Stock Return Serial Dependence and Out-of-Sample Portfolio Performance*

Victor DeMiguel
London Business School

Francisco J. Nogales
Universidad Carlos III de Madrid

Raman Uppal
Edhec Business School and CEPR

*Nogales is supported by the Spanish Government through project MTM2010-16519 and by the Comunidad de Madrid/Universidad Carlos III de Madrid through project CCG08-UC3M/ESP-4162. We are grateful to Eric Shirbini and Grigory Vilkov for helping us obtain the data we use, to John Campbell and Luis Viceira for making available their computer code, and to Francisco Barillas (discussant at the AFA), Jules van Binsbergen (discussant at the Empirical Asset Pricing Retreat), David Chapman (discussant at the UBC Summer Conference), Andrew Lo, Anthony Lynch, Peter Schotman, Piet Sercu (discussant at the the joint finance conference of the Leuven School of Business and Economics and the Louvain School of Management), and Mike Smith for detailed comments. We are also grateful for comments from seminar participants at London Business School, EDHEC Business School, the Catholic University of Leuven, the 2009 INFORMS Annual Meeting in San Diego, the 2010 Amsterdam Empirical Asset Pricing Retreat, the 2010 Summer Conference at the UBC Sauder School of Business, the 2011 AFA Annual Meeting, and the 2012 Computational Management Science Meeting. Send correspondence to Victor DeMiguel, London Business School, Regents Park, London NW1 4SA, UK; Telephone: +44 20 7000 8831; Email: avmiguel@london.edu.

Abstract

We study whether investors can exploit serial dependence in stock returns to improve out-of-sample portfolio performance. We show that a vector-autoregressive (VAR) model captures stock return serial dependence in a statistically significant manner. Analytically, we demonstrate that, unlike contrarian and momentum portfolios, an arbitrage portfolio based on the VAR model attains positive expected returns regardless of the sign of asset return cross-covariances and autocovariances. Empirically, we show, however, that both the arbitrage and mean-variance portfolios based on the VAR model outperform the traditional unconditional portfolios only for transaction costs below ten basis points. *JEL*: G11.

There is extensive empirical evidence that stock returns are serially dependent, and that this dependence can be exploited to produce abnormal positive expected returns. For instance, Jegadeesh and Titman (1993) find *momentum* in asset returns: "strategies which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past generate significant positive returns over 3- to 12-month holding periods." Lo and MacKinlay (1990) show that returns of large firms lead those of small firms, and a contrarian portfolio that takes advantage of this by being long past losing stocks and short past winners produces abnormal positive expected returns.[1]

Our objective is to study whether investors can exploit the stock return serial dependence to select portfolios of risky assets that perform well *out-of-sample*, both in the absence and presence of transaction costs. We tackle this task in three steps. First, we propose a vector autoregressive (VAR) model to capture stock return serial dependence, and test its statistical significance. Second, we characterize, both analytically and empirically, the expected return of an arbitrage (zero-cost) portfolio based on the VAR model, and compare it to those of contrarian and momentum arbitrage portfolios. Third, we evaluate empirically the out-of-sample gains from using investment (positive-cost) portfolios that exploit serial dependence in stock returns.

To identify the optimal portfolio weights, our work uses conditional forecasts of *expected returns* for *individual* stocks. This is in contrast to the recent literature on portfolio selection, which finds it optimal to *ignore* estimates of expected returns based only on historical return data, and tries to improve the estimation of the covariance matrix: see, for example, Chan et al. (1999), Ledoit and Wolf (2003, 2004), DeMiguel and Nogales (2009), and DeMiguel et al. (2009a).

Our paper makes three contributions to the literature on portfolio selection. First, we propose using a vector autoregressive (VAR) model to capture serial dependence in stock

[1] There is substantial empirical evidence of cross-covariances and autocovariances in returns. For example, there is a large body of research that documents momentum at the level of individual stocks (Jegadeesh (1990), Jegadeesh and Titman (1993), Moskowitz et al. (2012)), at the level of industries (Moskowitz and Grinblatt (1999)), and in size and book-to-market portfolios (Lewellen (2002)). There is also substantial evidence of autocovariances in stock returns (Fama and French (1988), Conrad and Kaul (1988, 1989, 1998), and reversal/overreaction (DeBondt and Thaler (1985)). Finally, a number of papers have documented the presence of cross-covariances, where the magnitude is related to factors such as firm size (Lo and MacKinlay (1990)), firm size within industries (Hou (2007)), trading volume (Chordia and Swaminathan (2000)), analyst coverage (Brennan et al. (1993)), and institutional ownership (Badrinath et al. (1995)).

returns. Our VAR model allows tomorrow's expected return on every stock to depend linearly on today's realized return on *every* stock, and therefore, it is general enough to capture any linear relation between stock returns in consecutive periods, irrespective of whether its origin can be traced back to cross-covariances, autocovariances, or both.[2] We verify the validity of the VAR model for stock returns by performing extensive statistical tests on five empirical datasets, and conclude that the VAR model is significant for all datasets. Moreover, we identify the origin of the predictability in the data and we find autocorrelation of portfolio and individual stock returns. We also find lead-lag relations between: big-stock portfolios and small-stock portfolios, growth-stock portfolios and value-stock portfolios, the HiTec industry portfolio and other industry portfolios, and big individual stocks and small individual stocks.

Our second contribution is to characterize, both analytically and empirically, the expected return of arbitrage (zero-cost) portfolios based on the VAR model, and to compare them to those of the contrarian and momentum arbitrage portfolios studied (among others) by Lo and MacKinlay (1990) and Pan et al. (2004). From the analytical perspective, Lo and MacKinlay (1990) show that the expected return of the contrarian arbitrage portfolio is positive if the stock return autocovariances are negative and the stock return cross-covariances are positive, and the expected return of the momentum arbitrage portfolio is positive if autocovariances are positive and cross-covariances negative. We demonstrate that the VAR arbitrage portfolio achieves a positive expected return in general, regardless of the sign of the autocorrelations and cross-correlations. Moreover, we find that the expected returns of the VAR arbitrage portfolio are large if the principal components of the covariance matrix provide a discriminatory forecast of tomorrow's asset returns; that is, if today's return on the principal components allow one to tell which assets will achieve high returns, and which low returns tomorrow.

[2]A broad variety of explanations have been offered for cross-covariances and autocovariances of asset returns. Some of these explanations are based on time-varying expected returns (Conrad and Kaul (1988)), with more recent work showing how to generate this variation in rational models (Berk et al. (1999) and Johnson (2002)). Other explanations rely on economic links, such as those between suppliers and customers (Cohen and Frazzini (2008)) and between upstream and downstream industries (Menzly and Ozbas (2010)). Then, there are explanations that are based on the slow transmission of information across investors (Hong and Stein (1999)). Finally, there are behavioral models of underreaction and overreaction, such as the ones by De Long et al. (1990), Daniel et al. (1998), and Barberis et al. (1998).

Empirically, we demonstrate that, while the profits of the contrarian and momentum arbitrage portfolios can be attributed mostly to their ability to exploit asset return autocovariances, the VAR arbitrage portfolio manages to exploit both cross-covariances and autocovariances. Moreover, the VAR arbitrage portfolio substantially outperforms the contrarian and momentum arbitrage portfolios, both in the absence and presence of transaction costs. We find, however, that the VAR arbitrage portfolio is profitable only with transaction costs smaller than 10 basis points, and the contrarian and momentum arbitrage portfolios only with transaction costs smaller than 5 basis points.

Our third contribution is to evaluate the out-of-sample gains associated with investing in two (positive-cost) portfolios that exploit stock return serial dependence. The first portfolio is the *conditional mean-variance portfolio* of a myopic investor who believes stock returns follow the VAR model. This portfolio relies on the assumption that stock returns in consecutive periods are linearly related. We consider a second portfolio that relaxes this assumption. Specifically, we consider the conditional mean-variance portfolio of a myopic investor who believes stock returns follow a nonparametric autoregressive (NAR) model, which does *not* require that the relation across stock returns be *linear*.[3] To control the high turnover associated with the conditional mean-variance portfolios, we focus on norm-constrained portfolios that are similar to those studied by DeMiguel et al. (2009a).

Our empirical results show that the norm-constrained conditional mean-variance portfolios outperform the traditional (unconditional) portfolios only for transaction costs below 10 basis points. Note that the VAR portfolios exploit serial dependence in a more general form than most of the previously documented approaches, and yet they are profitable only for transaction costs below 10 basis points. This implies that, when evaluating trading strategies that exploit serial dependence in stock returns, it is crucial to account for the frictions that exist in the real world.

To understand the origin of the predictability exploited by the conditional mean-variance portfolios, we consider the conditional mean-variance portfolios obtained from a lagged-

[3]We have also considered the *dynamic* portfolio of Campbell et al. (2003), which is the optimal portfolio of an intertemporally optimizing investor with Epstein-Zin utility, who believes that the returns follow a VAR model. We find that its performance is similar to that of the conditional mean-variance portfolios from VAR and thus to conserve space we do not report the results for the dynamic portfolios.

factor model based on the Fama-French and momentum factors (market, small minus big, high minus low book-to-market, and up-minus down), and we find that the market and high-minus-low factors drive most of the predictability exploited by the conditional portfolios. Moreover, we observe also that the gains from exploiting stock return serial dependence come in the form of higher expected return, because the out-of-sample variance of the conditional portfolios is higher than that of the unconditional (traditional) portfolios; that is, stock return serial dependence can be exploited to forecast stock mean returns much better than using the traditional (unconditional) sample estimator. Finally, we find that a substantial proportion of the gains associated with the conditional mean-variance portfolios arise from exploiting cross-covariances in stock returns, as opposed to just autocovariances.

The rest of this manuscript is organized as follows. Section 1 describes the datasets and the methodology we use for our empirical analysis. Section 2 states the VAR model of stock returns, tests its statistical significance, and uses the significance tests to identify the origin of the predictability in stock returns. Section 3 characterizes (analytically and empirically) the performance of a VAR arbitrage portfolio, and compares it to that of a contrarian and a momentum arbitrage portfolios. Section 4 describes the different investment portfolios we consider, and discusses their empirical performance. Section 5 concludes. Appendix A contains various robustness checks for our empirical findings with the supporting tables in an online appendix, and Appendix B contains the proofs for all propositions.

1. Data and Evaluation Methodology

In this section, we briefly describe our datasets and our methodology for evaluating the performance of portfolios.

1.1 Datasets

We consider five datasets for our empirical analysis: four datasets from Ken French's website, and one from CRSP, and for each dataset we report the results for close-to-close as well as open-to-close returns. The first two datasets contain the returns on 6 and 25 value-weighted portfolios of stocks sorted on size and book-to-market (6FF, 25FF). The third and fourth datasets contain the returns on the 10 and 48 industry value-weighted portfolios (10Ind, 48Ind). For close-to-close returns we use data from 1970 to 2011 downloaded from Ken French's website, while we build *open-to-close* returns from 1992 to 2011 using open-to-close data for individual stocks downloaded from the CRSP database, which records open-to-close returns only from 1992 onwards.

We also consider a fifth dataset containing individual stock returns from the CRSP database containing close-to-close and open-to-close returns on all stocks that were part of the S&P500 index at some point in time between 1992 and 2011 (100CRSP). To avoid any stock-survivorship bias, we randomly select 100 stocks every year using the following approach. At the beginning of each calendar year, we find the set of stocks for which we have returns for the entire period of our estimation window as well as for the next year. From those stocks, we randomly select 100 and use them for portfolio selection until the beginning of the next calendar year, when we randomly select stocks again.[4]

1.2 Evaluation methodology

We compare the performance of the different portfolios using four criteria, all of which are computed out of sample using a "rolling-horizon" procedure similar to that used by DeMiguel et al. (2009b):[5] (i) portfolio mean return; (ii) portfolio variance; (iii) Sharpe

[4]Observe that these five datasets are close to being tradable in practice, except for the illiquidity of the smaller stocks in the datasets from French's website. To see this, note first that for the French's datasets we use the value-weighted portfolios, which implies that no "internal" rebalancing is required for these portfolios. Second, the quantiles and industry definitions used to form French's datasets are updated only once a year, and thus the "internal" rebalancing due to this is negligible at the daily and weekly rebalancing periods that we consider. Therefore, the main barrier to the practical tradability of French's datasets is that these portfolios contain small illiquid stocks. This implies that when using daily return data and rebalancing, the historical portfolio return data may suffer from asynchronous trading at the end of the day. Regarding the CRSP datasets, we focus on stocks that are part of the S&P500 index, and thus, are relatively liquid. In order to understand whether our results are due to the effect of asynchronous trading, in Appendix A.2 we study the robustness of our results to the use of open-to-close returns (instead of close-to-close) and weekly returns (instead of daily), both of which suffer much less from the effects of asynchronous trading, and we find that indeed our results are robust.

[5]We use raw returns (that is, without subtracting the risk-free return) rather than returns in excess of the risk-free rate for our empirical evaluation, because we believe raw returns are more appropriate in the context of the portfolios that we consider, which are formed exclusively from risky assets. We have replicated our empirical evaluation using excess returns in Appendix A.1, and we find that the relative performance of the different portfolios is very similar, although the Sharpe ratios for the data with excess returns are smaller by about 0.3.

ratio, defined as the sample mean of out-of-sample returns divided by their sample standard deviation;[6] and, (iv) portfolio turnover (trading volume).

To measure the impact of proportional transactions costs on the performance of the different portfolios, we also compute the portfolio returns net of transactions costs as

$$r_{t+1}^k = \Big(1 - \kappa \sum_{j=1}^{N} \big|\mathrm{w}_{j,t}^k - \mathrm{w}_{j,(t-1)^+}^k\big|\Big)(\mathrm{w}_t^k)^\top r_{t+1},$$

where $\mathrm{w}_{j,(t-1)^+}^k$ is the portfolio weight in asset j at time t under strategy k *before* rebalancing, $\mathrm{w}_{j,t}^k$ is the desired portfolio weight at time t after rebalancing, κ is the proportional transaction cost, w_t^k is the vector of portfolio weights, and r_{t+1} is the vector of returns. We then compute the Sharpe ratio as described above, but using the out-of-sample returns net of transactions costs.

2. A Vector Autoregressive (VAR) Model of Stock Returns

We now introduce the VAR model. In Section 2.1, we describe the VAR model of stock returns, and in Section 2.2, we test the statistical significance of the VAR model for the five datasets described in Section 1.1. Finally, in Section 2.3, we use statistical tests to understand the nature of the relation between stock returns.

2.1 The VAR model

We use the following vector autoregressive (VAR) model to capture serial dependence in stock returns:

$$r_{t+1} = a + Br_t + \epsilon_{t+1}, \tag{1}$$

where $r_t \in \mathrm{I\!R}^N$ is the stock return vector for period t, $a \in \mathrm{I\!R}^N$ is the vector of intercepts, $B \in \mathrm{I\!R}^{N\times N}$ is the matrix of slopes, and ϵ_{t+1} is the error vector, which is independently

[6] To measure the statistical significance of the difference between the Sharpe ratios of two given portfolios, we use the (non-studentized) stationary bootstrap of Politis and Romano (1994) to construct a two-sided confidence interval for the difference between the Sharpe ratios (or certainty equivalents). We use 1,000 bootstrap resamples and an expected block size equal to 5. Then we use the methodology suggested in Ledoit and Wolf (2008, Remark 3.2) to generate the resulting bootstrap p-values.

and identically distributed as a multivariate normal with zero mean and covariance matrix $\Sigma_\epsilon \in \mathbb{R}^{N \times N}$, assumed to be positive definite.

Our VAR model considers multiple stocks and assumes that tomorrow's expected return on each stock (conditional on today's return vector) may depend linearly on today's return on any of the multiple stocks. This linear dependence is characterized by the slope matrix B (for instance, B_{ij} represents the marginal effect of $r_{j,t}$ on $r_{i,t+1}$ conditional on r_t). Thus, our model is sufficiently general to capture any linear relation between stock returns in consecutive periods, independent of whether its source is cross-covariances, autocovariances, or a combination of both.

VAR models have been used before for strategic asset allocation—see Campbell and Viceira (1999, 2002); Campbell et al. (2003); Balduzzi and Lynch (1999); Barberis (2000)—where the objective is to study how an investor should dynamically allocate her wealth across a few asset classes (e.g., a single risky asset (the index), a short-term bond, and a long-term bond), and the VAR model is used to capture the ability of certain variables (such as the dividend yield and the short-term versus long-term yield spread) to predict the returns on the single risky asset.[7] Our objective, on the other hand, is to study whether an investor can exploit stock return serial dependence to choose a portfolio of multiple risky assets with better out-of-sample performance, and thus, we use the VAR model to capture the ability of today's risky-asset returns to predict tomorrow's risky-asset returns. To the best of our knowledge, our paper is the first to investigate whether a VAR model at the individual risky-asset level can be used to choose portfolios with better out-of-sample performance.

[7] Lynch (2001) considers three risky assets (the three size and three book-to-market portfolios), but he does not consider the ability of each of these risky assets to predict the return on the other risky assets; instead, he considers the predictive ability of the dividend yield and the yield spread. The effectiveness of predictors such as size, value, and momentum in forecasting individual stock returns is examined in Section 4.4. The paper by Jurek and Viceira (2011) is a notable exception because it considers a VAR model that captures (among other things) the ability of the returns on the value and a growth portfolios to predict each other. VAR models have also been used to model serial dependence among individual stocks or international indexes. For instance, Tsay (2005, Chapter 8) estimates a VAR model for a case with only two risky assets, IBM stock and the S&P500 index, Eun and Shim (1989) estimate a VAR model for nine international markets, and Chordia and Swaminathan (2000) estimate a VAR model for two portfolios, one composed of high-trading-volume stocks and the other of low-trading-volume stocks.

