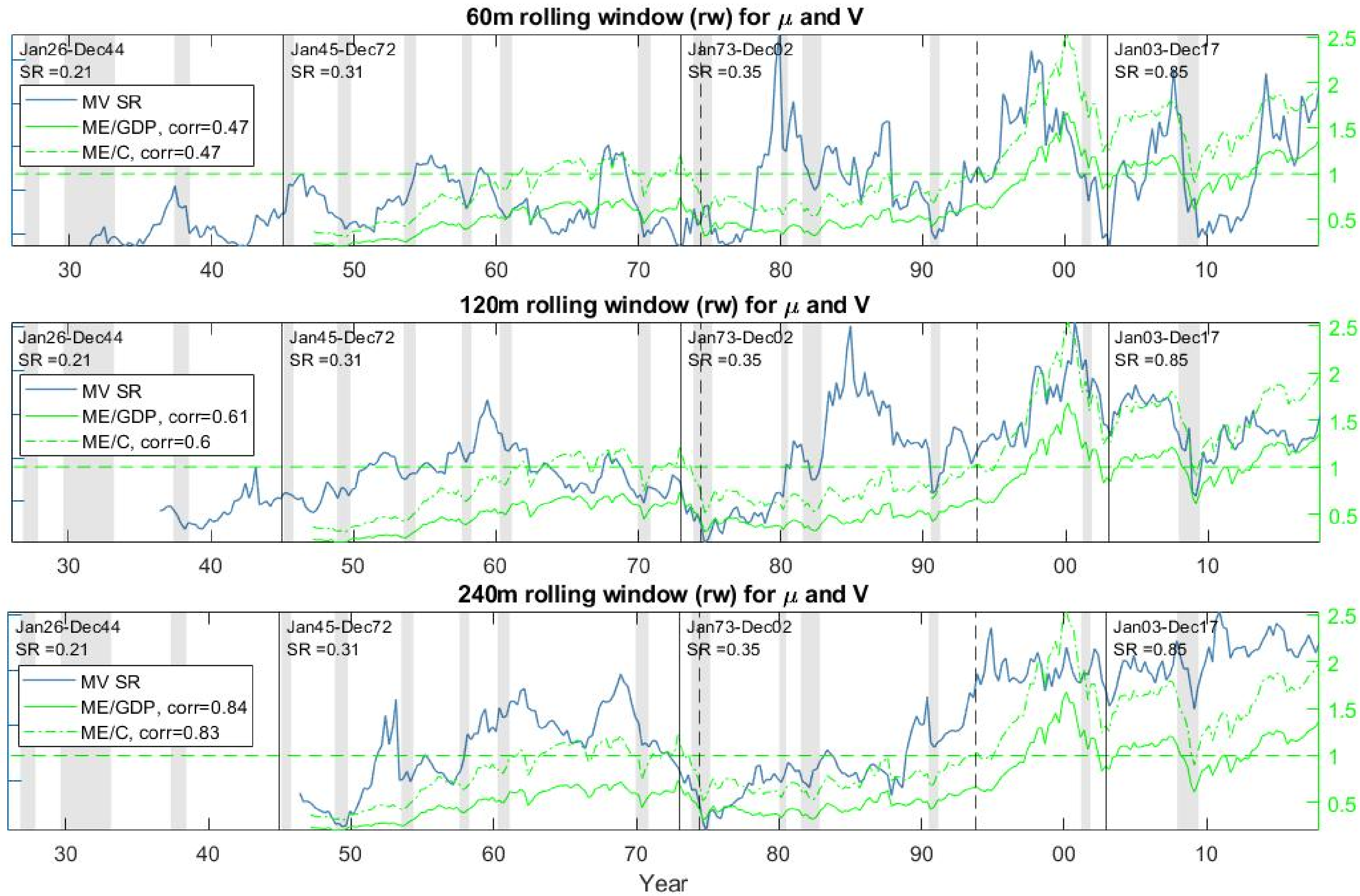

Figure 2: The slope of the Capital Allocation Line, the highest achievable Sharpe ratio, is a sufficient statistic to describe the Mean-Variance (MV) investment opportunity set. The figure shows how the U.S. MV opportunities over the monthly period January 1926 - December 2017 are volatile, pro-cyclical and increasing over time. The time-series of *ex-ante* MV Sharpe ratios $SR_t = \sqrt{\mu_t V_t^{-1} \mu_t}$ are plotted in blue and displayed on the left y-axes for different rolling windows used to estimate the vector of risk premia μ_t and the covariance matrix of returns V_t (previous 60,120 and 240 months in the top,middle and bottom graphs respectively). The *ex-post* MV Sharpe ratios, the fixed in-sample estimate of SR_t, are printed underneath each sub-sample date. The time series of the U.S. market growth relative to the economy are also super-imposed in the same graphs (green lines, right y-axes) and their correlation with the ex-ante MV Sharpe ratios reported in the legends. The solid lines represent the ratio between the U.S. market cap. and the GDP, while the dot-dashed series display the ratio between the U.S. market cap and the U.S. total personal consumption expenditures. Finally, the horizontal dashed green lines enable to quickly locate periods where market growth is faster/slower that that of the economy (when the solid and dot-dashed green time series are above/below it).

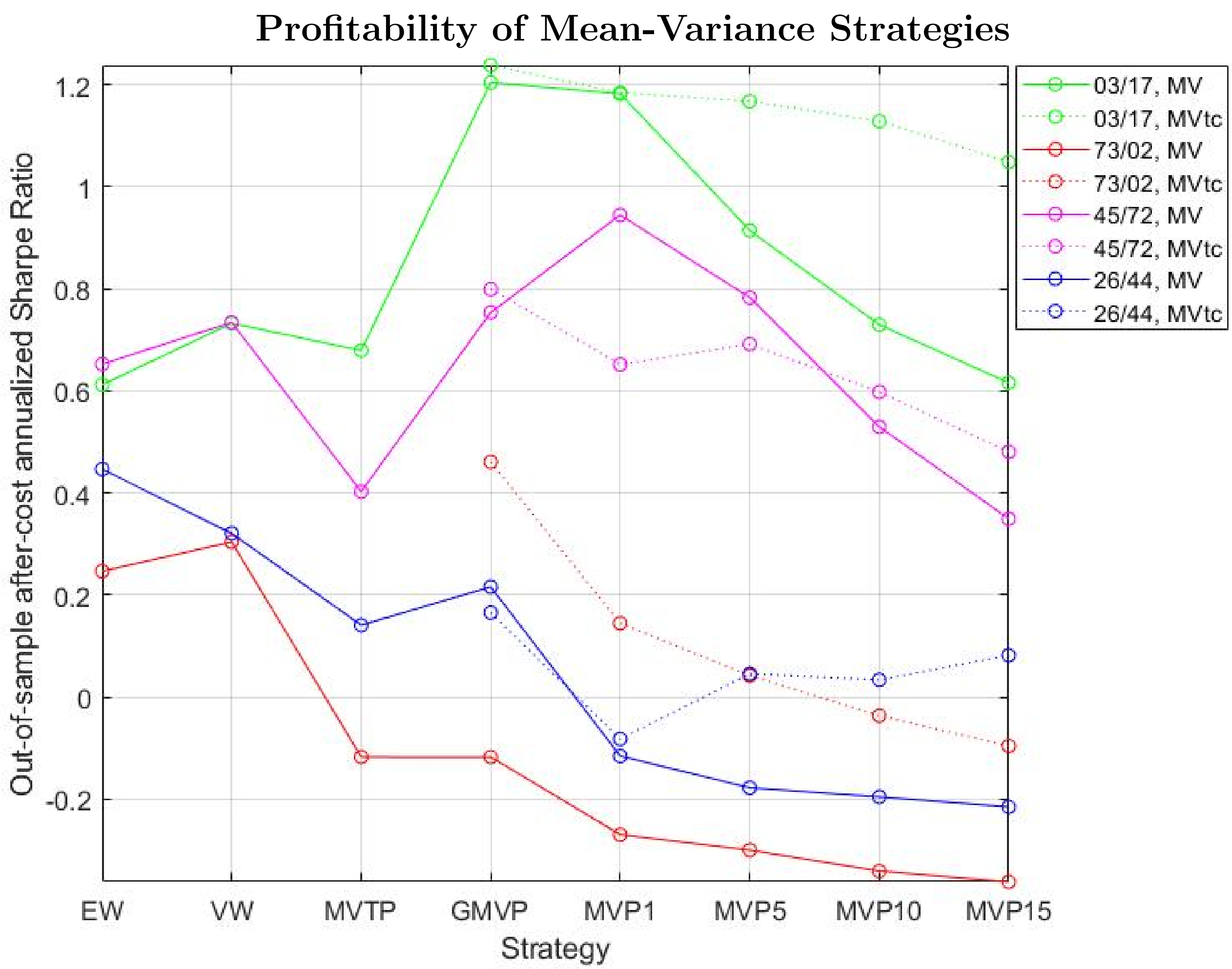

Figure 3: The profitability of mean-variance strategies in the U.S. stock market is pro-cyclical and has never been so high as recently. The graph plots the out-of-sample after cost annualized Sharpe ratios of several mean-variance (MV) strategies against two common benchmarks – the equally weighted (EW) and the value weighted market portfolio (VW) – over four consecutive non-overlapping sub-samples covering the period January 1926 - December 2017: the (mostly) recessionary periods 26/44 and 73/02 (in blue and red) and the (mostly) expansionary periods 45/72 and 03/17 (in magenta and green). All strategies are re-balanced monthly and invest in the entire universe of common stocks belonging to the NYSE, AMEX and NASDAQ. MVTP is the MV Tangency Portfolio, GMVP is the Global Minimum Variance Portfolio, while MVP1 through MVP15 are the frontier portfolios targeting an annualized risk premium of 1, 5, 10 and 15% respectively. The dotted lines report the Sharpe ratio of the cost optimized version of GMVP, MVP1, MVP5, MVP10 and MVP15 respectively as described in Section 2.1. Because any MV strategy requires a vector of risk premia and a covariance matrix, the reported Sharpe ratio of a given MV strategy is the average across those obtained by combining six different strategies only differing in the length of the rolling windows used to estimate the covariance matrix (employing either the past 6 or 12 months) and the premia (employing either the past 60, 120 or 240 months).

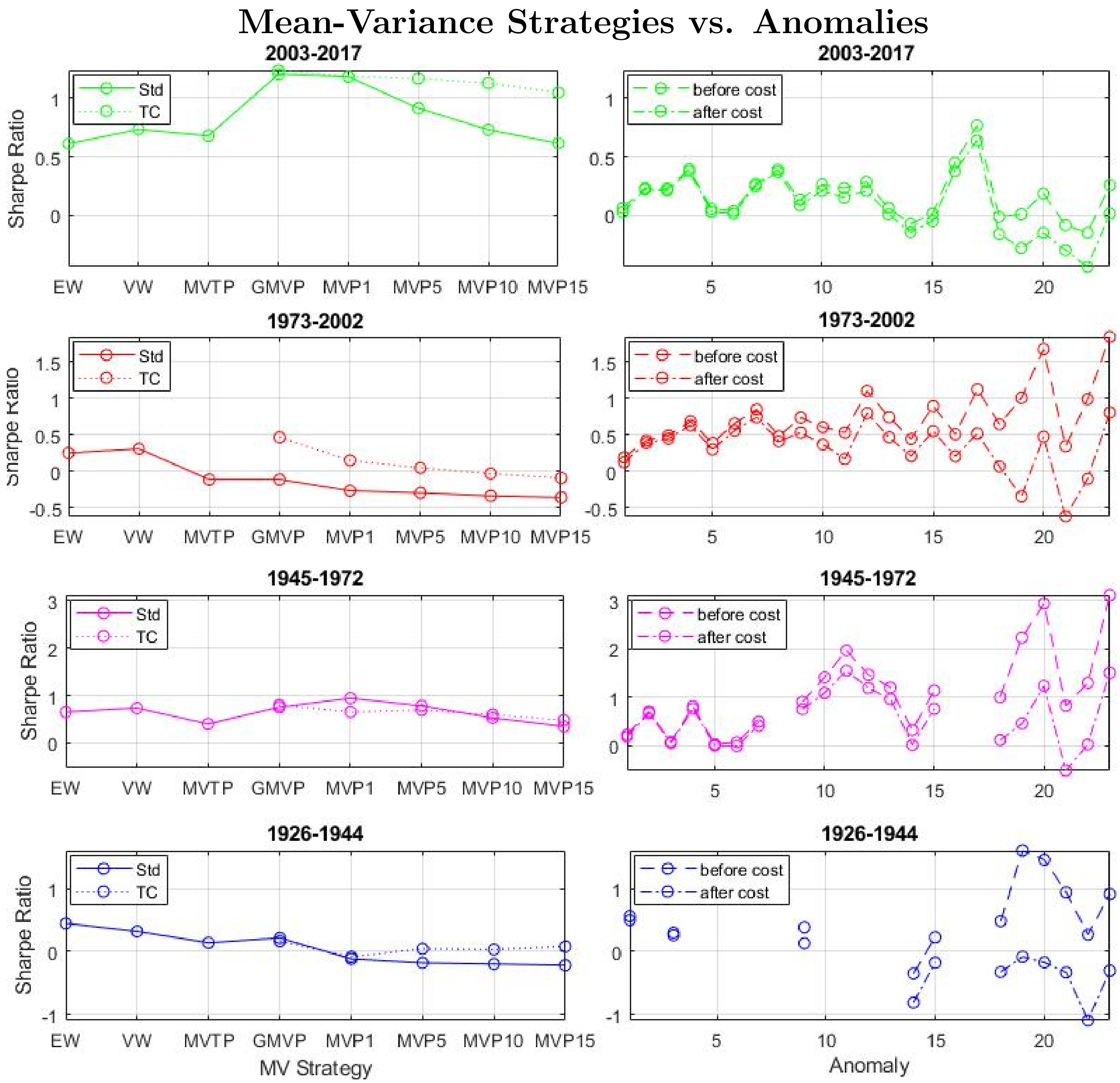

Figure 4: In contrast to the past, post-2002 investing in the U.S. stock market according to the Mean-Variance (MV) criterion is much more profitable than investing in anomalies. The graphs plot the out-of-sample after-cost annualized Sharpe ratios of our baseline set of analyzed strategies (left column) – the MV Tangency Portfolio (MVTP), the Global Minimum Variance Portfolio (GMVP), four frontier portfolios targeting an annualized risk premium of 1, 5, 10 and 15% (MVP1, MVP5, MVP10 and MVP15), and two common market benchmarks (the equally (EW) and value (VW) weighted portfolios) – against the leading 23 anomalies studied in Novy-Marx and Velikov (2016) (right column). All strategies are re-balanced monthly and invest in the entire universe of common stocks belonging to the NYSE, AMEX and NASDAQ. Solid lines report the Sharpe ratios of the standard textbook version of the strategies (here labeled "Std" strategies), dotted lines report those of the cost optimized version of GMVP, MVP1, MVP5, MVP10 and MVP15 as described in Section 2.1 (here labeled "TC" strategies), dashed lines those of the anomalies before cost and dash-dotted lines those of the anomalies after cost. The different rows display different sub-samples using the same colors used in Figure 3: from the most recent sample (03-17, top row, green) to the oldest one (26-44, bottom row, blue).

Recent Profitability of Mean-Variance Strategies

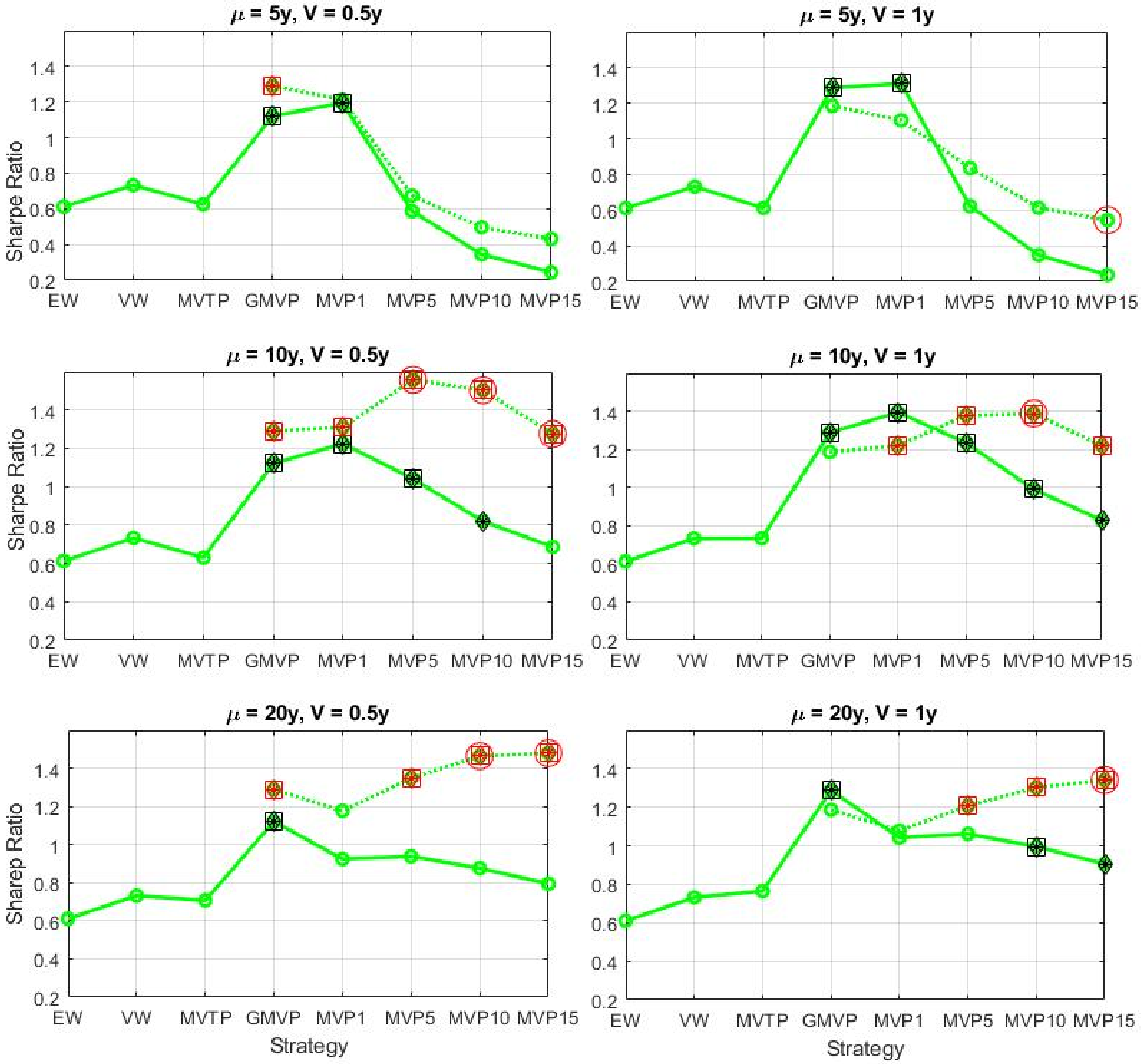

Figure 5: Post-2002, once stock risk premia are reliably estimated, cost-optimized Mean-Variance (MV) strategies targeting higher premia are very profitable in gaining exposure on the market, while the textbook Global Minimum Variance Portfolio (GMVP) remains an efficient alternative. The figure shows the annualized out-of-sample after-cost Sharpe ratios of our analyzed strategies (defined in the previous figures and in Section 2.1). The MV strategies are benchmarked against the equally (EW) and (VW) value weighted market portfolios as well as the SPDR and the IWM ETFs (mimicking the exposure to the $S\&P$ 500 and the Russell's 2000 indices), are re-balanced monthly and invest in the entire universe of common NYSE, AMEX and NASDAQ stocks over the period January 2003 - December 2017. Each graph differs in the length (in years) of the rolling windows used to estimate the risk premia vector μ_t (5,10,20) and the covariance matrix of return V_t (0.5,1). Sharpe ratios of given strategies that are better at the 10% level than: the EW, the VW, both the SPDR and the IWM ETFs, and their standard analogs (if they refer to cost-optimized strategies) are marked with a star, a diamond, a square and a circle respectively. Markers for standard(cost-optimized) strategies are reported in black(red). We test for Sharpe ratio differences using block bootstrapping (block sizes of 5 obs.) accounting for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)).

Behind recent profitability: Sharpe ratio decomposition

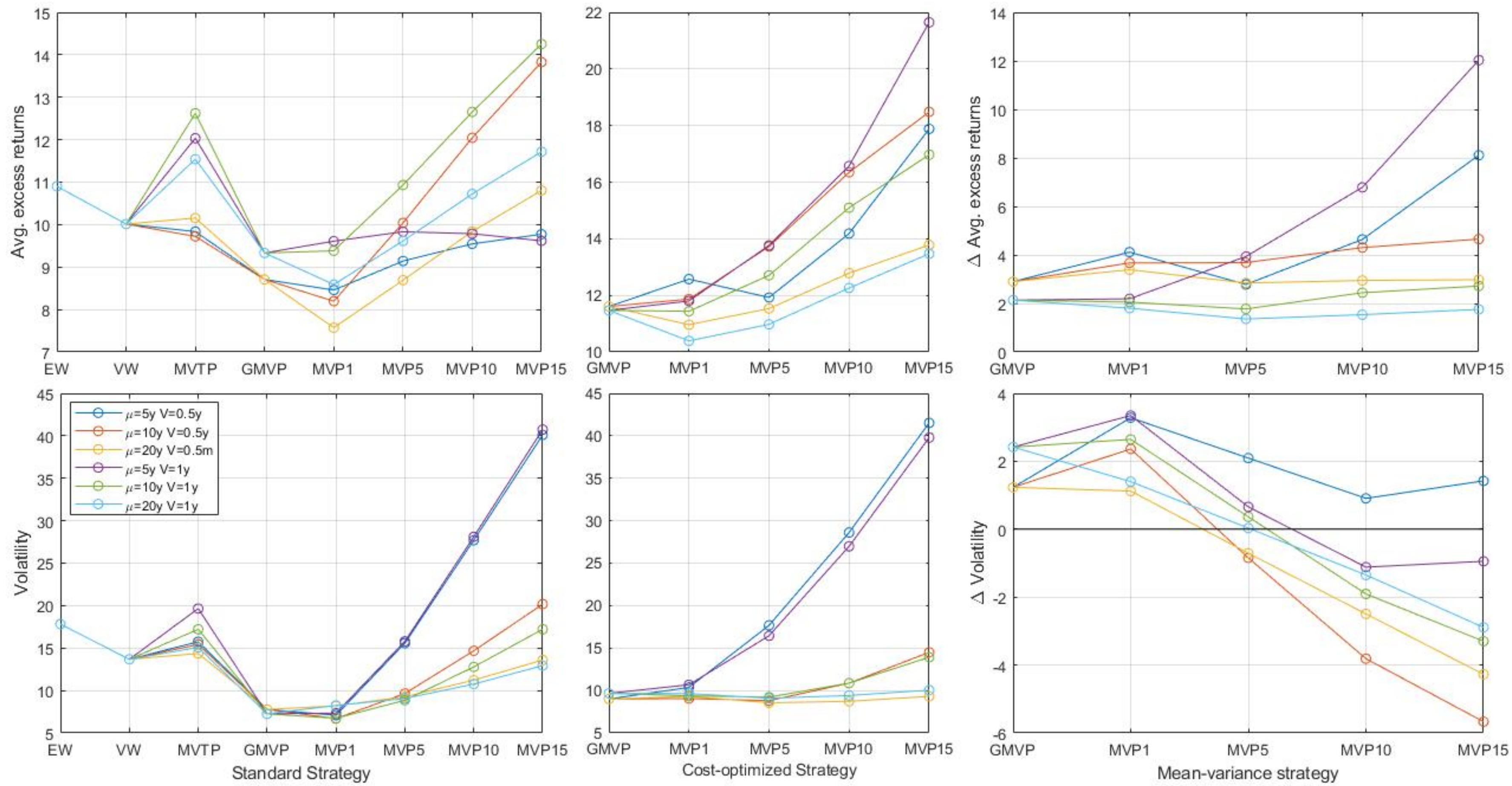

Figure 6: While the performance of $GMVP$ is driven by low volatility, the higher Sharpe ratios of cost-optimized strategies targeting higher premia come from both lower volatilities and higher average excess returns. The top graphs refer to the average excess returns (numerators of the Sharpe ratios reported in Figure 5) of our analyzed strategies while the bottom ones to their volatilitites (denominators of the Sharpe ratios reported in Figure 5) with the rightmost plots showing the differential between cost-optimized and standard strategies. The sample period is every month from January 2003 through December 2017 and the stock universe is represented by all common NYSE, AMEX and NASDAQ stocks. The analyzed strategies are three standard textbook types of MV portfolios: the MV Tangency Portfolio (MVTP), the GMVP and four frontier portfolios targeting an annualized risk premium of 1% through 15% (MVP1,MVP5,MVP10 and MVP15) – and their cost-optimized version when available: namely GMVPtc, MVP1tc,MVP5tc,MVP10tc and MVP15tc.

