DISCUSSION PAPER SERIES

No. 9456

STOCK RETURN SERIAL DEPENDENCE AND OUT-OF-SAMPLE PORTFOLIO PERFORMANCE

Victor DeMiguel, Francisco J. Nogales and Raman Uppal

FINANCIAL ECONOMICS

ISSN 0265-8003

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

Discussion Paper No. 9456
April 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9456

April 2013

ABSTRACT

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

We study whether investors can exploit stock return serial dependence to improve out-of- sample portfolio performance. To do this, we first show that a vector-autoregressive (VAR) model estimated with ridge regression captures daily stock return serial dependence in a stable manner. Second, we characterize (analytically and empirically) expected returns of VAR-based arbitrage portfolios, and show that they compare favorably to those of existing arbitrage portfolios. Third, we evaluate the performance of VAR-based investment (positive-cost) portfolios. We show that, subject to a suitable norm constraint, these portfolios outperform the traditional (unconditional) portfolios for transaction costs below 10 basis points.

JEL Classification: G11
Keywords: out-of-sample performance, portfolio choice, serial dependence and vector autoregression

Victor DeMiguel
London Business School
26 Sussex Place
London NW1 4SA

Email: avmiguel@london.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=162933

Francisco J. Nogales
Universidad Carlos III de Madrid
Av. Universidad 30
28911 Leganes (Madrid)
SPAIN

Email: fcojavier.nogales@uc3m.es

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=177158

Raman Uppal
Edhec Business School
10 Fleet Place, Ludgate
London EC4M 7RB

Email: raman.uppal@edhec.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=135807

*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 Grigory Vilkov and Eric Shirbini for helping us obtain the data we use, to John Campbell and Luis Viceira for making available their computer code, and to David Chapman, Andrew Lo, Anthony Lynch, Peter Schotman, Piet Sercu, and Jules van Binsbergen 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 Asset Pricing Conference, the 2010 Summer Conference at Sauder School of Business, and the 2011 AFA Annual Meeting, and the 2012 Computational Management Science Meeting.

Submitted 23 April 2013

1 Introduction

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. Specifically, they find that assets with high (low) returns over the last twelve months tend to have high (low) returns for the next six months, and that "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. Specifically, they estimate the cross-correlation matrices for the vector of returns on the five size-sorted quintiles of a sample of stocks from CRSP, and find that "current returns of smaller stocks are correlated with past returns of larger stocks, but not vice versa, a distinct lead-lag relation based on size." Moreover, they show that a contrarian portfolio that takes advantage of this lead-lag pattern in stock returns 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 that has been documented in the literature to select portfolios of risky assets that perform well *out-of-sample*. 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 that of other arbitrage portfolios from the literature. Third, we evaluate empirically the out-of-sample gains from using investment (positive-cost) portfolios that exploit serial dependence in stock returns.

[1] There is substantial empirical evidence of serial and cross-serial correlation in returns. For example, there is a large body of research that documents momentum at the level of individual firms in the U.S. (Jegadeesh (1990), Lehman (1990), Jegadeesh and Titman (1993)), at the level of industries (Moskowitz and Grinblatt (1999)), in size, book-to-market, and double-sorted size and book-to-market portfolios (Lewellen (2002)), and internationally (Rouwenhorst (1998)). There is also substantial evidence on serial dependence in stock returns (Fama and French (1988), Conrad and Kaul (1988, 1989, 1998), Lehman (1990), Boudoukh, Richardson, and Whitelaw (1994), Daniel, Hirshleifer, and Subrahmanyam (1998), and Ahn, Boudoukh, Richardson, and Whitelaw (2002)), and reversal/overreaction (DeBondt and Thaler (1985)). Finally, a number of papers have documented the presence of cross-correlations, 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, Jegadeesh, and Swaminathan (1993)), and institutional ownership (Badrinath, Kale, and Noe (1995)).

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. Merton (1980) explains the theoretical reason why it is much more difficult to get precise estimates of first moments than second moments. And, Jagannathan and Ma (2003) confirm this in practice: they find that the minimum-variance portfolio (which ignores estimates of expected returns) outperforms portfolios that rely on forecasts of expected returns, even when performance is measured using the Sharpe ratio, which depends on both the portfolio mean and variance. DeMiguel, Garlappi, and Uppal (2009) also find that portfolios that use estimated expected returns perform poorly out of sample, achieving substantially lower Sharpe ratios and higher turnovers compared to portfolios that ignore estimates of expected returns. Consequently, a large part of the literature on portfolio selection has focussed on improving the estimation of the covariance matrix: see, for example, Chan, Karceski, and Lakonishok (1999), Ledoit and Wolf (2003, 2004), DeMiguel and Nogales (2009), and DeMiguel, Garlappi, Nogales, and Uppal (2009). The focus on conditional expected returns for individual stocks distinguishes our work from these papers.

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 returns. Our VAR model allows tomorrow's expected return on every stock to depend linearly on today's realized return on *every* stock, and hence it is general enough to capture any linear relation between stock returns in consecutive periods, irrespective of whether its origin is momentum, lead-lag relations, or some other feature of the data.[2] We verify

[2]A broad variety of explanations have been offered for autocorrelations and cross-correlations 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, Green, and Naik (1999) and Johnson (2002)). Other explanations rely on economic links, such as those between suppliers and customers (Cohen and Frazzini (2008)), and upstream and downstream industries (Menzly and Ozbas (2010)). Then, there are explanations that are based on imperfections in markets, such as: segmentation of the market for securities (Merton (1987)); transaction costs (Mech (1993)); asymmetric information (Brennan, Jegadeesh, and Swaminathan (1993)); ambiguous (uncertain) information (Zhang (2006)); the slow transmission of information across investors (Hong and Stein (1999), Hong, Lim, and Stein (2000)); investor inattention that leads to a delay with which prices incorporate information (Ramnath (2002), Hirshleifer and Teoh (2003), Hou and Moskowitz (2005)), Peng and Xiong (2006), Hong, Torous, and Valkanov (2007), and Hou, Peng, and Xiong (2009)), and heterogeneity across investors in responding to information (Hong and Stein (1999) and DellaVigna and Pollet (2007)). Finally, there are behavioral explanations based on noise traders (De Long, Shleifer, Summers, and Waldmann (1990a)); herd behavior (Scharfstein and Stein (1990)); and the characteristics of investors such as overreaction (De Long, Shleifer, Summers, and Waldmann (1990b)), overconfidence (Daniel, Hirshleifer, and Subrahmanyam (1998)), and conservatism (Barberis, Shleifer, and Vishny (1998)).

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 use our significance tests to 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 zero-cost arbitrage portfolios based on the VAR model and to compare them to other arbitrage strategies. Analytically, we compare the expected return of the VAR arbitrage portfolio to that of the contrarian arbitrage portfolio studied by Lo and MacKinlay (1990), who show that the expected return on the contrarian arbitrage portfolio is positive if the stock return autocorrelations are negative and the stock return cross-correlations are positive. We show 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 stock returns; that is, if today's return on the principal components allow one to tell which stocks will achieve high returns, and which low returns tomorrow. Empirically, we show that the VAR arbitrage portfolio substantially outperforms (out of sample) the contrarian arbitrage portfolio and an arbitrage portfolio based on the unconditional sample mean.

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

[3]We have also considered the *dynamic* portfolio of Campbell, Chan, and Viceira (2003), which is the optimal portfolio of an intertemporally optimizing investor with Epstein-Zin utility, who believes that the

high turnover associated with the conditional mean-variance portfolios, we focus on norm-constrained portfolios similar to those studied by DeMiguel, Garlappi, Nogales, and Uppal (2009).