2.2 Significance of the VAR model

For this section, we assume that r_t is a jointly covariance-stationary process with finite mean $\mu = E[r_t]$, positive definite covariance matrix $\Gamma_0 = E[(r_t - \mu)(r_t - \mu)^\top]$, and finite cross-covariance matrix $\Gamma_1 = E[(r_{t-1} - \mu)(r_t - \mu)^\top]$. Estimating the VAR model in (1) requires estimating a large number of parameters, and thus standard ordinary least squares (OLS) estimators of the VAR model suffer from estimation error. To alleviate the impact of estimation error, we use *ridge regression*, see Hoerl and Kennard (1970), which is designed to give robust estimators even for models with a large numbers of parameters. Moreover, to test the statistical significance of the ridge estimator of the slope matrix, we use the stationary bootstrap method of Politis and Romano (1994).[8]

To verify the validity of the VAR model for stock returns, we perform the above test every month (roughly 22 trading days) of the time period spanned by each of the five datasets, using an estimation window of $\tau = 2000$ days each time. In all cases, the test rejects the null hypothesis of $B = 0$ at a 1% significance level; that is, for every period and each dataset, there exists at least one significant element in the matrix of slopes B. Hence, we infer that the VAR model is statistically significant for the five datasets we consider.

2.3 Interpretation of the VAR model

In this section, we test the significance of each of the elements of the estimated slope matrix B to improve our general understanding of the specific character of the serial dependence in stock returns present in the data. For ease of exposition, we first study two small datasets with only two assets each, and we then provide summary information for the full datasets.

[8]In particular, to test the null hypothesis $H_0\colon B = 0$, we first estimate equation (1) using ridge regression with an estimation window of $\tau = 2000$ days. Then, we propose the following test statistic $M = -(\tau - N)\ln(|\hat{\Omega}_1|/|\hat{\Gamma}_0|)$, where $\hat{\Omega}_1$ is the covariance matrix of the residuals $\hat{\epsilon}$ obtained after fitting the VAR equation (1) to the data. Because the distribution of M is not known when estimating the model using ridge regression, we approximate this distribution through a bootstrap procedure. To do that, we obtain $S = 100$ bootstrap errors from the residuals $\hat{\epsilon}$. Then, we generate recursively the bootstrap returns in equation (1) using the parameter estimates by ridge regression and the bootstrap errors. Then, we fit the VAR model to the bootstrap returns to obtain S bootstrap replicates of the covariance matrix of the residuals, $\hat{\Omega}_1$. Analogously, we repeat this procedure to generate recursively the bootstrap returns under the null hypothesis ($B = 0$) to obtain S bootstrap replicates of the covariance matrix of the returns, $\hat{\Gamma}_0$. Finally, we use these S bootstrap replicates to approximate the distribution of the test statistic M and the corresponding p-value for the hypothesis test that $B = 0$.

2.3.1 Results for two portfolios formed on size. We consider a dataset with one small-stock portfolio and one large-stock portfolio. The return on the first asset is the average equally-weighted return on the three small-stock portfolios in the 6FF dataset with six assets formed on size and book-to-market, and the return on the second asset is the average equally-weighted return on the three large-stock portfolios.

We first estimate the VAR model for the first 2,000-day estimation window (starting from the beginning of 1970), and test the significance of *each element* (i, j) of the matrix of slopes B with the null hypothesis: $H_0: \ B_{ij} = 0$. The estimated VAR model is:

$$\begin{aligned} r_{t+1,\text{small}} &= 0.0001 + 0.171 r_{t,\text{small}} + 0.151 r_{t,\text{big}}, \\ r_{t+1,\text{big}} &= 0.0002 + 0.076 r_{t,\text{small}} + 0.133 r_{t,\text{big}}. \end{aligned}$$

Both off-diagonal elements of the slope matrix are significant, but note that the B_{12} element 0.151 is substantially larger than the B_{21} element 0.076, which suggests that there is a lead-lag relation between big-stock and small-stock, with big-stock returns leading small-stock returns. This finding is consistent with the findings of Lo and MacKinlay (1990). Also, both small and large-stock portfolio returns have significant first-order autocorrelations.

To understand how the serial dependence in stock returns varies with time, we perform the above test for every trading day in our time series. Figure 1 shows the time evolution of the estimated diagonal and off-diagonal elements of the slope matrix, respectively. The solid lines give the estimated value of these elements, and we set the lines to be thicker for periods when the elements are statistically significant.

Our first observation is that the estimated elements of the slope matrix vary *smoothly* with time. Note that it is expected that the slope matrix will vary with time because market conditions changed substantially from 1970 to 2011, but the fact that they change smoothly suggests that ridge regression is effective at reducing the impact of estimation error.[9]

Our second observation is that the autocorrelations of both big- and small-stock returns decrease with time in our series, and the autocorrelation of small-stock returns actually

[9] Indeed, we have computed also the ordinary least squares estimators of the slope matrix and we observe that they do not vary as smoothly with time as those estimated using ridge regression.

becomes negative after the 2008 crisis year (for the performance of momentum strategies in extreme markets, see also Moskowitz et al. (2012, Section 4.3)). Also, the strength of the lead-lag relation between big- and small-stock returns decreases with time. Moreover, after the 2008 crisis year, the sign of the lead-lag relation between big- and small-stock returns changes, and although big-stocks continue to lead small stocks after the crisis, they lead with a negative sign. These two observations show that although there is serial dependence in stock returns both before and after the 2008 crisis, the nature of this serial dependence changed substantially.

2.3.2 Results for two portfolios formed on book-to-market ratio. We now study a second dataset with one low book-to-market stock portfolio (growth portfolio) and one high book-to-market stock portfolio (value portfolio). The return on the first asset is the average equally-weighted return on the two portfolios corresponding to low book-to-market stocks in the 6FF dataset, and the return on the second portfolio is the average equally-weighted return on the two portfolios corresponding to high book-to-market stocks.

The estimated VAR model for the first 2,000-day estimation window is:

$$\begin{aligned} r_{t+1,\text{growth}} &= 0.0007 + 0.176 r_{t,\text{growth}} + 0.079 r_{t,\text{value}}, \\ r_{t+1,\text{value}} &= 0.0006 + 0.141 r_{t,\text{growth}} + 0.119 r_{t,\text{value}}. \end{aligned}$$

Both off-diagonal elements of the slope matrix are significant, but the B_{21} element 0.141 is substantially larger than the B_{12} element 0.079, which indicates that there is a lead-lag relation between growth and value stocks, with the growth-stock returns leading value-stock returns. This finding is consistent with Hou and Moskowitz (2005) and Li (2011). Also, both growth- and value-stock portfolio returns have significant first-order autocorrelations.

To understand how the serial dependence in stock returns varies with time, we perform the above test for every trading day in our time series. Figure 2 shows the time evolution of the estimated diagonal and off-diagonal elements of the slope matrix, respectively. The solid lines give the estimated value of these elements, and we set the lines to be thicker for periods when the elements are statistically significant. We again observe that although the elements of the slope matrix vary with time, they vary smoothly, which suggests ridge regression deals well with estimation error. We observe also that the magnitude of the autocorrelations in growth and value stock returns decreases with time, and it becomes statistically indistinguishable from zero after the 2008 crisis year. Finally, the strength of the lead-lag relation between growth and value stock returns decreases with time, and in fact, after the 2008 crisis year the direction of the lead-lag relationship reverses, with value stocks leading growth stocks.

2.3.3 Results for the full datasets. We now summarize our findings for the five datasets described in Section 1.1. We start with the dataset with six portfolios of stocks sorted by size and book-to-market. Figure 3 shows the time evolution of the estimated diagonal and off-diagonal elements of the slope matrix. To make it easy to identify the most important elements of the slope matrix, we depict only those elements that are significant for long periods of time, and the legend labels are ordered in decreasing order of the length of the period when the element is significant. Note also that we number the different portfolios as follows: 1 = small-growth, 2 = small-neutral, 3 = small-value, 4 = big-growth, 5 = big-neutral, and 6 = big-value. We observe that the estimates of the elements of the slope matrix vary smoothly with time.

Figure 3a shows that there exist substantial and significant first-order autocorrelations in small-growth and big-growth portfolio returns, and smaller but also significant autocorrelations for all other portfolios. Figure 3b shows that there is strong evidence (in terms of magnitude and significance) that big-growth portfolios lead small-growth portfolios (element B_{14}) and big-neutral lead small-neutral (B_{25}); that is, the "big" portfolios lead the corresponding version of the "small" portfolios. Finally, we observe that small-growth portfolios lead both small-neutral (B_{21}) and small-value portfolios (B_{31}), and small-neutral lead small-value (B_{32}); that is, growth leads value among small-stock portfolios. We also observe that the autocorrelation in stock returns decreases with time, and for big-growth and big-neutral stocks it becomes negative after the 2008 year. We also observe that the magnitude of the lead-lag relations decreases with time. Moreover, after the 2008 crisis year, big-neutral stocks continue to lead small-neutral stocks, but with a negative sign. In summary, the patterns of serial dependence for the dataset with six portfolios of stocks sorted by size and book-to-market confirms our findings in Sections 2.3.1 and 2.3.2. We have obtained similar insights from the tests on the 25FF but to conserve space we do not report the results.

We now turn to the industry datasets, and to make the interpretation easier, we start with the dataset with five industry portfolios downloaded from Ken French's website, which contains the returns for the following industries: 1 = Cnsmr (Consumer Durables, Non-Durables, Wholesale, Retail, and Some Services), 2 = Manuf (Manufacturing, Energy, and Utilities), 3 = HiTec (Business Equipment, Telephone and Television Transmission), 4 = Hlth (Healthcare, Medical Equipment, and Drugs), and 5 = Other (Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment, Finance). Figure 4 shows the time evolution of the estimated diagonal and off-diagonal elements of the slope matrix, where the element numbers correspond to the industries as numbered above. Figure 4a shows that there exist strong first-order autocorrelations in Hlth, Other, and HiTec returns. Moreover, there is strong evidence that HiTec returns lead all other returns except Hlth (elements B_{23}, B_{53}, and B_{13}),[10] and that Hlth returns lead Cnsmr returns (B_{14}). Regarding the time variation of the serial dependence, we observe that the autocorrelation in industry portfolio returns decreases with time, and becomes negative for the Manuf and Other after the 2008 crisis year. Also, the magnitude of the lead-lag relations decreases with time. Moreover, although Hlth continues to lead Other and Cnsmr after the 2008 crisis year, it does so with the opposite sign. Similarly, although HiTec continues to lead Cnsmr and Manuf from 2000 to 2008, it does so with the opposite sign. The conclusions are similar for the 10Ind and 48Ind datasets, but to conserve space we do not report the results.

Finally, to understand the characteristics of the serial dependence in individual stock returns, we consider a dataset formed with the returns on individual stocks. For expositional purposes, we consider a dataset consisting of only four individual stocks. Two of these stocks correspond to relatively large companies (Exxon and General Electric) and two correspond to relatively small companies (Rowan Drilling and Genuine Parts). Figure 5 shows the time evolution of the estimated diagonal and off-diagonal elements of the slope matrix, where

[10] We find that the predictability associated with the HiTec industry returns is explained by their high correlation to two of the Fama-French factors: market (MKT) and high-minus-low book-to-market (HML).

we label the four companies as follows: 1 = Exxon, 2 = General Electric, 3 = Rowan, and 4 = Genuine Parts. We observe that Exxon and Genuine Parts both display significant negative autocorrelation. This is consistent with results in the literature that indicate that while portfolio returns are positively autocorrelated, individual stock returns are negatively autocorrelated. Also, there is evidence that both General Electric and Exxon lead Rowan Drilling. Note that Rowan Drilling is a supplier to Exxon, so the evidence that Exxon leads Rowan Drilling is consistent with the idea in Cohen and Frazzini (2008) that economic links between firms may lead to predictability in returns.

3. Analysis of VAR Arbitrage Portfolios

To gauge the potential of the VAR model to improve portfolio selection, we study the performance of an arbitrage (zero-cost) portfolio based on the VAR model, and compare it analytically and empirically to that of other arbitrage portfolios.

3.1 Analytical comparison

In this section, we compare *analytically* the expected return of the VAR arbitrage portfolio to that of a contrarian and a momentum arbitrage portfolios considered in the literature.

3.1.1 The contrarian and momentum arbitrage portfolios. Lo and MacKinlay (1990) consider the following contrarian ("c") arbitrage portfolio:

$$\mathrm{w}_{c,t+1} = -\frac{1}{N}(r_t - r_{et}e), \tag{2}$$

where $e \in \mathbb{R}^N$ is the vector of ones and $r_{et} = e^\top r_t/N$ is the return of the equally-weighted portfolio at time t. Note that the weights of this portfolio add up to zero, and thus it is an *arbitrage* portfolio. Also, the portfolio weight for every stock is equal to the negative of the stock return in excess of the return of the equally-weighted portfolio. That is, if a stock obtains a high return at time t, then the contrarian portfolio assigns a negative weight to it for period $t+1$, and hence this is a *contrarian* portfolio. Lo and MacKinlay (1990) show that the expected return of the contrarian arbitrage portfolio is:

$$E[\mathrm{w}_{ct}^{\top} r_t] = C + O - \sigma^2(\mu), \tag{3}$$

where

$$C = \frac{1}{N^2}(e^{\top}\Gamma_1 e - \mathrm{tr}(\Gamma_1)), \tag{4}$$

$$O = -\frac{N-1}{N^2}\mathrm{tr}(\Gamma_1), \tag{5}$$

$$\sigma^2(\mu) = \frac{1}{N}\sum_{i=1}^{N}(\mu_i - \mu_m)^2, \tag{6}$$

and where μ_i is the mean return on the ith stock, μ_m is the mean return on the equally-weighted portfolio, and "tr" denotes the trace of matrix. Equation (3) shows that the expected return of the contrarian arbitrage portfolio can be decomposed into three terms: a term C, which captures the impact of the asset return cross-covariances (off-diagonal elements of Γ_1); a term O, which captures the impact of autocovariances (diagonal elements of Γ_1); and, a term $\sigma^2(\mu)$, which captures the impact of unconditional mean returns.

Lo and MacKinlay (1990) use this decomposition to study whether contrarian profits can be attributed to overreaction, which they associate with the presence of negative stock return autocorrelations. Pan et al. (2004) use a similar approach to study the origin of momentum profits. Specifically, they consider the momentum arbitrage portfolio ("m") obtained by reversing the sign of the contrarian arbitrage portfolio ($\mathrm{w}_{m,t+1} = -\mathrm{w}_{c,t+1} = (r_t - r_{et}e)/N$), whose expected return is $E[\mathrm{w}_{mt}^{\top} r_t] = -E[\mathrm{w}_{ct}^{\top} r_t] = -C - O + \sigma^2(\mu)$. In Section 3.2, we estimate C and O using daily return data for size and book-to-market portfolios, industry portfolios, and individual stocks, and discuss how our findings relate to those by Lo and MacKinlay (1990) and Pan et al. (2004).

3.1.2 The VAR arbitrage portfolio. We consider the following VAR ("v") arbitrage portfolio:

$$\mathrm{w}_{v,t+1} = \frac{1}{N}(a + Br_t - r_{vt}e), \tag{7}$$

where $a + Br_t$ is the VAR model forecast of the stock return at time $t+1$ conditional on the return at time t, and $r_{vt} = (a + Br_t)^\top e/N$ is the VAR model prediction of the equally-weighted portfolio return at time $t+1$ conditional on the return at time t. Note that the weights of $\mathrm{w}_{v,t+1}$ add up to zero, and thus it is also an arbitrage portfolio. Also, the portfolio $\mathrm{w}_{v,t+1}$ assigns a positive weight to those assets whose VAR-based conditional expected return is above that of the equally-weighted portfolio, and a negative weight to the rest of the assets.

The following proposition gives the expected return of the VAR arbitrage portfolio, and shows that it is positive in general. For tractability, in the proposition we assume we can estimate the VAR model exactly, but in our empirical analysis in Section 3.2 we estimate the VAR model from empirical data.

Proposition 1. Assume that r_t is a jointly covariance-stationary process with mean $\mu = E[r_t]$, covariance matrix $\Gamma_0 = E[(r_t - \mu)(r_t - \mu)^\top]$, and cross-covariance matrix $\Gamma_1 = E[(r_{t-1} - \mu)(r_t - \mu)^\top]$. Assume also that the covariance matrix Γ_0 is positive definite. Finally, assume we can estimate the VAR model exactly; that is, let $B = \Gamma_1^\top \Gamma_0^{-1}$ and $a = (I - B)\mu$. Then,

1. The VAR arbitrage portfolio can be decomposed into the sum of an arbitrage portfolio $\mathrm{w}_{\Gamma t}$ that exploits the serial dependence captured by the slope matrix B, and an arbitrage portfolio $\mathrm{w}_{\mu t}$ that exploits the cross-variance in unconditional mean returns. Specifically,

$$\mathrm{w}_{vt} = \mathrm{w}_{\Gamma t} + \mathrm{w}_{\mu t},$$

where

$$\mathrm{w}_{\Gamma t} = \frac{1}{N}\left(B\left(r_{t-1} - \mu\right) - \frac{e^\top B(r_{t-1} - \mu)}{N} e\right), \tag{8}$$

$$\mathrm{w}_{\mu t} = \frac{1}{N}\left(\mu - \frac{e^\top \mu}{N} e\right). \tag{9}$$

2. The expected return of the VAR arbitrage portfolio is

$$E[\mathrm{w}_{vt}^\top r_t] = E[\mathrm{w}_{\Gamma t}^\top r_t] + E[\mathrm{w}_{\mu t}^\top r_t], \tag{10}$$

where

$$E[\mathrm{w}_{\Gamma t}^\top r_t] = \frac{\mathrm{tr}(\Gamma_1^\top \Gamma_0^{-1} \Gamma_1)}{N} - \frac{e^\top \Gamma_1^\top \Gamma_0^{-1} \Gamma_1 e}{N^2} \geq 0, \tag{11}$$

$$E[\mathrm{w}_{\mu t}^\top r_t] = \sigma^2(\mu) \geq 0. \tag{12}$$

Part 1 of Proposition 1 shows that the VAR arbitrage portfolio can be decomposed into a portfolio $\mathrm{w}_{\Gamma t}$ that exploits the serial dependence captured by the slope matrix B, and a portfolio $\mathrm{w}_{\mu t}$ that exploits the cross-variance in unconditional mean returns. To see this, note that the arbitrage portfolio $\mathrm{w}_{\Gamma t}$ assigns a positive weight to assets whose VAR-based estimate of conditional mean returns in excess of the unconditional mean return (the term $B(r_{t-1}-\mu)$)[11] is above that of the equally-weighted portfolio (the term $(e^\top B(r_{t-1}-\mu)/N)e$), and a negative weight to the rest of the assets. Also, the arbitrage portfolio $\mathrm{w}_{\mu t}$ assigns a positive weight to those assets whose unconditional mean return is above that of the equally-weighted portfolio.