Market Timing vs Estimation Error:
V_t rolling window of 6 months, sample 2003-2017

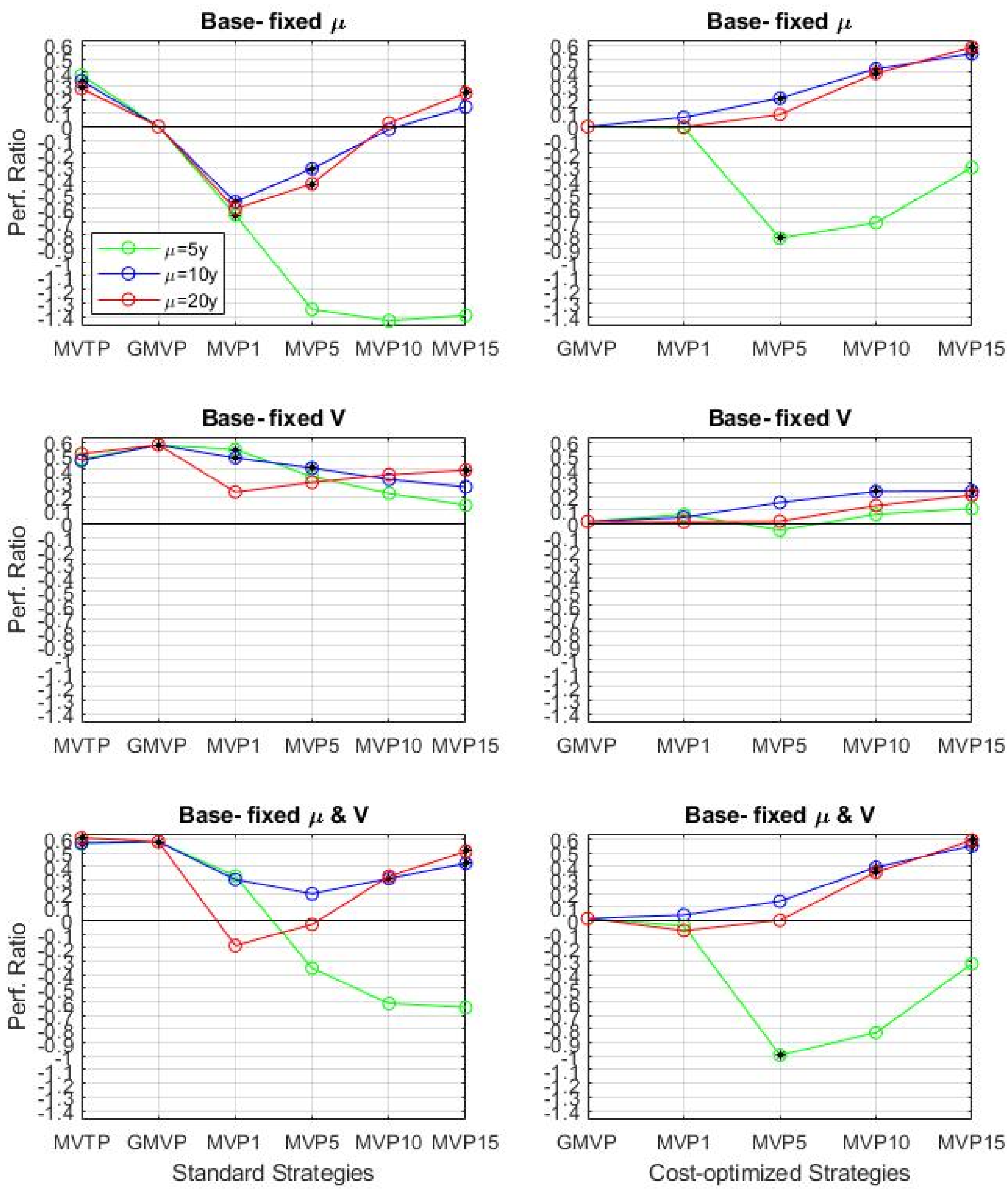

Figure 7: Standard strategies mostly time stock covariances and suffer from estimation error induced by stock premia. Cost-optimized strategies mostly time stock premia and suffers much less from estimation error. The performance ratios are computed for our strategies over the monthly period January 2003 - December 2017. Top/middle/bottom graphs isolate the effects coming from market timing (positive gap), or suffering from estimation error (negative gap) from the conditional stock premia/covariances/premia & covariances. The leftmost(rightmost) graphs refer to

Market Timing vs Estimation Error:
V_t rolling window of 1 year, sample 2003-2017

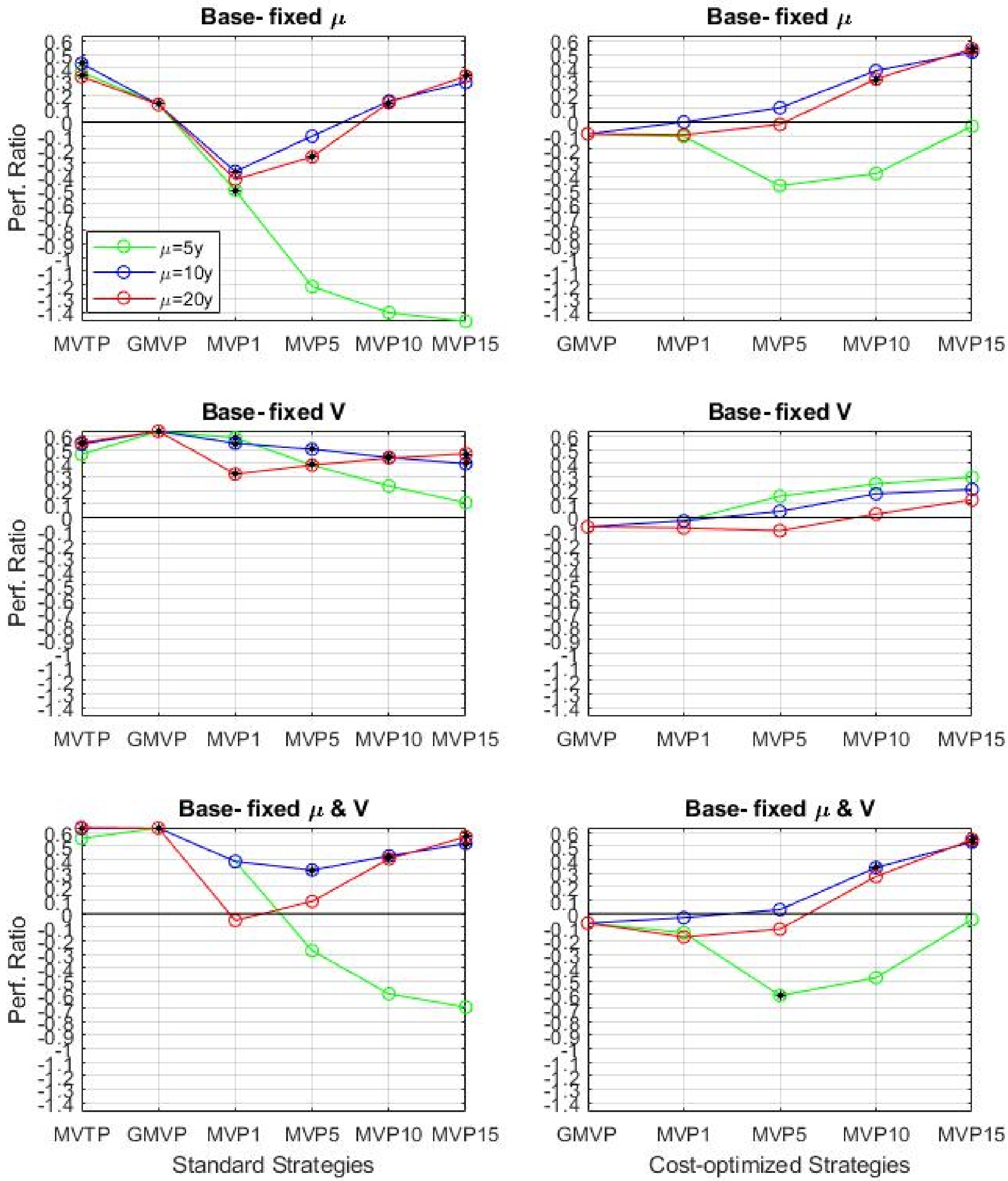

Figure 8: Standard strategies mostly time stock covariances and suffer from estimation error induced by stock premia. Cost-optimized strategies mostly time stock premia and suffers much less from estimation error. The performance ratios are computed for our strategies over the monthly period January 2003 - December 2017. Top/middle/bottom graphs isolate the effects coming from market timing (positive gap), or suffering from estimation error (negative gap) from the conditional stock premia/covariances/premia & covariances. The leftmost(rightmost) graphs refer to

Market αs and βs of cost-optimized and standard mean-variance strategies, sample 2003-3017

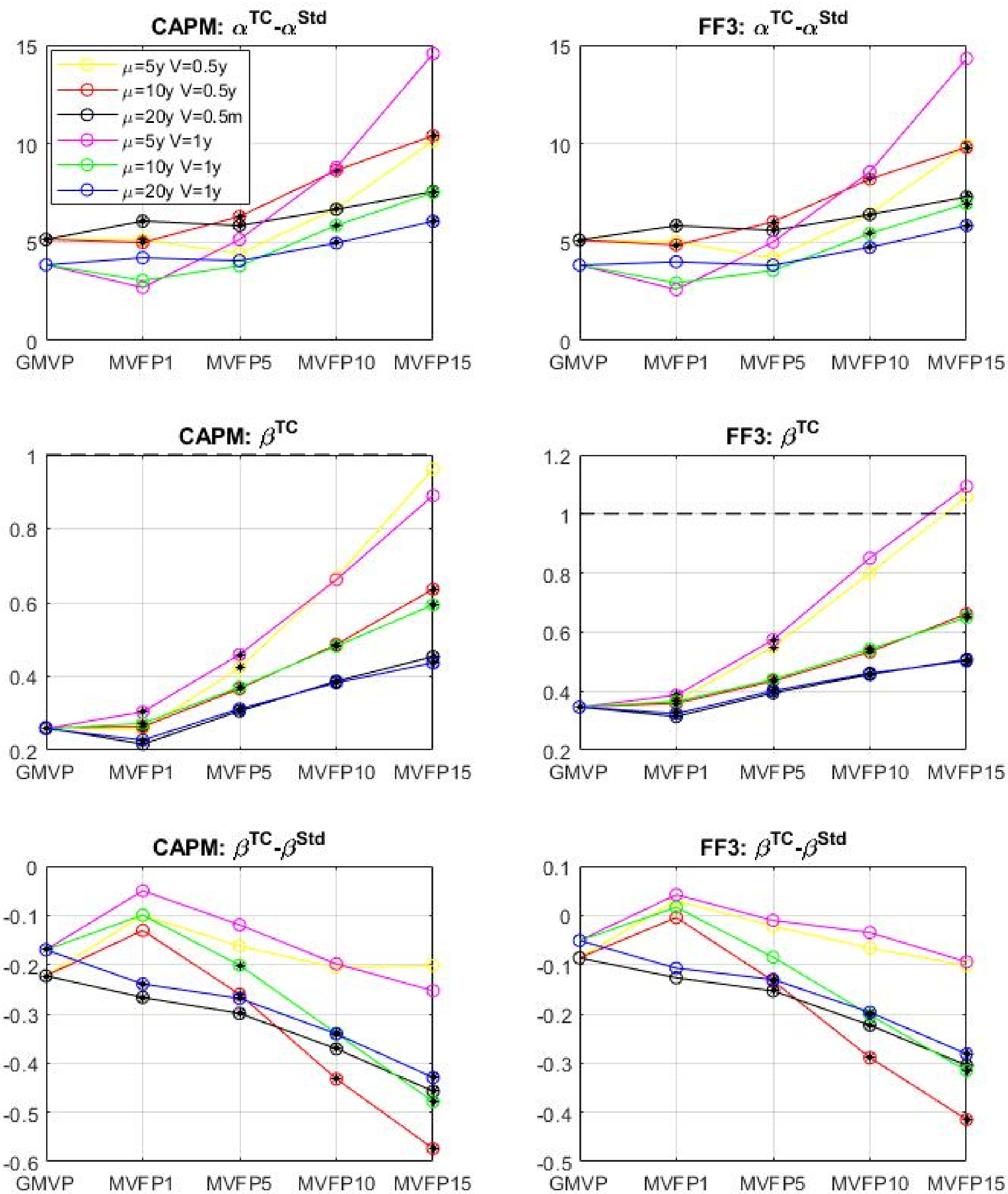

Figure 9: Frazzini et al. (2014) "betting against beta" theory can explain the patterns in our strategies (described in the previous figures) over the monthly sample 2003-2017. Cost-optimized strategies have higher(lower) alphas(betas) than their (less constrained) standard analogs as shown in the top(bottom) graphs. Market betas of cost optimized strategies, plotted in the middle graphs, are found smaller then 1. Leftmost(rightmost) plots assume the CAPM(Fama and French (1993) 3 factor) representation for the returns. Starred markers in top and bottom graphs highlight figures statistically significant at the 10% level, while those in middle graphs are significantly different

Downside risk: the stabilizing role of cost minimization

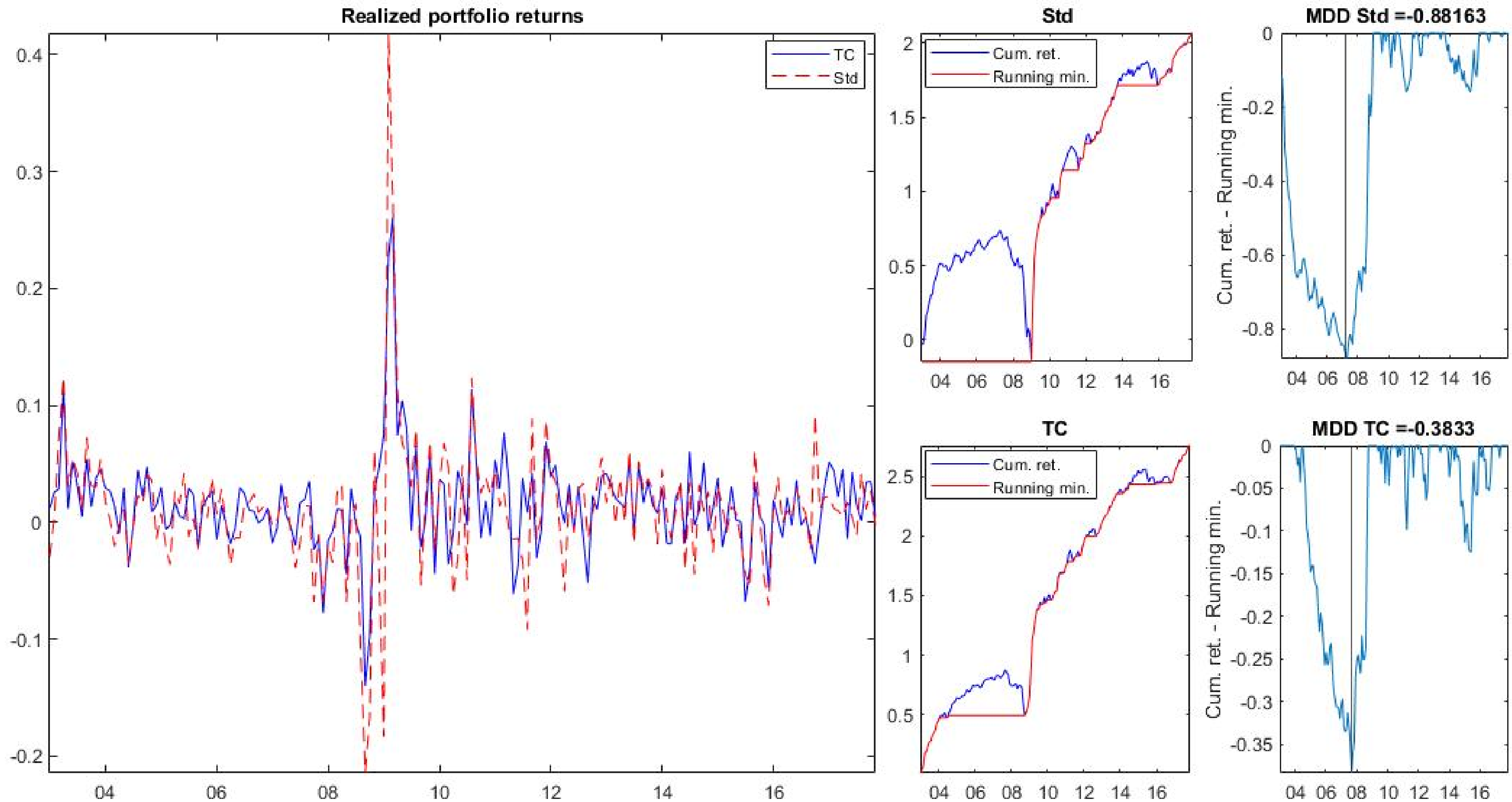

Figure 10: The existence of a no-trading region bounds the cost-optimized weights making the realized returns less volatile. This is crucial in dampening the negative realizations pushing down the Maximum Draw Downs. The figure decompose the Maximum Draw Downs (MDDs) of $MVP15$, our standard mean-variance frontier portfolio targeting an annualized premium of 15%, labeled as “Std”, as well as $MVP15tc$, its cost-optimized analog labeled as “TC”, over the monthly sample 2003-2017. MDD is a popular downside risk measure, representing the largest observed loss of a given strategy. It is computed as the minimum of the difference between the running minimum of the commutative after cost returns from any time t to the end of the sample and the current time t cumulative return level, as illustrated in the rightmost plots for the standard (top plot) and cost-optimized (bottom plot) strategies. The time series of the running minimum and the cumulative returns are plotted in the the two central graphs for the standard (top plot) and cost-optimized (bottom plot) strategies.

Downside risk: cost-optimized versus standard mean-variance strategies

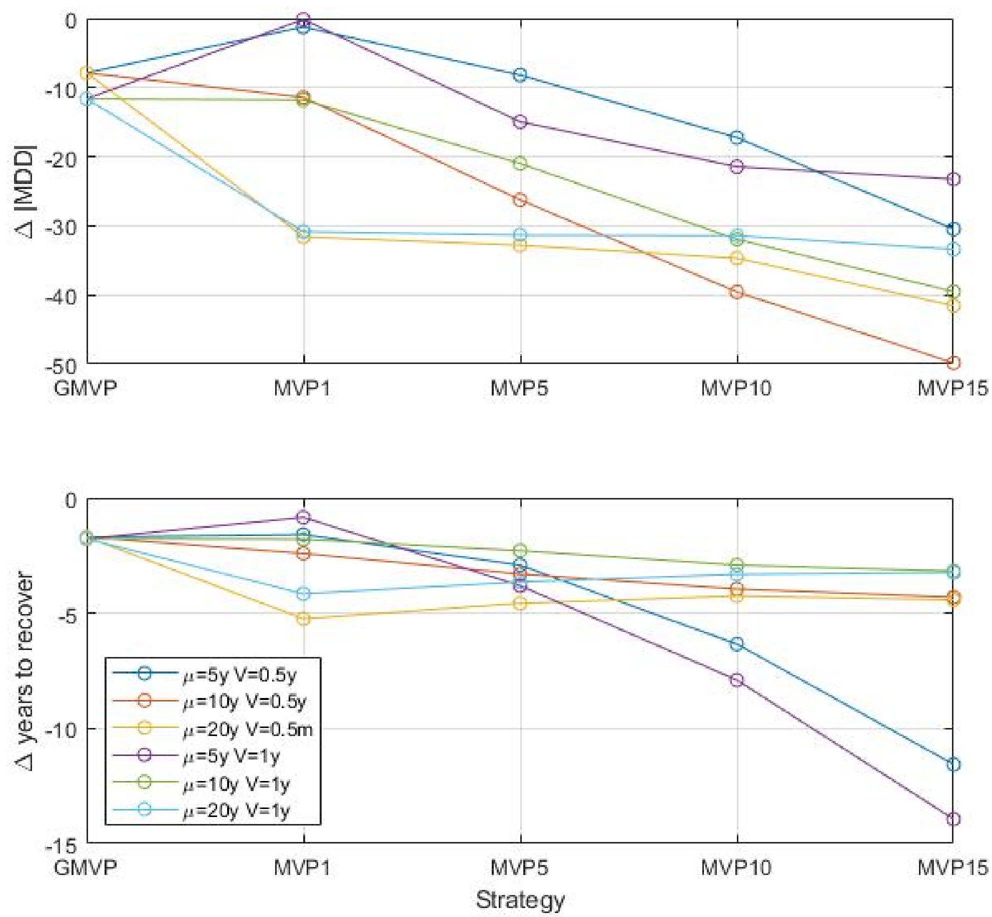

Figure 11: Cost-minimization is a useful stabilizing device to reduce downside risk as measured via the worst experienced loss (Maximum Draw Down or MDD) or by the average number of years to recover from it. The top graph plot the difference in absolute MDDs between our cost-optimized and standard strategies (which are described in Section 2.1) for the six different estimation configuration. Each configuration only differs in the rolling window lengths (expressed in years in the legend of the bottom figure) used to estimate the conditional vector of risk premia and the covariance matrix of returns. The bottom graph plots the difference in the average number of years required to recover from the worst observed loss between the cost-optimized and the standard strategies. The average number of years to recover from the worst loss is computed as the ratio of a strategy absolute MDD and its average excess return. The sample is every month from January 2003 to December 2017.

Price Impact: standard mean-variance strategies

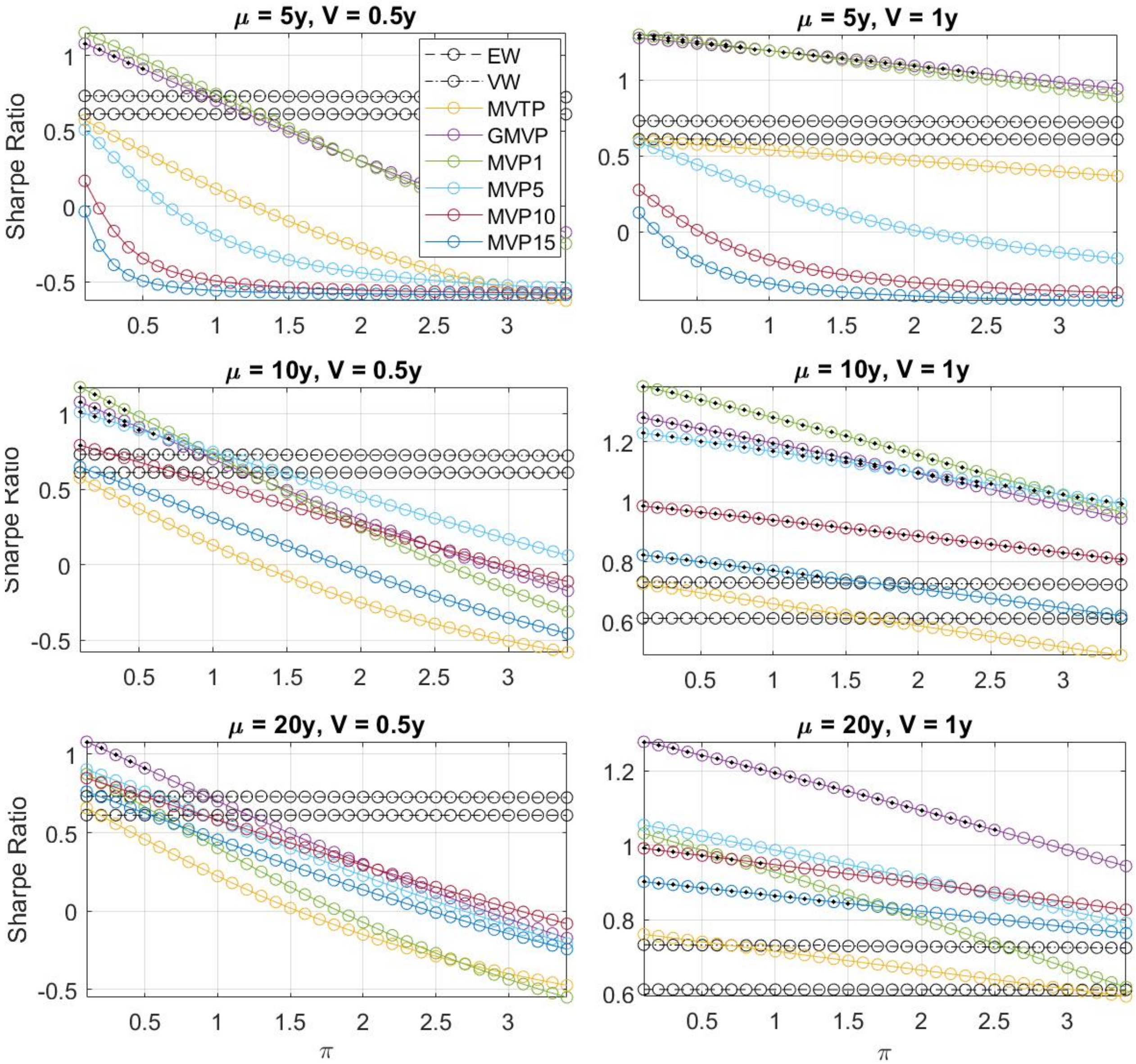

Figure 12: Standard mena-variance strategies are sensitive to price impact, however the Global Minimum Variance Portfolio (GMVP) can actually scale up if properly parametrized. The graphs plots the annualized after-cost out-of-sample Sharpe ratios of our standard strategies (described in Section 2.1 and printed in the legend) as a function of the price impact parameter π, which is inversely related to the liquidity and depth of the market and positively related to the portfolio size of the investors. Each graph differs in the combinations of rolling window lengths (in years) used to estimate the conditional risk premia μ_t and covariance matrix V_t as specified in the plots' sub-titles. Starred markers refer to ratios bigger than those of the EW and VW at the 10% level. We test for differences in Sharpe ratios using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). The sample is every month from January 2003 to December 2017.