Our empirical results show that, for the majority of the datasets we consider, the norm-constrained conditional mean-variance portfolios outperform the traditional (unconditional) portfolios out of sample for transaction costs below 10 basis points. Moreover, for a dataset containing high-turnover individual stocks, we find that the conditional portfolios from the VAR model substantially and significantly outperform the traditional portfolios even in the presence of transaction costs of 10 basis points. To understand the origin of the predictability exploited by the conditional portfolios, we consider the conditional mean-variance portfolios obtained from a lagged-factor model using as factors the Fama-French factors (market, small minus big, and high minus low book-to-market), and we find that the market and high-minus-low factors drive most of the predictability exploited by the conditional portfolios. Moreover, we also observe that the gains from exploiting stock return serial dependence come in the form of higher expected return, since the out-of-sample variances 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 from exploiting time serial dependence in stock returns is obtained by exploiting cross-covariances in stock returns, as opposed to just autocovariances.

The rest of this manuscript is organized as follows. Section 2 describes the datasets and the methodology we use for our empirical analysis. Section 3 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 4 characterizes (analytically and empirically) the performance of a VAR zero-cost arbitrage portfolio, and compares it to that of other arbitrage portfolios. Section 5 describes the different investment portfolios we consider, and discusses their empirical performance. Section 6 concludes. Robustness checks for our empirical findings and proofs for all propositions are relegated to the appendix.

---

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.

2 Data and Evaluation Methodology

2.1 Datasets

We consider five datasets for our empirical analysis: four datasets from Ken French's website, and one from CRSP, and for every 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.

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]

2.2 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

[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.1 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.

DeMiguel, Garlappi, and Uppal (2009): (i) portfolio mean return; (ii) portfolio variance; (iii) Sharpe ratio, defined as the sample mean of out-of-sample returns divided by their sample standard deviation;[5] 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

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

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.[6]

3 A Vector Autoregressive (VAR) Model of Stock Returns

We now introduce the VAR model. In Section 3.1, we describe the VAR model of stock returns, and in Section 3.2, we test the statistical significance of the VAR model for the five datasets described in Section 2.1. Finally, in Section 3.3, we use statistical tests to understand the nature of the relation between stock returns.

3.1 The VAR model

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

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

[5]Note that because we are considering investments in only risky assets, the numerator of the Sharpe ratio is the expected return, instead of the expected return in excess of the risk-free rate. 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.

[6]In Section A.3 we also consider the case where the investor's optimization problem incorporates transaction costs explicitly.

where $r_t \in \mathbb{R}^N$ is the stock return vector for period t, $a \in \mathbb{R}^N$ is the vector of intercepts, $B \in \mathbb{R}^{N \times N}$ is the matrix of slopes, and ϵ_{t+1} is the error vector, which is independently 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.[7]

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 for 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 momentum, lead-lag relations, or any other 1-lag time-series feature of the data.

VAR models have been used before for strategic asset allocation—see Campbell and Viceira (1999, 2002); Campbell, Chan, and Viceira (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.[8] 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 stocks with better out-of-sample performance, and thus, we use the VAR model to capture the ability of today's stock returns to predict tomorrow's stock returns.

VAR models have also been used before to model serial dependence among individual stocks or international indexes. For instance, Tsay (2005, Chapter 8) estimates a vector autoregressive 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

[7]To conserve space, we report only the results for the first-order vector autoregressive model, VAR(1), which is given in equation (1), but we have also estimated a general pth-order vector autoregressive model, VAR(p) using Schwarz's Bayesian criterion (Schwarz (1978)) to choose the order, and we have found the order $p = 1$ to be optimal.

[8]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 5.4. 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.

Chordia and Swaminathan (2000) estimate a vector autoregressive model for two portfolios, one composed of high-trading-volume stocks and the other of low-trading-volume stocks. However, to the best of our knowledge, our paper is the first to investigate whether a VAR model at the individual stock level can be used to choose portfolios with better out-of-sample performance.

3.2 Significance of the VAR model

Estimating the VAR model in (1) requires estimating a large number of parameters,[9] and thus standard ordinary least squares (OLS) estimators of the VAR model are noisy.[10] To obtain stable estimators, we use *ridge regression*, see Hoerl and Kennard (1970), which is designed to give stable 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).

In this section, we assume that r_t is a jointly covariance-stationary process with finite mean $\mu = E[r_t]$ and finite cross-covariance matrices $\Gamma_k = E[(r_{t-k} - \mu)(r_t - \mu)^\top]$ for $k = 0, 1$. We also assume that the covariance matrix Γ_0 is positive definite.

To test whether the VAR model is statistically significant, we propose a bootstrap test when the model is estimated by ridge regression. In particular, for the VAR model (1), to test the null hypothesis

$$H_0 : \ B = 0, \tag{2}$$

we first estimate equation (1) using ridge regression with an estimation window of $\tau = 2000$ days. Then, we propose the following test statistic

$$M = -(\tau - N) \ln \left(\frac{|\hat{\Omega}_1|}{|\hat{\Gamma}_0|} \right),$$

where $\hat{\Omega}_1$ is the covariance matrix of the residuals $\hat{\epsilon}$ obtained after fitting the VAR equation (1) to the data.

[9]One can show that the number of parameters to be estimated is $3(N^2 + N)/2$.

[10]For instance, the OLS estimator of the matrix of slopes is $\hat{B} = \Gamma_0^{-1}\Gamma_1$, where Γ_0 is the covariance matrix and Γ_1 is the lag-one cross-covariance matrix, and it is well-known that the covariance matrix of asset returns Γ_0 is ill conditioned, which implies that the OLS estimator of the slope matrix is likely to be very noisy. Indeed, our empirical results have demonstrated that the OLS estimator of the slope matrix is unstable.

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 (2).

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 in (2) 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.

3.3 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 exposition purposes, we first study two small datasets with only two assets each, and we then provide summary information for the full datasets.

3.3.1 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 a particular 2,000-day estimation window 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. Also, both small and large-stock portfolio returns have significant first-order autocorrelations.

To check whether the ridge estimator of the VAR model is stable, we perform the above test every trading day in our sample. 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 main observation is that the estimators of the slope matrix elements are reasonably stable, both in terms of magnitude and statistical significance. The stability of the estimated slope matrix shows that the ridge estimator of the VAR model deals well with estimation error. Note also that the estimators are time varying, which is to be expected as market conditions change during such a long period (1978–2011). The key is that the VAR model estimated with ridge regression is able to capture the current serial dependence of stock returns in a stable manner.

3.3.2 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 a particular 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 stock and value stock, with growth-stock returns leading value-stock returns. Also, both growth- and value-stock portfolio returns have significant first-order autocorrelations.

To check whether the ridge estimator of the VAR model is stable, we perform the previous test every trading day in our sample. Figure 2 shows the time evolution of the estimated diagonal and off-diagonal elements of the slope matrix. 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 the estimators of the slope matrix elements are stable, although they reflect the time-varying market conditions.

3.3.3 Results for the full datasets

We now summarize our findings for the five datasets described in Section 2.1. 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 estimators of the slope matrix elements are reasonably stable, although they reflect the time-varying market conditions.

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 on 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 have obtained similar insights from the tests on the 25FF but to conserve space we do not report the results in the manuscript.

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 five industries: 1 = Cnsmr (Consumer Durables, NonDurables, 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}), and that Hlth returns lead Cnsmr returns (B_{14}). The conclusions are similar for the 10Ind and 48Ind datasets, but we do not report the results in the manuscript to conserve space.

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 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.

4 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 considered in the literature.