Part 2 of Proposition 1 shows that the expected return of $\mathrm{w}_{\Gamma t}$ depends only on the on the covariance matrix Γ_0 and the cross-covariance matrix Γ_1, and the expected return of $\mathrm{w}_{\mu t}$ depends exclusively on the unconditional mean returns. Moreover, the result shows that both arbitrage portfolios ($\mathrm{w}_{\Gamma t}$ and $\mathrm{w}_{\mu t}$) make a nonnegative contribution to the expected returns of the VAR arbitrage portfolio, and that the expected return of the VAR arbitrage portfolio is strictly positive, except in the degenerate case in which the unconditional mean returns of all assets are identical.

Proposition 1 shows that the VAR arbitrage portfolio can always exploit the structure of the covariance and cross-covariance matrix, as well as that of the mean stock returns, to obtain a strictly positive expected return. This result contrasts with that obtained for the contrarian and momentum arbitrage portfolios. Essentially, the VAR arbitrage portfolio can exploit the autocorrelations and cross-correlations in stock returns regardless of their sign, whereas, as explained above, the expected return of the contrarian portfolio is positive if the autocorrelations are negative and the cross-correlations positive, and the expected

[11] To see this, note that the VAR model can be rewritten as $r_t - \mu = B(r_{t-1} - \mu) + \epsilon_t$.

return of the momentum portfolio is positive if the autocorrelations are positive and the cross-correlations negative.

The feature that distinguishes our VAR arbitrage portfolio from the contrarian and momentum arbitrage portfolios studied in the literature is that the VAR arbitrage portfolio exploits the full structure of the slope matrix B, whereas the contrarian and momentum arbitrage portfolios implicitly assume the slope matrix is a multiple of the identity matrix.[12] Proposition 1 demonstrates that this feature allows the VAR arbitrage to exploit asset return cross-covariances and autocovariances regardless of their sign.

Note also that the cross-sectional variance of mean stock returns has a negative impact on the expected return of the contrarian portfolio, but it has a positive impact on the expected return of the VAR and momentum portfolios. The reason for this is that the contrarian portfolio assigns a negative weight to those assets whose realized return at time t is above that of the equally-weighted portfolio and, as a result, the contrarian portfolio assigns a negative weight to assets with a mean return that is above average.

The following proposition shows that the portfolio $\mathrm{w}_{\Gamma t}$ can be further decomposed into a portfolio w_{Ct} that depends only on the off-diagonal elements of Γ_1 and a portfolio w_{Ot} that depends only on the diagonal elements of Γ_1.[13] We will use this result in Section 3.2 to test whether the gains associated with the VAR arbitrage portfolio arise from cross-covariances or autocovariances.

Proposition 2. Let the assumptions of Proposition 1 hold, then the arbitrage portfolio $\mathrm{w}_{\Gamma t}$ can be decomposed into the sum of two arbitrage portfolios: an arbitrage portfolio w_{Ct} that exploits cross-covariances of returns (off-diagonal elements of Γ_1), and an arbitrage portfolio w_{Ot} that exploits autocovariances of returns (diagonal elements of Γ_1). Specifically, $\mathrm{w}_{\Gamma t} = \mathrm{w}_{Ct} + \mathrm{w}_{Ot}$, where

$$\mathrm{w}_{Ct} = \frac{1}{N}\left(B_C\left(r_{t-1} - \mu\right) - \frac{e^\top B_C(r_{t-1} - \mu)}{N} e\right),$$

[12]To see this, note that the contrarian and momentum arbitrage portfolios can be obtained from the VAR arbitrage portfolio given in Proposition 1 by replacing the slope matrix B with $-I$ and I, respectively.

[13]Please note that $\mathrm{w}_{c,t}$ and w_{Ct} refer to different portfolios: $\mathrm{w}_{c,t}$ denotes the contrarian arbitrage portfolio of Lo and MacKinlay (1990), while w_{Ct} denotes the arbitrage portfolio that exploits the cross-covariances of returns (off-diagonal elements of Γ_1).

$$\mathrm{w}_{Ot} = \frac{1}{N}\left(B_O\left(r_{t-1} - \mu\right) - \frac{e^\top B_O(r_{t-1} - \mu)}{N} e\right),$$

where $B_C^\top = \Gamma_0^{-1}\left(\Gamma_1 - \mathrm{diag}(\Gamma_1)\right)$ depends only on the off-diagonal elements of Γ_1, $B_O^\top = \Gamma_0^{-1}\mathrm{diag}(\Gamma_1)$ depends only on the diagonal elements of Γ_1, and $B^\top = \Gamma_0^{-1}\Gamma_1 = B_C^\top + B_O^\top$.

3.1.3 Identifying the origin of predictability using principal components.

We now use principal-component analysis to identify the origin of the predictability in asset returns exploited by the VAR arbitrage portfolio. Specifically, we show that the ability of the VAR arbitrage portfolio to generate positive expected returns can be traced back to the ability of the principal components to forecast which assets will perform well and which poorly in the next period.

To see this, note that given a symmetric and positive definite covariance matrix Γ_0, we have that $\Gamma_0 = Q\Lambda_0 Q^\top$, where Q is an orthogonal matrix ($QQ^\top = I$) whose columns are the principal components of Γ_0, and Λ_0 is a diagonal matrix whose elements are the variances of the principal components. Therefore, the VAR model in Equation (1) can be rewritten as $r_{t+1} = a + BQQ^\top r_t + \epsilon_{t+1} = a + \hat{B}p_t + \epsilon_{t+1}$, where $p_t = Q^\top r_t \in \mathbb{R}^N$ is the return of the principal components at time t, and $\hat{B} = BQ$ is the slope matrix expressed in the reference frame defined by the principal components of the covariance matrix.

Proposition 3. Let the assumptions in Proposition 1 hold, then the expected return of the VAR arbitrage portfolio can be written as

$$E[\mathrm{w}_{vt}^\top r_t] = \frac{N-1}{N}\sum_j \lambda_j \mathrm{var}(\hat{B}_{\bullet j}) + \sigma^2(\mu),$$

where λ_j is the variance of the jth principal component of the covariance matrix Γ_0, and $\mathrm{var}(\hat{B}_{\bullet j})$ is the variance of the elements in the jth column of matrix $\hat{B}$.

Proposition 3 shows that the VAR arbitrage portfolio attains a high expected return when the variances of the columns of $\hat{B}$ multiplied by the variances of the corresponding principal components are high. The main implication of this result is that the information provided by today's return on the jth principal component is particularly useful when it has a *variable* impact on tomorrow's returns on the different assets; that is, when the variance of the jth column of $\hat{B}$ is high. Clearly, when this occurs, today's return on the jth principal component allows us to discriminate between assets we should go long and assets we should short. Moreover, if the variance of the jth principal component is high, then its realized values will lie in a larger range and this will also allow us to realize higher expected returns with the VAR arbitrage portfolios.

Finally, note that the results in Proposition 3 can be used to identify empirically the origin of the predictability exploited by the arbitrage VAR portfolio by estimating the principal components that contribute most to its expected return. For instance, for the size and book-to-market portfolio datasets we find that the principal components with the highest contributions are a portfolio long on big-stock portfolios and short on small-stock portfolios, and a portfolio long on value-stock portfolios and short on growth-stock portfolios; and for the industry dataset we find that the principal component with the highest contribution is long on the HiTec industry portfolio and short on the other industries.

3.1.4 Identifying the origin of predictability using factor models. Another approach to understand the origin of the predictability exploited by the VAR arbitrage portfolio is to consider a lagged-factor model instead of the VAR model. For instance, one could consider the following lagged-factor model:

$$r_{t+1} = a^f + B^f f_t + \epsilon^f_{t+1}, \tag{13}$$

where $a^f \in \mathbb{R}^N$ is the vector of intercepts, $B^f \in \mathbb{R}^{N\times F}$ is the matrix of slopes, $f_t \in \mathbb{R}^F$ is the factor return vector for period t, and ϵ^f_{t+1} is the error vector. This model will be particularly revealing when we choose factors that have a clear economic interpretation such as the Fama-French and momentum factors.

We then consider the following lagged-factor arbitrage portfolio:

$$\mathrm{w}_{f,t+1} = \frac{1}{N}(a^f + B^f f_t - r_{ft} e),$$

where $a^f + B^f f_t$ is the lagged-factor model forecast of the stock return at time $t+1$ conditional on the factor return at time t, and $r_{ft} = (a^f + B^f f_t)^\top e/N$ is the lagged-factor model prediction of the equally-weighted portfolio return at time $t+1$ conditional on the factor return at time t.

The following proposition gives the result corresponding to Proposition 3 in the context of the easier-to-interpret lagged-factor model.

Proposition 4. Assume that r_t is the jointly covariance-stationary process described in (13), the factor covariance matrix $\Gamma_0^f = E((f_t - \mu_f)^\top (f_t - \mu_f))$ is positive definite, and that we can estimate the lagged-factor model exactly. Then, the expected return of the lagged-factor arbitrage portfolio is

$$E[\mathrm{w}_{vt}^\top r_t] = \frac{N-1}{N} \sum_j \lambda_j^f \operatorname{var}(\hat{B}_{\bullet j}^f) + \sigma^2(\mu),$$

where λ_j^f is the variance of the jth principal component of the factor covariance matrix Γ_0^f, $\operatorname{var}(\hat{B}_{\bullet j}^f)$ is the variance of the elements in the jth column of matrix $\hat{B}^f$, and $\hat{B}^f$ is the slope matrix expressed in the frame of reference defined by the principal components for the factor covariance matrix; that is, $\hat{B}^f = B^f Q$, where Q is the matrix whose columns are the principal components of the factor covariance matrix.

Proposition 4 shows that the ability of the lagged-factor arbitrage portfolio to generate positive expected returns can be traced back to the ability of the principal components of the *factor* covariance matrix to forecast which stocks will perform well and which poorly in the next period. Moreover, because it is reasonable to expect that the factors will be relatively uncorrelated, in which case the principal components coincide with the factors, the predictability can be traced back to the ability of the *factors* to provide a discriminating forecast of which stocks will perform well and which poorly in the next period.

3.2 Empirical comparison

In this section, we compare empirically the performance of the conditional VAR arbitrage portfolio to those of the contrarian and momentum arbitrage portfolios. We first compare *in-sample* expected returns using the results in Sections 3.1.1 and 3.1.2. We then compare the *out-of-sample* expected returns and Sharpe ratios of the different arbitrage portfolios, using the rolling horizon methodology described in Section 1.2. Finally, we compare the out-of-sample performance in the presence of transaction costs.

3.2.1 In-sample comparison of performance. Panel A of Table 1 gives the in-sample expected return of the contrarian, momentum, and VAR arbitrage portfolios. For the contrarian and momentum portfolios, we report also the in-sample contribution to expected returns of the cross-covariances, C, which is defined in Equation (4); the in-sample contribution of autocovariances, O, defined in (5); and, the in-sample contribution of the cross-variance of mean returns, $\sigma^2(\mu)$, defined in (6). For the VAR arbitrage portfolio, based on the results in Propositions 1 and 2, we report the in-sample expected return of the arbitrage portfolio that exploits cross-covariances $E[\mathrm{w}_{Ct}^\top r_t]$, the arbitrage portfolio that exploits autocovariances $E[\mathrm{w}_{Ot}^\top r_t]$, and the arbitrage portfolio that exploits the cross-variance in unconditional mean returns $E[\mathrm{w}_{\mu t}^\top r_t]$.[14]

Comparing the term C, which captures the impact of cross-covariances on the expected return of the momentum and contrarian portfolios, with the term O, which captures the impact of autocovariances, we find that for all our datasets these two terms have opposite signs, and that the absolute impact of the autocovariances is substantially larger than that of the cross-covariances. In addition, we find that the cross-variance of mean returns $\sigma^2(\mu)$ is very small. This implies that *autocovariances* are mostly responsible for the profitability of the contrarian and momentum arbitrage portfolios for the datasets we consider. We also notice that for the datasets containing portfolio returns (6FF, 25FF, 10Ind, and 48 Ind), O is negative (that is, the autocovariances are positive), and for the dataset with individual stocks (100CRSP), O is positive (that is, the autocovariances are negative). Consequently, the momentum arbitrage portfolio attains a positive expected return for the four datasets containing portfolio returns, whereas it is the contrarian portfolio that has a positive expected return for the dataset containing individual stock returns.

Our results confirm the analysis of Pan et al. (2004, Table 3), who find that for *weekly* data on industry portfolio returns, the cross-covariances and autocovariances have offsetting

[14]To make a fair comparison (both in sample and out of sample) between the expected return of the different arbitrage portfolios, we normalize the arbitrage portfolios so that the sum of all positive weights equals one for all portfolios. We have tested also the raw (non-normalized) arbitrage portfolios, and the insights are similar.

impacts on the expected return of a momentum arbitrage portfolio, with the impact of the autocovariances being larger. We find that this holds also for *daily* return data for industry portfolios, as well as for size-and-book-to-market portfolios. Pan et al. (2004) suggest their results support the behavioral explanation for momentum profits based on investor underreaction (see Barberis et al. (1998), Daniel et al. (1998), and Hong and Stein (1999) for models of investor under reaction and overreaction). Our results provide some further support for this behavioral explanation in the context of daily return data for industry and size-and-book-to-market portfolios.

However, our results are in contrast to those of Lo and MacKinlay (1990, Table 3), who find that for *weekly* return data on individual stocks, the terms C and O have the same sign and similar magnitude. Panel A of Table 1 shows that for *daily* return data on individual stocks, these two terms have opposite signs, and that the autocovariances explain most of the profitability of a contrarian arbitrage portfolio. Moreover, for weekly return data on individual stocks (not reported in the paper), we find that C is close to zero, and thus the autocovariances again explain most of the profitability of the contrarian portfolio.

The difference between our results and those of Lo and MacKinlay (1990) may be explained by the fact that we consider only stocks that were part of the S&P500 index at some point between 1992 and 2011, whereas Lo and MacKinlay (1990) consider *all* stocks in NYSE-AMEX, including many small stocks. Consequently, their sample captures lead-lag relations between big and small stocks that we may not capture. Lo and MacKinlay (1990) interpret their findings as showing that it is not just overreaction that explains short-term contrarian profits in individual stock returns (see DeBondt and Thaler (1985) for a discussion of overreaction). Our results, however, suggest that overreaction is driving most of the profits when one applies the contrarian strategy to large stocks, such as the ones in the S&P500 index.

Panel A of Table 1 demonstrates that, unlike the contrarian and momentum arbitrage portfolio, the VAR arbitrage portfolio manages to exploit both the cross-covariances and autocovariances in asset returns. To see this, note that the expected return of both the arbitrage portfolio that exploits cross-covariances $E[\mathrm{w}_{Ct}^{\top} r_t]$, and the arbitrage portfolio that exploits autocovariances $E[\mathrm{w}_{Ot}^{\top} r_t]$, are strictly positive for all datasets. Finally, note that the in-sample expected return of the VAR arbitrage portfolio is larger than that of the contrarian portfolio for every dataset. Comparing the VAR arbitrage portfolio with the momentum portfolio, we observe that the VAR arbitrage portfolio outperforms the momentum arbitrage portfolio for the three datasets with large number of assets (25FF, 48Ind, 100CRSP).

3.2.2 Out-of-sample comparison of performance. Panel B of Table 1 reports the out-of-sample expected return of the contrarian, momentum, and VAR arbitrage portfolios as well as those of portfolios w_{Ct}, w_{Ot}, and $\mathrm{w}_{\mu t}$. Similar to the case of in-sample expected returns, we find that the out-of-sample expected return of the contrarian arbitrage portfolio is positive only for the dataset with individual stock returns, whereas the expected return of the momentum arbitrage portfolio is positive for the datasets with portfolio returns. We also find that the out-of-sample expected return of the VAR arbitrage portfolio is larger than those of the contrarian and momentum portfolios for every dataset. Moreover, the out-of-sample expected return of the portfolio that exploits autocovariances (w_O) is positive for every dataset, and that of the portfolio that exploits cross-covariances (w_C) is positive for the three datasets with the largest number of assets (25FF, 48Ind, 100CRSP). This again shows that, unlike the contrarian and momentum arbitrage portfolios, the VAR arbitrage portfolio generally manages to exploit both the autocovariances and cross-covariances of asset returns. Finally, we find that the out-of-sample expected return of the portfolio that exploits the cross-variance in unconditional mean returns is very small.

Panel C of Table 1 reports the out-of-sample Sharpe ratio of the contrarian, momentum, and VAR arbitrage portfolios computed using the rolling-horizon methodology described in Section 1.2, together with the p-values that the Sharpe ratios for the contrarian and unconditional arbitrage portfolios are different from that for the VAR arbitrage portfolio. The relative performance of the different arbitrage portfolios in terms of out-of-sample Sharpe ratios is similar to that in terms of expected returns. The VAR arbitrage portfolio outperforms the contrarian and momentum arbitrage portfolio in terms of out-of-sample Sharpe ratio for every dataset, with the difference being statistically significant. Finally, we find that the VAR arbitrage portfolio substantially outperforms the portfolio that exploits the cross-variance in unconditional mean returns ($\mathrm{w}_{\mu t}$) for every dataset.

Panels D and E of Table 1 report the out-of-sample Sharpe ratios of the different arbitrage portfolios in the presence of transaction costs of 5 and 10 basis points.[15] We observe that the VAR arbitrage portfolio outperforms the contrarian and momentum arbitrage portfolios even in the presence of transaction costs. We also observe that the VAR arbitrage portfolio is profitable in the presence of transaction costs of at most 5 basis points, whereas the contrarian and momentum arbitrage portfolios generally attain negative Sharpe ratios in the presence of transaction costs of 5 basis points. However, for a transaction cost of 10 basis points, none of these strategies are profitable.[16]

4. Analysis of VAR Mean Variance Portfolios

In this section, we describe the various investment (positive-cost) portfolios that we consider, and we compare their out-of-sample performance for the five datasets listed in Section 1.1. Section 4.1 discusses portfolios that ignore stock return serial dependence and Section 4.2 describes portfolios that exploit stock return serial dependence. Then, in Section 4.3 we characterize what proportion of the gains from exploiting serial dependence in stock returns comes from exploiting autocovariances and what proportion from exploiting cross-covariances, and in Section 4.4 we use a lagged-factor model to identify the origin of the predictability in stock returns exploited by the conditional portfolios.