Price Impact: cost-optimized mean-variance strategies

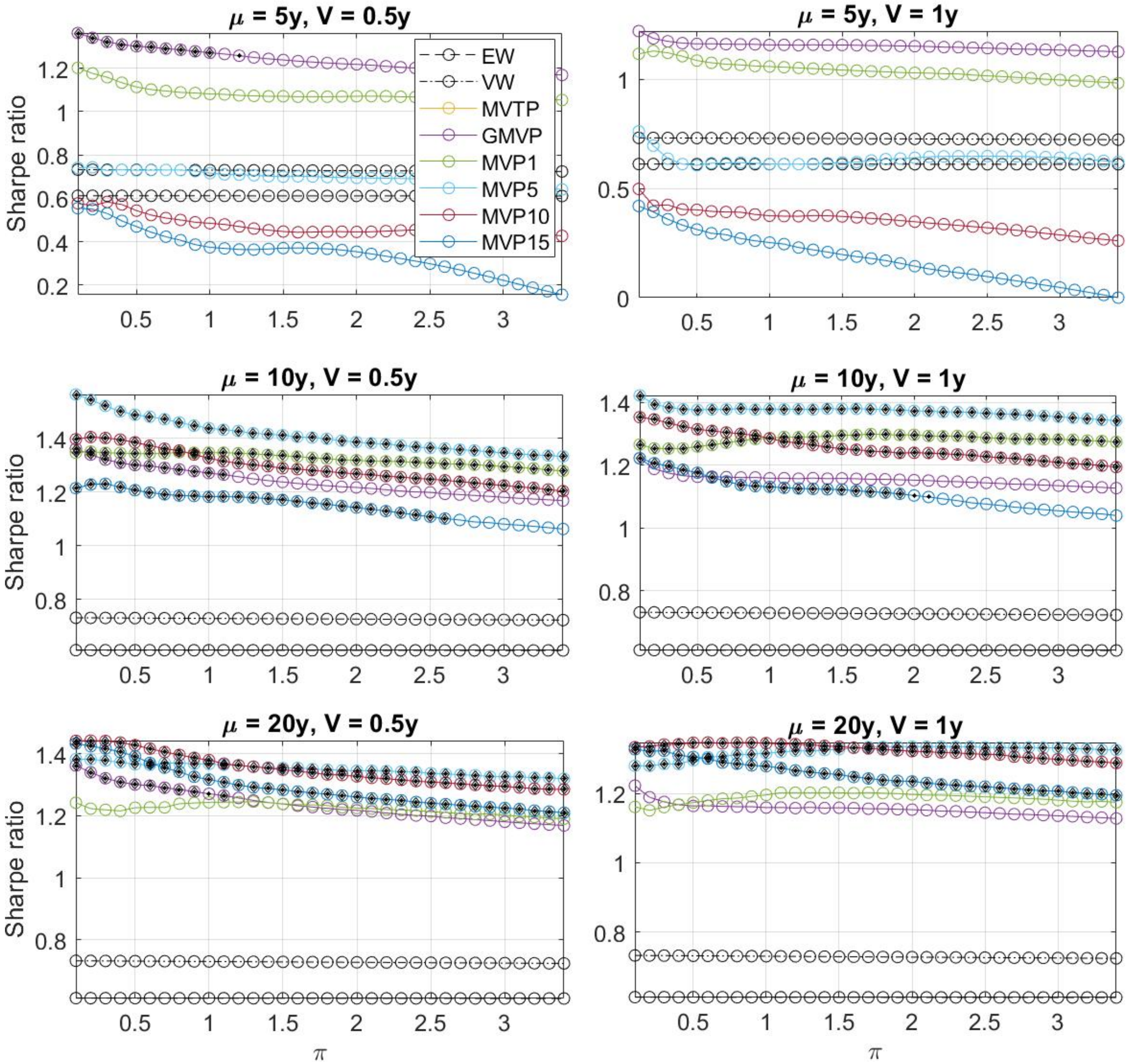

Figure 13: Cost-optimized strategies robustly scale up and are very profitable options to gain exposure on the market even for large institutional investors. The graphs plots the annualized after-cost out-of-sample Sharpe ratios of our cost-optimized strategies (described in Section 2.1 and printed in the legend) as a function of the price impact parameter π, which is inversely related to the liquidity and depth of the market and positively related to the portfolio size of the investors. Each graph differs in the combinations of rolling window lengths (in years) used to estimate the conditional risk premia μ_t and covariance matrix V_t as specified in the plots' sub-titles. Starred markers refer to ratios bigger than those of the EW and VW at the 10% level. We test for differences in Sharpe ratios using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). The sample is every month from January 2003 to December 2017.

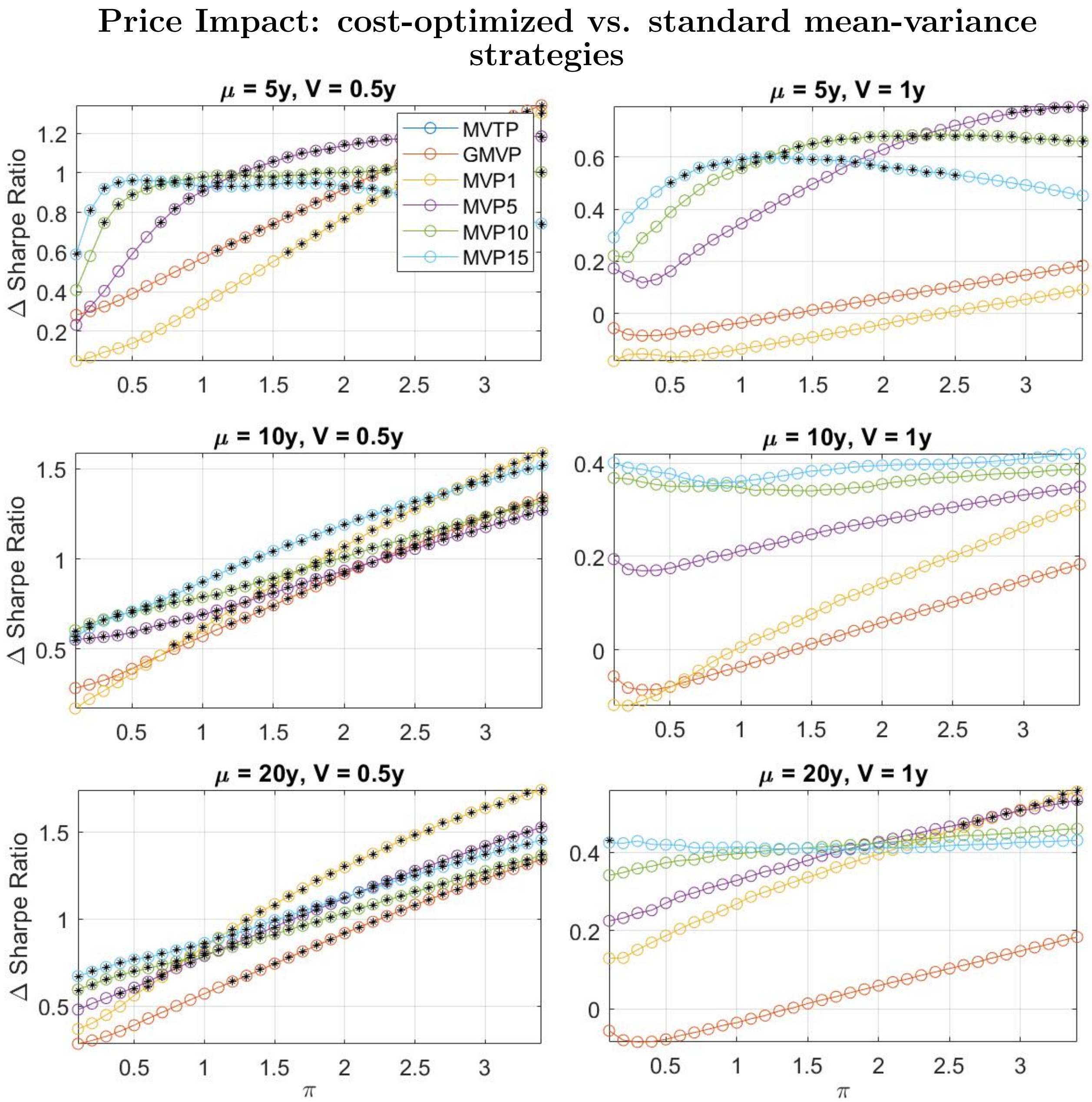

Figure 14: Cost-optimized mean-variance strategies consistently scale up more and are less sensitive to changes in price impact. The graphs plots the annualized after-cost out-of-sample Sharpe ratio differentials between our cost-optimized and standard strategies (described in Section 2.1 and printed in the legend) as a function of the price impact parameter π, which is inversely related to the liquidity and depth of the market and positively related to the portfolio size of the investors. Each graph differs in the combinations of rolling window lengths (in years) used to estimate the conditional risk premia μ_t and covariance matrix V_t as specified in the plots' sub-titles. Starred markers refer to ratio differential significant at the 10% level. We test for differences in Sharpe ratios using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). The sample is every month from January 2003 to December 2017.

Table 1: **Cost-minimization as an economically justified way to bound weights**

Strategy	GMVP	MVP1	MVP5	MVP10	MVP15
	ΔSR: Cost-Optimized - Standard (no-short selling)				
$\mu = 5y, V = 0.5y$	0.36	0.35	0.34	0.20	0.26
$\mu = 10y, V = 0.5y$	0.36	0.52	0.78***	0.76***	0.76***
$\mu = 20y, V = 0.5y$	0.36	0.72*	0.60*	0.76***	0.79***
$\mu = 5y, V = 1y$	0.17	0.10	0.46*	0.39	0.36
$\mu = 10y, V = 1y$	0.17	0.33	0.43*	0.56**	0.67***
$\mu = 20y, V = 1y$	0.17	0.53	0.41	0.50*	0.59**
	ΔSR: Cost-Optimized - Standard (no-short selling, 1% upper bound)				
$\mu = 5y, V = 0.5y$	0.50	0.55	-0.08	-0.20	-0.16
$\mu = 10y, V = 0.5y$	0.50	0.69**	0.79***	0.84***	0.70***
$\mu = 20y, V = 0.5y$	0.50	0.70	0.63*	0.77***	0.82***
$\mu = 5y, V = 1y$	0.35	0.36	0.05	-0.09	-0.07
$\mu = 10y, V = 1y$	0.35	0.57	0.56**	0.68***	0.61***
$\mu = 20y, V = 1y$	0.35	0.58	0.46	0.55*	0.64**
	ΔSR: Cost-Optimized - Standard (no-short selling, 2% upper bound)				
$\mu = 5y, V = 0.5y$	0.43	0.44	-0.03	-0.20	-0.17
$\mu = 10y, V = 0.5y$	0.43	0.56	0.78***	0.80***	0.66***
$\mu = 20y, V = 0.5y$	0.43	0.67	0.61*	0.77***	0.82***
$\mu = 5y, V = 1y$	0.26	0.22	0.12	-0.10	-0.12
$\mu = 10y, V = 1y$	0.26	0.41	0.48*	0.64***	0.59***
$\mu = 20y, V = 1y$	0.26	0.53	0.44	0.56*	0.65**

Notes: Cost minimization is an economically justified way to bound the mean-variance weights by only trading in some stocks to maximize the before-cost risk-return trade-off if the initial allocation is too displaced. The table reports the annualized after-cost out-of-sample Sharpe ratio differentials, ΔSR between our cost-optimized strategies (described in Section 2.1) and their standard analogs when their weights are constrained to: be non-negative (top panel), be non negative and additionally no more than 1% (middle panel) or 2% (bottom panel). Every raw in a given panel refers to a different rolling window length combination (expressed in years) used to estimate μ_t and V_t. The sample is every month from January 2003 to December 2017. Standard errors for ΔSR are estimated using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). ”,** and * highlight figures statistically significant at the 10%, 5% and 1% level.

Table 2: **Market time versus Estimation error: dynamics of average magnitudes over the last 100 years**

	Market timing				Estimation error			
	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$	**Mean**	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$	**Mean**
PANEL A: STANDARD MEAN-VARIANCE STRATEGIES								
sample 03-17:	0.2	0.51	0.42	**0.38**	-0.49	0	0	**-0.49**
sample 73-02:	0	0	0	**0**	-0.38	-0.58	-0.57	**-0.51**
sample 45-72:	0.2	0	0	**0.2**	-0.16	-0.19	-0.34	**-0.23**
sample 26-44:	0	0	0	**0**	-0.89	-0.33	-0.75	**-0.66**
PANEL B: STANDARD MEAN-VARIANCE STRATEGIES (EXCLUDING $MVTP$)								
sample 03-17:	0.19	0.53	0.42	**0.38**	-0.49	0	0	**-0.49**
sample 73-02:	0	0	0	**0**	0	-0.59	-0.55	**-0.57**
sample 45-72:	0.2	0	0	**0.2**	0	-0.17	-0.37	**-0.27**
sample 26-44:	0	0	0	**0**	-0.89	-0.35	-0.89	**-0.71**
PANEL C: COST-OPTIMIZED MEAN-VARIANCE STRATEGIES								
sample 03-17:	0.53	0.33	0.67	**0.51**	-0.55	0	-0.59	**-0.57**
sample 73-02:	0	0	0	**0**	-0.69	-0.24	-0.55	**-0.49**
sample 45-72:	0.41	0.29	0	**0.35**	0	0	-0.34	**-0.34**
sample 26-44:	0	0	0	**0**	0	0	-0.58	**-0.58**
PANEL D: ALL VARIANCE STRATEGIES (EXCLUDING $MVTP$)								
sample 03-17:	0.36	0.43	0.54	**0.45**	-0.52	0	-0.59	**-0.56**
sample 73-02:	0	0	0	**0**	-0.69	-0.41	-0.55	**-0.55**
sample 45-72:	0.3	0.29	0	**0.3**	0	-0.17	-0.35	**-0.26**
sample 26-44:	0	0	0	**0**	-0.89	-0.35	-0.73	**-0.66**

Notes: Market timing and estimation error have substantial impact on the performance of MV strategies over the last 100 years. The reported figures represent the average over strategies and across the six rolling-window-length configurations for the conditional vector of stock risk premia and covariance matrix ([0.5,5],[0.5,10],[0.5,15],[1,5],[1,10],[1,15] years) of all Sharpe ratio gaps (defined in Section 2.3) that are significant at least at the 10% level. The overall sample is every month from January 1926 through December 2017.

Table 3: **Market time versus Estimation error: dynamics of total amounts over the last 100 years**

	Market timing				Estimation error			
	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$	**Total**	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$	**Total**
PANEL A: STANDARD MEAN-VARIANCE STRATEGIES								
sample 03-17:	2.15	11.73	6.66	**20.54**	-3.62	0	0	**-3.62**
sample 73-02:	0	0	0	**0**	-0.75	-7.49	-11.22	**-19.46**
sample 45-72:	1.8	0	0	**1.8**	-0.48	-1.28	-3.38	**-5.14**
sample 26-44:	0	0	0	**0**	-1.78	-2.43	-5.66	**-9.87**
PANEL B: STANDARD MEAN-VARIANCE STRATEGIES (EXCLUDING $MVTP$)								
sample 03-17:	1.16	10.91	5.8	**17.87**	-3.62	0	0	**-3.62**
sample 73-02:	0	0	0	**0**	0	-6.44	-9.37	**-15.81**
sample 45-72:	1.8	0	0	**1.8**	0	-0.84	-2.84	**-3.68**
sample 26-44:	0	0	0	**0**	-1.78	-2.14	-5.31	**-9.23**
PANEL C: COST-OPTIMIZED MEAN-VARIANCE STRATEGIES								
sample 03-17:	5.4	0.67	5.38	**11.45**	-0.55	0	-1.18	**-1.73**
sample 73-02:	0	0	0	**0**	-8.68	-0.47	-1.64	**-10.39**
sample 45-72:	0.81	0.87	0	**1.68**	0	0	-1.35	**-1.35**
sample 26-44:	0	0	0	**0**	0	0	-1.16	**-1.16**
PANEL D: ALL VARIANCE STRATEGIES (EXCLUDING $MVTP$)								
sample 03-17:	6.56	11.58	11.18	**29.32**	-4.17	0	-1.18	**-5.35**
sample 73-02:	0	0	0	**0**	-8.68	-6.91	-11.01	**-26.2**
sample 45-72:	2.61	0.87	0	**3.48**	0	-0.84	-4.19	**-5.03**
sample 26-44:	0	0	0	**0**	-1.78	-2.14	-6.47	**-10.39**

Notes: The market timing abilities display a pro-cyclical pattern with magnitudes in expansions increasing over time. Moreover, estimation error is counter-cyclical with magnitudes in recessions decreasing over time. The reported figures represent the sum over strategies and across the six rolling-window-length configurations for the conditional vector of stock risk premia and covariance matrix ([0.5,5],[0.5,10],[0.5,15],[1,5],[1,10],[1,15] years) of all Sharpe ratio gaps (defined in Section 2.3) that are significant at least at the 10% level. The overall sample is every month from January 1926 through December 2017.

Table 4: **Correlations between Mean-Variance Profitability and Market timing: Standard Strategies**

	Rolling window length V	$\bar{\mu}$		$\bar{V}$		$\bar{\mu}$ & $\bar{V}$	
Sample	μ	0.5 yrs	1 yr	0.5 yrs	1 yr	0.5 yrs	1 yr
03-17	5 yrs	-0.32	0.14	0.99	1	0.85	0.89
	10 yrs	-0.99	-0.8	0.91	0.93	0.36	0.47
	15 yrs	-0.86	-0.3	0.59	0.7	0.07	0.19
	Mean	**-0.72**	**-0.32**	**0.83**	**0.88**	**0.43**	**0.52**
73-02	5 yrs	0.43	0.64	0.42	-0.22	0.31	0.24
	10 yrs	0.5	0.66	0.06	-0.49	0.1	0.61
	15 yrs	0.49	0.82	-0.42	0.89	0.82	0.89
	Mean	**0.47**	**0.71**	**0.02**	**0.06**	**0.41**	**0.58**
45-72	5 yrs	0.79	0.9	0.89	0.9	0.73	0.84
	10 yrs	0.6	0.76	0.45	0.68	0.24	0.57
	15 yrs	0.18	0.29	-0.55	-0.11	-0.53	-0.2
	Mean	**0.52**	**0.65**	**0.26**	**0.49**	**0.15**	**0.4**
26-44	5 yrs	0.95	0.97	0.82	0.96	0.89	0.93
	10 yrs	0.76	0.44	0.74	0.68	0.23	0.12
	15 yrs	-	-	-	-	-	-
	Mean	**0.85**	**0.71**	**0.78**	**0.82**	**0.56**	**0.53**
	Overall mean	**0.28**	**0.44**	**0.47**	**0.56**	**0.39**	**0.51**

Notes: There exist a substantial positive correlation between the profitability of our analyzed standard mean-variance strategies (detailed in Section 2.1) and the level of market timing / estimation error. Profitability is measured by the after-cost out-of-sample Sharpe ratios while the level of market timing / estimation error for each given strategy is represented by the Sharpe ratio differential between the real-time version of the given strategy and its unimplementable analog where at least one of the first two moments of the return distribution is(are) kept fixed at the look-ahead biased in-sample estimate(s)(with high, or positive, values capturing actual market timing and low, or negative, values actual estimation error as explained in Section 2.3). The sample covers every month from January 1926 through December 2017 and it is divided in 4 sub-samples: 26-44,45-72,73-02,03-17.

Table 5: **Correlations between Mean-Variance Profitability and Market timing: Cost-optimized Strategies**

		$\bar{\mu}$		$\bar{V}$		$\bar{\mu}$ & $\bar{V}$	
	Rolling window length V	0.5 yrs	1 yr	0.5 yrs	1 yr	0.5 yrs	1 yr
Sample	μ						
03-17	5 yrs	0.46	0.04	0.24	-0.96	0.59	0.1
	10 yrs	0.03	0.36	0.52	0.46	0.13	0.22
	15 yrs	0.9	0.89	0.85	0.81	0.88	0.89
	Mean	**0.46**	**0.43**	**0.54**	**0.1**	**0.53**	**0.4**
73-02	5 yrs	0.94	0.98	0.54	0.77	1	1
	10 yrs	0.99	0.96	0.69	0.54	0.99	0.98
	15 yrs	0.93	-0.26	0.83	-0.37	0.94	-0.03
	Mean	**0.95**	**0.56**	**0.68**	**0.31**	**0.98**	**0.58**
45-72	5 yrs	0.72	0.44	0.76	0.5	0.86	0.86
	10 yrs	0.73	-0.36	0.83	-0.05	0.88	0.29
	15 yrs	0.72	0.06	0.83	0.46	0.79	-0.01
	Mean	**0.73**	**0.04**	**0.8**	**0.3**	**0.85**	**0.38**
26-44	5 yrs	0.53	0.99	0.11	0.65	0.92	0.96
	10 yrs	0.9	0.93	0.87	0.97	0.94	0.98
	15 yrs	-	-	-	-	-	-
	Mean	**0.71**	**0.96**	**0.49**	**0.81**	**0.93**	**0.97**
	Overall mean	**0.71**	**0.5**	**0.63**	**0.38**	**0.82**	**0.58**

Notes: There exist a substantial positive correlation between the profitability of our analyzed cost-optimized mean-variance strategies (detailed in Section 2.1) and the level of market timing / estimation error. Profitability is measured by the after-cost out-of-sample Sharpe ratios while the level of market timing / estimation error for each given strategy is represented by the Sharpe ratio differential between the real-time version of the given strategy and its unimplementable analog where at least one of the first two moments of the return distribution is(are) kept fixed at the look-ahead biased in-sample estimate(s)(with high, or positive, values capturing actual market timing and low, or negative, values actual estimation error as explained in Section 2.3). The sample covers every month from January 1926 through December 2017 and it is divided in 4 sub-samples: 26-44,45-72,73-02,03-17.