4.1 Analytical comparison

In this section, we compare *analytically* the expected return of the VAR arbitrage portfolio to that of the contrarian arbitrage portfolio studied by Lo and MacKinlay (1990).[11]

4.1.1 The contrarian arbitrage portfolio

To study whether contrarian profits are due exclusively to market overreaction, Lo and MacKinlay (1990) consider the following contrarian ("c") arbitrage portfolio:

$$\mathrm{w}_{c,t+1} = -\frac{1}{N}(r_t - r_{et}e), \tag{3}$$

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:

$$E[\mathrm{w}_{ct}^\top r_t] = C + O - \sigma^2(\mu), \tag{4}$$

[11] Note that Lo and MacKinlay (1990) did not propose the contrarian strategy as a practical investment strategy for choosing portfolios of stocks, but rather to show that contrarian profits are not necessarily due to stock market overreaction. We, however, find the comparison between the VAR and contrarian arbitrage portfolios helpful in the context of testing the potential of the VAR model for portfolio selection.

where

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

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

$$\sigma^2(\mu) = \frac{1}{N}\sum_{i=1}^{N}(\mu_i - \mu_m)^2, \tag{5}$$

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. Note that C is a positive multiple of the sum of the cross-covariances of stock returns, O is a negative multiple of the sum of the autocovariances, and $\sigma^2(\mu)$ is the cross-sectional variance of expected stock returns. Therefore, equation (4) shows that the contrarian arbitrage portfolio has a positive expected return if the cross-covariances are positive, the autocovariances are negative, and their combined effect on the expected return, measured through the sum $C + O$, is larger than the cross-sectional variance of expected stock returns; that is, if $C + O > \sigma^2(\mu)$.

4.1.2 The VAR arbitrage portfolio

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

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

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 stocks whose VAR-based conditional expected return is above that of the equally-weighted portfolio, and a negative weight to the rest of the stocks.

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, and hence we set $B = \Gamma_1^\top \Gamma_0^{-1}$ and $a = (I - B)\mu$, which are the VAR parameters that result in a stock return process with an expected return equal to

μ, a covariance matrix equal to Γ_0, and a lag-1 cross-covariance matrix equal to Γ_1. Note that we do *not* make this assumption in our empirical analysis in Section 4.2, and instead estimate the VAR model from empirical data.

Proposition 1 *Assume that r_t is a jointly covariance-stationary process with mean $\mu = E[r_t]$ and cross-covariance matrices $\Gamma_k = E[(r_{t-k} - \mu)(r_t - \mu)^\top]$ for $k = 0, 1$. 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 the expected return of the VAR arbitrage portfolio is*

$$E[w_{vt}^\top r_t] = G + \sigma^2(\mu) \geq 0, \tag{6}$$

where

$$G = \frac{\mathrm{tr}(\Gamma_1^\top \Gamma_0^{-1} \Gamma_1)}{N} - \frac{e^\top \Gamma_1^\top \Gamma_0^{-1} \Gamma_1 e}{N^2} \geq 0. \tag{7}$$

Proposition 1 shows that the expected return of the VAR arbitrage portfolio is the sum of two terms, $G + \sigma^2(\mu)$. From (7) we see that G depends only on the covariance matrix Γ_0 and the lag-one cross-covariance matrix Γ_1, while $\sigma^2(\mu)$ depends exclusively on the stock mean returns. Moreover, the proposition shows that each of these two terms makes a *nonnegative* contribution to the expected return of the VAR arbitrage portfolio. Furthermore, Proposition 1 also shows that the expected return of the VAR arbitrage portfolio is *strictly* positive in general because $\sigma^2(\mu) > 0$ except for the degenerate case where all assets have the same expected return.

4.1.3 Comparing the contrarian and VAR arbitrage portfolios

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 arbitrage portfolio. 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 positive and the cross-correlations negative.

Note also that the cross-sectional variance of mean stock returns enters the expression for the contrarian portfolio expected return as a negative term, but it enters the expression for the VAR portfolio's expected return as a positive term. 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 tends to assign a negative weight to assets with a mean return that is above average. This results in the negative contribution of the cross-variance of mean stock returns to the expected return of the contrarian arbitrage portfolio.

4.1.4 Identifying the origin of predictability using principal components

We now use principal component analysis to identify the origin of the predictability in stock 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 stocks will perform well and which poorly in the next period.

To see this, note first 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 we can rewrite the VAR model in Equation (1) as

$$r_{t+1} = a + BQQ^\top r_t + \epsilon_{t+1}$$

$$r_{t+1} = a + \hat{B}p_t + \epsilon_{t+1},$$

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

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

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

Proposition 2 shows that the VAR arbitrage portfolio attains 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 stocks we should go long and stocks we should short tomorrow. 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 2 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 highest contribution 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 datasets we find that the principal component with highest contribution is long on the HiTec industry portfolio and short on the other industries.

4.1.5 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:

$$r_{t+1} = a^f + B^f f_t + \epsilon_{t+1}^f \tag{8}$$

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_{t+1}^f 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 2 in the context of the easier-to-interpret lagged-factor model.

Proposition 3 *Assume that r_t is the jointly covariance-stationary process described in* (8)*, and the factor covariance matrix $\Gamma_0^f = E((f_t - \mu_f)^\top (f_t - \mu_f))$ is positive definite. Moreover, assume we can estimate the lagged-factor model exactly, then the expected return of the lagged-factor arbitrage portfolio is*

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

where λ_j^f is the variance of the jth principal component of the factor covariance matrix Γ_0^f, $var(\hat{B}_{\bullet j}^f)$ 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 3 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.

4.2 Empirical comparison

In this section, we compare *empirically* the performance of the VAR arbitrage portfolio to those of the contrarian arbitrage portfolio and an arbitrage portfolio based on sample mean returns. We first compare the *in-sample* expected return of the contrarian and VAR arbitrage portfolios by using the analytical expressions in Equation (4) and Proposition 1. We then compare the *out-of-sample* expected return and Sharpe ratio of the different arbitrage portfolios, using the rolling horizon methodology described in Section 2.2.

4.2.1 In-sample comparison of performance

The first panel of Table 1 gives the in-sample values of C, O, $\sigma^2(\mu)$, G, as well as the in-sample expected returns of the contrarian and VAR arbitrage portfolios, which are calculated using equations (4)–(5) and (6)–(7), for the five datasets considered.[12] The results show that the contrarian portfolio achieves a positive in-sample expected return only for the 100CRSP dataset. This is not surprising because the contrarian strategy makes sense in the context of individual stocks, as is the case for the CRSP datasets. The rest of the datasets we consider consist of assets that are portfolios of stocks, and it is well known—see Campbell, Lo, and MacKinlay (1997)—that portfolio returns have positive autocorrelation, which implies that O is negative, and hence the contrarian strategy has a negative expected return. Finally, note from the second panel of Table 1 that the in-sample expected return of the VAR arbitrage portfolio is positive for *all* datasets, and it is larger than that of the contrarian portfolio for the 100CRSP dataset.

4.2.2 Out-of-sample comparison of performance

We now compare the out-of-sample expected return and Sharpe ratio of the VAR arbitrage portfolio to that of two other arbitrage portfolios: (i) the contrarian arbitrage portfolio given in (3); and, (ii) an arbitrage portfolio based on the *unconditional* sample mean return, which

[12] 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.

we compute as:

$$\mathrm{w}_{s,t+1} = \frac{1}{N}\left(\hat{\mu} - \frac{\hat{\mu}^\top e}{N} e\right),$$

where $\hat{\mu}$ is the sample mean return vector, and $\hat{\mu}^\top e/N$ is the equally-weighted portfolio sample mean return; that is, this portfolio assigns a positive weight to stocks that have a larger sample mean return than the equally-weighted portfolio, and a negative weight to the rest.