4.1 Portfolios that *ignore* stock return serial dependence

We describe below three portfolios that do not take into account serial dependence in stock returns: the equally-weighted ($1/N$) portfolio, the shortsale-constrained minimum-variance portfolio, and the norm-constrained mean-variance portfolio.

[15] French (2008, p. 1553) estimates that the trading cost in 2006, including "total commissions, bid-ask spreads, and other costs investors pay for trading services," and finds that these costs have dropped significantly over time: "from 146 basis points in 1980 to a tiny 11 basis points in 2006." His estimate is based on stocks traded on NYSE, Amex, and NASDAQ, while the stocks that we consider in our CRSP datasets are limited to those that are part of the S&P500 index. Note also that the trading cost in French, and in earlier papers estimating this cost, is the cost paid by the average investor, while what we have in mind is a professional trading firm that presumably pays less than the average investor.

[16] Note that in the presence of transaction costs, the out-of-sample Sharpe ratio of the momentum arbitrage portfolio is no longer the negative of the out-of-sample Sharpe ratio of the contrarian portfolio.

4.1.1 The $1/N$ portfolio.

The $1/N$ portfolio studied by DeMiguel et al. (2009b) is simply the portfolio that assigns an equal weight to all N stocks. In our evaluation, we consider the $1/N$ portfolio with rebalancing; that is, we rebalance the portfolio every day so that the weights for every asset are equal.

4.1.2 The shortsale-constrained minimum-variance portfolio.

The shortsale-constrained minimum-variance portfolio is the solution to the problem

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \Sigma \mathrm{w}, \tag{14}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1, \tag{15}$$

$$\mathrm{w} \geq 0, \tag{16}$$

where $\Sigma \in \mathbb{R}^{N \times N}$ is the covariance matrix of stock returns, $\mathrm{w}^\top \Sigma \mathrm{w}$ is the portfolio return variance, the constraint $\mathrm{w}^\top e = 1$ ensures that the portfolio weights sum up to one, and the constraint $\mathrm{w} \geq 0$ precludes any short positions.[17] For our empirical evaluation, we use the shortsale-constrained minimum-variance portfolio computed by solving problem (14)–(16) after replacing the covariance matrix by the shrinkage estimator proposed by Ledoit and Wolf (2003).[18]

4.1.3 The norm-constrained mean-variance portfolio.

The mean-variance portfolio is the solution to:

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \Sigma \mathrm{w} - \frac{1}{\gamma} \mathrm{w}^\top \mu, \tag{17}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1, \tag{18}$$

where μ is the mean stock return vector and γ is the risk-aversion parameter. Because the weights of the unconstrained mean-variance portfolio estimated from empirical data tend to take extreme values that fluctuate over time and result in poor out-of-sample performance

[17] We focus on the *shortsale-constrained* minimum-variance portfolio because the *unconstrained* minimum-variance portfolio for our datasets typically includes large short positions that are associated with high costs. Nevertheless, we have replicated all of our analysis using also the *unconstrained* minimum-variance portfolio and the relative performance of the different portfolios is similar.

[18] We use an estimation window of 1000 days, which results in reasonably stable estimators, while allowing for a reasonably long time series of out-of-sample returns for performance evaluation.

(see DeMiguel et al. (2009b)), we report the results only for constrained mean-variance portfolios.

Specifically, we consider a 1-norm-constraint on the difference between the mean-variance portfolio and the benchmark shortsale-constrained minimum-variance portfolio; see DeMiguel et al. (2009a) for an analysis of norm constraints in the context of portfolio selection.[19] Specifically, we compute the norm-constrained mean-variance portfolios by solving problem (17)–(18) after imposing the additional constraint that the norm of the difference between the mean-variance portfolio and the shortsale-constrained minimum-variance portfolio is smaller than a certain threshold δ; that is, after imposing that $\|\mathrm{w} - \mathrm{w}_0\|_1 = \sum_{i=1}^{N} |\mathrm{w}_i - (\mathrm{w}_0)_i| \leq \delta$, where w_0 is the shortsale-constrained minimum-variance portfolio. We use the shortsale-constrained minimum-variance portfolio as the target because of the stability of its portfolio weights. We consider three values of the threshold parameter: $\delta_1 = 2.5\%$, $\delta_2 = 5\%$, and $\delta_3 = 10\%$. Thus, for the case where the norm constraint has a threshold of 2.5% and the benchmark is the shortsale-constrained minimum-variance portfolio, the sum of all negative weights in the norm-constrained conditional portfolios must be smaller than 2.5%.

For our empirical evaluation, we compute the norm-constrained (unconditional) mean-variance portfolio by solving problem (17)–(18) after replacing the mean stock return vector by its sample estimate, and the covariance matrix by the shrinkage estimator of Ledoit and Wolf (2003). We consider values of the risk aversion parameter $\gamma = \{1, 2, 10\}$, but our main insights are robust to the value of the risk aversion parameter and thus to conserve space we report the results for only $\gamma = 2$.

4.1.4 Empirical performance. Panel A of Table 2 gives the out-of-sample Sharpe ratio of the portfolios that ignore serial dependence in stock returns together with the p-value that the Sharpe ratio is different from that of the shortsale-constrained minimum-variance portfolio. We observe that the minimum-variance portfolio attains a substantially higher out-of-sample Sharpe ratio than the equally-weighted portfolio for all datasets except the

[19] We have also considered imposing shortsale constraints, instead of norm-constraints, on the conditional mean-variance portfolio, but we find that the resulting conditional portfolios have very high turnover, so we do not report the results to conserve space.

100CRSP dataset, where the two portfolios achieve a similar Sharpe ratio. The explanation for the good performance of the shortsale-constrained minimum-variance portfolio is that the estimator of the covariance matrix we use (the shrinkage estimator of Ledoit and Wolf (2003)) is a very accurate estimator and, as a result, the performance of the minimum-variance portfolio is very good.[20]

We also observe that the norm-constrained unconditional mean-variance portfolio outperforms the shortsale-constrained minimum-variance portfolio for two of the five datasets (6FF, 25FF), but the difference in performance is neither substantial nor significant. Finally, the turnover of the different portfolios is reported in Table 3. We observe from this table that the turnover of the different portfolios that ignore stock return serial dependence is moderate ranging for the different portfolios and datasets from 0.2% to 3% per day.

Hereafter, we use the *shortsale-constrained minimum-variance portfolio* as our main benchmark because of its good out-of-sample performance, reasonable turnover, and absence of shortselling.[21]

4.2 Portfolios that exploit stock return serial dependence

We consider two portfolios that exploit stock return serial dependence. The first portfolio is the *conditional mean-variance portfolio* of an investor who believes stock returns follow the VAR model. This portfolio relies on the assumption that stock returns in consecutive periods are linearly related. We also consider a portfolio that relaxes this assumption. Specifically, we consider the conditional mean-variance portfolio of an investor who believes stock returns follow a nonparametric autoregressive (NAR) model, which does not require that stock returns be linearly related.

[20]We have also computed the Sharpe ratio of the market portfolio, using the market factor data from Ken French's website, and we find that for the first four datasets, which cover the period 1970—2011, the market portfolio attains a Sharpe ratio of 0.69, and for the last dataset, which covers the period 1992–2011, it achieves a Sharpe ratio of 0.53; that is, the market portfolio is outperformed by both the 1/N and minimum-variance portfolio.

[21]Note that one could also use the norm-constrained unconditional mean-variance portfolio as the benchmark, but because our norm-constraints impose a restriction on the difference between the computed portfolio weights and the weights of the shortsale-constrained minimum-variance portfolio, it makes more sense to use the shortsale-constrained minimum-variance portfolio as the benchmark. However, in our discussion below we also explain how the norm-constrained conditional portfolios perform compared to the norm-constrained unconditional mean-variance portfolios.

Because it is well-known that conditional mean-variance portfolios estimated from historical data have extreme weights that fluctuate substantially over time and have poor out-of-sample performance, we will consider only norm-constrained conditional mean-variance portfolios. Specifically, we consider a 1-norm-constraint on the difference between the conditional mean-variance portfolio and the benchmark shortsale-constrained minimum-variance portfolio.

4.2.1 The conditional mean-variance portfolio from the VAR model. One way to exploit serial dependence in stock returns is to use the *conditional mean-variance portfolios based on the VAR model.* These portfolios are optimal for a myopic investor (who cares only about the returns tomorrow) who believes stock returns follow a linear VAR model. They are computed by solving problem (17)–(18) after replacing the mean and covariance matrix of asset returns with their conditional estimators obtained from the VAR model. Specifically, these portfolios are computed from the mean of tomorrow's stock return conditional on today's stock return: $\mu_V = a + Br_t$, where a and B are the ridge-regression estimators of the coefficients of the VAR model obtained from historical data, and the conditional covariance matrix of tomorrow's stock returns:

$$\Sigma_V = \frac{1}{\tau} \sum_{i=t-\tau+1}^{t} (r_i - a - Br_{i-1})(r_i - a - Br_{i-1})^\top.$$

In addition, we apply the shrinkage approach of Ledoit and Wolf (2003) to obtain a more stable estimator of the conditional covariance matrix. Moreover, to control the turnover of the resulting portfolios, we impose a 1-norm constraint on the difference with the weights of the shortsale-constrained minimum-variance portfolio. As in the case of the unconditional portfolios, we evaluate the performance of the conditional portfolios for values of the risk aversion parameter $\gamma = \{1, 2, 10\}$, but the insights are robust to the value of the risk aversion parameter, and thus, we report the results only for the case of $\gamma = 2$.

4.2.2 The conditional mean-variance portfolio from the NAR model. One assumption underlying the VAR model is that the relation between stock returns in consecutive periods is linear. To gauge the effect of this assumption, we consider a nonparametric autoregressive (NAR) model.[22] We focus on the nonparametric technique known as *nearest-neighbor regression.* Essentially, we find the set of, say, 50 historical dates when asset returns were closest to today's asset returns, and we term these 50 historical dates the "*nearest neighbors.*" We then use the empirical distribution of the 50 days following the 50 nearest-neighbor dates as our conditional empirical distribution of stock returns for tomorrow, conditional on today's stock returns. The main advantage of this nonparametric approach is that it does not assume that the time serial dependence in stock returns is of a linear type, and in fact, it does not make any assumptions about the type of relation between them. The conditional mean-variance portfolios from NAR are the optimal portfolios of a myopic investor who believes stock returns follow a nonparametric autoregressive (NAR) model.

The *conditional mean-variance portfolios based on the NAR model* are obtained by solving the problem (17)–(18) after replacing the mean and covariance matrix of asset returns with their conditional estimators obtained from the NAR model. That is, we use the mean of tomorrow's stock return conditional on today's return:

$$\mu_N = \frac{1}{k}\sum_{i=1}^{k} r_{t_i+1},$$

where t_i for $i = 1, 2, ..., k$ are the time indexes for the k nearest neighbors in the historical time series of stock returns, and the covariance matrix of tomorrow's stock return conditional on today's return:

$$\Sigma_N = \frac{1}{k-1}\sum_{i=1}^{k}(r_{t_i+1} - \mu_N)(r_{t_i+1} - \mu_N)^\top.$$

In addition, we apply the shrinkage approach of Ledoit and Wolf (2003) to obtain a more stable estimator of the conditional covariance matrix. Moreover, to control the turnover of the resulting portfolios, we focus on the case 1-norm-constraints on the difference with the weights of the shortsale-constrained minimum-variance portfolio. As before, we report results for the risk aversion parameter $\gamma = 2$.

[22] See Gyorfi et al. (1987) for an in-depth discussion of nonparametric regression and Gyorfi et al. (2008, 2007) for an application to portfolio selection, and Mizrach (1992) for an application to exchange rate forecasting.

Sections 4.2.3 and 4.2.4 discuss the performance of the portfolios described above in the absence and presence of proportional transaction costs, respectively.

4.2.3 Empirical performance. Panel B in Table 2 gives the out-of-sample Sharpe ratios of the portfolios that exploit serial dependence in stock returns. Our main observation is that both the VAR and NAR portfolios that exploit stock return serial dependence substantially outperform the three traditional (unconditional) portfolios in terms of the out-of-sample Sharpe ratio. For instance, the norm-constrained conditional mean-variance portfolio from VAR substantially outperforms the shortsale-constrained minimum-variance portfolio for all datasets, and the difference in performance widens as we relax the norm constraint from $\delta_1 = 2.5\%$ to $\delta_3 = 10\%$. We also note that the performance of the conditional portfolios from the VAR and NAR models is similar for the datasets with a small number of assets, but the portfolios from the VAR model outperform the portfolios from the nonparametric approach for the largest datasets (48Ind and 100CRSP). This is not surprising as it is well known that the performance of the nonparametric nearest-neighbor approach relative to that of the parametric linear approach deteriorates with the number of explanatory variables; see Hastie et al. (2005, Section 7.3).

Table 3 gives the turnover of the various portfolios we study. We observe that imposing norm-constraints is an effective approach for reducing the turnover of the conditional mean-variance portfolios from VAR while preserving their good out-of-sample performance. Specifically, although the Sharpe ratio of the conditional mean-variance portfolios from VAR decreases generally, when we make the norm constraint tighter (decrease δ), it stays substantially larger than the Sharpe ratio of the shortsale-constrained minimum-variance and norm-constrained unconditional mean-variance portfolios for all datasets. Moreover, the turnover of the norm-constrained conditional mean-variance portfolios from VAR decreases drastically as we make the norm constraint tighter. For the case with $\delta_1 = 2.5\%$, the turnover of the conditional mean-variance portfolio from VAR stays below 3% for all datasets, for the case with $\delta_2 = 5\%$, it stays below 6%, and for the case with $\delta_3 = 10\%$, it stays below 15%. The effect of the norm constraints on the conditional mean-variance portfolios from NAR is similar to that on the conditional portfolios from VAR.

We observe from our empirical results on out-of-sample mean and variance (not reported) that the gains from using the norm-constrained portfolios come in the form of higher expected return, because the out-of-sample variance of these portfolios is much higher than that of the unconditional (traditional) portfolios; that is, stock return serial dependence can be used to obtain stock mean return forecasts that are much better than those from the traditional sample mean estimator based on historical data.

4.2.4 Empirical performance in the presence of transaction costs. We now evaluate the relative performance of the different portfolios in the presence of proportional transactions costs. Tables 4 and 5 give the out-of-sample Sharpe ratio of the different portfolios after imposing transaction costs of 5 and 10 basis points, respectively. We observe from Table 4 that, in the presence of a proportional transactions cost of 5 basis points, the norm-constrained conditional portfolios from the VAR model outperform the benchmark minimum-variance portfolio for all five datasets, and the differences increase as we relax the norm constraint from $\delta_1 = 2.5\%$ to $\delta_3 = 10\%$. The norm-constrained conditional portfolios from NAR perform similarly to those from VAR except for the largest datasets (48Ind and 100CRSP), where their performance is worse—again, this is to be expected when we use the nonparametric nearest-neighbor approach. Table 5 demonstrates that, in the presence of a transactions cost of 10 basis points, the conditional portfolios from the VAR outperform the shortsale-constrained minimum-variance portfolio for only three of the five datasets (25FF, 48Ind, and 100CRSP), which have a larger number of assets. We conclude that the conditional portfolios from the VAR model outperform the shortsale-constrained minimum-variance portfolio only for transaction costs below 10 basis points. Thus, it is clear that to take advantage of the VAR-based strategies, efficient execution of trades is important.

4.3 Exploiting autocovariances versus cross-covariances

In this section, we investigate what proportion of the gains associated with the conditional mean-variance portfolios is obtained by exploiting autocovariances, and what proportion is obtained by exploiting cross-covariances. To do this, we compare the performance of the conditional mean-variance portfolios from VAR defined in Section 4.2.1, with that of a conditional mean-variance portfolio obtained from a *diagonal* VAR model, which is a VAR model estimated under the additional restriction that only the diagonal elements of the slope matrix B are different from zero.[23]

Our empirical analysis shows that a substantial part of the gains comes from exploiting cross-covariances in stock returns. We find that for the 6FF dataset, 99% of the gains come from exploiting cross-covariances; for the 25FF dataset, 72% of the gains come from exploiting cross-covariances; for the 10Ind dataset, 25% of the gains come from cross-covariances; for the 48Ind dataset, 29% of the gains come from cross-covariances; and finally, for the 100CRSP dataset, 19% of the gains come from cross-covariances.[24]

4.4 Origin of the predictability exploited by conditional portfolios

To understand the origin of the predictability exploited by the conditional portfolios from the VAR model, we compare the performance of the conditional portfolios based on the VAR model to that of conditional portfolios based on the lagged-factor model defined in Equation (13). To identify the origin of the predictability exploited by the conditional portfolios, we first consider a four-factor model including the Fama-French and momentum factors (MKT, SMB, HML, and UMD), and then four separate one-factor models, each of them including only one of the four factors listed above.

Table 6 reports the performance of the conditional portfolios from these five models, the first with four factors, and the rest with a single factor. First, we observe that the conditional portfolios from the four-factor model outperform the benchmark shortsale-constrained minimum-variance portfolio for all datasets except 100CRSP. Second, comparing the Sharpe ratios for the portfolios based on the factor model in Table 6 to the Sharpe ratios for the conditional portfolios based on the full VAR model in Table 2, we notice that the performance of the conditional portfolios from the four-factor model is similar to that of the conditional portfolios from the VAR model for the 6FF and 25FF datasets, a bit worse for the 10Ind

[23]To make this comparison we relax the norm constraint so that we can disentangle the effect of the diagonal versus off-diagonal elements of the slope matrix, without the confounding effect of the norm constraints.

[24]Note that the proportion of the gains from using the VAR conditional mean-variance portfolios that are associated with cross-covariances and autocovariances differs from that for the VAR arbitrage portfolios reported in Table 1. This is not surprising because the arbitrage and mean-variance portfolios are constructed using very different procedures. Nevertheless, the interpretation of both sets of results is consistent: the VAR-based portfolios generally manage to exploit both cross-covariances and autocovariances in asset returns.

and 48Ind datasets, and substantially worse for the 100CRSP dataset. The reason for this is that the Fama-French and momentum factors capture most of the predictability in the datasets of portfolios of stocks sorted by size and book-to-market, but reflect only part of the predictability captured by the full VAR model for the datasets of industry portfolios and individual stocks. These results justify the importance of considering the full VAR model.