Table 6: **Most/least profitable standard mean-variance strategies and associated market timing percentiles**

		Roll. wind. length V	0.5 yrs				1 yr			
Sample	Benchmark	μ	Strategy	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$	Strategy	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$
03-17	MOST profitable	5 yrs	$MVP1$	16.67$^-$	95.46$^+$	86.36$^-$	$MVP1$	13.64$^-$	86.36$^+$	86.36$^+$
		10 yrs	$MVP1$	5.56$^-$	86.36$^+$	68.18$^-$	$MVP1$	4.55$^-$	86.36$^+$	77.27$^+$
		15 yrs	$GMVP$	–	95.46$^+$	86.36$^+$	$GMVP$	59.09$^+$	95.46$^+$	95.46$^+$
		Mean	–	**11.11**	**92.42**	**80.3**	–	**25.76**	**89.39**	**86.36**
	LEAST profitable	5 yrs	$MVP15$	61.11$^-$	22.73$^+$	40.09$^-$	$MVP15$	40.09$^-$	22.73$^+$	40.09$^-$
		10 yrs	$MVTP$	61.11$^+$	50$^+$	59.09$^+$	$MVP1$	77.27$^+$	59.09$^+$	59.09$^+$
		15 yrs	$MVTP$	72.22$^+$	86.36$^+$	68.18$^+$	$GMVP$	68.18$^+$	68.18$^+$	68.18$^+$
		Mean	–	**64.82**	**53.03**	**56.06**	–	**62.12**	**50**	**56.06**
73-02	MOST profitable	5 yrs	$MVTP$	83.33$^+$	68.18$^+$	68.18$^+$	$GMVP$	95.46$^+$	13.64$^-$	68.18$^-$
		10 yrs	$GMVP$	–	13.64$^-$	40.90	$MVTP$	77.27$^+$	77.27$^+$	68.18$^-$
		15 yrs	$MVP5$	72.22$^+$	13.64$^-$	40.90$^+$	$MVTP$	77.27$^+$	50$^-$	31.81$^-$
		Mean	–	**77.78**	**31.81**	**50**	–	**83.33**	**46.97**	**56.06**
	LEAST profitable	5 yrs	$MVP1$	94.44$^+$	13.64$^-$	86.36	$^+$ $MVP1$	86.36$^+$	22.73$^-$	86.36$^+$
		10 yrs	$MVTP$	38.89$^-$	4.55$^-$	4.54$^-$	$MVP15$	50$^-$	50$^-$	4.54$^-$
		15 yrs	$MVP15$	50$^-$	31.82$^-$	4.54$^-$	$MVP15$	50$^-$	22.73$^-$	4.54$^-$
		Mean	–	**61.11**	**16.67**	**31.81**	–	**62.12**	**31.82**	**31.82**
45-72	MOST profitable	5 yrs	$MVP1$	94.44$^+$	59.09$^+$	68.18$^+$	$MVP1$	95.46$^+$	68.18$^+$	68.18$^+$
		10 yrs	$MVP1$	94.44$^+$	50$^-$	59.09$^+$	$MVP1$	95.46$^+$	31.82$^+$	59.09$^+$
		15 yrs	$MVP1$	83.33$^+$	4.55$^-$	13.64$^+$	$MVP1$	86.36$^+$	4.55$^-$	22.73$^+$
		Mean	–	**90.74**	**37.88**	**46.97**	–	**92.42**	**34.85**	**50**
	LEAST profitable	5 yrs	$MVP15$	27.78$^-$	4.55$^-$	13.63$^-$	$MVP15$	13.64$^-$	13.64$^-$	4.54$^-$
		10 yrs	$MVP15$	61.11$^+$	22.73$^-$	50$^+$	$MVTP$	4.55$^-$	4.55$^-$	4.54$^-$
		15 yrs	$MVTP$	16.17$^-$	50$^-$	22.73$^-$	$MVTP$	4.55$^+$	40.91$^+$	13.64$^+$
		Mean	–	**35.19**	**25.76**	**28.79**	–	**7.58**	**19.7**	**7.58**
26-44	MOST profitable	5 yrs	$GMVP$	–	77.27$^-$	86.36	$GMVP$	86.36$^-$	77.27$^-$	86.36$^-$
		10 yrs	$MVTP$	83.33$^+$	95.46$^+$	86.36$^+$	$MVTP$	86.36$^+$	95.46$^+$	86.36$^+$
		15 yrs	$MVTP$	–	–	–	$MVTP$	–	–	–
		Mean	–	**83.33**	**86.36**	**86.36**	–	**86.36**	**86.36**	**86.36**
	LEAST profitable	5 yrs	$MVP15$	16.67$^-$	50$^-$	4.55$^-$	$MVP15$	13.64$^-$	4.55$^-$	4.55$^-$
		10 yrs	$MVP1$	27.78$^-$	31.82$^-$	68.18$^-$	$MVP1$	50$^+$	40.09$^-$	68.18$^-$
		15 yrs	$MVPT$	–	–	–	$MVP1$	–	–	–
		Mean	–	**22.22**	**40.90**	**36.36**	–	**31.81**	**22.73**	**36.36**

Notes: The majority of best(worst) strategies are associated with high(low) levels of market timing (estimation error). For additional details see Section 6

Table 7: **Most/least profitable cost-optimized strategies and associated market timing percentiles**

		Rolling window length	V	0.5 yrs				1 yr			
Sample	Benchmark	μ		Strategy	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$	Strategy	$\bar{\mu}$	$\bar{V}$	$\bar{\mu}$ & $\bar{V}$
03-17	MOST profitable	5 yrs		$GMVPtc$	–	13.64^{+}	68.18	$GMVPtc$	68.18^{-}	4.54^{-}	59.09^{-}
		10 yrs		$MVP5tc$	72.22^{+}	32.82^{+}	32.82^{+}	$MVP10tc$	86.36^{+}	31.81^{+}	68.18^{+}
		15 yrs		$MVP15tc$	94.44^{+}	59.09^{+}	95.46^{+}	$MVP15tc$	95.45^{+}	40.90^{+}	86.36^{+}
		Mean		–	**83.33**	**34.84**	**65.15**	–	**83.33**	**25.75**	**71.21**
	LEAST profitable	5 yrs		$MVP15tc$	72.22^{-}	40.91^{+}	50^{-}	$MVP15tc$	77.27^{-}	59.09^{+}	68.18^{-}
		10 yrs		$MVP15tc$	94.44^{+}	59.09^{+}	95.46^{+}	$GMVPtc$	22.73^{-}	4.55^{-}	4.55^{-}
		15 yrs		$MVP1tc$	27.78^{-}	4.54^{+}	13.64^{-}	$MVP1tc$	31.81^{-}	13.64^{-}	4.55^{-}
		Mean		–	**64.82**	**34.84**	**53.03**	–	**43.93**	**25.76**	**25.76**
73-02	MOST profitable	5 yrs		$GMVPtc$	–	95.46^{+}	95.46	$GMVPtc$	50^{-}	95.46^{+}	95.46^{+}
		10 yrs		$GMVPtc$	–	95.46^{+}	95.46	$GMVPtc$	59.09^{-}	95.46^{+}	95.46^{+}
		15 yrs		$GMVPtc$	–	95.46^{+}	86.36	$GMVPtc$	13.64^{-}	68.18^{+}	68.18^{+}
		Mean		–	–	**95.46**	**92.42**	–	**40.91**	**86.36**	**86.36**
	LEAST profitable	5 yrs		$MVP15tc$	5.56^{-}	59.09^{+}	31.81^{-}	$MVP15tc$	13.63^{-}	31.82^{-}	4.54^{-}
		10 yrs		$MVP15tc$	5.56^{-}	86.36^{+}	50^{-}	$MVP15tc$	4.55^{-}	86.36^{+}	31.82^{-}
		15 yrs		$MVP15tc$	5.56^{-}	59.09^{+}	59.09^{-}	$MVP15tc$	22.73^{-}	86.36^{+}	77.27^{+}
		Mean		–	**5.56**	**68.18**	**46.97**	–	**13.63**	**68.18**	**37.88**
45-72	MOST profitable	5 yrs		$GMVPtc$	–	95.46^{+}	95.46	$GMVPtc$	59.09^{-}	77.27^{+}	77.27^{+}
		10 yrs		$GMVPtc$	–	86.36^{+}	95.46	$MVP5tc$	59.09^{+}	86.36^{+}	77.27^{+}
		15 yrs		$MVP5tc$	50^{+}	86.36^{+}	86.36^{+}	$MVP5tc$	31.82^{+}	68.18^{+}	68.18^{+}
		Mean		–	**50**	**89.39**	**92.42**	–	**50**	**77.27**	**74.24**
	LEAST profitable	5 yrs		$MVP15tc$	38.89^{-}	68.18^{+}	40.91^{-}	$MVP15tc$	40.91^{-}	59.09^{+}	50^{-}
		10 yrs		$MVP15tc$	5.56^{-}	13.64^{-}	13.64^{-}	$MVP15tc$	50^{+}	50^{+}	50^{-}
		15 yrs		$MVP15tc$	5.56^{-}	13.64^{-}	31.82^{-}	$MVP15tc$	40.91^{+}	50^{+}	77.27^{+}
		Mean		–	**16.68**	**31.82**	**28.79**	–	**43.94**	**53.03**	**59.09**
26-44	MOST profitable	5 yrs		$GMVPtc$	–	86.36^{-}	95.46	$GMVPtc$	95.46^{-}	95.46^{-}	95.46^{-}
		10 yrs		$MVP15tc$	94.44^{+}	68.18^{-}	95.46^{+}	$MVP15tc$	95.46^{+}	68.18^{+}	95.46^{+}
		15 yrs		$GMVPtc$	–	–	–	$GMVPtc$	–	–	–
		Mean		–	**94.44**	**77.27**	**95.46**	–	**95.46**	**81.82**	**95.46**
	LEAST profitable	5 yrs		$MVP15tc$	61.11^{-}	68.18^{-}	40.91^{-}	$MVP15tc$	40.91^{-}	50^{-}	40.91^{-}
		10 yrs		$MVP1tc$	5.56^{-}	4.55^{-}	4.55^{-}	$MVP1tc$	4.55^{-}	4.55^{-}	4.55^{-}
		15 yrs		$GMVPtc$	–	–	–	$GMVPtc$	–	–	–
		Mean		–	**33.33**	**36.36**	**22.73**	–	**22.73**	**22.73**	**22.73**

Notes: The majority of best(worst) strategies are associated with high(low) levels of market timing (estimation error). For additional details see Section 6.

Internet Appendix

A Robustness: The alternative scenarios

In this section we detail the list of alternative scenarios which we use to test the robustness of our analysis conducted in the main text.

A.1 SP500 universe

This scenario is the same as the baseline with the exception that the stock universe considered only includes stocks that belongs to the $SP500$ index. Before March 1957, date in which the $SP500$ index was established, we consider the sample in which in each month we retain the 500 highest stocks by end-of-month market capitalization among all the common stocks traded in AMEX/NYSE. This scenario will enable us to gauge the effect of small stocks (which are now not in the sample) as well as more broadly the effect of diversification (since this is a sub-sample of our baseline universe) in our analysis.

A.2 Quarterly frequency

This scenario only differs from the baseline in that we re-balance our strategies quarterly rather than monthly. Given our monthly holding period returns from CRSP we obtain the quarterly analogs by compounding in each quarter the monthly returns: i.e. the quarter-end q realized return of stock i is defined as $ret_q \equiv \prod_{i=1}^{3}(1 + ret_{m_i})$ where ret_{m_i} is the CRSP monthly holding period return realized at the end of month m_i. In a similar fashion we obtain the quarterly excess returns and the market excess returns from the monthly analogs. To estimate the conditional vector of risk premia following the baseline approach (i.e. via the missing-at-random two-step Fama and MacBeth (1973) procedure developed in Gagliardini et al. (2016)) we feed the algorithm with quarterly data and obtain quarterly estimates. Finally, to estimates the conditional covariance matrix of returns we use the same baseline approach (i.e. we compute the covariance matrix implied by the 1-factor market model, also (improperly) referred to as the CAPM) using all business days available over the previous 6 or 12 months and then transforming the estimates to the quarterly frequency by multiplying by 252 and dividing by 4. This scenario will enable us to gauge the effect of the re-balancing frequency in our analysis.

A.3 $[V, \mu] = [1, 20]y$ universe

This scenario only differs from the baseline in that it restricts the universe of stocks available for trading to those available to a MV investor who estimates the conditional vector of risk premia μ_t and the covariance matrix V_t using the previous 20 and 1 year of data respectively. This universe coincides with one of our 6 combination for the rolling window lengths used to compute μ_t and V_t. Under our filters, this combination requires stocks to have non-missing estimates in 80% of the past 20 and 1 years respectively. Any stock that does not satisfy the former(latter) requirement is excluded in the formation of MV strategies that use stock premia (covariance) information. Therefore, each of the 6 different $\mu_t - V_t$ scenarios has a different implied stock universe. In order to isolate the effect of this artifact we perform our analysis in all the six different scenarios by fixing (restricting) the stock universe to the smallest implied one.

A.4 Without TAQ data

In the baseline scenario, we estimate the conditional vector of proportional costs C_t as in Chen and Velikov (2020) using TAQ data (and ISSM for data in the 80s).[32] The ISSM-TAQ data covers the period 1983-2016. For 2017 and for 1926 through 1983 we use low-frequency spread measurements from the literature as detailed below. In this scenario we estimate these costs using only the low-frequency spread measurements. Specifically again following Chen and Velikov (2020) we compute four different proxies and use the simple average as our spread. The proxies are the Hasbrouck (2009)'s Gibbs sampler estimates, the Corwin and Schultz (2012)'s high-low spread, the Abdi and Ranaldo (2017)'s close-high-low spreads and the volume-over-volatility estimate based on Kyle and Obizhaev (2016)'s micro-structure invariant hypothesis.[33] This scenario mainly allows us to gauge the effect of TAQ data on

[32]Please refer to Chen and Velikov (2020) for additional details.
[33]Please refer to Chen and Velikov (2020) for additional details.

our analysis.

A.5 CAPM μ (Fama and MacBeth (1973) approach)

In this scenario we depart from the baseline only with respect to the way we estimate the vector of conditional risk premia. In particular, instead of using the missing-at-random two-step Fama and MacBeth (1973) procedure developed in Gagliardini et al. (2016), we simply use the classical Fama and MacBeth (1973) procedure without the statistical filter applied in the second step following Gagliardini et al. (2016). This scenario mainly allows us to gauge the effect of the statistical filter applied in the second step in the Gagliardini et al. (2016) methodology.

A.6 FF3 μ and V (Gagliardini et al. (2016) approach)

In this scenario we estimates the conditional vector of risk premia and covariance matrix using the Fama and French (1993) 3-factor model representation instead of the CAPM (or market model) one from the baseline scenario. Ceteris paribus, all estimations are carried out exactly as in the baseline scenario, we are just adding HML and SMB as two additional factors on top of the market factor. This scenario allows us to gauge the effect of a richer model for estimating the first two conditional moments of the returns' distribution.

A.7 Shrinkaged V

In this scenario we depart from the baseline by estimating the conditional covariance matrix of returns via the shrinkage estimator developed in Ledoit and Wolf (2017, 2020). In contrast to the baseline scenario, we use a statistical nonparametric approach for estimating V instead of a model-based approach. This scenario allows us to gauge the effect of the trade-off between estimation error and bias implied in the estimation of V. While shrinkage methods are mostly designed to minimize the bias issue, model-based estimator mostly minimizes estimation error.

A.8 Unconditional universe

This scenario similarly to the "$[V, \mu] = [1, 20]y$ universe" scenario restricts the implied universe of tradeble stocks. This time we look at the stocks that would be unrealistically (i.e. not in real time) available to an MV investor who would have access to the overall in-sample estimates of the first two moments of the returns' distribution. This is a stricter subset of any of the real-time universes implied by our 6 baseline combinations of different rolling window lengths for the estimation of the conditional vector of risk premia μ_t and covariance matrix V_t. This is because in this scenario we require stocks that are not present in the entire sample of reference (mainly the sample 2003-2017) at least 80% of the time to be discarded. The purpose of this scenario is primarily the one of gauging the effect of market timing and estimation error by offsetting the potential bias due to change in the implied stock universe composition which is potentially affecting our baseline analysis as explained at the end of Section 5.2.

B Robustness: Recent mean-variance profitability

We confirm under our alternative scenarios the robustness of the insights from the main text that: i) the profitability of MV strategies has never been as high as recently, ii) $GMVP$ remains among the most profitable textbook strategies, and iii) the cost-optimized strategies remain the best way to efficiently target (higher) risk premia once stock risk premia are estimated using at least 10 years of data. We also find that: 1) there are economically sizable diversification benefits in investing in all common stocks, 2) at least in recent times and for portfolios of thousands of stocks, the trade-off between bias and estimation error in the covariance estimator is clearly in favor of a reduction of the latter, and 3) the second step in Gagliardini et al. (2016) approach that we use in the baseline scenario to estimate the conditional stock risk premia is useful in filtering out a portion of the noise present in the estimates.

As Figure 3 in the main text, Figure 15 plots the average annualized out-of-sample after-cost Sharpe ratios of our analyzed strategies in the eight alternative scenarios described in Section A. As in the main text, the after-cost performances of the MV strategies post 2002 (displayed in green) is found at record heights, and cost-optimized strategies (in dotted green) dominates textbook analogs (in solid green) at targeting higher returns.

As in Figure 5 in the main text, Figure 16 through 23 plot the actual out-of-sample performance for each risk premium - covariance rolling-window-length combination behind the averages displayed in Figure 15 in the eight alternative scenarios described in Section A. In economic terms $GMVP$ remains among the most profitable textbook strategies,[34] and the cost-optimized strategies remain the best way to efficiently target (higher) risk premia once stock risk premia are estimated using at least 10 years of data. The Sharpe ratios of $MVTP5tc$ through $MVP15tc$ are never below 1 and most of the times at least 1.2. The Sharpe ratios of $GMVP$ are also mostly above 1 except when we estimate the conditional covariance matrix via the Ledoit and Wolf (2017, 2020) shrinkage estimator instead of imposing a factor structure: in this case the ratios are between 0.73 and 0.78. Therefore we conclude that, at least in recent times and for portfolios of thousands of stocks, the trade-off between bias and estimation error in the covariance estimator is clearly in favor of a reduction of the latter.

From a statistical point of view, $GMVP$ is not better then the SPDR and IWM ETFs (i.e. the markers are not surrounded by a black square) when the covariance matrix is estimated over the last 6 months, we invest in SP500 or in the unconditional stock universe, re-balance quarterly and do not use TAQ data to estimate transaction costs. Moreover $GMVP$ is not significantly better then any of our benchmarks (EW, VW and the SPDR and IWM ETFs at least at the 10% level) when the covariance estimator is the Ledoit and Wolf (2017, 2020) shrinkage estimator and when we re-balance quarterly and the covariance matrix is our baseline estimated over the past 12 months (i.e. when markers are not surrounded by anything). The performance of our cost-optimized strategies targeting higher returns, once stock premia are estimated over at least the last 10 years, is more stable. $MVP5tc$ through $MVP15tc$ are generally statistically better then our benchmarks (i.e. markers are surrounded by a red star, a diamond and a square for significance with respect to EW, VW and the ETFs respectively) and often statistically better then their analog counterparts (markers surrounded by red circles). Notable exceptions are represented by the scenario

[34] A notable exception might be the unconditional stock universe scenario when risk premia are estimated with a rolling window of 10 years or less. In these cases the performance gap with $MVP1$ appears substantial, more then 0.2 when V_t is estimated over the past 6 months and approximately 0.4 when V_t is estimated annually. However, even in such circumstances the Sharpe ratio of $GMVP$ is quite high and almost 1.

where we re-balance quarterly, in the 20-1 year rolling-window-length combination for the stock premia and their covariances, and under the market model assumption for the return generation process when we estimate the conditional vector of risk premia via the canonical Fama and MacBeth (1973) procedure and use the last 12 months to estimate the covariance matrix. The latter finding confirm the usefulness of the second step of the Gagliardini et al. (2016) in filtering out a portion of the noise present in the estimates.

The last noteworthy insight is about the role of diversification. If we compare the performance in our baseline scenario (Figure 5) where we invest in all common stocks in the NYSE/AMEX and NASDAQ, with the sub-universe of stocks only belonging to the S&P 500 (Figure 16) we notice how investing in $GMVP$ and $MVP5tc$ trough $MVTP15tc$ is consistently more profitable when small stocks are included. That is the Sharpe ratios are all higher under the baseline scenario. Remember that the performance is after cost so it already accounts for the higher cost of trading small stocks. We conclude that there are economically sizable diversification benefits in investing in all common stocks.

C Robustness: Cost-minimization to unlock stock premia information

In Section 5.2 we have argued that the stabilizing role of cost-optimized weights allows to market time stock risk premia by smoothing out the extra noise that prevents MV standard strategies to exploit the information contained in the stock premia. In this section we confirm the robustness of our claim to the evaluated alternative scenarios with the exception of the scenarios where our strategies are implemented quarterly. In such scenario our cost-optimized strategies still outperform but no market timing is detected. We ascribe this lack of findings to the small sample size. We also notice how allowing for a richer model for the risk premia and the covariance structure (specifically the Fama and French (1993) 3-factor model versus the single-factor market model, or CAPM) allows our cost-optimized strategies to additionally engage in covariance timing, while preserving their ability to time risk premia (coming from the stabilizing role played by the cost minimization). Finally, as in the previous section we find additional evidence in the ability of the Gagliardini et al. (2016) procedure to filter out noise in the risk premia estimates through the second step.

As in Section 5.2, our claim is based on the following findings: i) our most profitable cost-optimized strategies, those targeting higher returns, dominates their textbook analogs in terms of after-cost out-of-sample Sharpe ratios, ii) their risk primia timing abilities are substantially higher, iii) the contribution of risk premia timing is substantially higher than that from timing the covariances for our most profitable cost-optimized strategies, and iv) our most profitable textbook strategies market time stock covariances. For our baseline setup findings ii) through iv) were displayed in Figure 7 and 8 via performance ratios,[35] while finding i) was displayed in Figure 5 via our canonical metric: the out-of-sample after-cost Sharpe ratio. In an analog fashion Figure 24 through 31 present the robustness for findings ii) through iv), while Figure 16 through 23 report the robustness for i) relative to our alternative scenarios described in Section A.

Findings i) through iv) are in general robust with one major exception: the scenario where we re-balance at the quarterly frequency. In such scenario, despite finding that our cost-optimized strategies dominates their standard analogs (especially when we estimate the covariance matrix using the last 6 months of data), we do not detect any statistically significant sign of market timing from them. This lack of results might be due to the reduced sample size which under our quarterly re-balancing scheme is now a third of the main sample with only 60 observations. Other noticeble patterns are those emerging from the scenario where risk premia are estimated via Fama and MacBeth (1973) imposing a CAPM representation and the scenario where we estimate the return moments under the Fama and French (1993) structure. In the latter scenario, while cost-optimized strategies targeting higher returns still dominates their textbook analogs (with performances in the same order of magnitude of those from the baseline scenario) and display substantially higher risk premia timing activity, they also display a considerable amount of covariance timing (in the same order of magnitude of that from risk premia). Therefore, allowing for a richer covariance structure allows our cost-optimized strategies to additionally engage in covariance timing, while preserving their ability to time risk premia (coming from the stabilizing role played by the cost minimization). Finally, the noise filtering second step in the Gagliardini et al. (2016) used in our baseline also has an impact on the stock premia timing abilities of cost-optimizing strategies: while the risk premia timing abilities from $MVP10tc$ and $MVP15tc$

[35] Defined as the Sharpe ratio differentials described in Section 2.3 divided by the real-time performance of our strategies.

are substatially higher then those (if any) form $MVP10$ and $MVP15$ when the covariance matrix is estimating using the previous 6 months of data, unlike in our baseline scenario such difference disappears when the covariance matrix uses the past year of data. In line with what discussed in Section B, this finding corroborates the ability of the Gagliardini et al. (2016) procedure to filter out noise in the risk premia estimates through the second step.

D Robustness: Cost-minimization to reduce downside risk

As in the main text (see Section 5.3), we confirm how cost minimization is also effective in reducing downside risk as measured via the Maximum Draw Down (MDD), or the average time needed to recover from the worst incurred loss, in virtually all the alternative scenarios detailed in Section A.

Analogously from the top graph of Figure 11, Figure 32 reports the differences in absolute MDDs between mean-variance cost-optimized and standard analog strategies for the alternative scenarios. Except from $MVP1$ under the scenario where the conditional return moments are those implied by the Fama and French (1993) 3-factor model and μ_t is computed using the past 5 years of data, cost optimized strategies enjoy (much) lower worst losses as measure by their MDDs.

Similar unreported graphs, available upon requests, also show how cost optimized strategies continue to br able to recover faster from their worst incurred losses. Here the only exception is represented by the scenario where we restrict the stock universe to the $SP500$ constituents for the case where μ_t is estimated using the past 5 years of data.

E Robustness: Cost-minimization to scale up

In Section 5.4 we have shown that minimizing over costs enhances the scalability of our mean-variance strategies. In particular, handling price impact as a type of cost, quadratic in the size of the trades, consistently delivers higher Sharpe ratios at any analyzed (average) level of price pressure. Since quadratic costs are a popular why to handle price pressure,[36] we concluded that cost optimized mean variance strategies appear profitable options to gain exposure on the market even for large institutional investors.

In this section we show that this insight is robust to: the way we measure costs, the presence of small stocks, and the rebalancing frequency of our portfolios. We do so by focusing on the scenarios where: costs are estimated without using TAQ data, stocks are restricted to be $SP500$ constituents, and our strategies are rebalanced quarterly. As in the baseline scenario, the performance gap between cost-optimized and standard analogs is on average monotonically increasing in the price pressure severity (or equivalently in the dollar size of the strategies).[37]

As Figure 14 in the main text, Figure 33-35 plot the differences in annualized after-cost Sharpe ratios between our cost-optimized and standard analog strategies as a function of the price impact parameter π. Starred markers highlight differences that are significant at the 10% level. The range for π covers the interval $[0, 1.6]$; this is in contrast with the interval $[0, 3.4]$ adopted in the main analysis. For the case of the $SP500$ scenario there is a precise reason,[38] while for the other cases we saved on computational time.[39]

Estimating costs without using TAQ data, discarding small stocks or rebalance quarterly do not change the qualitative dynamics for the price impact found in the main analysis. Adopting cost-optimized strategies enhance the after-cost performance at any level of π (all the plotted Sharpe ratio differentials are positive), the benefits are increasing in the price pressure severity (most differentials are monotonically increasing in π), and the effects are more marked when we estimate the conditional covariance matrix of stock returns over the last 6 rather 12 months.

[36] See Dybvig and Pezzo (2020) and references therein.

[37] We did not perform the robustness on the remaining scenarios detailed in Section A due to the computationally intensive nature of such checks. We are willing to perform such robustness upon requests.

[38] Analogously to the baseline scenario but only using stocks that belongs to the $SP500$, we linearize the square-root price impact function from Figure 2 of Frazzini et al. (2015) and use the trade sizes of our analyzed strategies in the absence of price impact to derive $\overline{\pi}$ as a function of portfolio sizes ranging from 1 million to 10000 billions USD. Inspecting such function reveals how after a size of 10 billions USD, $\overline{\pi}$ caps at 1.55. The entire procedure is detailed in footnote 24 in the main text.

[39] The interval length can be broaden upon request.

F Robustness: Market timing and estimation error as drivers of mean-variance profitability

In this section we extend the analysis conducted in Section 6 to our alternative scenarios described in Section A. We confirm the main insights that market timing and estimation error have non-negligible impacts on the mean-variance performance and share very similar dynamics. We also learn how estimating risk premia through a parsimonious specification for the return representation while filtering out part of the noise following Gagliardini et al. (2016) helps identifying such patterns.

F.1 Magnitudes of market timing and estimation error

There is quite some variability in the magnitudes of market timing and estimation error across our scenarios, however the general tendency that emerges is in line with the insights from the main text. Most importantly their economic impact is quite substantial, ranging form average Sharpe ratios increments between 0.08 and 0.50 due to market impact and average decrements between 0.35 and 0.61 due to estimation error. Moreover average market timing magnitudes are consistently high across scenarios in expansions (almost always above 0.16 with an average of 0.27 for the subsample 45-72 and never below 0.37 with an average of 0.50 for the most recent subsample 03-17) and low in recessions - signaling a substantial impact of estimation error (with Sharpe ratio decrements of at least 0.28 in the subsample 26-44 and at least 0.43 in the subsample 73-02).

In line with the stabilizing role of cost optimization found in the main analysis, we also find that market timing is (somewhat) higher for transaction-cost optimized strategies while the opposite is true for estimation error (with magnitudes higher in absolute terms for standard strategies).