The third and fourth panels in Table 1 give the out-of-sample expected returns and Sharpe ratios, respectively, of the contrarian, VAR, and unconditional arbitrage portfolios computed using the rolling-horizon methodology described in Section 2.2. We first compare the VAR and contrarian arbitrage portfolios. Note that similar to the in-sample results, the contrarian arbitrage portfolio attains a negative out-of-sample expected return for all datasets except the 100CRSP dataset. On the other hand, the VAR arbitrage portfolios attains positive out-of-sample expected returns for all datasets, which are also substantially larger than the expected returns of the contrarian arbitrage portfolio in absolute value.[13]

The relative performance of the arbitrage portfolios in terms of Sharpe ratios is similar to that in terms of expected returns. The VAR arbitrage portfolio attains positive Sharpe ratios for all datasets, while the contrarian arbitrage portfolio attains a negative Sharpe ratio for all datasets except the 100CRSP dataset, where its Sharpe ratio is still substantially lower than that of the VAR arbitrage portfolio. As with the in-sample results in the previous subsection, the reason for the negative value of the out-of-sample expected return and Sharpe ratio of the arbitrage contrarian portfolio is that the assets in the Fama-and-French and industry datasets are portfolios of stocks, which tend to be positively autocorrelated, and it is intuitively clear that contrarian portfolios will, in general, have negative returns when applied to datasets with positively autocorrelated assets. We also observe that the out-of-sample expected return and Sharpe ratio of the VAR arbitrage portfolio are much larger than those of the arbitrage portfolio based on the unconditional sample mean.

Observe that the VAR arbitrage portfolio attains surprisingly high out-of-sample Sharpe ratios (ranging from 3.32 for the 100CRSP dataset to 4.90 for the 25FF dataset). We must

[13]Therefore the VAR arbitrage portfolio outperforms also the momentum arbitrage portfolio obtained by reversing the sign of the contrarian portfolio weights.

note, however, that these high Sharpe ratios are associated with very high trading volumes, and hence it is not clear whether the VAR arbitrage portfolios can be implemented in the presence of transaction costs, and especially the costs entailed when shorting assets. We study this issue in the next section, where we evaluate the performance of the conditional mean-variance portfolios based on the VAR model with shortsales prohibited, both without and with transaction costs.

5 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 on the five datasets listed in Section 2.1. Section 5.1 discusses portfolios that ignore stock return serial dependence and Section 5.2 describes portfolios that exploit stock return serial dependence. Then, in Section 5.3 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 5.4 we use a lagged-factor model to trace the origin of the predictability in stock returns exploited by the conditional portfolios.

5.1 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.

5.1.1 The $1/N$ portfolio

The $1/N$ portfolio studied by DeMiguel, Garlappi, and Uppal (2009) 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.

5.1.2 The shortsale-constrained minimum-variance portfolio

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

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \Sigma \mathrm{w}, \tag{9}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1, \tag{10}$$

$$\mathrm{w} \geq 0, \tag{11}$$

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, and 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.[14] For our empirical evaluation, we use the shortsale-constrained minimum-variance portfolio computed by solving problem (9)–(11) after replacing the covariance matrix by the shrinkage estimator proposed by Ledoit and Wolf (2003).[15]

5.1.3 The norm-constrained mean-variance portfolio

The mean-variance portfolio is the solution to:

$$\min_{\mathrm{w}} \quad \mathrm{w}^\top \Sigma \mathrm{w} - \frac{1}{\gamma} \mathrm{w}^\top \mu, \tag{12}$$

$$\text{s.t.} \quad \mathrm{w}^\top e = 1, \tag{13}$$

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 (see DeMiguel, Garlappi, and Uppal (2009)), 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, Garlappi, Nogales, and Uppal (2009) for an analysis of norm constraints in the context of portfolio selection.[16] Specifically, we compute

[14] 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.

[15] 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.

[16] 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.

the norm-constrained mean-variance portfolios by solving problem (12)–(13) 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 (12)–(13) 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 $\gamma = 2$.

5.1.4 Empirical performance

The top panel in 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 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.

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.[17]

5.2 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.

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.[18]

5.2.1 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

[17]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.

[18]We also considered imposing a shortsale-constraint on the conditional mean-variance portfolios, but we find that the daily turnover of the resulting portfolios is still too large to give meaningful results, and thus we report the performance of only the norm-constrained conditional mean-variance portfolios.

cares only about the returns tomorrow) who believes stock returns follow a linear VAR model. They are computed by solving problem (12)–(13) 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 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 focus on the case with 1-norm-constraints on the difference with the weights of the shortsale-constrained minimum-variance portfolio. As for 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 from the results are robust to the value of the risk aversion parameter, and thus, we report the results only for the case of $\gamma = 2$.

5.2.2 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.[19] 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

[19] See Gyorfi, Kohler, Krzyzak, and Walk (1987) for an in-depth discussion of nonparametric regression and Gyorfi, Udina, and Walk (2008, 2007) for an application to portfolio selection, and Mizrach (1992) for an application to exchange rate forecasting.

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 (12)–(13) 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 $\gamma = 2$.

Sections 5.2.3 and 5.2.4 discuss the performance of the portfolios described above in the absence and presence of proportional transaction costs, respectively.

5.2.3 Empirical performance

The last panel 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 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, Tibshirani, Friedman, and Franklin (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, in general, 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 in the tables to conserve space) that the gains from using the norm-constrained portfolios come in the form of higher expected return, since 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.

5.2.4 Empirical performance in the presence of transaction cost

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 a transaction costs of 5 and 10 basis points, respectively.

From Table 4 we observe that, in the presence of a proportional transactions cost of 5 basis points, the norm-constrained conditional portfolios from the VAR model substantially 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 similar 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 generally outperform the shortsale-constrained minimum-variance portfolio for transaction costs below 10 basis points.

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. From the above results it is clear that to take advantage of the VAR-based strategies, efficient execution of trades will be important.

5.3 Exploiting autocovariances versus cross-covariances

In this section, we investigate what proportion of the gains from exploiting time serial dependence in stock returns is obtained by exploiting autocovariances in stock returns, 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 5.2.1, 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 can be different from zero.[20]

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, most 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. This is not surprising, because we already found in Section 4 that there exist statistically significant lead-lag relations between the assets in our datasets.

5.4 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 8. 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

[20]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.

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 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 individual stock returns tomorrow. Note that this is very different from the type of predictability exploited in the literature before, where typically today's dividend yield and today's short-term versus long-term yield spread 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 individual stocks will have high returns and which individual stocks will have low returns tomorrow.

6 Conclusion

In this paper, we have investigated whether investors can use a vector autoregressive (VAR) model to exploit the autocorrelation and cross-correlation documented in the literature to improve the out-of-sample performance of static and dynamic portfolios. Our VAR model allows tomorrow's expected return on every stock to depend linearly on today's realized return on *every* stock, and hence it is general enough to capture any linear relation between stock returns in consecutive periods, irrespective of whether its origin is momentum, lead-lag relations, or some other feature of the data. We also consider a nonparametric autoregressive (NAR) model, which does not require that the relation across stock returns be linear.

We find that the VAR model is statistically significant for all five datasets that we consider, which include four datasets from Ken French's website (consisting of daily returns on 6 and 25 value-weighted portfolios of stocks sorted on size and book-to-market, and the 10 and 48 industry value-weighted portfolios) and a dataset containing individual stock returns from the CRSP database. For all these datasets, we consider two versions: one that has close-to-close returns and a second that has open-to-close returns, with the results for the latter reported in the robustness section.

Next, we characterize, both analytically and empirically, the expected return of an *arbitrage* (zero-cost) portfolio based on the VAR model, and show that it compares favorably to that of other arbitrage portfolios in the literature, such as the contrarian portfolio considered in Lo and MacKinlay (1990) and Khandani and Lo (2010). In contrast to the contrarian arbitrage portfolio, whose expected return is positive if the stock return autocorrelations are negative and the stock return cross-correlations are positive, the VAR arbitrage portfolio achieves a positive expected return in general, regardless of the sign of the autocorrelations and cross-correlations. Empirically, we show that the VAR arbitrage portfolio outperforms (out of sample) the contrarian arbitrage portfolio and an arbitrage portfolio based on the unconditional sample mean.