Moreover, comparing the performance of the conditional portfolios from the four different one-factor models, we observe that most of the predictability in all datasets comes from the MKT and HML factors. The implication is that the conditional portfolios are exploiting the ability of today's return on the MKT and HML factors to forecast the returns of multiple risky assets tomorrow. Note that this is very different from the type of predictability exploited in the literature before, where typically today's dividend yield and spread between the short-term and long-term yields have been used to predict tomorrow's return on a single risky index. The conditional portfolios we study exploit the ability of today's return on the MKT and HML factors to forecast which risky assets will have high returns and which will have low returns tomorrow.

5. Conclusion

We conclude with a word of caution. In this paper, we have developed a methodology that exploits serial dependence of a more general form than previously documented and have shown that exploiting this serial dependence in zero-cost arbitrage portfolios and positive-cost investment portfolios leads to high Sharpe ratios even out-of-sample. However, taking advantage of this serial dependence requires strategies that have high turnover. Consequently, even these stronger serial dependence results are exploitable only if one can trade for less than 10 basis points. This implies that the standard serial dependence results may not be exploitable at even lower transaction costs. Thus, when evaluating trading strategies proposed in the literature it is important to account for the frictions that exist in the real world, and to keep in mind the experience described in Roll (1994): "Over the past decade, I have attempted to exploit many of the seemingly most promising 'inefficiencies' by actually trading significant amounts of money. . . . Many of these effects are surprisingly strong in reported empirical work, but I have never yet found one that worked in practice."

A. Robustness Checks

In this appendix, we report the results of several additional analysis that we have undertaken to test the robustness of our findings.

A.1 Use of excess returns

We have decided to report in the manuscript the results for the case with raw returns because we think this is more appropriate in the context of portfolios formed exclusively with risky assets. However, we have repeated the out-of-sample evaluation of the conditional and unconditional mean-variance portfolios using excess returns instead of raw returns. The results, reported in Tables A1–A4 of the online appendix, show that the relative performance of the different portfolios is very similar for the case with excess returns and the case with raw returns (that is, when the risk-free return is not subtracted). The Sharpe ratios of portfolio returns for the case with excess returns are smaller by around 0.3 compared to those for the case with raw returns.

A.2 Robustness to asynchronous trading

To check whether our results are driven by asynchronous trading, we evaluate the performance of the different portfolios on open-to-close and weekly return versions of all five datasets we consider, as well as a dataset containing open-to-close industry ETF returns.

We find that the results are generally robust to using open-to-close and weekly return data. This shows that there is serial dependence in open-to-close and weekly return data, which are much less likely to be affected by asynchronous or infrequent trading than the close-to-close daily data. This result is consistent with the findings of Lo and MacKinlay (1990, p. 197) and Anderson et al. (2005) that the lead-lag relations in stock returns cannot be attributed entirely to asynchronous or infrequent trading.

A.2.1 Open-to-close return data. We evaluate the performance of the different portfolios on open-to-close return versions of all five datasets we consider, which are less likely to be affected by the effects of asynchronous trading. The out-of-sample Sharpe ratios for the different portfolios for open-to-close return data are reported in Tables A5, A7, and A8 in the appendix, for transaction costs of 0, 5, and 10 basis points, respectively, with the turnover reported in Table A6. We find that the conditional portfolios from the VAR model outperform the shortsale-constrained minimum-variance portfolio for transaction costs below 5 basis points.

A.2.2 Open-to-close industry ETF return data. We evaluate the performance of the different portfolios on a dataset with open-to-close returns for nine industry ETFs for which we have obtained daily return data from 1998 to 2013 from Bloomberg.[25] The results (not reported) show that the conditional portfolios outperform the benchmark substantially and significantly for transaction costs of 5 basis points, and their performance is similar to that of the benchmark for transaction costs of 10 basis points.

A.2.3 Weekly return data and rebalancing. We evaluate the performance of the different portfolios on weekly return data for the five datasets we consider in the manuscript. The results are reported in Tables A9 and A10 in the appendix for the cases with transaction costs of 0 and 5 basis points, respectively.

We find that our results are generally robust to the use of weekly data. For instance, we find that even with weekly data the norm-constrained conditional mean-variance portfolios with $\delta_1 = 2.5\%$ generally outperform the minimum-variance portfolios in terms of Sharpe ratio for all datasets.[26] Comparing the performance of the conditional portfolios for daily and weekly return data, we find that the conditional portfolios perform slightly better with daily than with weekly data. We believe the reason for this is that the magnitude of the serial dependence that the VAR model captures is larger for higher frequency data.

Table A10 shows that the norm-constrained conditional portfolios with $\delta_1 = 2.5\%$ tend to outperform the minimum-variance portfolio for most weekly datasets even in the presence of proportional transactions costs of 5 basis points, although the differences are not substantial. This is a bit surprising because as one decreases the amount of trading, one

[25]The nine US equity ETFs we consider have tickers XLY, IYZ, XLP, XLE, XLF, XLV, XLB, XLK, XLU. We selected these nine ETFs because they are the ETFs for which data is available for a reasonably long time period (1998–2013) and they also have large trading volumes.

[26]We use an estimation window of 260 weeks.

would expect that the transactions costs associated with the conditional mean-variance portfolios would be smaller, and hence these portfolios would perform better than their daily-rebalanced counterparts. But as discussed above, the degree of predictability decreases with data frequency, and hence the advantage of trading less frequently (and thus incurring lower transactions costs) is offset by the lower degree of predictability in the lower frequency data.

A.3 High turnover, size, and price stocks and Dow Jones stocks

We have evaluated the performance of the conditional portfolios on the 100CRSP dataset where at the beginning of each calendar year we choose the 100 stocks with highest turnover, size, or price as our investment universe, and also where we choose the stocks in the Dow Jones index.

Table A11 in the appendix reports the results for the sample of stocks with high turnover. We find that the conditional portfolios outperform the benchmark for transaction costs of 10 basis points, and the difference in Sharpe ratios is both substantial and statistically significant; that is, the performance of the conditional portfolios is better for high turnover stocks that for our base case with stocks selected from the S&P500 index. This results is particularly relevant as high turnover stocks are unlikely to suffer from the effects of asynchronous or infrequent trading.

The results for stocks with large size, high price, and stocks in the Dow-Jones, not reported to conserve space, show that the conditional portfolios outperform the benchmark for transaction costs of up to 5 basis points. They also outperform the benchmark for transaction costs below 10 basis points, when the threshold of the norm constraint is sufficiently low ($\delta_2 = 5\%$). Summarizing, we find that our results are better for stocks with large turnover, and robust for stocks with large size and price, and for stocks in the Dow Jones index.

A.4 In-sample optimal portfolios with proportional transactions costs

We have used norm constraints to control the high turnover of the conditional mean-variance portfolios and reduce the impact of transactions costs. An alternative approach is to impose the transactions costs explicitly in the mean-variance portfolio optimization problem, and thus, obtain a portfolio that is optimal (at least in-sample) in the presence of proportional transactions costs. In particular, one could solve the following mean-variance problem with proportional costs:

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \Sigma \mathrm{w} - \frac{1}{\gamma}\mathrm{w}^\top \mu + \kappa \|\mathrm{w} - \mathrm{w}_0\|_1, \tag{A1}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1, \tag{A2}$$

where κ is the rate of proportional transactions cost, w_0 is the portfolio before trading, $\|\mathrm{w} - \mathrm{w}_0\|_1$ is the one norm of the difference between the portfolio weights before and after trading, and hence, $\kappa\|\mathrm{w} - \mathrm{w}_0\|_1$ is the transactions cost.

To understand whether this alternative approach is effective, we have evaluated the out-of-sample performance of the conditional portfolios from VAR and NAR computed by solving the problem in (A1)–(A2). Surprisingly, we find that their *out-of-sample* performance in the presence of transaction costs is only slightly better than that of the *unconstrained* conditional mean variance portfolios, which are computed ignoring transaction costs. Moreover, we find that the performance of the conditional portfolios computed by solving (A1)–(A2) is much worse in the presence of transaction costs than that of the *norm-constrained* conditional mean-variance portfolios studied in Section 4.2.

The explanation for this is that the portfolios computed by solving the problem in (A1)–(A2) are much more sensitive to estimation error than the norm-constrained conditional portfolios that we consider. To illustrate this, we consider the following simple two-asset example adapted from the example in Footnote 8 of DeMiguel et al. (2009b). Suppose that the true per annum conditional mean and conditional volatility of returns for both assets are the same, 8% and 20%, respectively, and that the conditional correlation is 0.99. In this case, because the two assets are identical, the optimal conditional mean-variance weights for the two assets would be 50%. Moreover, assume that there are transaction costs of 5 basis points, the starting portfolio w_0 is equal to the optimal equal-weighted portfolio, and the benchmark portfolio for the norm constraints is also equal to the optimal equal-weighted portfolio.

Then it is straightforward to see that if all conditional moments where estimated without error, all three conditional portfolios (the unconstrained conditional portfolio that ignores transaction costs, the conditional portfolio computed by solving the problem in (A1)–(A2), and the norm-constrained conditional portfolio) would be equal to the optimal equal-weighted portfolio. If, on the other hand, the conditional mean return on the first asset is estimated with error to be 9% instead of 8%, then simple computations show that the unconstrained conditional mean-variance portfolio that ignores transaction costs would recommend a weight of 635% in the first asset and −535% in the second asset; the conditional portfolio computed by solving (A1)–(A2) would recommend a weight of 612% in the first asset and −512% in the second asset; and the norm-constrained conditional portfolio with $\delta = 5\%$ would recommend a weight of 52.5% in the first asset and 47.5% in the second asset. That is, the norm-constrained conditional portfolios would be much closer to the optimal portfolio than the conditional portfolio computed by solving (A1)–(A2).

Roughly speaking, the advantage of the norm-constraint is that it imposes an *absolute* limit on trading (a limit of δ around the benchmark portfolio), whereas the transaction costs in the objective function (A1) do not impose a limit, but rather induce a comparison between the size of the *estimated* conditional utility and the size of the transaction costs, where the conditional utility is estimated with error. As a result, we observe that the weights of the portfolios computed solving (A1)–(A2) fluctuate excessively from one period to the next due to estimation error, and their performance is quite poor in the presence of transaction costs.

B. Proofs for all the Propositions

B.1 Proof for Proposition 1

Part 1. From equation (7), we have that the VAR arbitrage portfolio is

$$\mathrm{w}_{vt} = \frac{1}{N}\left(a + Br_{t-1} - \frac{e^\top (a + Br_{t-1})}{N} e\right).$$

Because $a = (I - B)\mu$, we have that

$$\begin{aligned} \mathrm{w}_{vt} &= \frac{1}{N}\left(B(r_{t-1} - \mu) - \frac{e^\top B (r_{t-1} - \mu)}{N} e\right) + \frac{1}{N}\left(\mu - \frac{e^\top \mu}{N} e\right), \\ &= \mathrm{w}_{\Gamma t} + \mathrm{w}_{\mu t}. \end{aligned}$$

Part 2. We first compute $E[\mathrm{w}_{\Gamma t}^\top r_t]$. Note that $E[\mathrm{w}_{\Gamma t}^\top] = 0$, and thus $E[\mathrm{w}_{\Gamma_1 t}^\top r_t] = E[\mathrm{w}_{\Gamma_1 t}^\top (r_t - \mu)]$. From the definition of $\mathrm{w}_{\Gamma_1 t}$ we have that

$$E[\mathrm{w}_{\Gamma_1 t}^\top (r_t - \mu)] = \frac{1}{N} E\left[(r_{t-1} - \mu)^\top B^\top (r_t - \mu)\right] + \frac{1}{N^2} E\left[e^\top \left((r_{t-1} - \mu)^\top B^\top e\right)(r_t - \mu)\right].$$

We now manipulate each of the two expectations on the right-hand side of the equation above. The first expectation can be rewritten as follows:

$$\begin{aligned} E[(r_{t-1} - \mu)^\top B^\top (r_t - \mu)] &= E\left[\mathrm{tr}\left((r_{t-1} - \mu)^\top B^\top (r_t - \mu)\right)\right], \\ &= E\left[\mathrm{tr}\left(B^\top (r_t - \mu)(r_{t-1} - \mu)^\top\right)\right], \\ &= \mathrm{tr}\left(B^\top \Gamma_1^\top\right) = \mathrm{tr}\left(\Gamma_1^\top B^\top\right). \end{aligned}$$

The second expectation can be rewritten as follows:

$$\begin{aligned} E\left[e^\top ((r_{t-1} - \mu)^\top B^\top e)(r_t - \mu)\right] &= \mathrm{tr}(B^\top e e^\top E[(r_t - \mu)(r_{t-1} - \mu)^\top]), \\ &= \mathrm{tr}(B^\top e e^\top \Gamma_1^\top) = \mathrm{tr}(e^\top \Gamma_1^\top B^\top e). \end{aligned}$$

Equation (11) follows from the fact that $B = \Gamma_1^\top \Gamma_0^{-1}$, and clearly $E[\mathrm{w}_{\Gamma_1 t}^\top (r_t - \mu)] \geq 0$.

Finally, we now compute

$$
\begin{aligned}
E[\mathrm{w}_{\mu t}^{\top} r_t] &= E[\frac{1}{N}\left(\mu - \frac{e^{\top}\mu}{N}e\right)^{\top} r_t], \\
&= \frac{1}{N}\left(\mu - \frac{e^{\top}\mu}{N}e\right)^{\top} \mu, \\
&= \sigma^2(\mu) \geq 0,
\end{aligned}
$$

which concludes the proof. ■

B.2 Proof for Proposition 2

The result follows from Part 1 of Proposition 1 and the fact that $B = B_C + B_O$. ■

B.3 Proof for Proposition 3

From Proposition 1 we have that

$$
E[\mathrm{w}_{vt}^{\top} r_t] = G + \sigma^2(\mu) = \frac{\mathrm{tr}(B\Gamma_0 B^{\top})}{N} - \frac{e^{\top}B\Gamma_0 B^{\top}e}{N^2} + \sigma^2(\mu).
$$

Because the covariance matrix Γ_0 is symmetric and positive definite, we know that we can write $\Gamma_0 = Q\Lambda_0 Q^{\top}$, where Q is an orthogonal matrix ($Q^{\top}Q = I$) whose columns are the principal components of Γ_0, and Λ_0 is a diagonal matrix whose elements are the eigenvalues of Γ_0, which are equal to the variances of the principal components of Γ_0. Hence,

$$
E[\mathrm{w}_{vt}^{\top} r_t] = \frac{\mathrm{tr}(BQ\Lambda_0 Q^{\top}B^{\top})}{N} - \frac{e^{\top}BQ\Lambda_0 Q^{\top}B^{\top}e}{N^2} + \sigma^2(\mu).
$$

Let $\hat{B} = BQ$, then

$$
\begin{aligned}
E[\mathrm{w}_{vt}^{\top} r_t] &= \frac{\mathrm{tr}(\hat{B}\Lambda_0 \hat{B}^{\top})}{N} - \frac{e^{\top}\hat{B}\Lambda_0 \hat{B}^{\top}e}{N^2} + \sigma^2(\mu) \\
&= \frac{\sum_{i,j}\lambda_j \hat{B}_{ij}^2}{N} - \frac{\sum_{i,j,k}\lambda_j \hat{B}_{ij}\hat{B}_{kj}}{N^2} + \sigma^2(\mu) \\
&= \frac{1}{N^2}\sum_{i,j,k}\lambda_j \hat{B}_{ij}(\hat{B}_{ij} - \hat{B}_{kj}) + \sigma^2(\mu)
\end{aligned}
$$

$$= \frac{1}{N} \sum_{i,j} \lambda_j \hat{B}_{ij} (\hat{B}_{ij} - \hat{B}_{-j}) + \sigma^2(\mu),$$

where $\hat{B}_{-j} = \sum_i \hat{B}_{ij}/N$. Moreover, because $\sum_i \hat{B}_{-j}(\hat{B}_{ij} - \hat{B}_{-j}) = 0$, we have that

$$E[\mathrm{w}_{vt}^\top r_t] = \frac{1}{N} \sum_j \lambda_j \sum_i (\hat{B}_{ij} - \hat{B}_{-j})^2 + \sigma^2(\mu),$$

$$= \frac{N-1}{N} \sum_j \lambda_j \mathrm{var}(\hat{B}_{\bullet j}) + \sigma^2(\mu),$$

where $\mathrm{var}(\hat{B}_{\bullet j})$ is the variance of the elements in the jth column of matrix $\hat{B}$.

B.4 Proof for Proposition 4

The proof is very similar to those for Propositions 1 and 3.

Table 1: Empirical results for arbitrage (zero-cost) portfolios

This table reports the in-sample and out-of-sample characteristics of the different arbitrage portfolios for the five datasets considered. Panel A gives the in-sample expected returns of the contrarian, momentum, and VAR arbitrage portfolios, which are calculated using the results in Sections 3.1.1 and 3.1.2. Panel B reports the out-of-sample expected returns of the different portfolios. Panels C–D report the out-of-sample Sharpe ratio of the different arbitrage portfolios for the five datasets considered, in the presence of transaction costs of 0, 5, and 10 basis points, respectively, together with the p-values that the Sharpe ratios for the different portfolios are different from those for the corresponding VAR arbitrage portfolio.