This results are shown in Table 9 for the first 7 scenarios detailed in Section A (with the exception of the "unconditional universe" scenario).[40] The table reports for each of the analyzed scenario the average market timing and estimation error magnitudes across the

[40] Which is very similar in nature to the scenario "$[V, \mu] = [1, 20]y$" and which results, when analyzed in Section C in the most recent subsample 03-17, are found very similar to the baseline.

three types of comparisons where the real-time strategies are benchmarked against their unimplementable analogs with fixed in-sample first, second and combined first and second moments of the returns' distribution.[41]

F.2 Pro-cyclical increasing profitability

Once established that market timing and estimation error robustly have non-negligible impact on the performance of our strategies over the last 100 years, we analyze in the alternative scenarios described in Section A the characteristics of such performance.

Figure 15 confirms how, as in the baseline scenario presented in Figure 3 and correctly anticipated by the ex-ante dynamics of the MV investment opportunities discussed in Section 3.1, the performance of our analyzed strategies appears pro-cyclical with magnitudes increasing in expansions. For extra details on the graphs please refer to Section 6.2.

We therefore retain the baseline conclusion that the performance of our strategies is high in expansionary periods where trading frictions (measured by median bid-ask spreads) are low and low in recessionary periods where the same trading frictions are high.

F.3 Pro-cyclical increasing market timing and counter-cyclical estimation error

We confirm in our analyzed alternative scenarios described in Section A that: 1) aggregate market timing, defined as the sum of all the individual strategies' impacts, is in general pro-cyclical and increasing over time and 2) aggregate estimation error, defined as the sum of all the individual strategies' impacts, is counter-cyclical with magnitudes decreasing over time.

Table 9 displays the results for the first 7 scenarios detailed in Section A (with the exception of the "unconditional universe" scenario).[42] The table reports for each of the analyzed scenario the sum of market timing and estimation error magnitudes across the

[41] That is, for each analyzed scenario we only report the columns labeled "Mean" in Table 2 from the main text.

[42] The last scenario is by construction very similar to the third and thus left out for brevity from the current analysis. Results upon request are available but in line with all the other presented in this section.

three types of comparisons where the real-time strategies are benchmarked against their unimplementable analogs with fixed in-sample first, second and combined first and second moments of the returns' distribution.[43] At it is immediately apparent by looking at the "Total" columns, market timing across scenarios is higher in the expansionary periods 45-72 and 03-17 with the level being higher in the most recent sample and estimation error is lower in the recessionary periods 26=44 and 73-02 with the level being lower in the most recent sample. These patterns, when detected (i.e. when different from 0) are reflected in the majority of the analyzed scenarios for both cost-optimized and traditional mean-variance strategies.[44]

As in the baseline scenario, this patterns suggest that market timing and estimation error are important drivers behind mean-variance profitability over the last 100 years in the U.S. stock market.

F.4 Correlation analysis

In this subsection we estimate the correlation between market timing (or negative estimation error) and mean-variance profitability for our alternative scenarios. With the exception of the "$CAPM\ \mu$" and "$FF3$" scenarios, we confirm the existence of a meaningful positive relationship. The exceptions highlight the delicate role played by the risk premia estimation which are more transparent when parsimoniously estimated as a one factor (market) model via the noise filtering procedure of Gagliardini et al. (2016).

As in the baseline scenario, we measure the correlation of the Sharpe ratio differential between a given strategy and its unimplementable analog where at least one of the first two moments of the returns' distribution are fixed at their in-sample levels with the Sharpe ratio level of the implementable strategy. Table 10 shows the average correlation across risk premia estimates (with rolling windows of 5,10 and 20 years) and sample periods (26-44 through 03-17) for each analyzed alternative scenario in the case we fix the first, second or both moments of the returns' distribution under the estimates for the covariance matrix V

[43] That is, for each analyzed scenario we only report the columns labeled "Total" in Table 3 from the main text.

[44] Notable exceptions are the "*quarterly*" and "$CAPM\ \mu$" scenarios for the cost-optimized strategies and the estimates from the sample period 26-44 in the "Shrinkaged V" scenario for the standard strategies.

using rolling windows of either 6 or 12 months.

Concentrating on the rightmost panel reporting the combined effects of the first and second moments of the returns distribution, i.e. the correlation between market timing (negative estimation error) induced by both risk premia and covariances and the mean-variance strategies' performance, we mostly observe a positive relationship. In particular, whenever we estimate risk premia through the market model (also improperly referred to as the CAPM model) under the noise filtering procedure of Gagliardini et al. (2016) as in the baseline setup, we find a positive association between market timing and performance no matter if: we discard small stocks ("$SP500$" scenario), we trade quarterly ("*quarterly*" scenario), we reduce the implied stock universe to that available if we estimate V and μ using a 1 and 20 year rolling window respectively ("$[V\mu] = [1, 20]y$" scenario), we discard TAQ data when measuring the costs ("*No TAQ*" scenario), or we estimate V through the shrinkaged estimator developed in Ledoit and Wolf (2017, 2020) ("*Shrinkaged V*" scenario). While such combined contributions predominantly come from timing risk premia in the "$SP500$" scenario and from using a rolling window of 1 year for V for the case of standard strategies in the "$[V\mu] = [1, 20]y$" scenario, they more generally come from both risk premia and covariance timing in isolation for the rest of the cases (see the leftmost and the middle panels). Moreover, as in the main scenario we notice how cost-optimized strategies displays a higher correlation than their standard analogs.

Lastly, it is interesting to notice how negative the correlation is for the "$CAPM\ \mu$" and the "$FF3$" scenarios where risk premia are *not* estimated via the market model following the Gagliardini et al. (2016) procedure. A negative relationship for the "$CAPM\ \mu$" scenario is consistent with the pattern of the aggregate market timing and estimation error found in Table 9 where for standard strategies we found a pro-cyclical pattern for market timing but partially fail to find a counter cyclical pattern for estimation error (due to the measurements in the 26-44 subsample), and for cost-optimized strategies we find an unusually high amount of market timing in the recessionary 73-02 period and fail to detect estimation error everywhere except in the 26-44 period. For the case "$FF3$" scenarios a negative correlation appears instead in contrast with the results from Table 9. These findings are in line with the well-known presence of noise in the risk premia estimates obscuring the underlying patterns. It therefore seems from our analysis that adopting a parsimonious specification for the return representation and filtering out part of the noise through the Gagliardini et al. (2016) is the most efficient approach among the one analyzed in this paper.

Robustness: Profitability of Mean-Variance Strategies

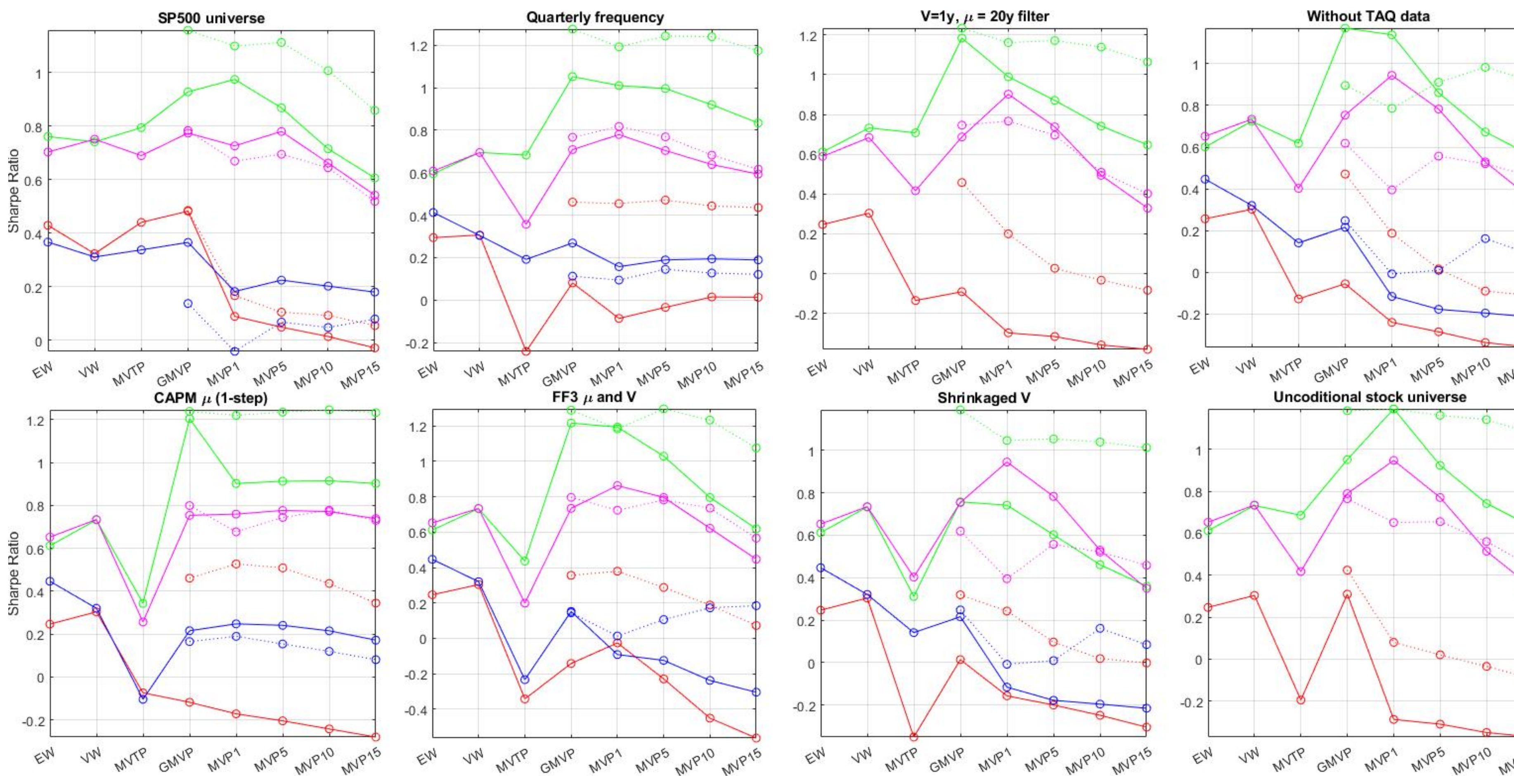

Figure 15: As in Figure 3 the profitability of our MV strategies (described in Figure 3 and in Section 2.1) is pro-cyclical and never so high as recently. Solid(dashed) lines refer to textbook (cost-optimized) strategies. The sample is every month from January 1926 to December 2017. The average annualized out-of-sample Sharpe ratios are plotted over four nonoverlapping subsamples: 1926-44 (blue), 1945-72 (magenta), 1973-02 (red) and 2003-17 (green).

Recent Profitability of Mean-Variance Strategies: S&P 500 stocks

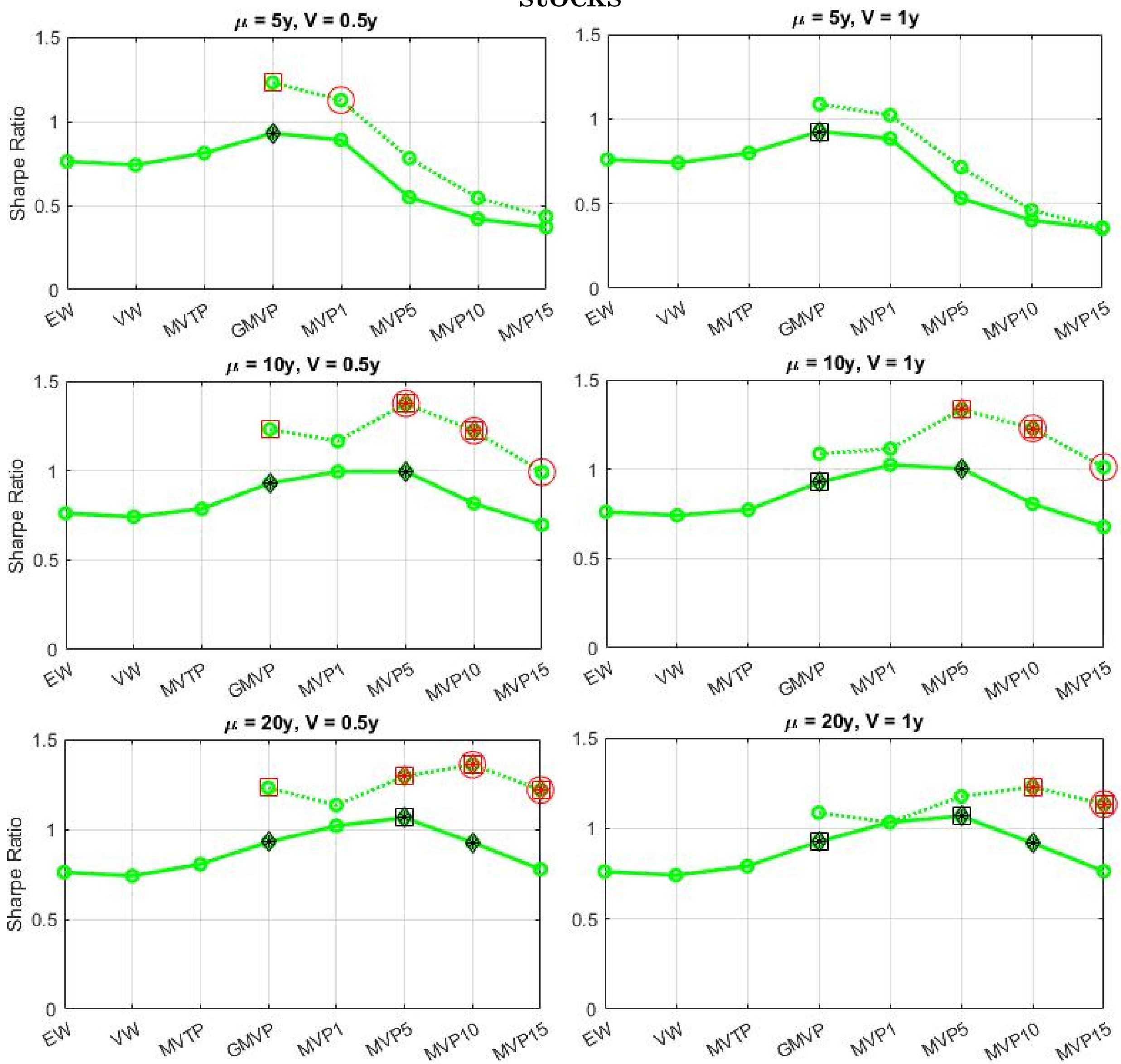

Figure 16: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: quarterly re-balancing

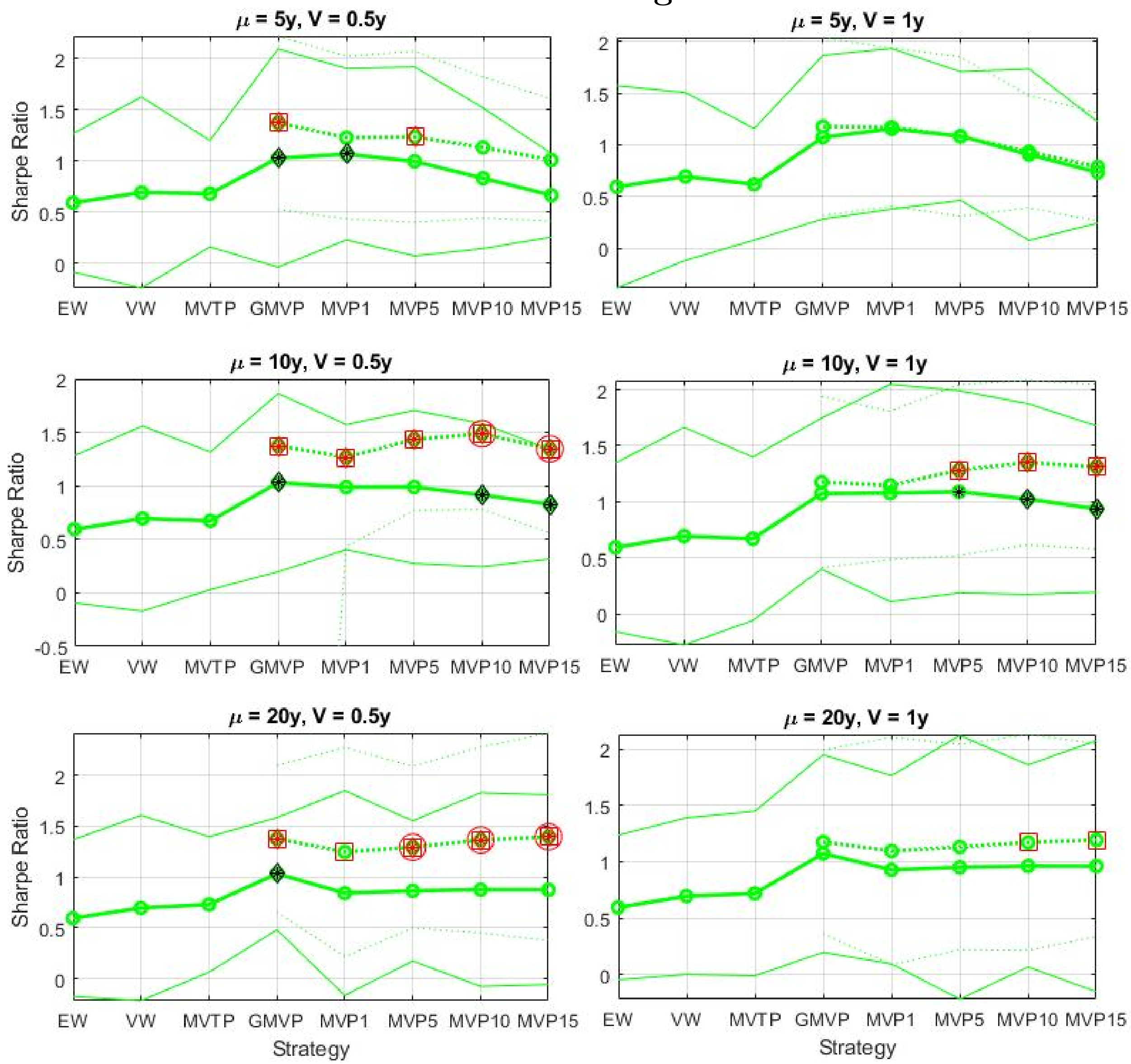

Figure 17: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: 20-1 years risk premia - covariance rolling-window-length combination implied universe

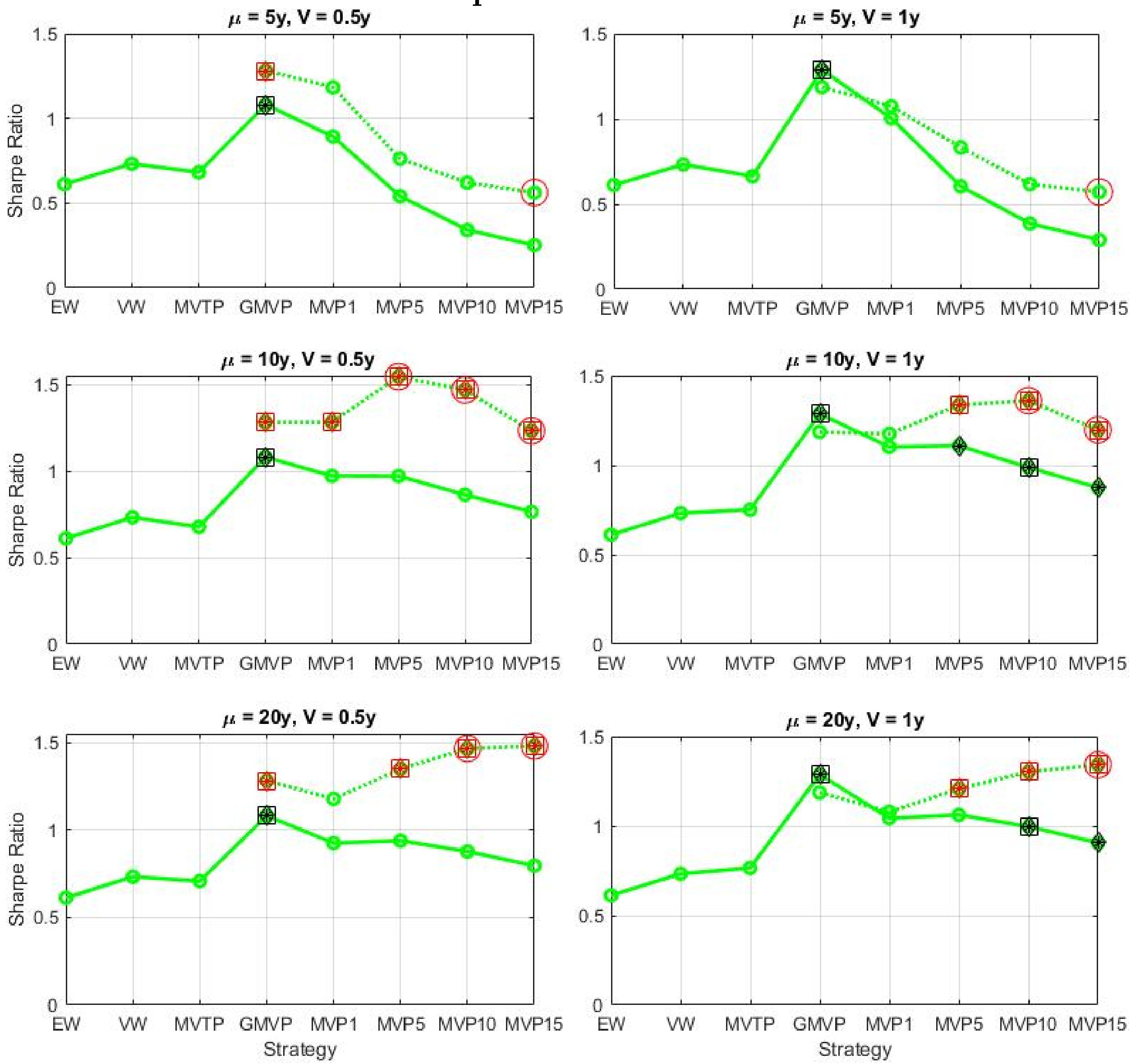

Figure 18: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: estimating costs without TAQ data

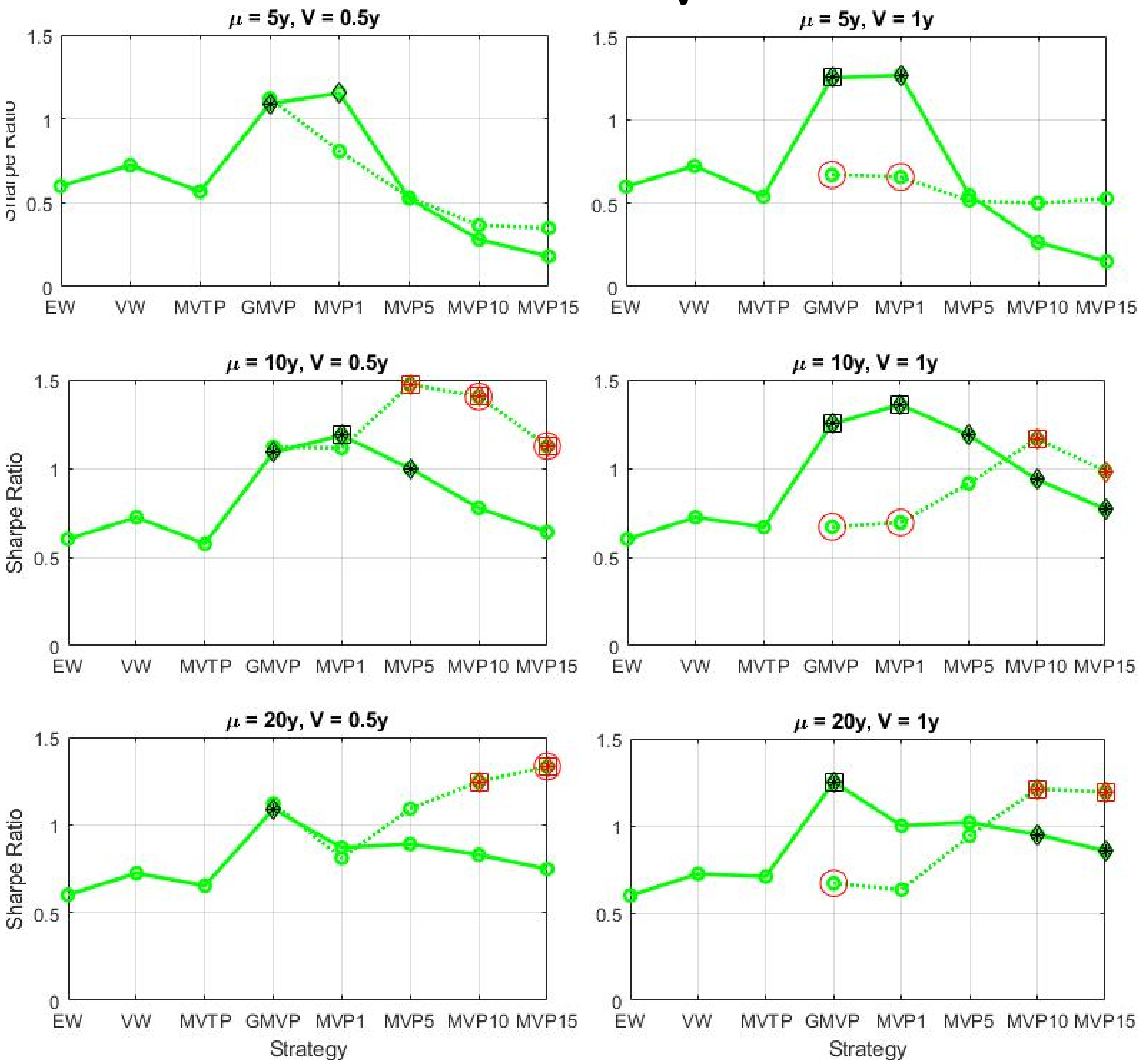

Figure 19: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: estimating CAPM risk premia via Fama and MacBeth (1973)