Finally, we evaluate the performance of two investment (positive-cost) portfolios: a *conditional* mean-variance myopic portfolio based on the linear VAR model, and a conditional mean-variance portfolio using a *nonparametric* autoregressive (NAR) model. We find that, subject to a norm constraint on the portfolio weights, these conditional investment portfolios outperform the traditional (unconditional) mean-variance portfolio and the shortsale-constrained minimum-variance portfolio for transaction costs below 10 basis points. We show that a substantial part of the gains from using the VAR model arise from exploiting cross-covariances in stock returns.

In order 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 the conditional portfolios based on a lagged-factor model, where the factors are 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. We find that most of the predictability in all datasets comes from the MKT and HML factors. Note that this is very different from the type of predictability exploited in the literature before, where typically today's dividend yield and today's short-term versus long-term yield spread have been used to predict tomorrow's return on a single risky index. We also find that, while the conditional portfolios from the four-factor model typically outperform the benchmark shortsale-constrained minimum-variance portfolio, they do not perform as well as the portfolios based on the full VAR model. 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.

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 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 in agreement with the observation by Lo and MacKinlay (1990, p. 197) and Anderson, Eom, Hahn, and Park (2005) that the lead-lag relations in stock returns they document cannot be completely attributed to asynchronous or infrequent trading.

A.1.1 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 A1, A3, and A4 in the appendix, for transaction costs of 0, 5, and 10 basis points, respectively, with the turnover reported in Table A2. 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.1.2 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.[21] The results, not reported in the manuscript to conserve space, show that the conditional portfolios outperform the benchmark substantially and significantly for

[21] 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.

transaction costs of 5 basis points, and their performance is similar to that of the benchmark for transaction costs of 10 basis points.

A.1.3 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 A5 and A6 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 on all datasets.[22] 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 A6 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, but the differences are not substantial; that is, the insight that the conditional portfolios outperform the benchmark for transaction costs of 5 basis points is generally robust to the use of weekly return data. This is a bit surprising because as one decreases the amount of trading, one 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 we discussed previously, 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.

Summarizing, we find that the overall relative performance in the presence of transactions costs is roughly the same independent of the frequency.

A.2 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.

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

Table A7 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.3 In-sample optimal portfolios with proportional transactions costs

In this manuscript, 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 transactions 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 Problem (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 Problem (A1)–(A2) is much worse in the presence of transaction costs than that of the *norm-constrained* conditional mean-variance portfolios studied in Section 5.2.[23]

The explanation for this is that the portfolios computed by solving Problem (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, Garlappi, and Uppal (2009). 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 Problem (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 Problem (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 Problem (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 of Problem (A1)–(A2) 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 Problem (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.

[23] To conserve space we have not included the tables with these results, but they are available upon request.

B Proofs for all the Propositions

Proof for Proposition 1

We first show that $E[\mathrm{w}_{vt}^\top r_t] = G + \sigma^2(\mu)$. To see this, note that $E[\mathrm{w}_{vt}^\top r_t] = c_1 - c_2$, where $c_1 = \frac{1}{N}E[(a + Br_{t-1})^\top r_t]$ and $c_2 = \frac{1}{N}E[(r_{v,t-1}e)^\top r_t]$.

We start by computing c_1:

$$\begin{aligned}
Nc_1 &= E((a + Br_{t-1})^\top r_t) \\
&= E(a^\top r_t) + E(r_{t-1}^\top B^\top r_t) \\
&= a^\top \mu + E(\mathrm{tr}(r_{t-1}^\top B^\top r_t)) \\
&= \mu^\top (1 - B^\top)\mu + E(\mathrm{tr}(B^\top r_t r_{t-1}^\top)) \\
&= \mu^\top \mu - \mu^\top B^\top \mu + \mathrm{tr}(B^\top \; E(r_t r_{t-1}^\top)) \\
&= \mu^\top \mu - \mu^\top B^\top \mu + \mathrm{tr}(B^\top \; (\Gamma_1^\top + \mu\mu^\top)) \\
&= \mu^\top \mu - \mu^\top B^\top \mu + \mathrm{tr}(B^\top \; \Gamma_1^\top) + \mu^\top B^\top \mu \\
&= \mu^\top \mu + \mathrm{tr}(\Gamma_1^\top \Gamma_0^{-1} \Gamma_1).
\end{aligned}$$

We now compute c_2:

$$\begin{aligned}
c_2 &= \frac{1}{N}E(r_{v,t-1} \; e^\top r_t) \\
&= \frac{1}{N^2}E(e^\top a^\top e r_t + e^\top \; (r_{t-1}^\top B^\top e) \; r_t) \\
&= \frac{1}{N^2}e^\top \; (a^\top) \; e\mu + \frac{1}{N^2}(\mathrm{tr}(B^\top e e^\top E(r_t r_{t-1}^\top))) \\
&= \frac{1}{N^2}((\mu^\top e)^2 - e^\top(\mu^\top B^\top e)\mu) + \frac{1}{N^2}(e^\top \Gamma_1^\top \Gamma_0^{-1}\Gamma_1 e + e^\top(\mu^\top B^\top e)\mu).
\end{aligned}$$

Hence

$$\begin{aligned}
c_1 - c_2 &= \frac{1}{N}(\mathrm{tr}(\Gamma_1^\top \Gamma_0^{-1}\Gamma_1) - \frac{1}{N}e^\top \Gamma_1^\top \Gamma_0^{-1}\Gamma_1 e) + \frac{1}{N}(\mu^\top \mu - \frac{1}{N}(\mu^\top e)^2) \\
&= \frac{1}{N}(\mathrm{tr}(\Gamma_1^\top \Gamma_0^{-1}\Gamma_1) - \frac{1}{N}e^\top \Gamma_1^\top \Gamma_0^{-1}\Gamma_1 e) + \frac{1}{N}\sum_{i=1}^{N}(\mu_i - \mu_m)^2, \\
&= G + \frac{1}{N}\sum_{i=1}^{N}(\mu_i - \mu_m)^2 = G + \sigma^2(\mu),
\end{aligned}$$

which proves that equation (6) holds.

We now show that the two terms on the right-hand side of (6) are nonnegative. The term $\sigma^2(\mu)$ is obviously greater than or equal to zero, so it only remains to show that G is nonnegative. To see this, note that

$$\begin{aligned}
G &= \frac{1}{N}\big(\mathrm{tr}(\Gamma_1^\top \Gamma_0^{-1}\Gamma_1) - \frac{1}{N}e^\top \Gamma_1^\top \Gamma_0^{-1}\Gamma_1 e\big) \\
&= \frac{1}{N}(\|\Gamma_0^{-1/2}\Gamma_1\|_F^2 - \frac{1}{N}\|\Gamma_0^{-1/2}\Gamma_1 e\|_2^2),
\end{aligned} \tag{B1}$$

where $\Gamma_0^{-1/2}$ is the Cholesky factor of the positive definite matrix Γ_0, and $\|A\|_F = \sum_{i=1}^N \sum_{j=1}^N a_{ij}^2$ is the Frobenius norm of matrix A. Also, note that

$$\|\Gamma_0^{-1/2}\Gamma_1 e\|_2 \leq \|\Gamma_0^{-1/2}\Gamma_1\|_F \|e\|_2 = \sqrt{N}\|\Gamma_0^{-1/2}\Gamma_1\|_F, \tag{B2}$$

where the inequality follows from basic matrix norm properties; see Golub and Loan (1996, Ch. 2). Equation (B1) together with inequality (B2) imply that $G \geq 0$.