Quantity/strategy	6FF	25FF	10Ind	48Ind	100CRSP
Panel A. In-sample expected returns					
C	1.1817	1.4719	0.4569	0.7531	-0.1763
O	-1.4050	-1.7076	-0.8773	-1.0163	0.4915
$\sigma^2(\mu)$	0.0017	0.0019	0.0003	0.0007	0.0028
Contrarian	-0.2250	-0.2376	-0.4207	-0.2640	0.3124
Momentum	0.2250	0.2376	0.4207	0.2640	-0.3124
VAR cross ($E[\mathrm{w}_{Ct}^\top r_t]$)	0.0304	0.2657	0.0608	0.3057	0.8922
VAR auto ($E[\mathrm{w}_{Ot}^\top r_t]$)	0.0342	0.0448	0.0297	0.0277	0.0556
VAR mean ($E[\mathrm{w}_{\mu t}^\top r_t]$)	0.0017	0.0019	0.0003	0.0007	0.0028
VAR total ($E[\mathrm{w}_{vt}^\top r_t]$)	0.0662	0.3125	0.0907	0.3341	0.9506
Panel B. Out-of-sample expected returns					
Contrarian	-0.1932	-0.1865	-0.4002	-0.2796	0.2289
Momentum	0.1932	0.1865	0.4002	0.2796	-0.2289
VAR cross ($E[\mathrm{w}_{Ct}^\top r_t]$)	-0.3982	0.2088	-0.0675	0.2135	0.4848
VAR auto ($E[\mathrm{w}_{Ot}^\top r_t]$)	0.6898	0.1818	0.5024	0.2310	0.0326
VAR mean ($E[\mathrm{w}_{\mu t}^\top r_t]$)	0.0114	0.0086	-0.0028	0.0029	-0.0071
VAR total ($E[\mathrm{w}_{vt}^\top r_t]$)	0.3031	0.3992	0.4321	0.4475	0.5102
Panel C. Out-of-sample Sharpe ratios, no transaction costs					
Contrarian	-2.0804	-2.2784	-3.0846	-2.1028	0.9719
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Momentum	2.0804	2.2784	3.0846	2.1028	-0.9719
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VAR mean ($E[\mathrm{w}_{\mu t}^\top r_t]$)	0.2639	0.6049	-0.0092	0.2522	-0.5281
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VAR total ($E[\mathrm{w}_{vt}^\top r_t]$)	3.5979	4.8768	3.9208	4.0031	3.1908
Panel D. Out-of-sample Sharpe ratios, transaction costs of 5 bps					
Contrarian	-5.4935	-6.3778	-5.5978	-4.6862	-0.5437
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Momentum	-1.3347	-1.8089	0.5702	-0.4803	-2.4901
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VAR mean ($E[\mathrm{w}_{\mu t}^\top r_t]$)	-5.5331	-13.0008	-6.0236	-11.7469	-14.9267
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VAR total ($E[\mathrm{w}_{vt}^\top r_t]$)	0.4241	1.5140	1.2890	1.2275	1.0878
Panel E. Out-of-sample Sharpe ratios, transaction costs of 10 bps					
Contrarian	-8.8374	-10.4510	-8.0969	-7.2657	-2.0578
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
Momentum	-4.6822	-5.8457	-1.9324	-3.0595	-4.0085
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VAR mean ($E[\mathrm{w}_{\mu t}^\top r_t]$)	-11.0857	-25.0183	-11.9030	-23.4676	-29.1792
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
VAR total ($E[\mathrm{w}_{vt}^\top r_t]$)	-2.5957	-1.6899	-1.3234	-1.5232	-1.0077

Table 2: Sharpe ratios for investment (positive-cost) portfolios

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Panel A. Portfolios that ignore stock return serial dependence					
1/N	0.8100	0.8458	0.7669	0.7690	0.6244
	(0.00)	(0.00)	(0.01)	(0.00)	(0.90)
Minimum variance	1.0697	1.0331	0.9507	1.0165	0.6132
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	1.0696	1.0332	0.9506	1.0165	0.6132
	(0.89)	(0.00)	(0.90)	(0.68)	(0.16)
norm cons. (δ_2)	1.0766	1.0334	0.9522	1.0165	0.5657
	(0.00)	(0.00)	(0.76)	(0.61)	(0.09)
norm cons. (δ_3)	1.0898	1.0454	0.9519	1.0152	0.4456
	(0.00)	(0.01)	(0.92)	(0.93)	(0.02)
Panel B. Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	1.0956	1.0460	0.9766	1.0251	0.6273
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	1.1308	1.0957	1.0285	1.1092	0.7381
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	1.2037	1.2250	1.1350	1.2879	0.9666
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	1.1012	1.0433	0.9853	1.0215	0.6245
	(0.00)	(0.00)	(0.00)	(0.00)	(0.22)
norm cons. (δ_2)	1.1394	1.0971	1.0367	1.0719	0.6772
	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)
norm cons. (δ_3)	1.2206	1.2185	1.1367	1.2039	0.7233
	(0.00)	(0.00)	(0.00)	(0.00)	(0.11)

Table 3: Turnovers for investment (positive-cost) portfolios

This table reports the daily turnovers for the different investment portfolios and datasets.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Panel A. Portfolios that ignore stock return serial dependence					
1/N	0.0027	0.0031	0.0044	0.0065	0.0144
Minimum variance	0.0042	0.0097	0.0049	0.0196	0.0232
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.0043	0.0097	0.0049	0.0196	0.0232
norm cons. (δ_2)	0.0049	0.0097	0.0068	0.0196	0.0251
norm cons. (δ_3)	0.0067	0.0122	0.0106	0.0228	0.0310
Panel B. Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.0209	0.0135	0.0168	0.0223	0.0261
norm cons. (δ_2)	0.0479	0.0363	0.0476	0.0593	0.0562
norm cons. (δ_3)	0.1059	0.1075	0.1100	0.1487	0.1237
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.0263	0.0131	0.0286	0.0213	0.0333
norm cons. (δ_2)	0.0594	0.0436	0.0733	0.0560	0.0828
norm cons. (δ_3)	0.1308	0.1395	0.1606	0.1713	0.2011

Table 4: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 5 basis points

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets in the presence of a proportional transactions cost of 5 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Panel A. Portfolios that ignore stock return serial dependence					
1/N	0.8079	0.8433	0.7634	0.7641	0.6161
	(0.00)	(0.00)	(0.01)	(0.00)	(0.86)
Minimum variance	1.0659	1.0246	0.9460	0.9968	0.5943
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	1.0658	1.0246	0.9459	0.9968	0.5943
	(0.55)	(0.00)	(0.79)	(0.56)	(0.21)
norm cons. (δ_2)	1.0723	1.0249	0.9456	0.9969	0.5454
	(0.00)	(0.00)	(0.96)	(0.50)	(0.09)
norm cons. (δ_3)	1.0838	1.0347	0.9417	0.9926	0.4210
	(0.00)	(0.03)	(0.74)	(0.84)	(0.01)
Panel B. Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	1.0769	1.0341	0.9603	1.0028	0.6060
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	1.0881	1.0639	0.9823	1.0497	0.6927
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	1.1091	1.1305	1.0280	1.1390	0.8681
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	1.0777	1.0318	0.9575	1.0001	0.5974
	(0.00)	(0.00)	(0.00)	(0.00)	(0.74)
norm cons. (δ_2)	1.0864	1.0587	0.9655	1.0158	0.6101
	(0.00)	(0.00)	(0.00)	(0.01)	(0.62)
norm cons. (δ_3)	1.1037	1.0958	0.9810	1.0328	0.5639
	(0.00)	(0.00)	(0.00)	(0.07)	(0.62)

Table 5: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 10 basis points

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets in the presence of a proportional transactions cost of 10 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Panel A. Portfolios that ignore stock return serial dependence					
1/N	0.8058	0.8409	0.7600	0.7591	0.6077
	(0.00)	(0.00)	(0.01)	(0.00)	(0.78)
Minimum variance	1.0622	1.0160	0.9413	0.9771	0.5754
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	1.0619	1.0161	0.9411	0.9771	0.5754
	(0.35)	(0.00)	(0.80)	(0.42)	(0.23)
norm cons. (δ_2)	1.0679	1.0164	0.9389	0.9772	0.5250
	(0.01)	(0.00)	(0.71)	(0.39)	(0.10)
norm cons. (δ_3)	1.0778	1.0240	0.9314	0.9699	0.3963
	(0.00)	(0.12)	(0.51)	(0.70)	(0.00)
Panel B. Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	1.0583	1.0222	0.9440	0.9804	0.5848
	(0.01)	(0.00)	(0.14)	(0.03)	(0.01)
norm cons. (δ_2)	1.0453	1.0320	0.9360	0.9901	0.6472
	(0.00)	(0.00)	(0.26)	(0.13)	(0.00)
norm cons. (δ_3)	1.0144	1.0359	0.9211	0.9901	0.7695
	(0.00)	(0.02)	(0.07)	(0.55)	(0.00)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	1.0543	1.0204	0.9298	0.9786	0.5703
	(0.00)	(0.00)	(0.00)	(0.00)	(0.56)
norm cons. (δ_2)	1.0334	1.0204	0.8943	0.9596	0.5429
	(0.00)	(0.19)	(0.00)	(0.01)	(0.28)
norm cons. (δ_3)	0.9869	0.9730	0.8252	0.8616	0.4046
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)

Table 6: Sharpe ratios for conditional portfolios based on lagged-factor models with zero transaction costs

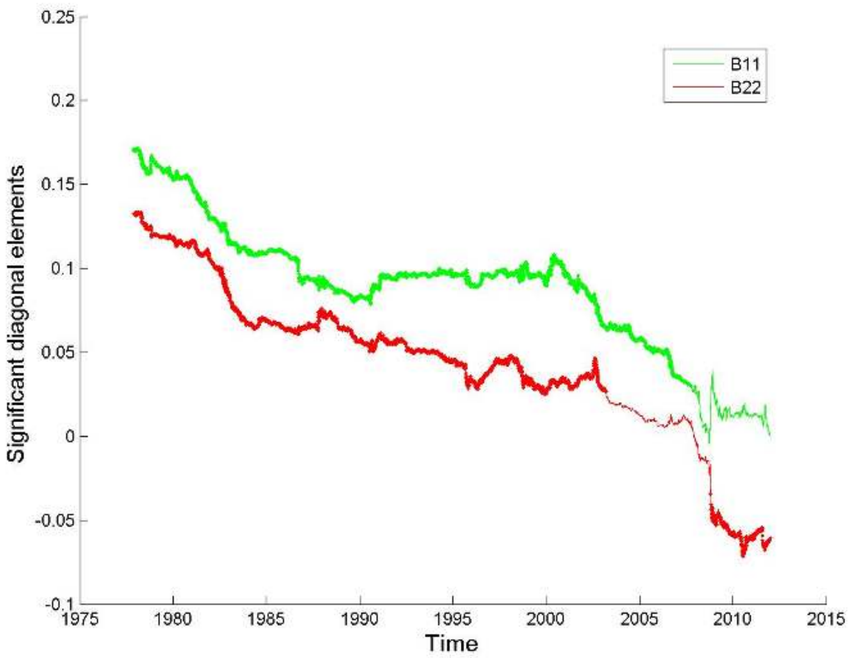
This table reports the annualized out-of-sample Sharpe ratios for the different constrained portfolios and datasets, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Panel A. Portfolios that ignore stock return serial dependence					
1/N	0.8518	0.8947	0.7654	0.7740	0.6389
	(0.00)	(0.00)	(0.02)	(0.00)	(0.22)
Minimum variance	1.1087	1.0809	0.9498	1.0153	0.7456
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Panel B. Portfolios that exploit stock return serial dependence					
Four factors					
norm cons. (δ_1)	1.1400	1.0977	0.9662	1.0238	0.7491
	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
norm cons. (δ_2)	1.1790	1.1581	1.0004	1.0786	0.7911
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	1.2621	1.2805	1.0819	1.2145	0.9255
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Market factor					
norm cons. (δ_1)	1.1279	1.0912	0.9613	1.0199	0.7462
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	1.1531	1.1257	0.9855	1.0659	0.7493
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	1.2059	1.1990	1.0370	1.1833	0.7911
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
SMB factor					
norm cons. (δ_1)	1.1122	1.0825	0.9537	1.0160	0.7455
	(0.00)	(0.00)	(0.00)	(0.00)	(0.73)
norm cons. (δ_2)	1.1171	1.0883	0.9613	1.0262	0.7439
	(0.00)	(0.00)	(0.00)	(0.00)	(0.20)
norm cons. (δ_3)	1.1278	1.1121	0.9820	1.0660	0.7404
	(0.00)	(0.00)	(0.00)	(0.00)	(0.48)
HML factor					
norm cons. (δ_1)	1.1313	1.0939	0.9612	1.0178	0.7471
	(0.00)	(0.00)	(0.00)	(0.00)	(0.22)
norm cons. (δ_2)	1.1588	1.1339	0.9838	1.0468	0.7648
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
norm cons. (δ_3)	1.2171	1.2232	1.0334	1.1219	0.7900
	(0.00)	(0.00)	(0.00)	(0.00)	(0.13)
UMD factor					
norm cons. (δ_1)	1.1087	1.0820	0.9475	1.0154	0.7448
	(0.96)	(0.01)	(0.03)	(0.80)	(0.26)
norm cons. (δ_2)	1.1093	1.0835	0.9477	1.0161	0.7203
	(0.72)	(0.15)	(0.42)	(0.84)	(0.00)
norm cons. (δ_3)	1.1095	1.0830	0.9466	1.0057	0.6804
	(0.86)	(0.76)	(0.63)	(0.37)	(0.01)

Figure 1: Two Size-Sorted Portfolios

Figure (a) depicts the time evolution of the diagonal elements of the slope matrix, while Figure (b) depicts the time evolution of the off-diagonal elements of the slope matrix. The solid lines give the estimated value of these elements, and we use thicker lines for periods when the elements are statistically significant.

(a) Diagonal elements of the slope matrix



(b) Off-diagonal elements of the slope matrix

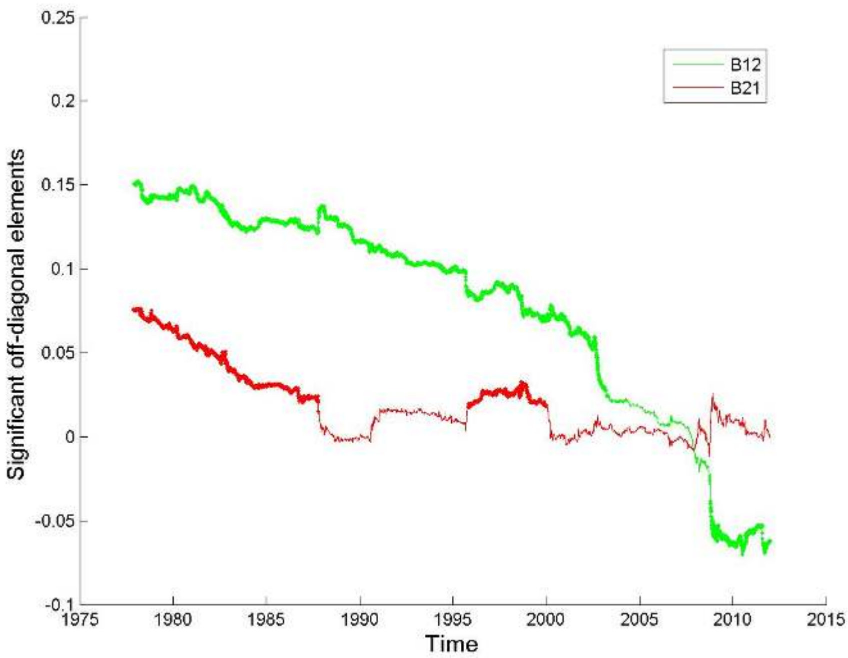
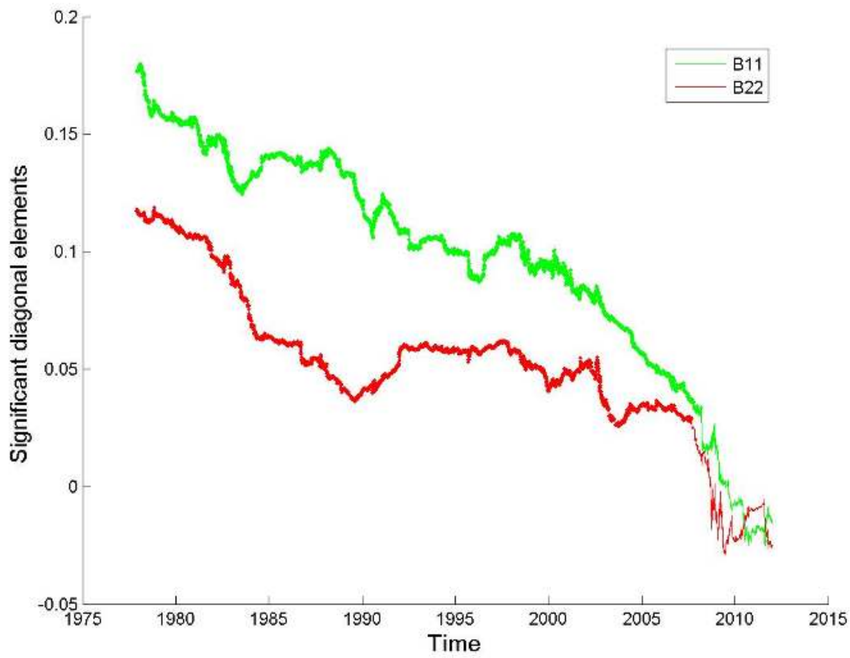


Figure 2: Two Book-to-Market-Sorted Portfolios

Figure (a) depicts the time evolution of the diagonal elements of the slope matrix, while Figure (b) depicts the time evolution of the off-diagonal elements of the slope matrix. The solid lines give the estimated value of these elements, and we use thicker lines for periods when the elements are statistically significant.