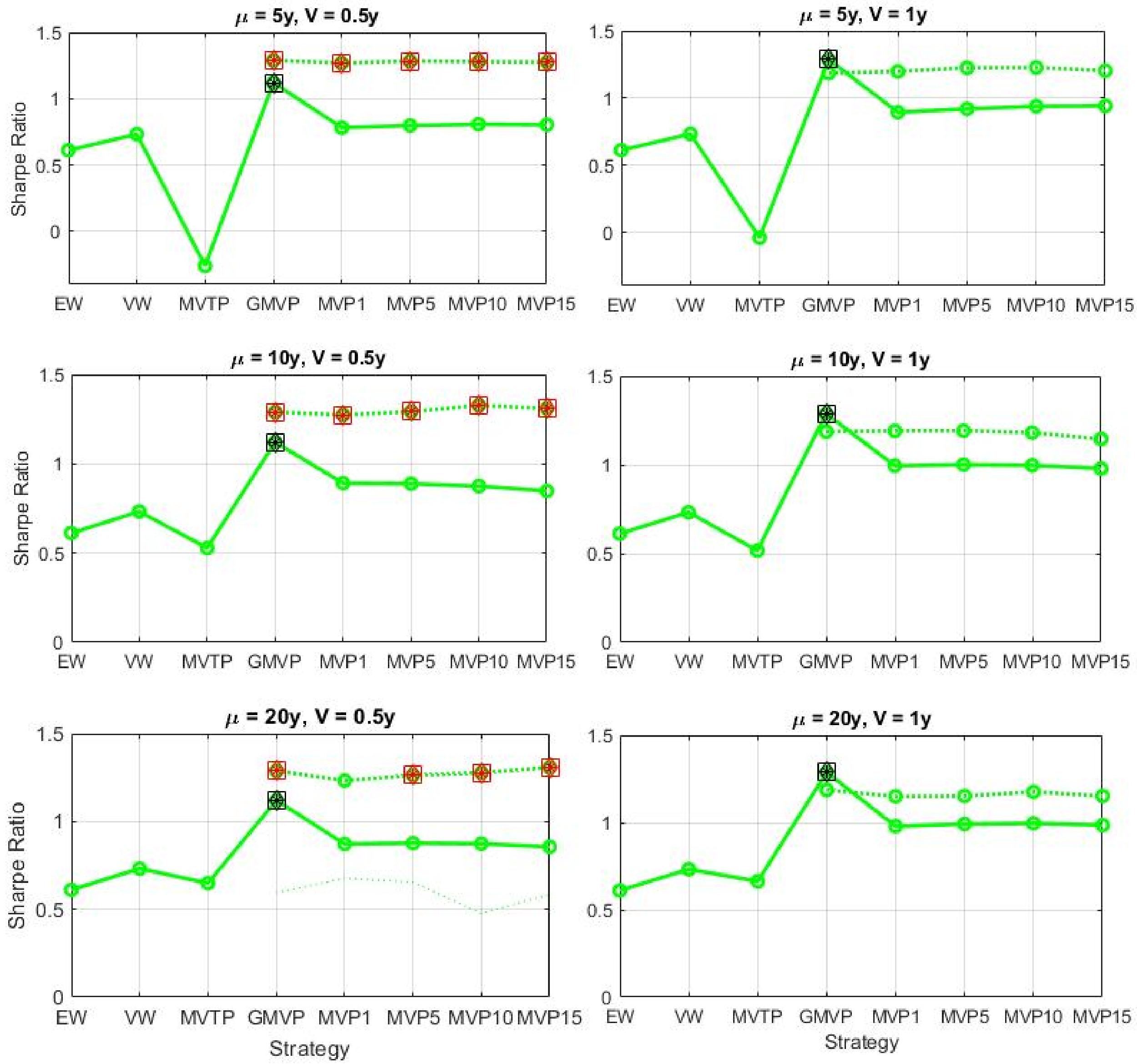

Figure 20: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: estimating Fama and French (1993) risk premia and covariances

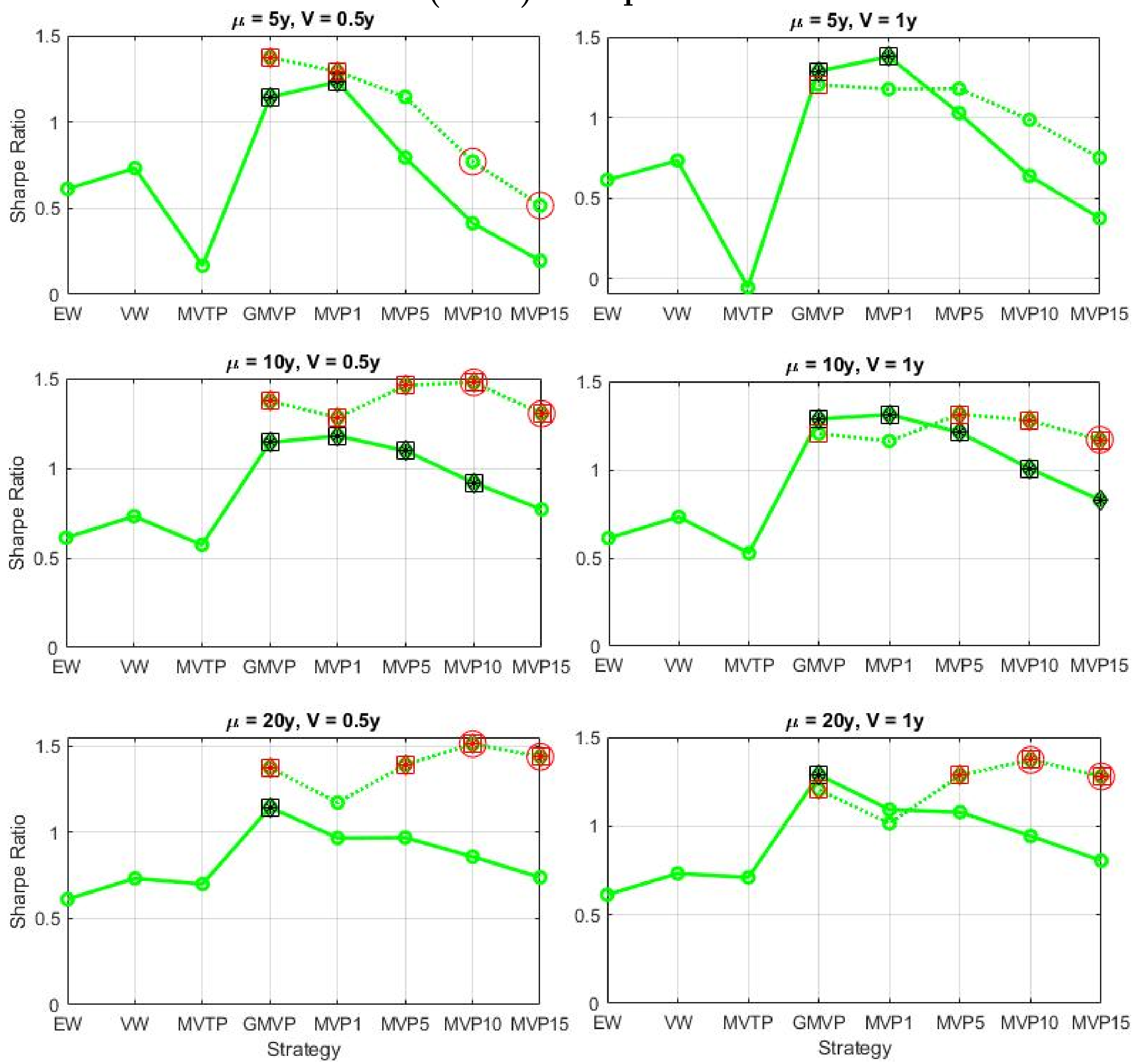

Figure 21: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: estimating the covariance matrix via the Ledoit and Wolf (2017, 2020) shrinkage estimator

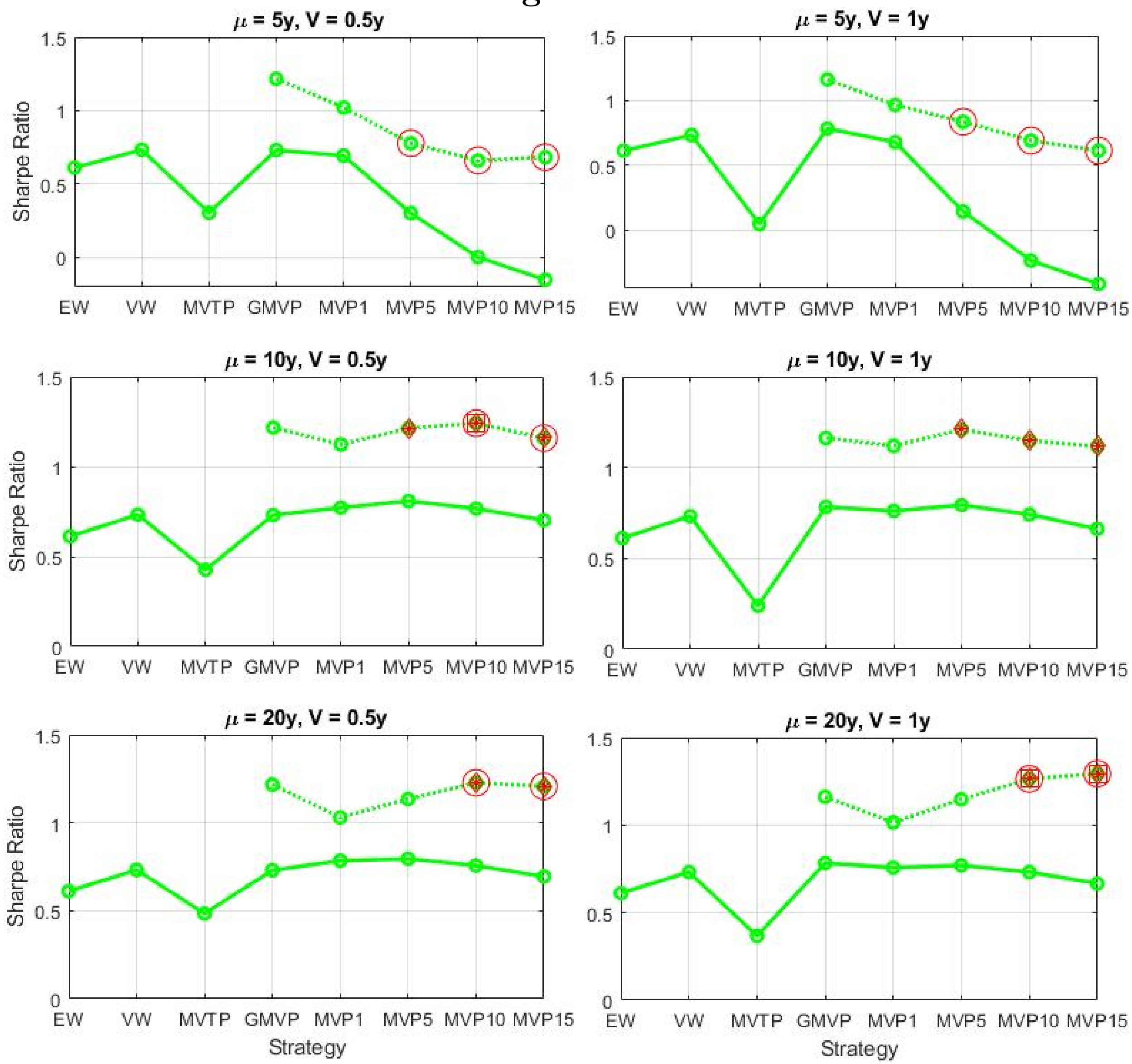

Figure 22: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Recent Profitability of Mean-Variance Strategies: unconditional stock universe

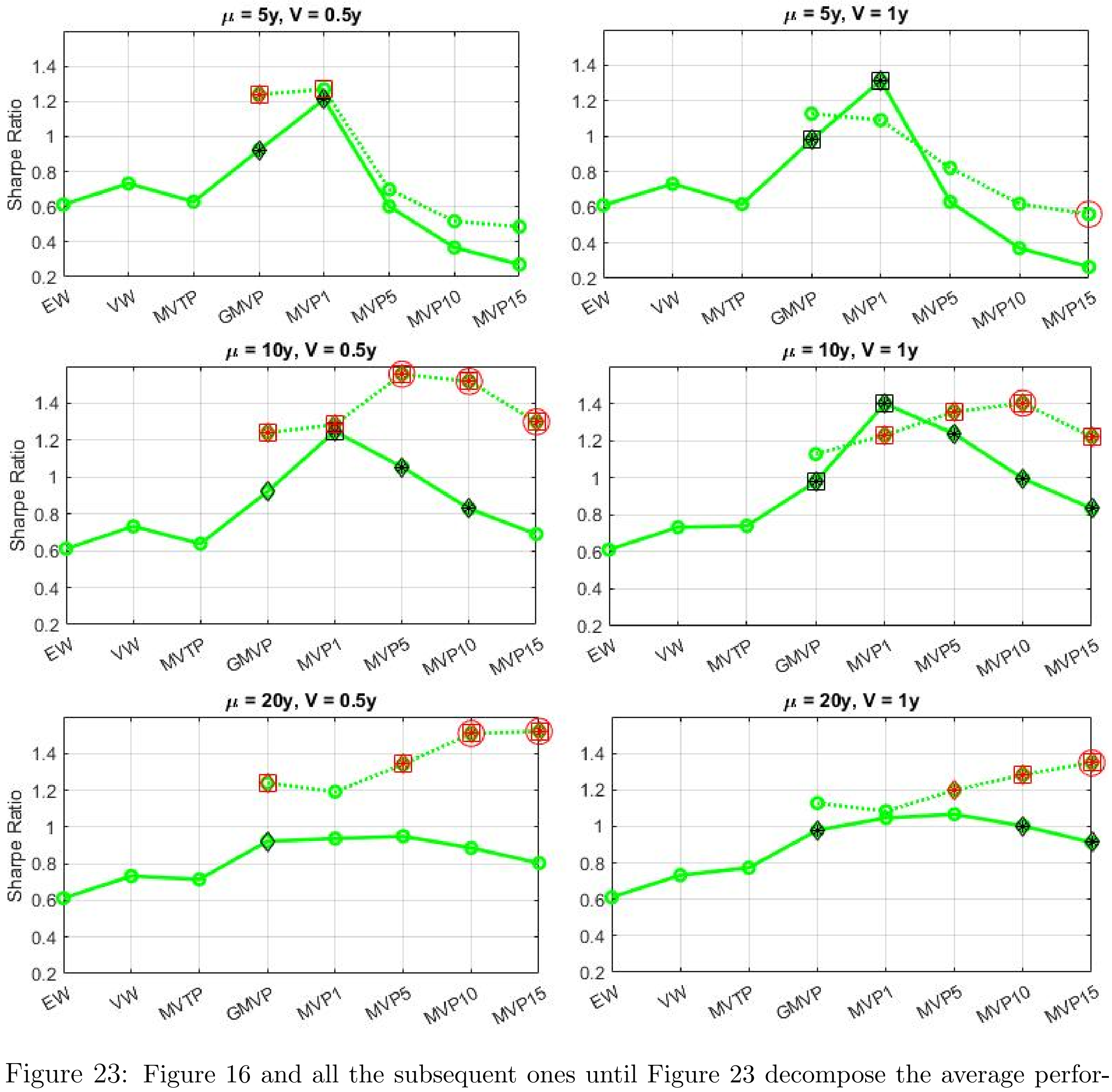

Figure 23: Figure 16 and all the subsequent ones until Figure 23 decompose the average performance shown in Figure 15 across the six baseline risk premia - covariance rolling window-length configurations over every month (or quarter for Figure 17) between January 2003 and December 2017. Each of these 8 figure refer to a different alternative scenario as detailed in Section A.

Robustness: cost-minimization to unlock stock premia information. Scenario: S&P500 universe

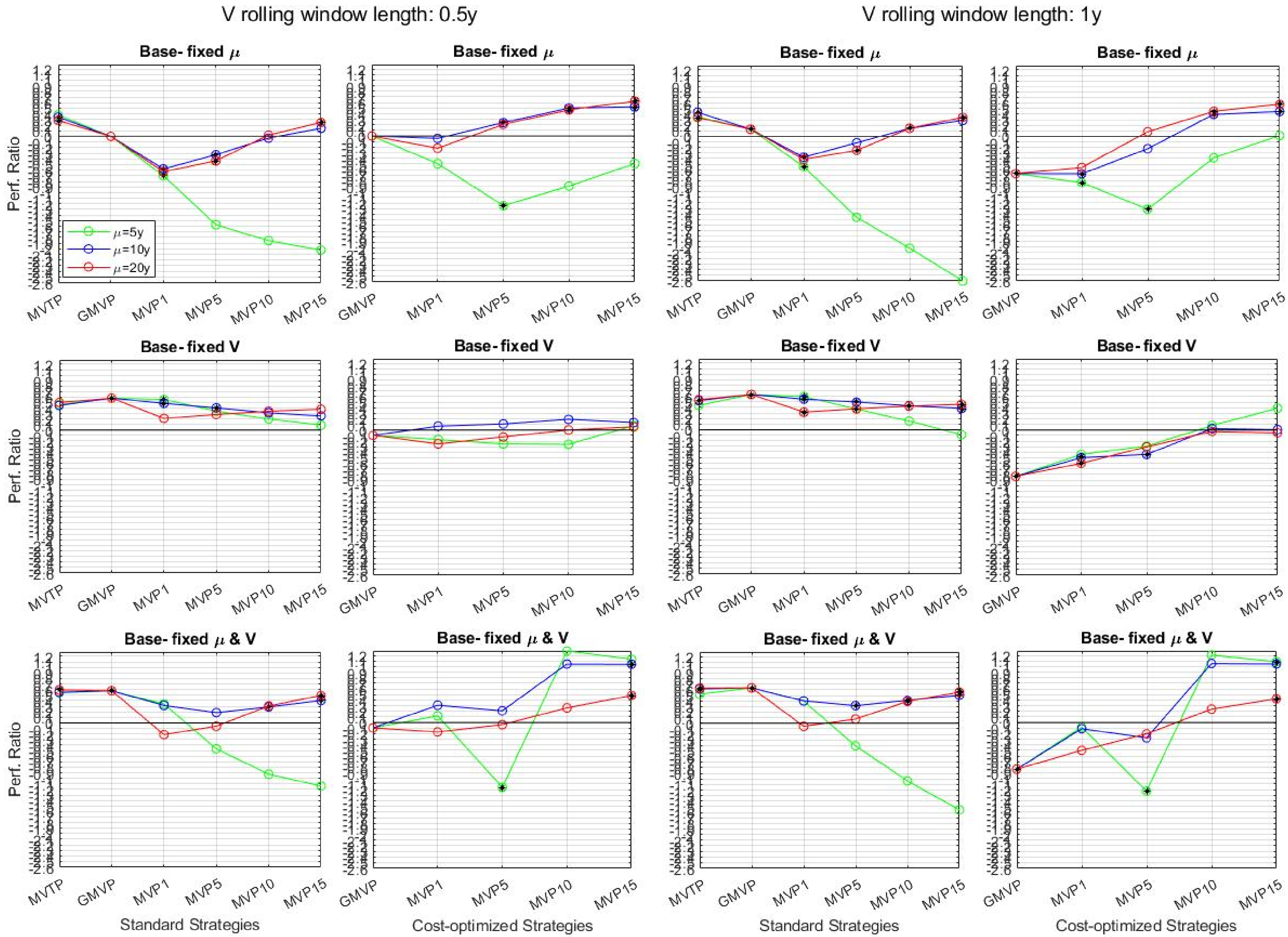

Figure 24: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: quarterly re-balancing

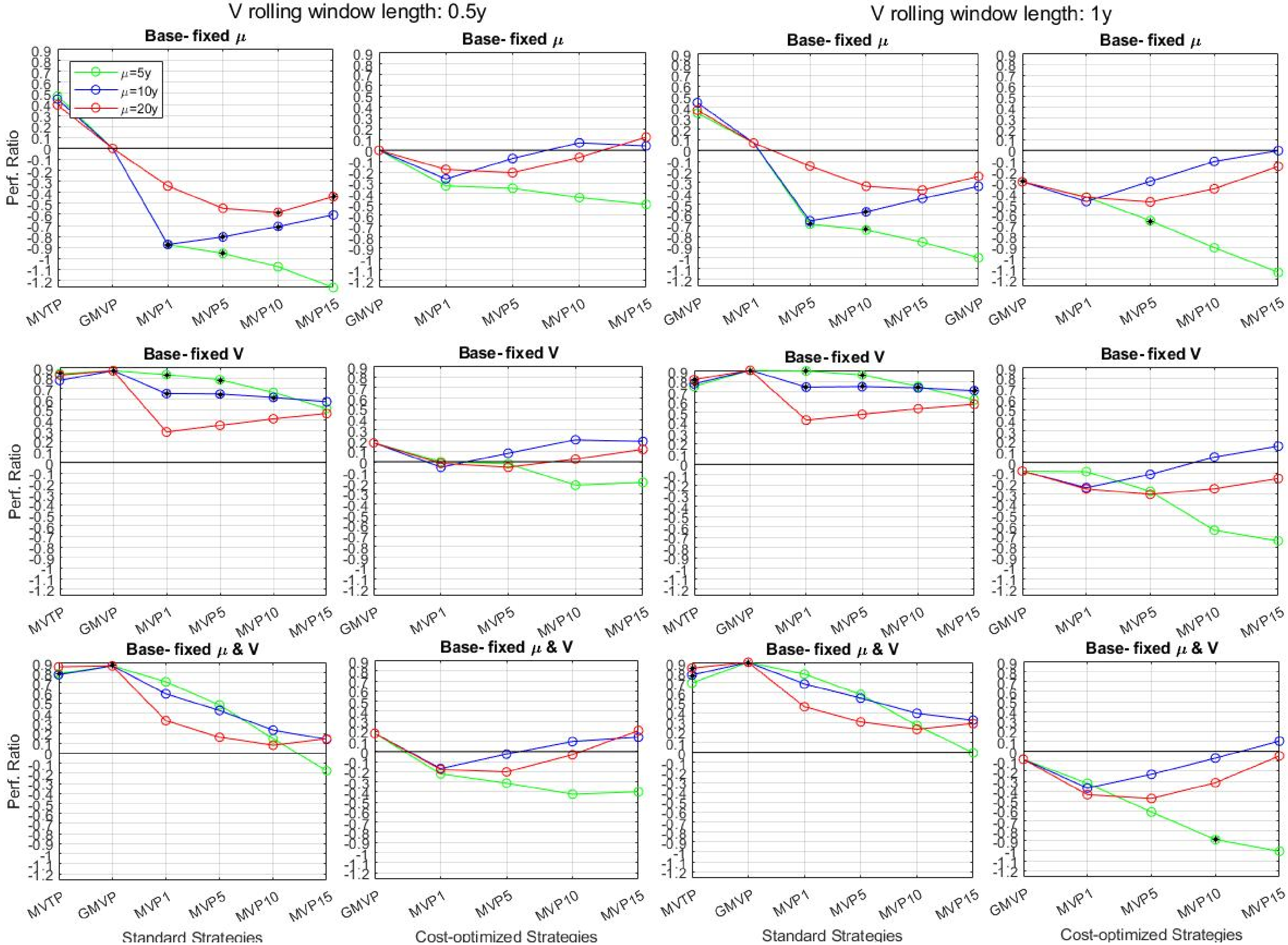

Figure 25: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: 20-1 years risk premia - covariance rolling-window-length combination