Proof for Proposition 2

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

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

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

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

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

Proof for Proposition 3

The proof is very similar to those of Propositions 1 and 2.

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

This table reports the in- and out-of-sample characteristics of the contrarian, VAR, and unconditional arbitrage portfolios for the five datasets considered. The first and second panels give the in-sample values of C, O, $\sigma^2(\mu)$, G, as well as the in-sample expected returns of the contrarian and VAR arbitrage portfolios, which are calculated using equations (4)–(5) and (6)–(7). The third and fourth panels give the out-of-sample expected return and Sharpe ratios, respectively, of the contrarian, VAR, and unconditional arbitrage portfolios computed using the rolling-horizon methodology described in Section 2.2.

Quantity/strategy	6FF	25FF	10Ind	48Ind	100CRSP
In-sample values of C, O, $\sigma^2(\mu)$, G					
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
G	0.0704	0.1685	0.0842	0.1679	0.8178
In-sample expected returns					
Contrarian	-0.2250	-0.2376	-0.4207	-0.2640	0.3124
Conditional VAR	0.0721	0.1705	0.0845	0.1686	0.8206
Out-of-sample expected returns					
Contrarian	-0.1932	-0.1865	-0.4002	-0.2796	0.2289
Conditional VAR	0.3080	0.3983	0.4308	0.4514	0.5298
Unconditional	0.0383	0.0446	-0.0099	0.0255	-0.0598
Out-of-sample Sharpe ratios					
Contrarian	-2.0804	-2.2784	-3.0846	-2.1028	0.9719
Conditional VAR	3.6806	4.9082	3.8043	4.0190	3.3238
Unconditional	0.4830	0.6554	-0.0758	0.2055	-0.2675

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
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)
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
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
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
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)
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
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)
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 tct = 0bps

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
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)
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)

Table A1: 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 A2: 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 A3: 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 A4: 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 A5: 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 A6: 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 A7: Sharpe ratios for dataset with returns on the 100 stocks with highest turnover in the S&P500, for different levels of transaction costs

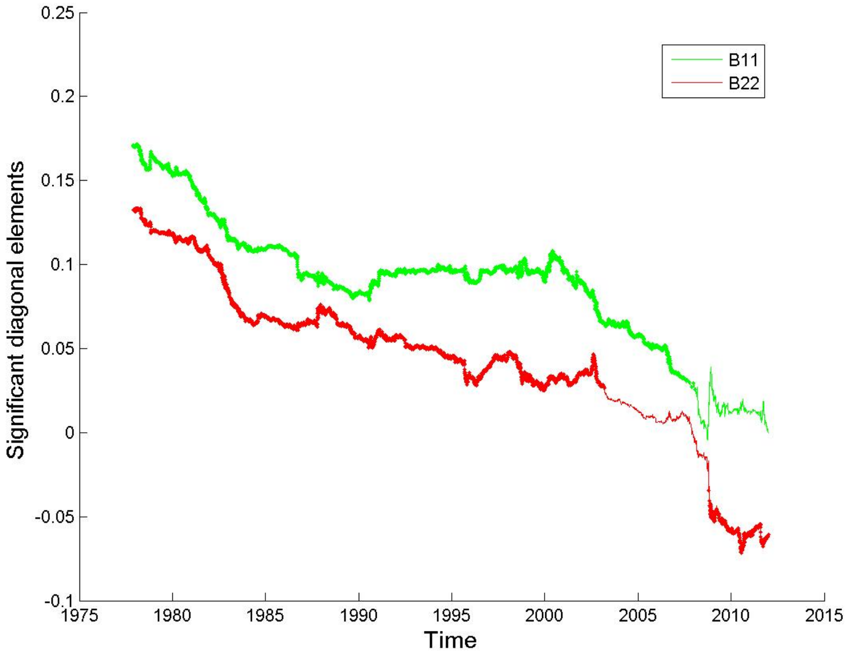
This table reports the annualized out-of-sample Sharpe ratios for the different portfolios and for the dataset with returns on the 100 stocks with highest turnover in the S&P500, for different levels of transaction costs, together with the P-value that the Sharpe ratio for a strategy is different from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

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

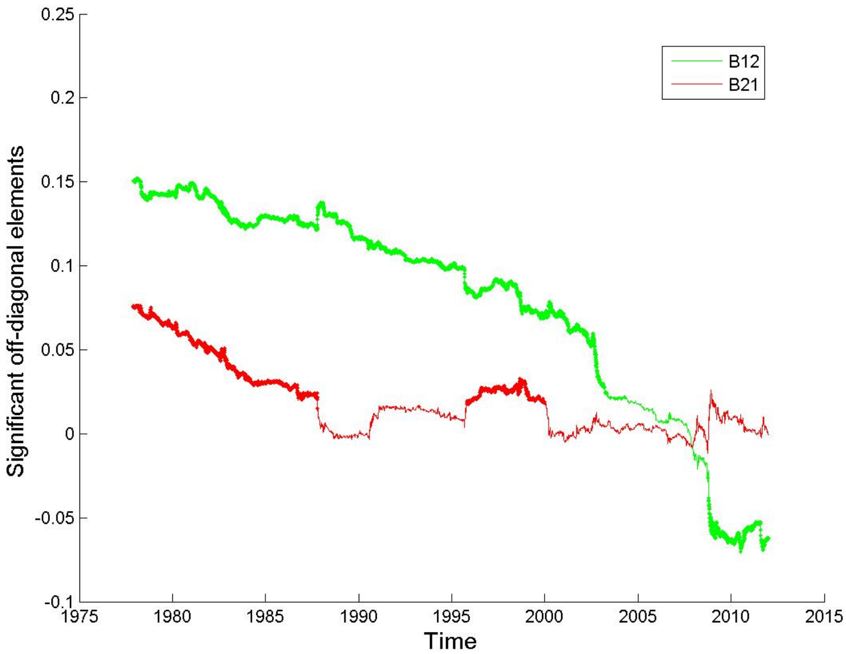
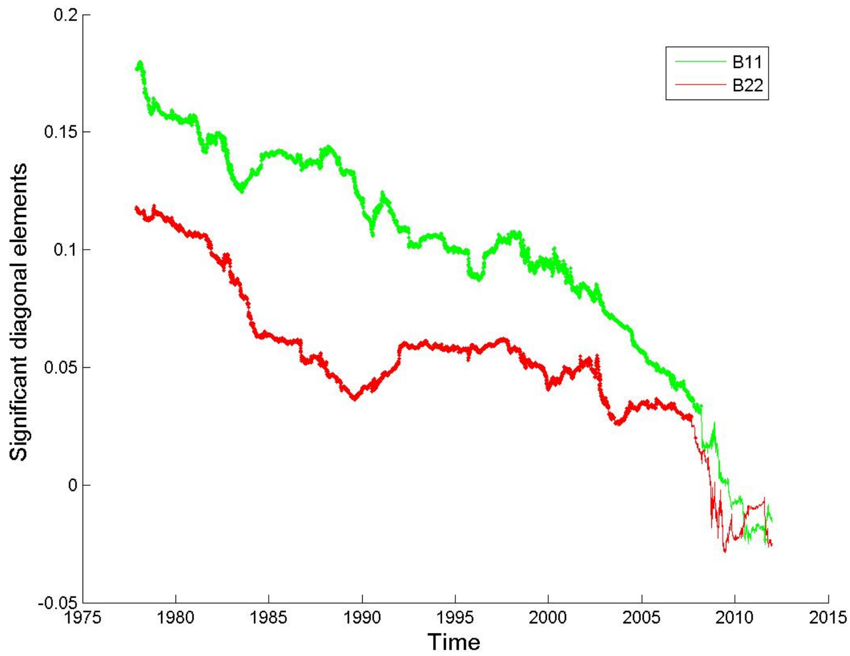