(a) Diagonal elements of the slope matrix



(b) Off-diagonal elements of the slope matrix

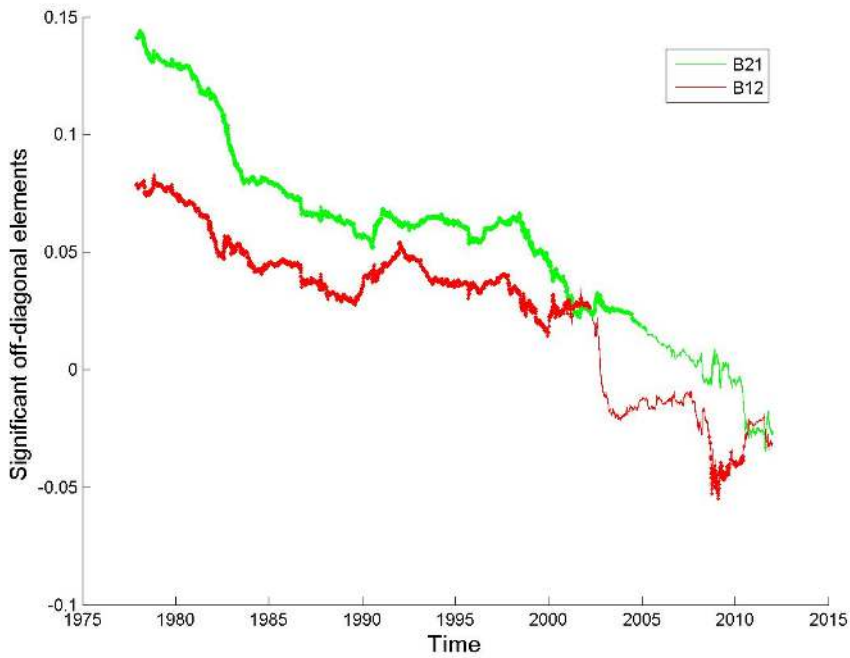
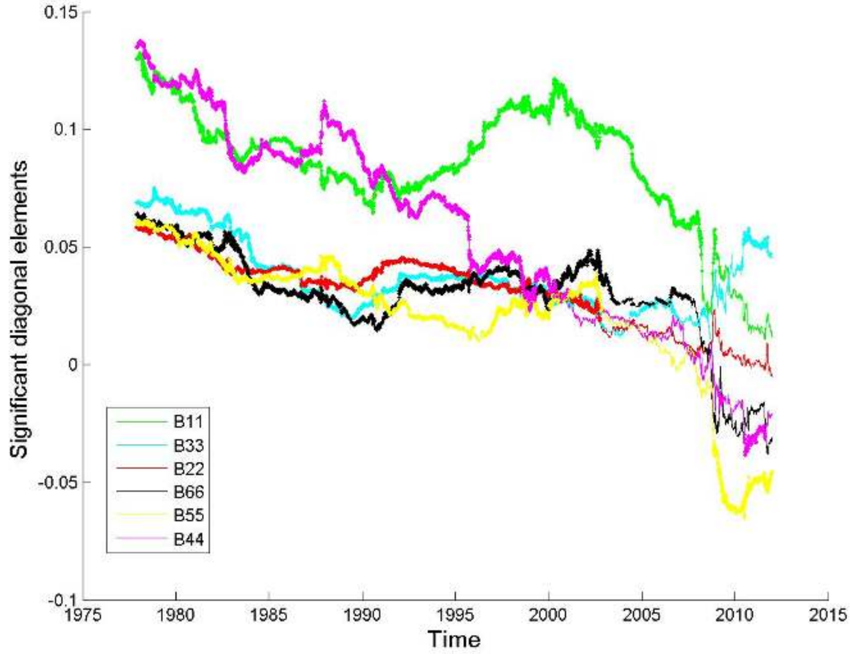


Figure 3: Six Size- and Book-to-Market-Sorted Portfolios

Figure (a) depicts the time evolution of the diagonal elements of the slope matrix, while Figure (b) depicts the time evolution of the off-diagonal elements of the slope matrix. The solid lines give the estimated value of these elements, and we use thicker lines for periods when the elements are statistically significant.

(a) Diagonal elements of the slope matrix



(b) Off-diagonal elements of the slope matrix

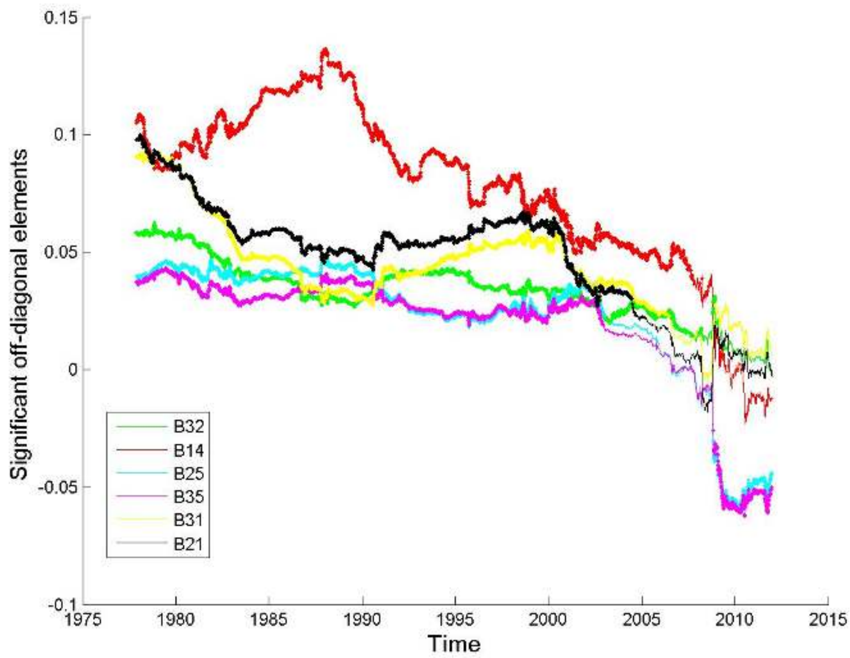
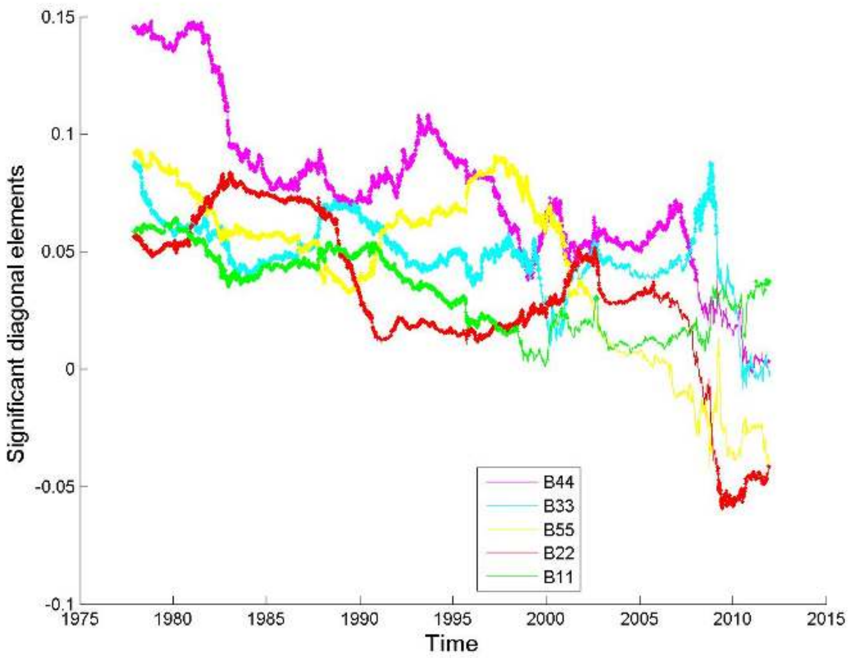


Figure 4: Five Industry Portfolios

Figure (a) depicts the time evolution of the diagonal elements of the slope matrix, while Figure (b) depicts the time evolution of the off-diagonal elements of the slope matrix. The solid lines give the estimated value of these elements, and we use thicker lines for periods when the elements are statistically significant.

(a) Diagonal elements of the slope matrix



(b) Off-diagonal elements of the slope matrix

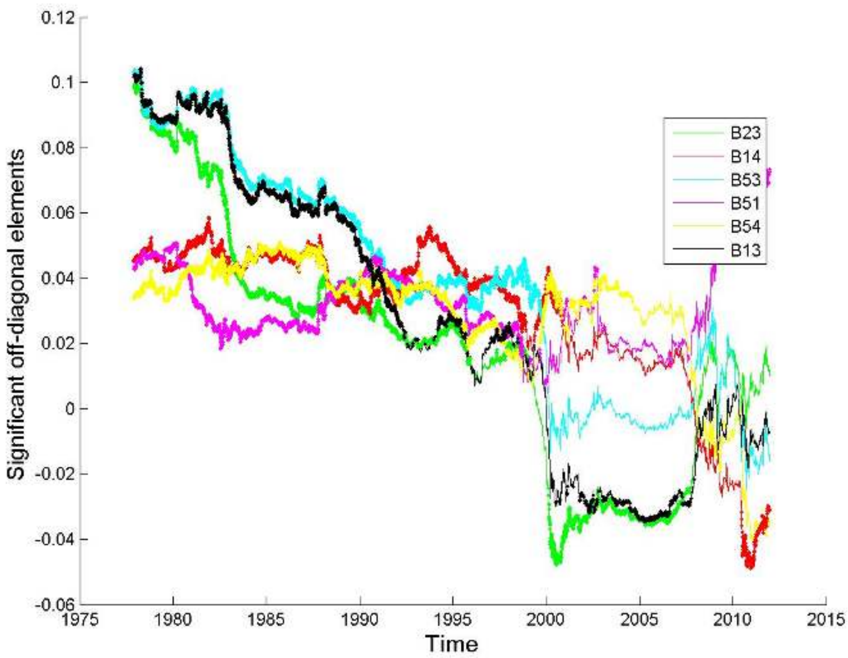
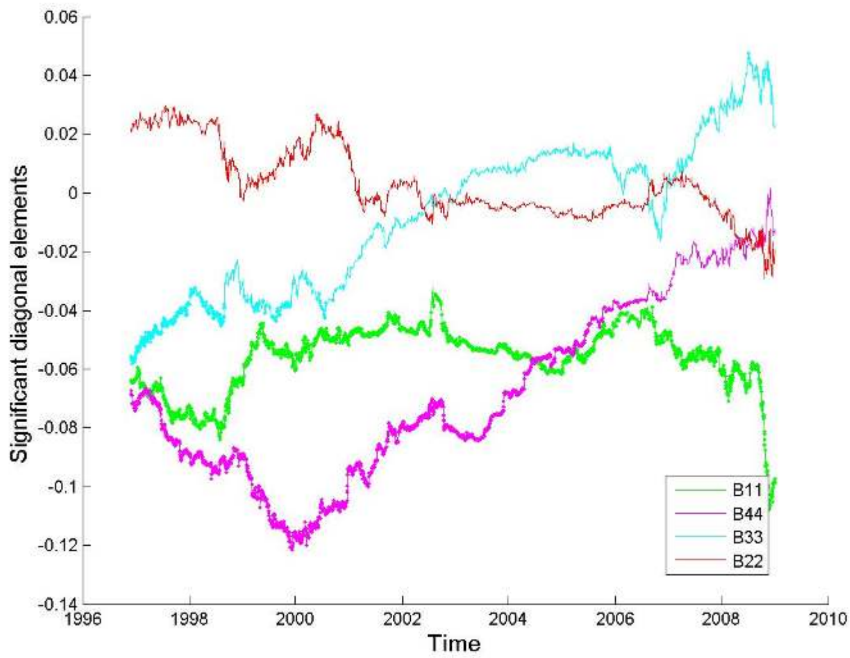


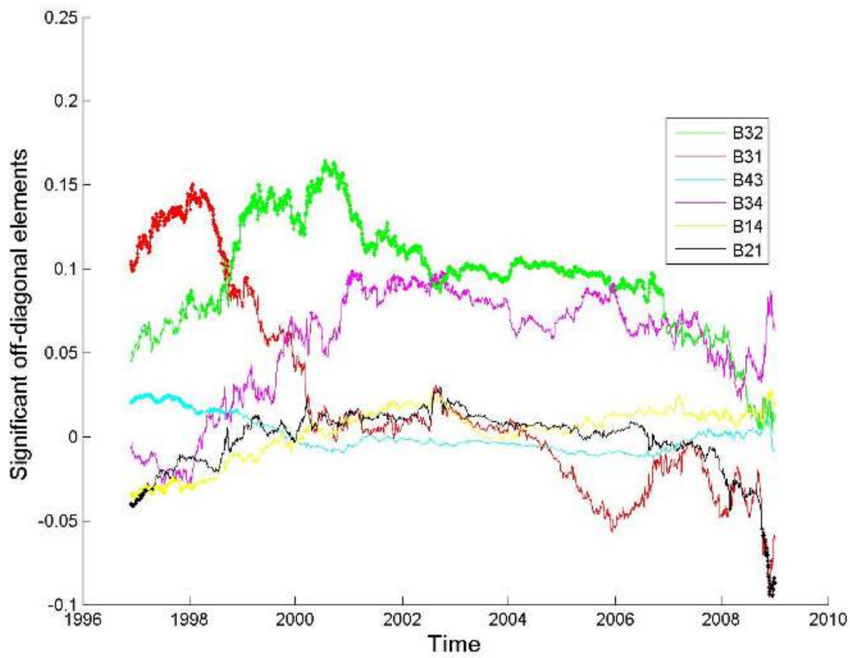
Figure 5: Four Individual Stocks

Figure (a) depicts the time evolution of the diagonal elements of the slope matrix, while Figure (b) depicts the time evolution of the off-diagonal elements of the slope matrix. The solid lines give the estimated value of these elements, and we use thicker lines for periods when the elements are statistically significant.

(a) Diagonal elements of the slope matrix



(b) Off-diagonal elements of the slope matrix



References

Anderson, R., K. Eom, S. Hahn, and J. Park. 2005. Stock Return Autocorrelation is Not Spurious. Preprint.

Badrinath, S. G., J. R. Kale, and T. H. Noe. 1995. Of Shepherds, Sheep, and the Cross-autocorrelations in Equity Returns. *Review of Financial Studies* 8:401–30.

Balduzzi, P., and A. W. Lynch. 1999. Transaction Costs and Predictability: Some Utility Cost Calculations. *Journal of Financial Economics* 52:47–78.

Barberis, N. 2000. Investing for the Long Run When Returns Are Predictable. *Journal of Finance* 55:225–64.

Barberis, N., A. Shleifer, and R. Vishny. 1998. A Model of Investor Sentiment. *Journal of Financial Economics* 49:307–43.

Berk, J. B., R. C. Green, and V. Naik. 1999. Optimal Investment, Growth Options, and Security Returns. *Journal of Finance* 54:1553–1607.

Brennan, M. J., N. Jegadeesh, and B. Swaminathan. 1993. Investment Analysis and the Adjustment of Stock Prices to Common Information. *Review of Financial Studies* 6:799–824.

Campbell, J. Y., Y. L. Chan, and L. M. Viceira. 2003. A Multivariate Model of Strategic Asset Allocation. *Journal of Financial Economics* 67:41–80.

Campbell, J. Y., and L. M. Viceira. 1999. Consumption and Portfolio Decisions When Expected Returns Are Time Varying. *Quarterly Journal of Economics* 114:433–95.

Campbell, J. Y., and L. M. Viceira. 2002. *Strategic Asset Allocation*. New York: Oxford University Press.

Chan, L. K. C., J. Karceski, and J. Lakonishok. 1999. On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *Review of Financial Studies* 12:937–974.

Chordia, T., and B. Swaminathan. 2000. Trading Volume and Cross-Autocorrelations in Stock Returns. *The Journal of Finance* LV:913–935.

Cohen, L., and A. Frazzini. 2008. Economic Links and Predictable Returns. *Journal of Finance* 63:1977–2011.

Conrad, J., and G. Kaul. 1988. Time-Variation in Expected Returns. *Journal of Business* 61:409–425.

———. 1989. Mean Reversion in Short-Horizon Expected Returns. *Review of Financial Studies* 2:225–240.

———. 1998. An Anatomy of Trading Strategies. *Review of Financial Studies* 11:489–519. URL `http://rfs.oxfordjournals.org/content/11/3/489.abstract`.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam. 1998. Investor Psychology and Security Market Under- and Overreactions. *Journal of Finance* 53:1839–85.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann. 1990. Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *Journal of Finance* 45:379–395.

DeBondt, W., and R. H. Thaler. 1985. Does the Stock Market Overreact? *Journal of Finance* 40:793–808.

DeMiguel, V., L. Garlappi, F. Nogales, and R. Uppal. 2009a. A Generalized Approach to Portfolio Optimization: Improving Performance By Constraining Portfolio Norms. *Management Science* 55:798–812.

DeMiguel, V., L. Garlappi, and R. Uppal. 2009b. Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies* 22:1915–1953.

DeMiguel, V., and F. J. Nogales. 2009. Portfolio Selection with Robust Estimation. *Operations Research* 57:560–577.

Eun, C. S., and S. Shim. 1989. International Transmission of Stock Market Movements. *Journal of Financial and Quantitative Analysis* 24:241–256.

Fama, E. F., and K. R. French. 1988. Permanent and Temporary Components of Stock Prices. *Journal of Political Economy* 96:246–73.

French, K. R. 2008. Presidential Address: The Cost of Active Investing. *Journal of Finance* 63:1537–1573.

Gyorfi, L., M. Kohler, A. Krzyzak, and H. Walk. 1987. *A Distribution-Free Theory of Nonparametric Regression*. New York: Springer.

Gyorfi, L., F. Udina, and H. Walk. 2007. Experiments on Universal Portfolio Selection Using Data From Real Markets. UPF working paper 2007.

———. 2008. Nonparametric Nearest Neighbor Based Empirical Portfolio Selection Strategies. *Statistics and Decisions* 26:145–157.

Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin. 2005. The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer* 27:83–85.

Hoerl, A. E., and R. W. Kennard. 1970. Ridge Regression: Applications to Nonorthogonal Problems. *Technometrics* 12:69–82.

Hong, H., and J. C. Stein. 1999. A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *Journal of Finance* 54:2143–84.

Hou, K. 2007. Industry Information Diffusion and the Lead-lag Effect in Stock Returns. *Review of Financial Studies* 20:1113–1138.

Hou, K., and T. J. Moskowitz. 2005. Market Frictions, Price Delay, and the Cross-Section of Expected Returns. *Review of Financial Studies* 18:981–1020.

Jegadeesh, N. 1990. Evidence of Predictable Behavior of Security Returns. *Journal of Finance* 45:881–98.

Jegadeesh, N., and S. Titman. 1993. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance* 48:65–91.

Johnson, T. C. 2002. Rational Momentum Effects. *Journal of Finance* 57:585–608.

Jurek, J. W., and L. M. Viceira. 2011. Optimal Value and Growth Tilts in Long-Horizon Portfolios. *Review of Finance* 15:29–74.

Ledoit, O., and M. Wolf. 2003. Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection. *Journal of Empirical Finance* 10:603–621.

———. 2004. A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices. *Journal of Multivariate Analysis* 88:365–411.

———. 2008. Robust Performance Hypothesis Testing with the Sharpe Ratio. *Journal of Empirical Finance* 15:850–859.

Lewellen, J. 2002. Momentum and Autocorrelation in Stock Returns. *Review of Financial Studies* 15:533–63.

Li, M. 2011. Value vs. Growth: Who Leads the Cyclical Stock Market? *Banking and Finance Review* 3.

Lo, A., and A. C. MacKinlay. 1990. When Are Contrarian Profits Due to Stock Market Overreaction? *The Review of Financial Studies* 3:175–205.

Lynch, A. W. 2001. Portfolio Choice and Equity Characteristics: Characterizing the Hedging Demands Induced by Return Predictability. *Journal of Financial Economics* 62:67–130.