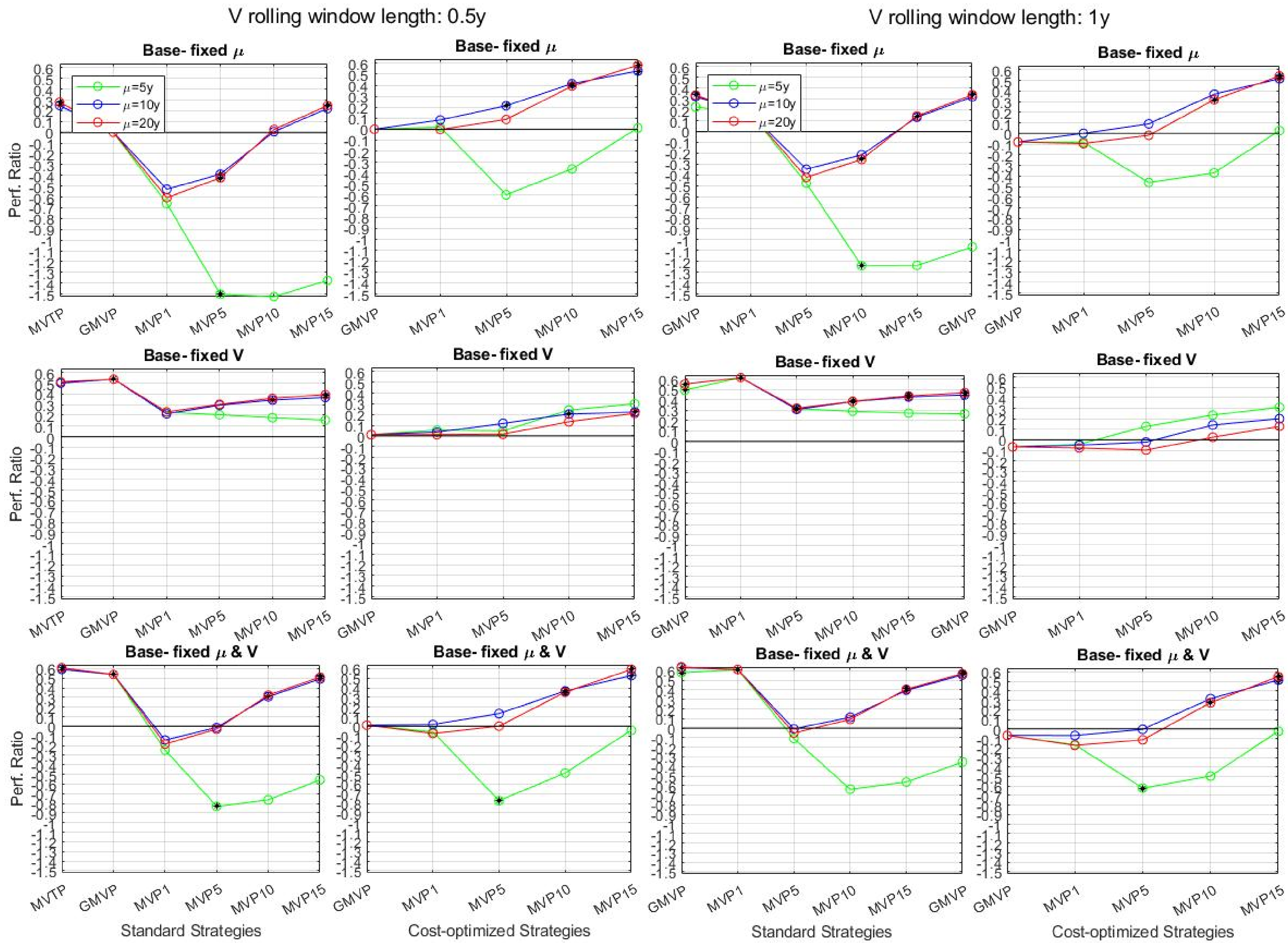

Figure 26: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: transaction cost estimates without TAQ data

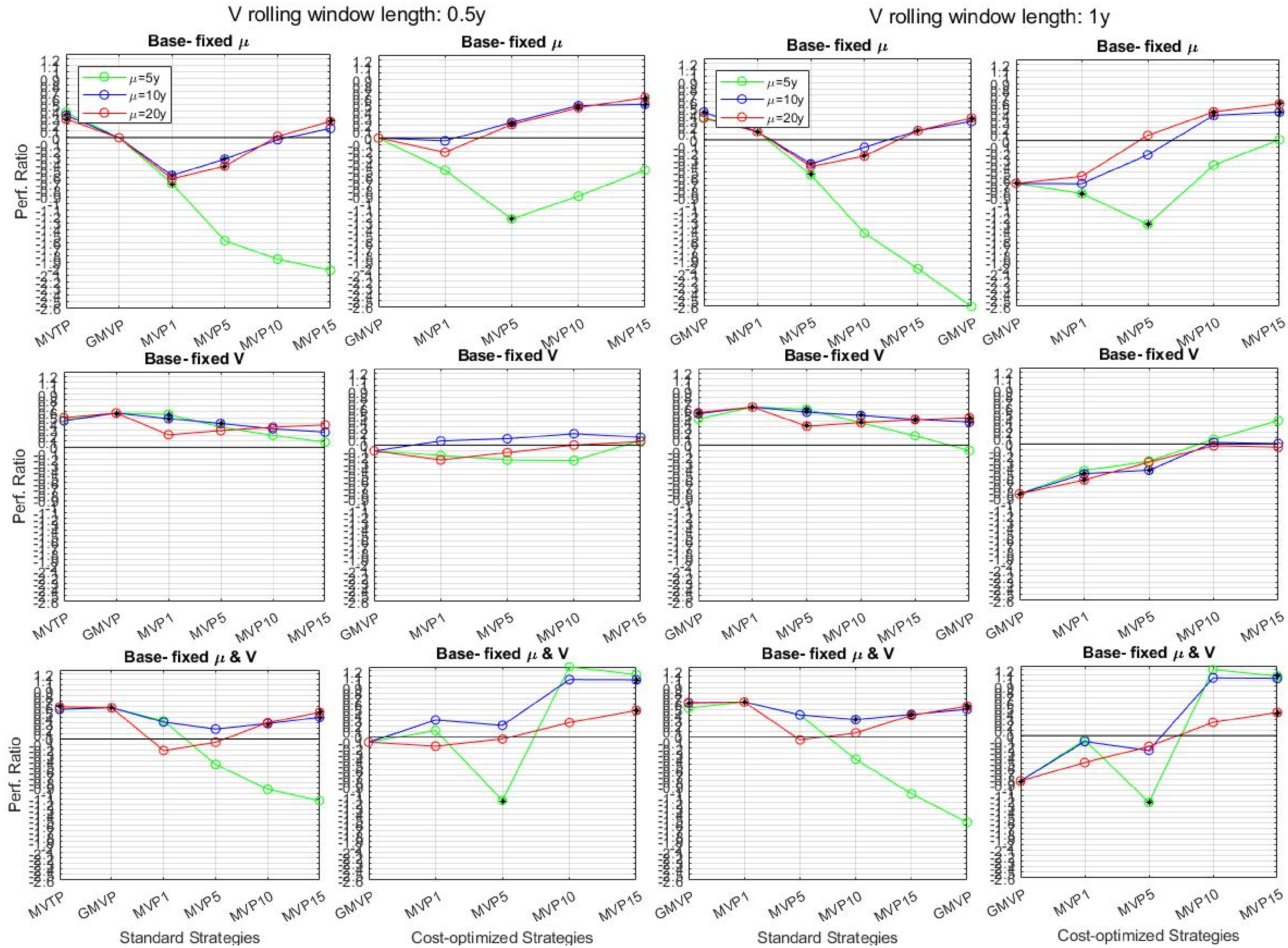

Figure 27: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: CAPM risk premia via Fama and MacBeth (1973)

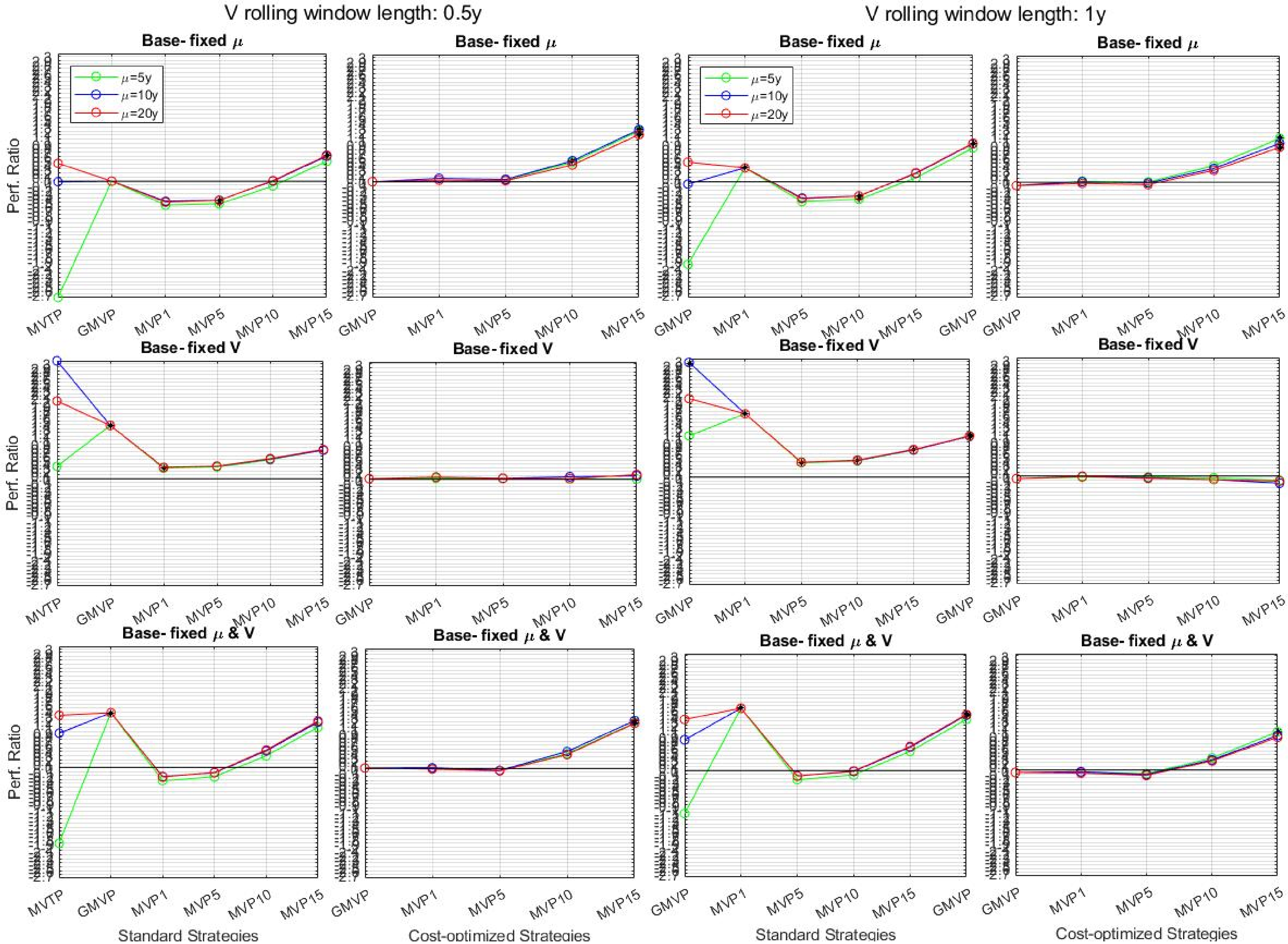

Figure 28: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: Fama and French (1993) risk premia and implied covariances

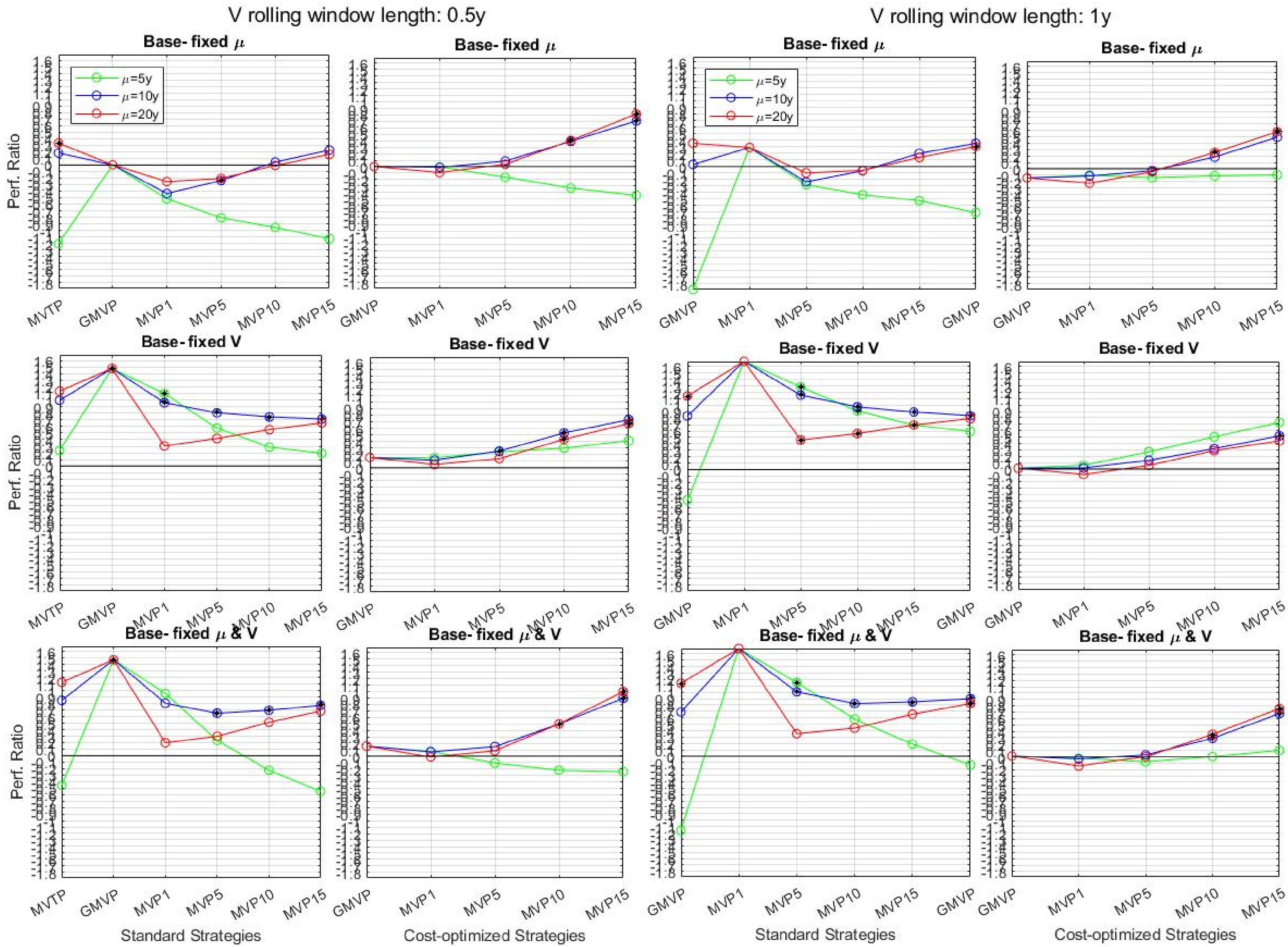

Figure 29: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: Ledoit and Wolf (2017, 2020) shrinkaged covariances

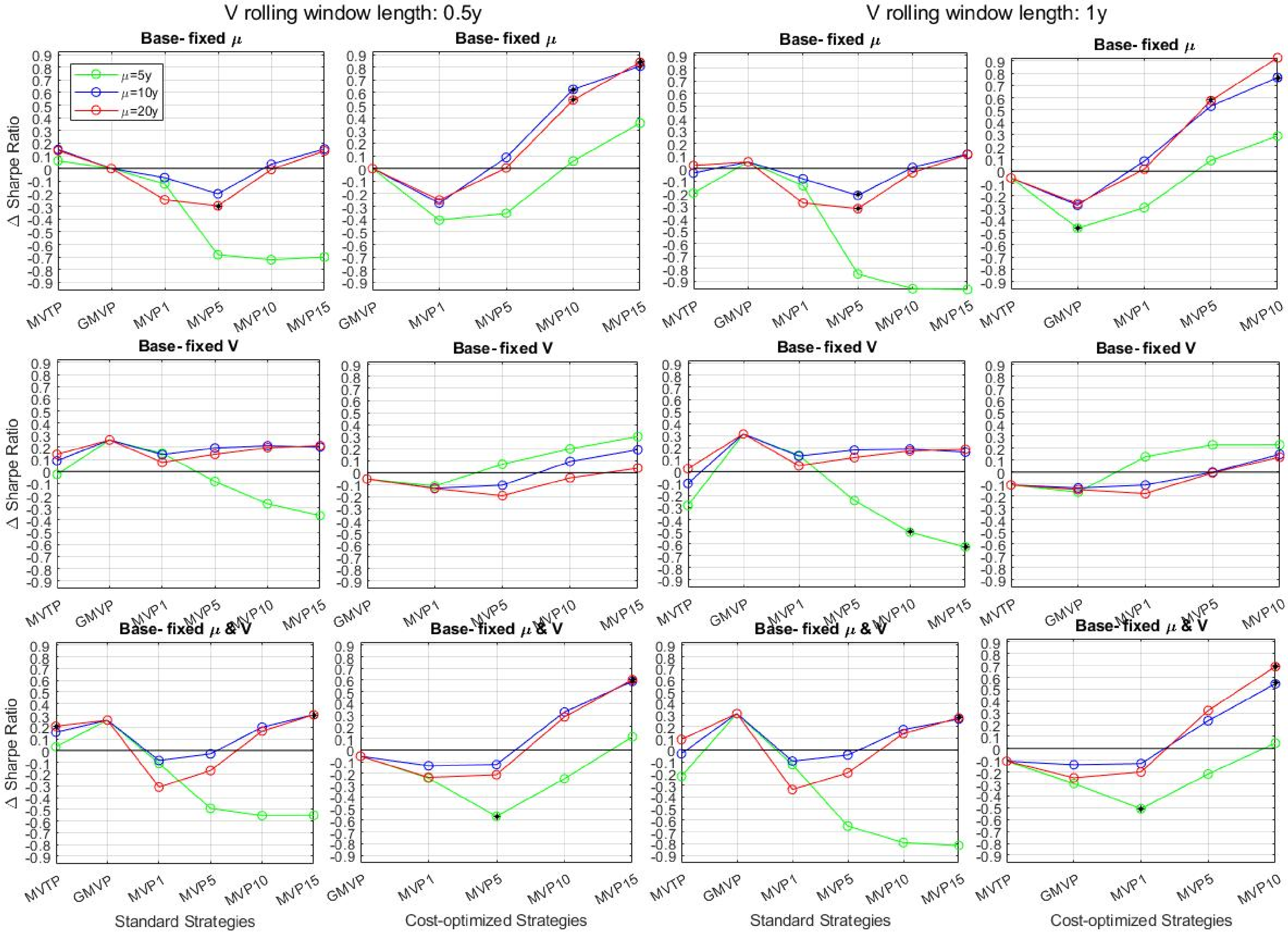

Figure 30: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: cost-minimization to unlock stock premia information.
Scenario: unconditional universe

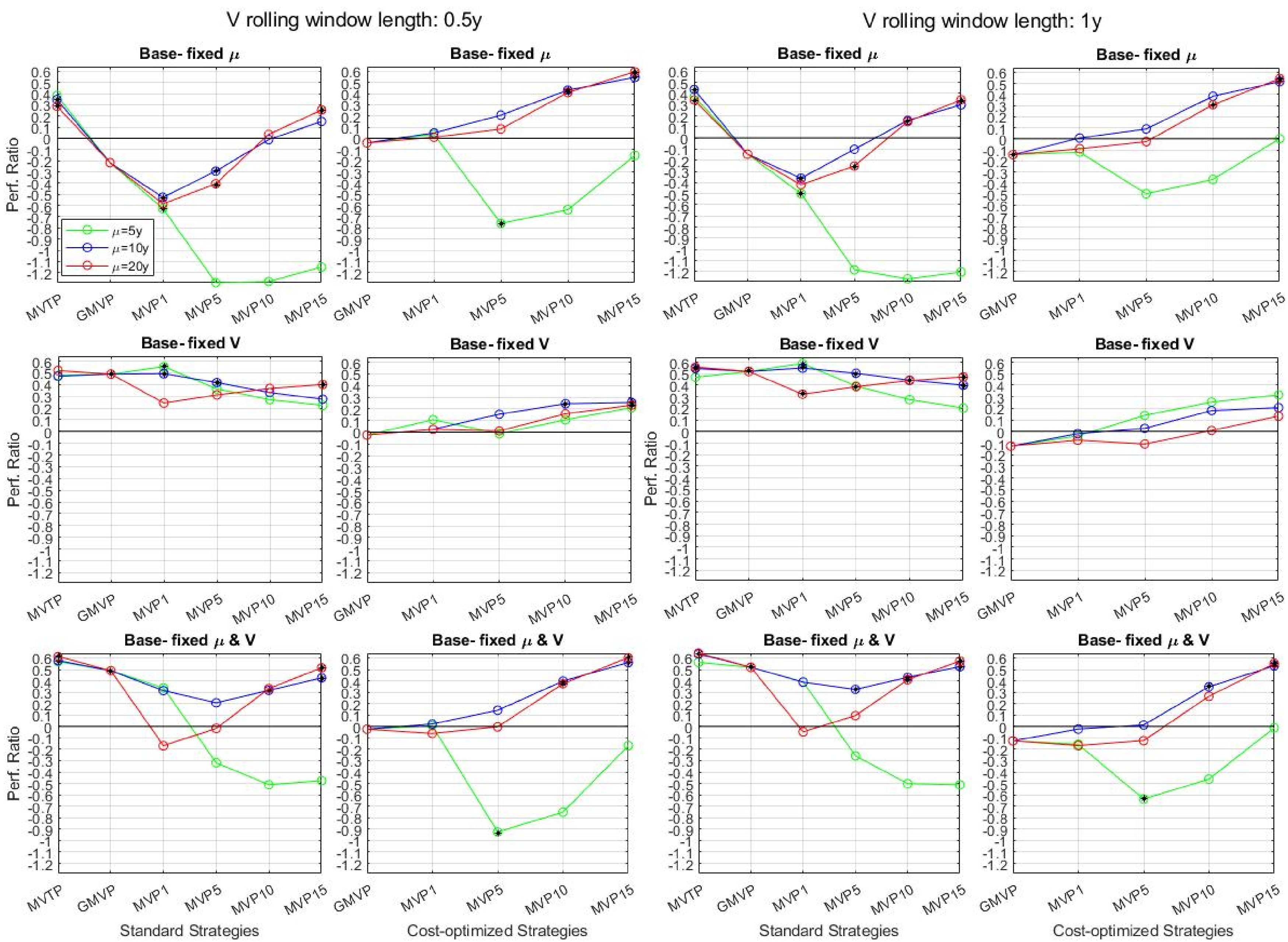

Figure 31: As in Figure 7 and 8, Figure 24 through 31 display the performance ratios for our alternative scenarios described in Section A. The performance ratios are defined as the out-of-sample after-cost Sharpe ratio differential described in Section 2.3 divided by the real-time performance of our strategies.

Robustness: Cost-minimization to reduce downside risk

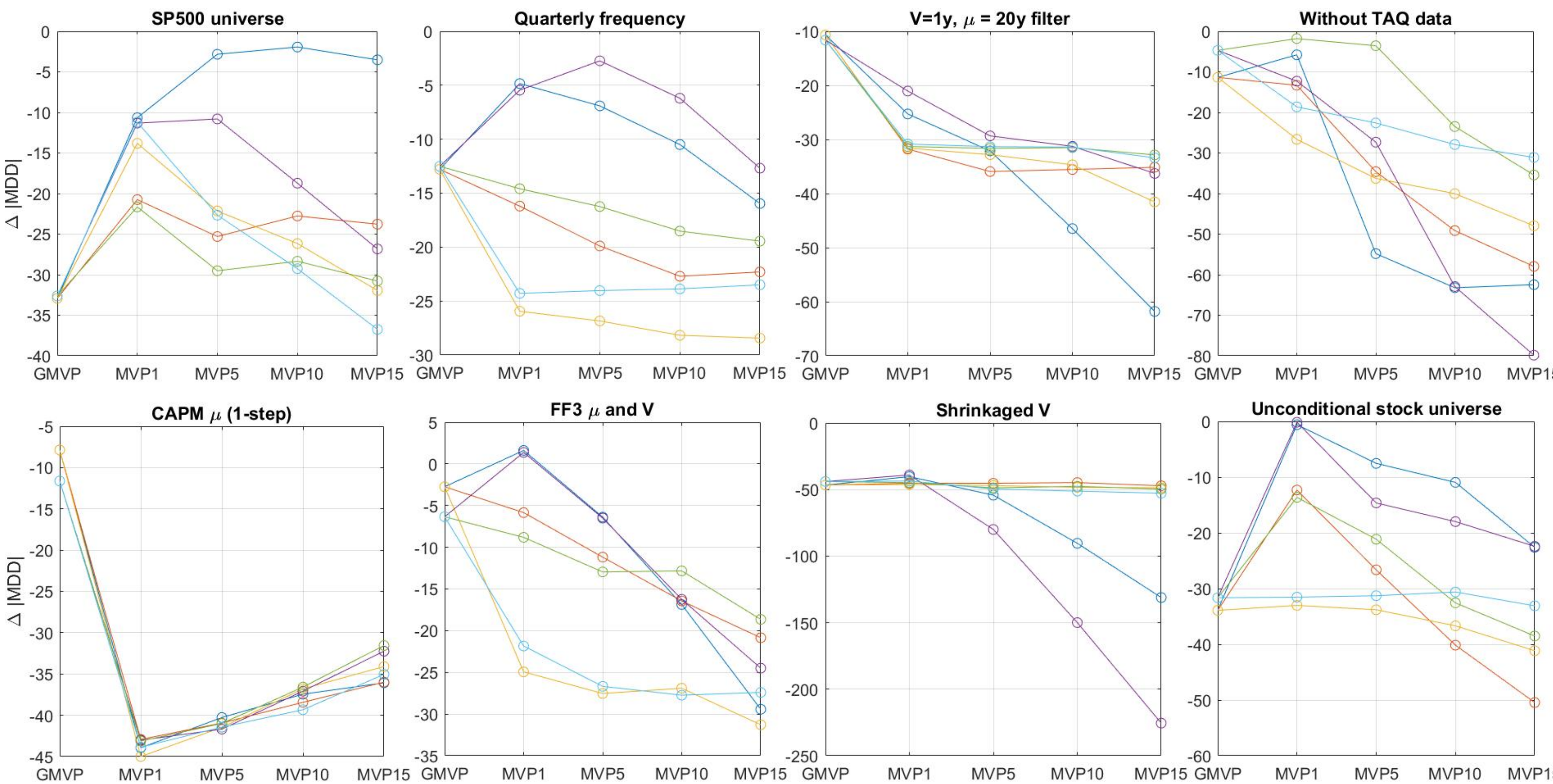

Figure 32: The fact that mean-variance cost-optimized strategies are less prone to downside risk as measured via Maximum Draw Down (MDD),or worst incurred loss, vis-a-vis their standard analogs is robust to our alternative scenarios. Each figure plots the absolute difference in MDDs between cost-optimized and standard analog strategies over the monthly period January 2003 - December 2017. The alternative scenarios, titled in each graph, are described in Section A.