Figure 2: Two Book-to-Market-Sorted Portfolios

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

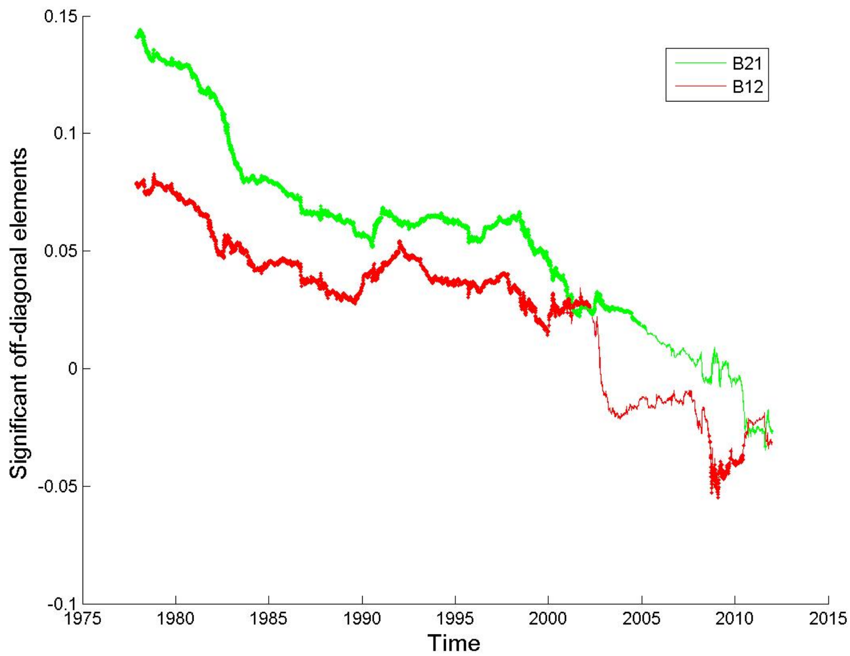
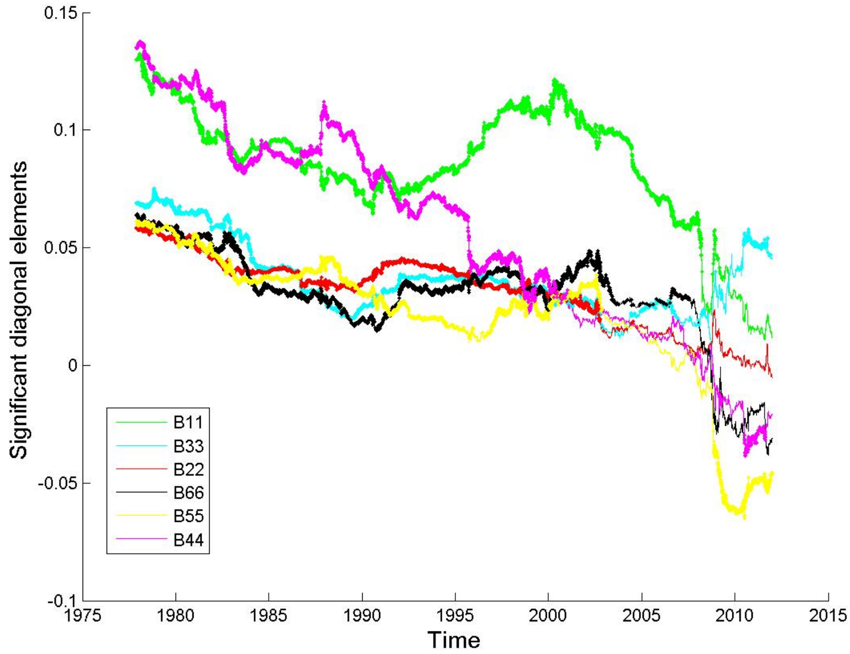


Figure 3: Six Size- and Book-to-Market-Sorted Portfolios

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

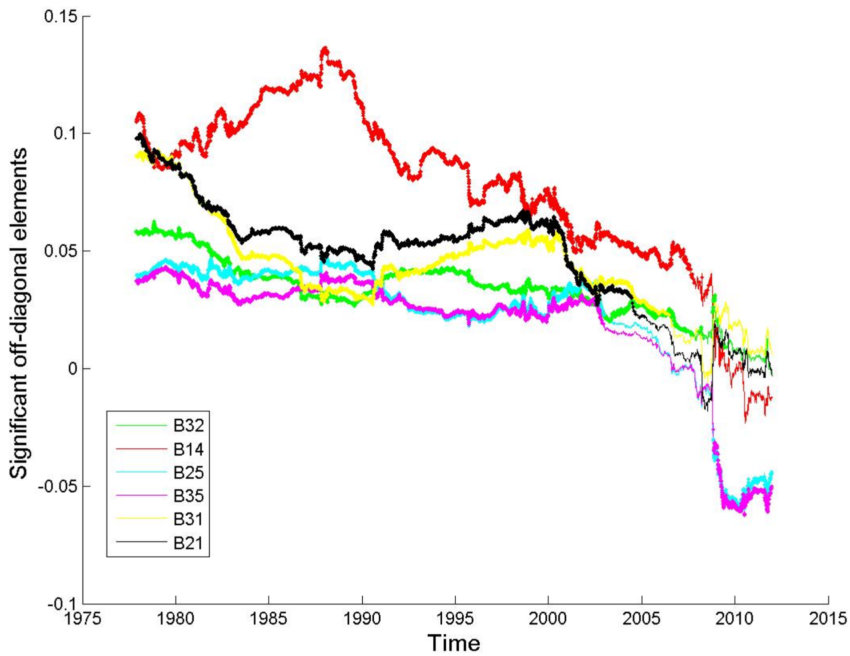
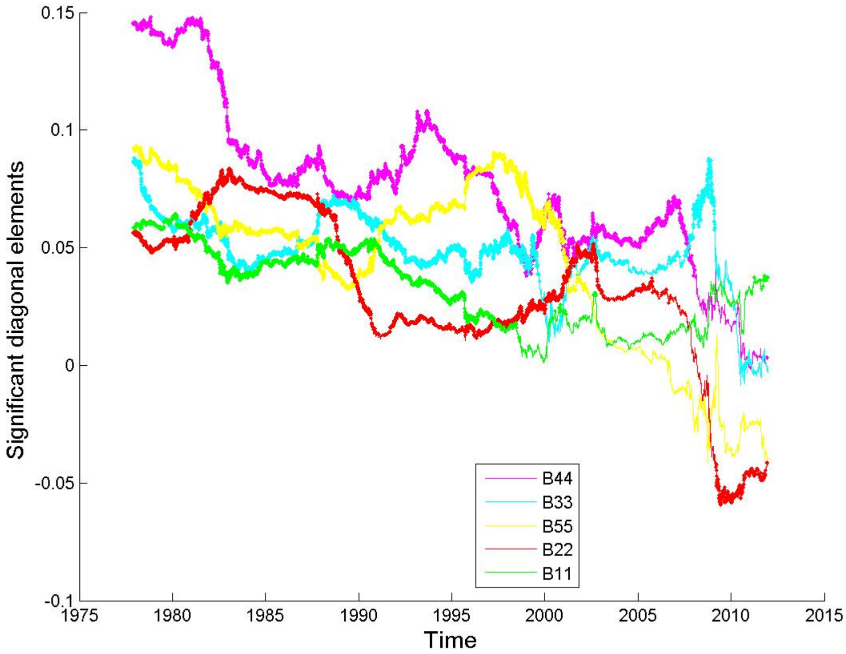