Menzly, L., and O. Ozbas. 2010. Market Segmentation and Cross-predictability of Returns. *Journal of Finance* 65:1555–1580.

Mizrach, B. 1992. Multivariate Nearest-Neighbour Forecasts of EMS Exchange Rates. *Journal of Applied Econometrics* 7:S151–S163.

Moskowitz, T. J., and M. Grinblatt. 1999. Do Industries Explain Momentum? *Journal of Finance* 54:1249–90.

Moskowitz, T. J., Y. H. Ooi, and L. H. Pedersen. 2012. Time Series Momentum. *Journal of Financial Economics* 104:228–250.

Pan, M.-S., K. Liano, and G.-C. Huang. 2004. Industry momentum strategies and autocorrelations in stock returns. *Journal of Empirical Finance* 11:185–202.

Politis, D., and J. Romano. 1994. The Stationary Bootstrap. *Journal of the American Statistical Association* 89:1303–1313.

Roll, R. 1994. What Every CEO Should Know About Scientific Progress in Economics: What is Known and What Remains to be Resolved. *Financial Management* 23:69–75.

Tsay, R. S. 2005. *Analysis of Financial Time Series.* 2nd ed. Wiley-Interscience.

Material for on-line appendix

Table A1: Sharpe ratios for investment (positive-cost) portfolios using excess returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.4879	0.5207	0.4436	0.4541	0.5106
	(0.00)	(0.00)	(0.19)	(0.07)	(0.67)
Minimum variance	0.6966	0.6659	0.5443	0.5956	0.4533
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.6966	0.6659	0.5442	0.5956	0.4533
	(0.96)	(0.00)	(0.77)	(0.63)	(0.20)
norm cons. (δ_2)	0.7033	0.6662	0.5456	0.5957	0.4068
	(0.00)	(0.00)	(0.75)	(0.47)	(0.11)
norm cons. (δ_3)	0.7160	0.6778	0.5461	0.5975	0.2891
	(0.00)	(0.01)	(0.90)	(0.89)	(0.01)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.7214	0.6763	0.5734	0.6051	0.4689
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	0.7573	0.7221	0.6263	0.6924	0.5568
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	0.8266	0.8474	0.7309	0.8826	0.7384
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.7279	0.6759	0.5790	0.6006	0.4645
	(0.00)	(0.00)	(0.00)	(0.00)	(0.24)
norm cons. (δ_2)	0.7659	0.7292	0.6304	0.6507	0.5177
	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
norm cons. (δ_3)	0.8466	0.8501	0.7312	0.7839	0.5673
	(0.00)	(0.00)	(0.00)	(0.00)	(0.10)

Table A2: Turnovers for investment (positive-cost) portfolios using excess returns

This table reports the daily turnovers for the different investment portfolios and datasets.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.0027	0.0031	0.0044	0.0065	0.0144
Minimum variance	0.0042	0.0095	0.0047	0.0193	0.0232
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.0044	0.0095	0.0048	0.0193	0.0232
norm cons. (δ_2)	0.0049	0.0095	0.0067	0.0192	0.0251
norm cons. (δ_3)	0.0068	0.0119	0.0104	0.0224	0.0310
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.0202	0.0132	0.0187	0.0220	0.0265
norm cons. (δ_2)	0.0465	0.0353	0.0488	0.0614	0.0550
norm cons. (δ_3)	0.1024	0.1045	0.1103	0.1554	0.1213
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.0263	0.0129	0.0285	0.0210	0.0333
norm cons. (δ_2)	0.0594	0.0435	0.0732	0.0557	0.0829
norm cons. (δ_3)	0.1308	0.1393	0.1605	0.1710	0.2011

Table A3: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 5 basis points using excess returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets in the presence of a proportional transactions cost of 5 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.4858	0.5183	0.4402	0.4492	0.5023
	(0.00)	(0.00)	(0.18)	(0.10)	(0.58)
Minimum variance	0.6929	0.6575	0.5398	0.5763	0.4344
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.6927	0.6575	0.5396	0.5764	0.4344
	(0.66)	(0.00)	(0.70)	(0.49)	(0.22)
norm cons. (δ_2)	0.6989	0.6578	0.5390	0.5764	0.3865
	(0.00)	(0.00)	(0.93)	(0.39)	(0.11)
norm cons. (δ_3)	0.7099	0.6673	0.5360	0.5752	0.2645
	(0.00)	(0.03)	(0.88)	(0.96)	(0.02)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.7034	0.6647	0.5552	0.5831	0.4474
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	0.7158	0.6911	0.5789	0.6309	0.5123
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	0.7352	0.7555	0.6237	0.7270	0.6415
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.7044	0.6646	0.5513	0.5796	0.4374
	(0.00)	(0.00)	(0.00)	(0.00)	(0.78)
norm cons. (δ_2)	0.7128	0.6910	0.5593	0.5950	0.4505
	(0.00)	(0.00)	(0.00)	(0.01)	(0.63)
norm cons. (δ_3)	0.7297	0.7275	0.5755	0.6132	0.4080
	(0.00)	(0.00)	(0.00)	(0.09)	(0.62)

Table A4: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 10 basis points using excess returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets in the presence of a proportional transactions cost of 10 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.4837	0.5159	0.4367	0.4443	0.4940
	(0.00)	(0.00)	(0.18)	(0.15)	(0.56)
Minimum variance	0.6891	0.6491	0.5352	0.5570	0.4155
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.6888	0.6492	0.5349	0.5571	0.4155
	(0.35)	(0.00)	(0.65)	(0.37)	(0.25)
norm cons. (δ_2)	0.6946	0.6494	0.5325	0.5571	0.3662
	(0.00)	(0.00)	(0.55)	(0.31)	(0.10)
norm cons. (δ_3)	0.7039	0.6569	0.5258	0.5529	0.2398
	(0.00)	(0.11)	(0.54)	(0.82)	(0.01)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.6854	0.6530	0.5370	0.5610	0.4258
	(0.00)	(0.00)	(0.37)	(0.01)	(0.01)
norm cons. (δ_2)	0.6744	0.6600	0.5314	0.5693	0.4678
	(0.00)	(0.00)	(0.41)	(0.12)	(0.03)
norm cons. (δ_3)	0.6437	0.6636	0.5164	0.5713	0.5445
	(0.00)	(0.09)	(0.06)	(0.42)	(0.02)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.6809	0.6533	0.5237	0.5585	0.4103
	(0.00)	(0.00)	(0.00)	(0.00)	(0.58)
norm cons. (δ_2)	0.6598	0.6528	0.4883	0.5392	0.3834
	(0.00)	(0.31)	(0.00)	(0.02)	(0.31)
norm cons. (δ_3)	0.6128	0.6048	0.4198	0.4424	0.2487
	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)

Table A5: Sharpe ratios for investment (positive-cost) portfolios with open-to-close returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.5107	0.5768	0.3603	0.3810	0.3504
	(0.00)	(0.00)	(0.01)	(0.01)	(0.08)
Minimum variance	0.8712	0.8983	0.6772	0.7288	0.5573
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.8816	0.8984	0.6777	0.7288	0.5579
	(0.00)	(0.00)	(0.82)	(0.00)	(0.00)
norm cons. (δ_2)	0.8952	0.9080	0.7022	0.7293	0.6181
	(0.00)	(0.00)	(0.01)	(0.00)	(0.05)
norm cons. (δ_3)	0.9271	0.9738	0.7414	0.7917	0.7477
	(0.00)	(0.00)	(0.00)	(0.01)	(0.02)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.8754	0.9031	0.6800	0.7325	0.5577
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	0.8823	0.9284	0.6908	0.7675	0.5745
	(0.00)	(0.00)	(0.00)	(0.00)	(0.05)
norm cons. (δ_3)	0.8978	0.9762	0.7071	0.8411	0.5973
	(0.00)	(0.00)	(0.00)	(0.00)	(0.06)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.8931	0.9071	0.6933	0.7337	0.5886
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_2)	0.9179	0.9532	0.7160	0.7953	0.6495
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	0.9667	1.0363	0.7513	0.9407	0.6186
	(0.00)	(0.00)	(0.00)	(0.00)	(0.57)

Table A6: Turnovers for investment (positive-cost) portfolios with open-to-close returns

This table reports the daily turnovers for the different investment portfolios and datasets.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.0029	0.0037	0.0047	0.0066	0.0131
Minimum variance	0.0056	0.0115	0.0071	0.0149	0.0287
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.0056	0.0115	0.0074	0.0149	0.0287
norm cons. (δ_2)	0.0059	0.0124	0.0082	0.0148	0.0302
norm cons. (δ_3)	0.0074	0.0151	0.0106	0.0179	0.0351
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.0096	0.0138	0.0093	0.0163	0.0288
norm cons. (δ_2)	0.0202	0.0279	0.0216	0.0407	0.0335
norm cons. (δ_3)	0.0497	0.0674	0.0507	0.1123	0.0513
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.0280	0.0163	0.0288	0.0170	0.0391
norm cons. (δ_2)	0.0603	0.0591	0.0768	0.0572	0.0827
norm cons. (δ_3)	0.1226	0.1550	0.1650	0.1852	0.2025

Table A7: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 5 basis with open-to-close returns points

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets in the presence of a proportional transactions cost of 5 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.5086	0.5741	0.3566	0.3759	0.3411
	(0.00)	(0.00)	(0.01)	(0.00)	(0.07)
Minimum variance	0.8668	0.8889	0.6707	0.7138	0.5325
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.8771	0.8891	0.6709	0.7138	0.5331
	(0.00)	(0.00)	(0.96)	(0.00)	(0.01)
norm cons. (δ_2)	0.8904	0.8978	0.6946	0.7143	0.5921
	(0.00)	(0.00)	(0.02)	(0.00)	(0.04)
norm cons. (δ_3)	0.9212	0.9614	0.7315	0.7735	0.7180
	(0.00)	(0.00)	(0.00)	(0.01)	(0.02)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.8677	0.8919	0.6715	0.7161	0.5329
	(0.14)	(0.00)	(0.47)	(0.00)	(0.00)
norm cons. (δ_2)	0.8661	0.9058	0.6709	0.7266	0.5456
	(0.66)	(0.00)	(0.96)	(0.24)	(0.11)
norm cons. (δ_3)	0.8579	0.9213	0.6604	0.7287	0.5531
	(0.04)	(0.00)	(0.31)	(0.53)	(0.35)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.8706	0.8938	0.6668	0.7167	0.5548
	(0.06)	(0.00)	(0.26)	(0.00)	(0.01)
norm cons. (δ_2)	0.8695	0.9051	0.6451	0.7378	0.5783
	(0.52)	(0.00)	(0.00)	(0.03)	(0.13)
norm cons. (δ_3)	0.8682	0.9101	0.5995	0.7546	0.4560
	(0.88)	(0.06)	(0.00)	(0.21)	(0.58)

Table A8: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 10 basis points with open-to-close returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets in the presence of a proportional transactions cost of 10 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.5066	0.5715	0.3529	0.3708	0.3318
	(0.00)	(0.00)	(0.00)	(0.00)	(0.13)
Minimum variance	0.8623	0.8795	0.6642	0.6989	0.5077
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.8726	0.8797	0.6641	0.6989	0.5083
	(0.00)	(0.00)	(0.98)	(0.00)	(0.01)
norm cons. (δ_2)	0.8857	0.8877	0.6870	0.6994	0.5661
	(0.00)	(0.00)	(0.03)	(0.00)	(0.07)
norm cons. (δ_3)	0.9153	0.9490	0.7216	0.7553	0.6882
	(0.00)	(0.00)	(0.01)	(0.02)	(0.02)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.8600	0.8807	0.6629	0.6997	0.5081
	(0.00)	(0.12)	(0.25)	(0.34)	(0.02)
norm cons. (δ_2)	0.8499	0.8831	0.6510	0.6857	0.5166
	(0.00)	(0.40)	(0.00)	(0.23)	(0.28)
norm cons. (δ_3)	0.8180	0.8664	0.6137	0.6162	0.5088
	(0.00)	(0.18)	(0.00)	(0.00)	(0.98)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	0.8481	0.8805	0.6403	0.6996	0.5210
	(0.00)	(0.42)	(0.00)	(0.18)	(0.15)
norm cons. (δ_2)	0.8211	0.8569	0.5743	0.6803	0.5070
	(0.00)	(0.00)	(0.00)	(0.13)	(0.99)
norm cons. (δ_3)	0.7698	0.7839	0.4478	0.5686	0.2933
	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)

Table A9: Sharpe ratios for investment (positive-cost) portfolios with weekly returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets using weekly returns, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.8304	0.8486	0.8351	0.7917	0.6020
	(0.00)	(0.00)	(0.07)	(0.01)	(0.05)
Minimum variance	0.9955	1.0166	1.0182	1.0315	0.9495
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	1.0001	1.0245	1.0193	1.0353	0.9545
	(0.01)	(0.00)	(0.77)	(0.44)	(0.84)
norm cons. (δ_2)	1.0056	1.0376	1.0196	1.0456	0.9711
	(0.00)	(0.00)	(0.89)	(0.23)	(0.73)
norm cons. (δ_3)	1.0156	1.0573	1.0192	1.0560	0.9561
	(0.00)	(0.00)	(0.92)	(0.32)	(0.95)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	1.0038	1.0274	1.0194	1.0377	0.9505
	(0.00)	(0.00)	(0.62)	(0.07)	(0.94)
norm cons. (δ_2)	1.0171	1.0444	1.0239	1.0521	0.9636
	(0.00)	(0.00)	(0.31)	(0.02)	(0.59)
norm cons. (δ_3)	1.0498	1.0866	1.0285	1.0696	0.9804
	(0.00)	(0.00)	(0.39)	(0.08)	(0.57)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	1.0119	1.0429	1.0252	1.0409	0.9369
	(0.00)	(0.00)	(0.03)	(0.09)	(0.52)
norm cons. (δ_2)	1.0283	1.0718	1.0315	1.0510	0.9194
	(0.00)	(0.00)	(0.06)	(0.11)	(0.44)
norm cons. (δ_3)	1.0610	1.1269	1.0432	1.0647	0.8746
	(0.00)	(0.00)	(0.08)	(0.16)	(0.32)

Table A10: Sharpe ratios for investment (positive-cost) portfolios and transactions costs of 5 basis points with weekly returns

This table reports the annualized out-of-sample Sharpe ratios for the different investment portfolios and datasets with weekly returns, in the presence of a proportional transaction cost of 5 basis points, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy	6FF	25FF	10Ind	48Ind	100CRSP
Portfolios that ignore stock return serial dependence					
1/N	0.8294	0.8475	0.8333	0.7894	0.5982
	(0.00)	(0.00)	(0.07)	(0.03)	(0.04)
Minimum variance	0.9921	1.0098	1.0140	1.0226	0.9393
	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio					
norm cons. (δ_1)	0.9965	1.0174	1.0146	1.0260	0.9435
	(0.03)	(0.00)	(0.83)	(0.54)	(0.84)
norm cons. (δ_2)	1.0018	1.0301	1.0143	1.0355	0.9591
	(0.00)	(0.00)	(0.95)	(0.27)	(0.76)
norm cons. (δ_3)	1.0113	1.0490	1.0125	1.0447	0.9425
	(0.02)	(0.00)	(0.91)	(0.33)	(1.00)
Portfolios that exploit stock return serial dependence					
Conditional mean variance portfolio from VAR					
norm cons. (δ_1)	0.9986	1.0189	1.0134	1.0265	0.9370
	(0.00)	(0.00)	(0.79)	(0.26)	(0.80)
norm cons. (δ_2)	1.0088	1.0324	1.0145	1.0358	0.9434
	(0.00)	(0.00)	(0.94)	(0.16)	(0.86)
norm cons. (δ_3)	1.0346	1.0656	1.0126	1.0428	0.9497
	(0.00)	(0.00)	(0.91)	(0.34)	(0.81)
Conditional mean variance portfolio from NAR					
norm cons. (δ_1)	1.0036	1.0305	1.0137	1.0259	0.9181
	(0.00)	(0.00)	(0.87)	(0.52)	(0.30)
norm cons. (δ_2)	1.0141	1.0516	1.0115	1.0243	0.8918
	(0.00)	(0.00)	(0.71)	(0.88)	(0.23)
norm cons. (δ_3)	1.0347	1.0922	1.0060	1.0195	0.8303
	(0.00)	(0.00)	(0.50)	(0.93)	(0.18)

Table A11: Sharpe ratios for dataset with returns on the 100 stocks with highest turnover in the S&P500, for different levels of transaction costs

This table reports the annualized out-of-sample Sharpe ratios for the different portfolios and for the dataset with returns on the 100 stocks with highest turnover in the S&P500, for different levels of transaction costs, together with the p-value that the Sharpe ratio for a strategy is different from that for the shortsale-constrained minimum-variance portfolio.

Strategy Strategy	100CRSP 0 bp	100CRSP 5 bp	100CRSP 10 bp
Portfolios that ignore stock return serial dependence			
1/N	0.4175	0.4095	0.4015
	(0.80)	(0.82)	(0.93)
Minimum variance	0.4580	0.4368	0.4156
	(1.00)	(1.00)	(1.00)
Unconditional mean variance portfolio			
norm cons. (δ_1)	0.4622	0.4409	0.4197
	(0.35)	(0.38)	(0.38)
norm cons. (δ_2)	0.3958	0.3723	0.3488
	(0.10)	(0.07)	(0.06)
norm cons. (δ_3)	0.3103	0.2815	0.2527
	(0.05)	(0.05)	(0.03)
Portfolios that exploit stock return serial dependence			
Conditional mean variance portfolio from VAR			
norm cons. (δ_1)	0.4879	0.4614	0.4350
	(0.00)	(0.00)	(0.05)
norm cons. (δ_2)	0.6176	0.5645	0.5114
	(0.00)	(0.00)	(0.00)
norm cons. (δ_3)	0.8335	0.7302	0.6267
	(0.00)	(0.00)	(0.00)
Conditional mean variance portfolio from NAR			
norm cons. (δ_1)	0.5154	0.4802	0.4450
	(0.00)	(0.00)	(0.02)
norm cons. (δ_2)	0.5987	0.5213	0.4438
	(0.00)	(0.01)	(0.39)
norm cons. (δ_3)	0.7364	0.5853	0.4340
	(0.00)	(0.02)	(0.74)