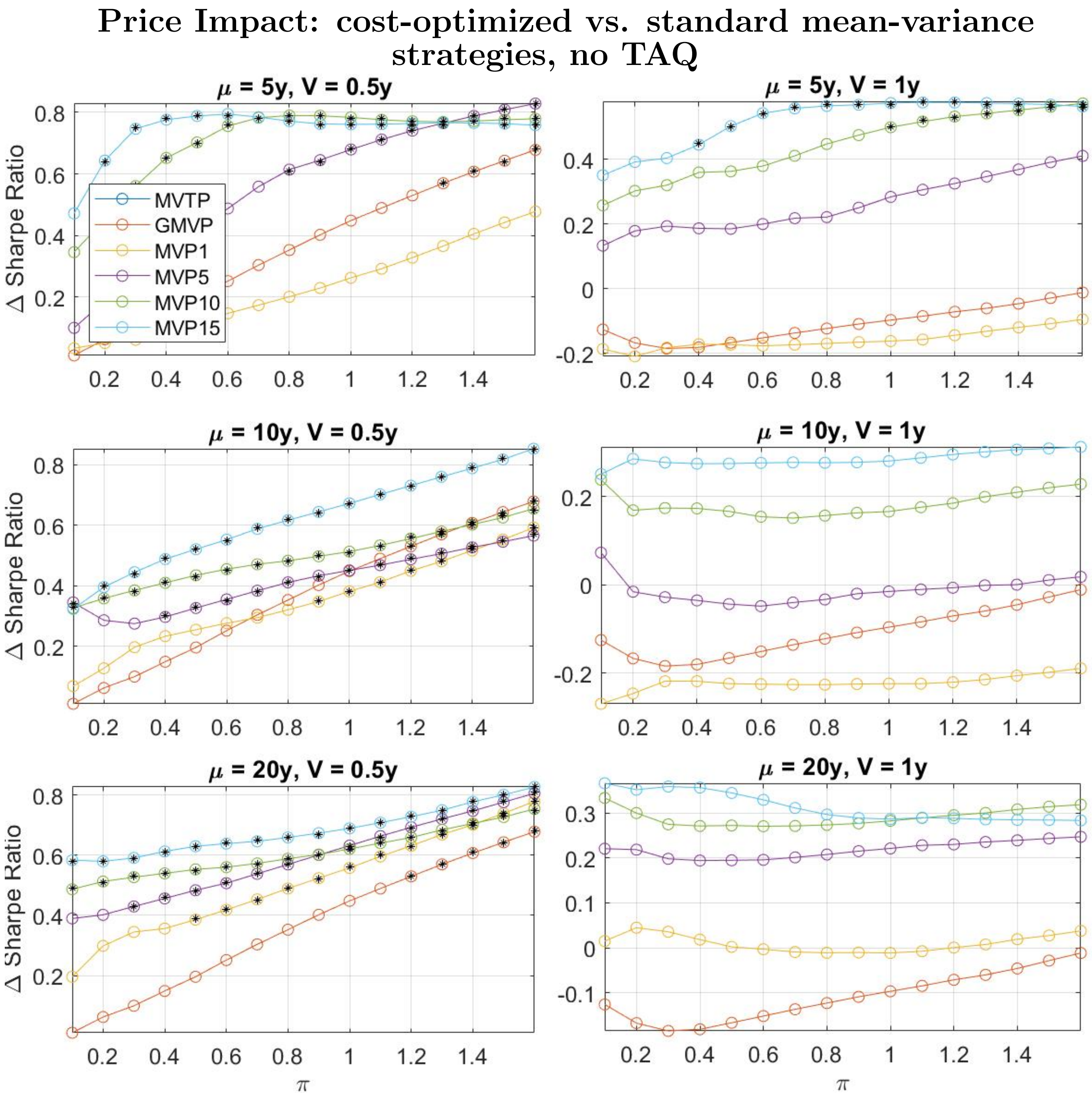

Figure 33: Cost-optimized mean-variance strategies continue to consistently scale up more in presence of price impact when cost are estimated without using TAQ data. The graphs plots the annualized after-cost out-of-sample Sharpe ratio differentials between our cost-optimized and standard strategies (described in Section 2.1 and printed in the legend) as a function of the price impact parameter π, which is inversely related to the liquidity and depth of the market and positively related to the portfolio size of the investors. Each graph differs in the combinations of rolling window lengths (in years) used to estimate the conditional risk premia μ_t and covariance matrix V_t as specified in the plots' sub-titles. Starred markers refer to ratio differential significant at the 10% level. We test for differences in Sharpe ratios using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). The sample is every month from January 2003 to December 2017.

Price Impact: cost-optimized vs. standard mean-variance strategies, $SP500$ universe

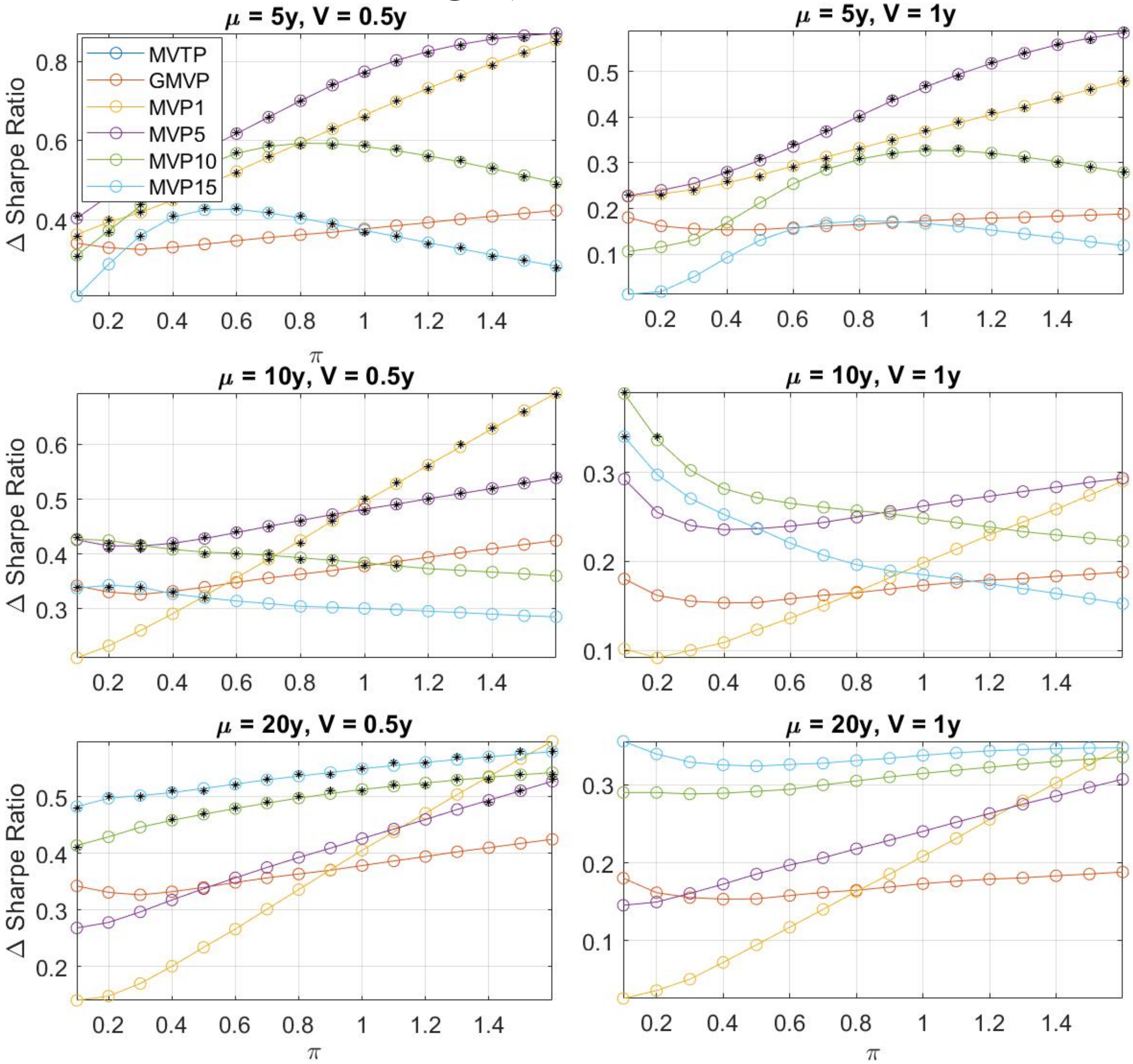

Figure 34: Cost-optimized mean-variance strategies continue to consistently scale up more in presence of price impact when we restrict the stocks universe to the $SP500$ constituents. The graphs plots the annualized after-cost out-of-sample Sharpe ratio differentials between our cost-optimized and standard strategies (described in Section 2.1 and printed in the legend) as a function of the price impact parameter π, which is inversely related to the liquidity and depth of the market and positively related to the portfolio size of the investors. Each graph differs in the combinations of rolling window lengths (in years) used to estimate the conditional risk premia μ_t and covariance matrix V_t as specified in the plots' sub-titles. Starred markers refer to ratio differential significant at the 10% level. We test for differences in Sharpe ratios using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). The sample is every month from January 2003 to December 2017.

Price Impact: cost-optimized vs. standard mean-variance strategies, quarterly re-balancing

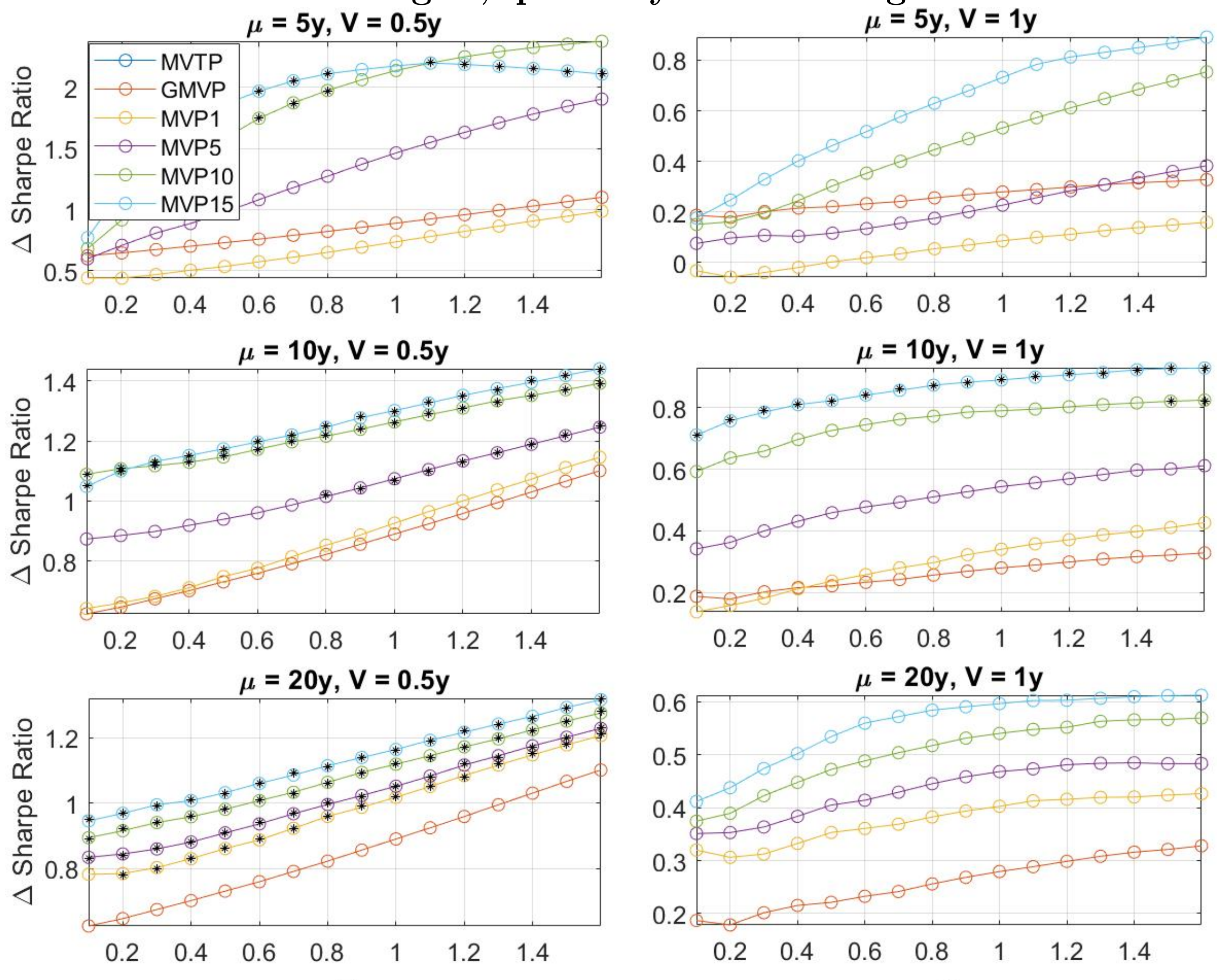

Figure 35: Cost-optimized mean-variance strategies continue to consistently scale up more in presence of price impact when we re-balance at the quarterly (rather then at the monthly) frequency. The graphs plots the annualized after-cost out-of-sample Sharpe ratio differentials between our cost-optimized and standard strategies (described in Section 2.1 and printed in the legend) as a function of the price impact parameter π, which is inversely related to the liquidity and depth of the market and positively related to the portfolio size of the investors. Each graph differs in the combinations of rolling window lengths (in years) used to estimate the conditional risk premia μ_t and covariance matrix V_t as specified in the plots' sub-titles. Starred markers refer to ratio differential significant at the 10% level. We test for differences in Sharpe ratios using block bootstrapping with block sizes of 5 observations to account for heteroskedasticity, cross- and auto-correlation (Ledoit and Wolf (2008)). The sample is every month from January 2003 to December 2017.

Table 8: **Market time versus Estimation error: dynamics of average magnitudes over the last 100 years**

	Market timing								Estimation error							
Scenario	1	2	3	4	5	6	7	**Mean**	1	2	3	4	5	6	7	**Mean**
PANEL A: STANDARD MEAN-VARIANCE STRATEGIES																
sample 03-17:	0.12	0.73	0.35	0.35	0.48	0.37	0.20	**0.37**	-0.50	-0.64	-0.49	-0.50	-0.44	-0.16	-0.42	**-0.45**
sample 73-02:	0.03	0.16	0	0	0.60	0.42	0.30	**0.22**	-0.44	-0.77	-0.48	-0.58	-0.54	-0.52	-0.43	**-0.54**
sample 45-72:	0.11	0.16	0.21	0.20	0.21	0.31	0.20	**0.20**	-0.06	-0.81	-0.12	-0.23	-0.49	-0.25	-0.24	**-0.31**
sample 26-44:	0	0	-	0	0.79	0	0.47	**0.21**	-0.13	0	-	-0.66	-0.72	-0.73	-0.55	**-0.47**
PANEL B: STANDARD MEAN-VARIANCE STRATEGIES (EXCLUDING $MVTP$)																
sample 03-17:	0.15	0.84	0.35	0.36	0.46	0.38	0.22	**0.39**	-0.50	-0.64	-0.49	-0.50	-0.44	-0.16	-0.42	**-0.45**
sample 73-02:	0.02	0.16	0	0	0.36	0.42	0.30	**0.18**	-0.44	-1.12	-0.59	-0.60	-0.58	-0.61	-0.40	**-0.62**
sample 45-72:	0.12	0.16	0.21	0.20	0.21	0.31	0.20	**0.20**	-0.08	-1.12	-0.15	-0.27	-0.23	-0.18	-0.27	**-0.33**
sample 26-44:	0	0	-	0	0	0	0.47	**0.08**	-0.13	0	-	-0.71	0	-0.67	-0.62	**-0.36**
PANEL C: COST-OPTIMIZED MEAN-VARIANCE STRATEGIES																
sample 03-17:	0.45	0	0.48	0.57	0.63	0.48	0.65	**0.47**	-0.34	-0.68	-0.55	-0.53	0	0	-0.50	**-0.37**
sample 73-02:	0	0.62	0	0	0.41	0.37	0.34	**0.25**	-0.64	0	-0.50	-0.46	0	0	-0.57	**-0.31**
sample 45-72:	0.24	0.62	0.37	0.35	0.31	0.41	0	**0.33**	0	0	-0.33	-0.34	0	-0.43	-0.35	**-0.21**
sample 26-44:	0	0	-	0	0	0	0	**0.00**	-0.37	-0.59	-	-0.58	-0.40	-0.52	-0.60	**-0.51**
PANEL D: ALL VARIANCE STRATEGIES (EXCLUDING $MVTP$)																
sample 03-17:	0.37	0.84	0.42	0.45	0.52	0.43	0.44	**0.50**	-0.42	-0.71	-0.51	-0.52	-0.44	-0.16	-0.49	**-0.46**
sample 73-02:	0.02	0.55	0	0	0.42	0.38	0.31	**0.24**	-0.45	-1.12	-0.54	-0.53	-0.58	-0.61	-0.43	**-0.61**
sample 45-72:	0.16	0.55	0.31	0.03	0.24	0.37	0.20	**0.27**	-0.08	-1.12	-0.27	-0.26	-0.23	-0.26	-0.26	**-0.35**
sample 26-44:	0	0	-	0	0	0	0.47	**0.08**	-0.28	-0.59	-	-0.66	-0.40	-0.58	-0.61	**-0.52**

Notes: Market timing and estimation error have non-negligible impacts on the performance of our analyzed strategies in the alternative scenarios. The table reports for each of the analyzed scenarios (1:$SP500$, 2:*Quarterly*, 3:$[V, \mu] = [1, 20]y$, 4:*No TAQ*, 5:$CAPM$ μ, 6:$FF3$, 7:*Shrinkaged* V) the average of market timing and estimation error magnitudes across the three types of comparisons where the real-time strategies are benchmarked against their unimplementable analogs with fixed in-sample first, second and combined first and second moments of the returns' distribution. That is, for each analyzed scenario we only report the columns labeled "Mean" in Table 2 from the main text.

Table 9: **Market time versus Estimation error: dynamics of total amounts over the last 100 years**

	Market timing								Estimation error							
Scenario	1	2	3	4	5	6	7	**Mean**	1	2	3	4	5	6	7	**M**
PANEL A: STANDARD MEAN-VARIANCE STRATEGIES																
03-17:	0.64	21.54	18.83	19.15	15.12	19.00	1.76	**13.72**	-7.11	-7.53	-3.06	-3.67	-2.62	-0.16	-1.96	**-3**
73-02:	0.09	0.48	0	0	1.19	2.22	2.49	**0.92**	-7.30	-5.46	-21.69	-25.21	-20.87	-27.45	-25.82	**-1**
45-72:	1.98	0.48	1.81	1.80	3.80	5.50	2.12	**2.50**	-0.54	-5.15	-0.43	-5.14	-4.92	-6.76	-4.84	**-3**
26-44:	0	0	-	0	1.58	0	6.09	**1.28**	-1.73	0	-	-9.87	-1.44	-9.99	-8.14	**-5**
PANEL B: STANDARD MEAN-VARIANCE STRATEGIES (EXCLUDING $MVTP$)																
03-17:	0.30	18.77	15.38	16.73	14.27	18.51	1.34	**12.19**	-7.11	-7.53	-3.06	-3.67	-2.62	-0.16	-1.96	**-3**
73-02:	0.06	0.48	0	0	0.36	2.22	2.49	**0.80**	-7.30	-3.33	-17.91	-21.07	-19.13	-22.25	-15.93	**-1**
45-72:	1.77	0.48	1.81	1.80	3.80	5.50	2.12	**2.47**	-0.51	-3.33	-0.15	-3.68	-1.41	-2.39	-3.68	**-2**
26-44:	0	0	-	0	0	0	6.09	**1.02**	-1.73	0	-	-9.23	0	-7.32	-7.50	**-4**
PANEL C: COST-OPTIMIZED MEAN-VARIANCE STRATEGIES																
03-17:	6.84	0	10.25	6.69	8.58	9.87	7.46	**7.10**	-0.96	-2.56	-1.11	-9.44	0	0	-1.54	**-2**
73-02:	0	13.47	0	0	6.46	1.07	1.02	**3.15**	-7.79	0	-10.18	-7.08	0	0	-5.22	**-4**
45-72:	1.40	13.47	3.78	1.88	2.36	7.72	0	**4.37**	0	0	-0.94	-1.35	0	-5.59	-1.11	**-1**
26-44:	0	0	-	0	0	0	0	**0**	-6.75	-8.16	-	-1.16	-2.94	-1.50	-3.67	**-4**
PANEL D: ALL VARIANCE STRATEGIES (EXCLUDING $MVTP$)																
03-17:	7.14	18.77	25.63	23.42	22.85	28.38	8.80	**19.28**	-8.07	-10.09	-4.17	-13.11	-2.62	-0.16	-3.50	**-5**
73-02:	0.06	13.95	0	0	6.82	3.29	3.51	**3.95**	-15.09	-3.33	-28.09	-28.15	-19.13	-22.25	-21.15	**-1**
45-72:	3.17	13.95	5.59	3.48	6.26	13.22	2.12	**6.83**	-0.51	-3.33	-1.09	-5.03	-1.41	-7.98	-4.79	**-3**
26-44:	0	0	-	0	0	0	6.09	**1.02**	-8.48	-8.16	-	-10.39	-2.94	-8.82	-11.17	**-8**

Notes: As in the baseline scenario, market timing tends to be pro-cyclical and increasing over time while estimation error counter-cyclical and decreasi time. The table reports for each of the analyzed scenarios (1:$SP500$, 2:*Quarterly*, 3:$[V, \mu] = [1, 20]y$, 4:*No TAQ*, 5:$CAPM\ \mu$, 6:$FF3$, 7:*Shrinkaged* sum of market timing and estimation error magnitudes across the three types of comparisons where the real-time strategies are benchmarked agains unimplementable analogs with fixed in-sample first, second and combined first and second moments of the returns' distribution. That is, for each ar scenario we only report the columns labeled "Total" in Table 3 from the main text.

Table 10: **Overall Correlations between Mean-Variance Profitability and Market timing**

Scenario	Strategies	Roll. wind. length V	$\bar{\mu}$ 0.5 yrs	$\bar{\mu}$ 1 yr	$\bar{V}$ 0.5 yrs	$\bar{V}$ 1 yr	$\bar{\mu}$ & $\bar{V}$ 0.5 yrs	$\bar{\mu}$ & $\bar{V}$ 1 y
1	*Std*	Overall mean	0.13	0.19	-0.21	0.04	0.21	0.2
*SP*500	*TC*	Overall mean	0.43	0.53	-0.18	0	0.54	0.5
2	*Std*	Overall mean	0.13	0.27	0.50	0.59	0.27	0.4
Quarterly	*TC*	Overall mean	0.50	0.75	0.60	0.61	0.32	0.5
3	*Std*	Overall mean	-0.08	0.45	0	0.43	0.22	0.4
$[V\,\mu] = [1, 20]y$	*TC*	Overall mean	0.72	0.52	0.58	0.33	0.76	0.5
4	*Std*	Overall mean	0.24	0.42	0.41	0.52	0.38	0.5
No TAQ	*TC*	Overall mean	0.67	0.60	0.58	0.54	0.74	0.6
5	*Std*	Overall mean	-0.37	-0.31	0.05	-0.04	-0.67	-0.5
CAPM μ	*TC*	Overall mean	-0.66	-0.57	-0.09	0.02	-0.67	-0.7
6	*Std*	Overall mean	-0.81	-0.49	-0.33	-0.12	-0.57	-0.4
*FF*3	*TC*	Overall mean	-0.76	-0.72	-0.60	-0.82	-0.88	-0.8
7	*Std*	Overall mean	0.26	0.39	0.41	0.62	0.39	0.5
Shrinked V	*TC*	Overall mean	0.48	0.39	0.46	0.30	0.72	0.5

Notes: In line with the baseline scenario, market timing abilities are found positively correlated with mean-variance profitability, especially for cost-opt strategies. As in the baseline scenario, we measure the correlation of the Sharpe ratio differential between a given strategy and its unimplementable analog at least one of the first two moments of the returns' distribution are fixed at their in-sample levels with the Sharpe ratio level of the implementable st The table shows the average correlation across risk premia estimates (with rolling windows of 5,10 and 20 years) and sample periods (26-44 through 03- each analyzed alternative scenario in the case we fix the first, second or both moments of the returns' distribution (column $\bar{\mu}, \bar{V}$ and $\bar{\mu}\&\bar{V}$) under the est for the covariance matrix V using rolling windows of either 6 or 12 months.