Figure 4: Five Industry Portfolios

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

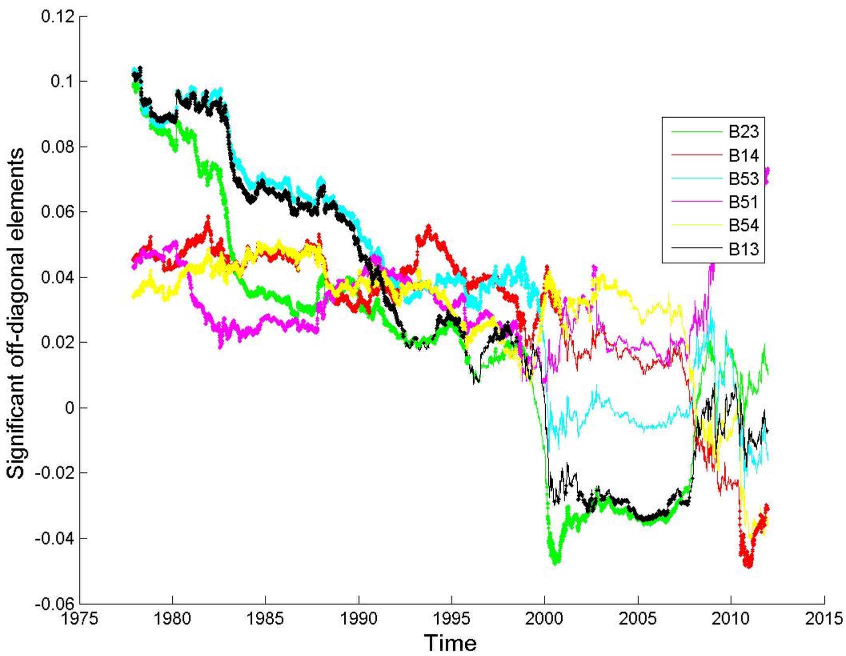
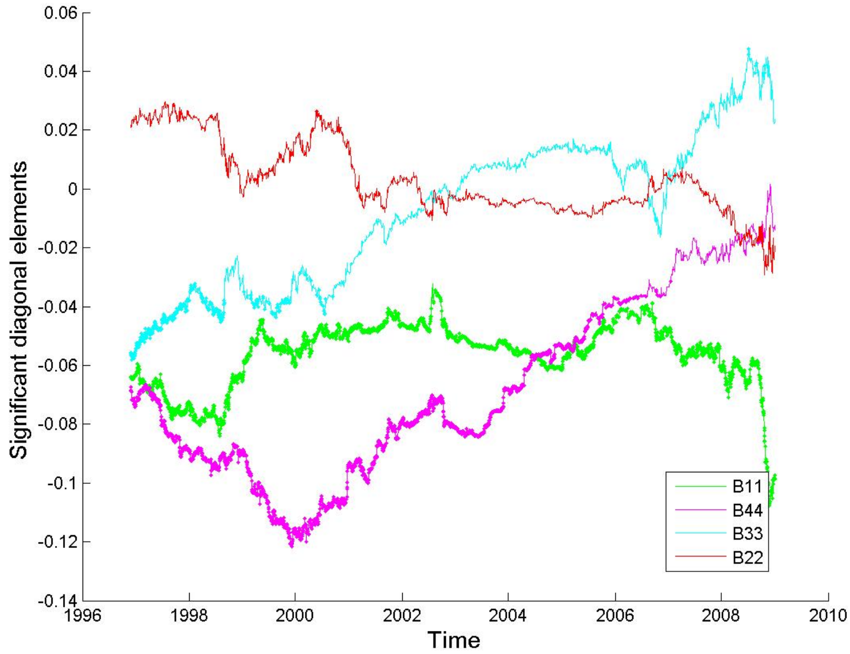


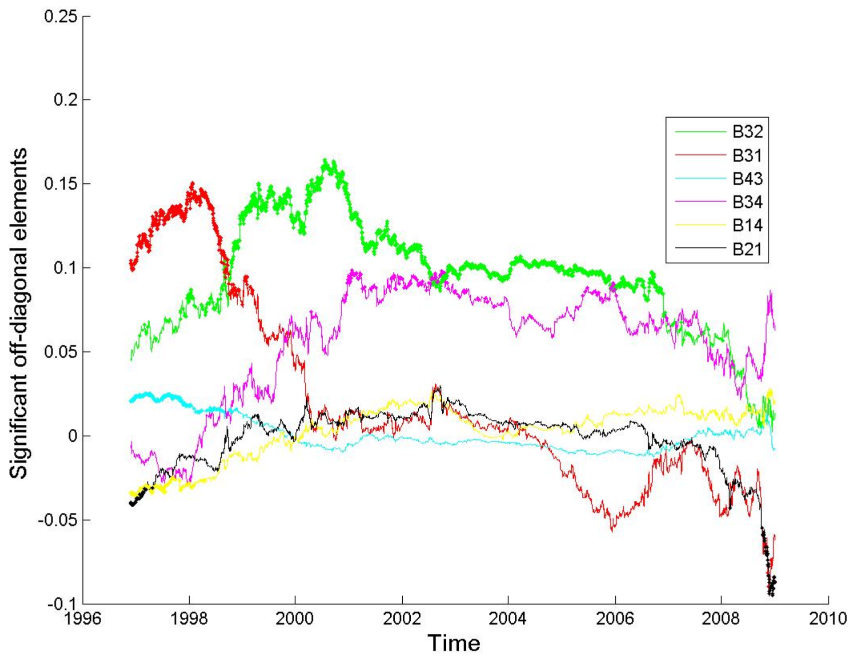
Figure 5: Four Individual Stocks

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

Ahn, D., J. Boudoukh, M. Richardson, and R. Whitelaw, 2002, "Partial adjustment or stale prices? Implications From Stock Index and Futures Return Autocorrelations," *Review of Financial Studies*, 15, 655–689.

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.

Boudoukh, J., M. Richardson, and R. Whitelaw, 1994, "A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Stock Returns," *Review of Financial Studies*, 7, 539–573.

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., A. W. Lo, and A. G. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

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*, Oxford University Press, New York.

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.

Conrad, J., and G. Kaul, 1989, "Mean Reversion in Short-Horizon Expected Returns," *Review of Financial Studies*, 2, 225–240.

Conrad, J., and G. Kaul, 1998, "An Anatomy of Trading Strategies," *Review of Financial Studies*, 11, 489–519.

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, 1990a, "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98, 703–738.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann, 1990b, "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.

DellaVigna, S., and J. Pollet, 2007, "Demographics and Industry Returns," *American Economic Review*, 97, 1667–1702.

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

DeMiguel, V., L. Garlappi, and R. Uppal, 2009, "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.

Golub, G. H., and C. F. V. Loan, 1996, *Matrix Computations*, The Johns Hopkins University Press, Baltimore.

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

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

Gyorfi, L., F. Udina, and H. Walk, 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.

Hirshleifer, D., and S. H. Teoh, 2003, "Limited Attention, Information Disclosures, and Financial Reporting," *Journal of Accounting and Economics*, 36, 337–386.

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

Hong, H., T. Lim, and J. C. Stein, 2000, "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies," *Journal of Finance*, 55, 265–95.

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.

Hong, H., W. Torous, and R. Valkanov, 2007, "Do Industries Lead the Stock Market," *Journal of Financial Economics*, 83, 367–396.

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.

Hou, K., L. Peng, and W. Xiong, 2009, "A Tale of Two Anomalies: The Implications of Investor Attention for Price and Earnings Momentum," Working paper, Princeton University.

Jagannathan, R., and T. Ma, 2003, "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58, 1651–1684.

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.

Khandani, A. E., and A. W. Lo, 2010, "What Happened to the Quants in August 2007? Evidence from Factors and Transactions Data," forthcoming in *Journal of Financial Markets*.

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.

Ledoit, O., and M. Wolf, 2004, "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices," *Journal of Multivariate Analysis*, 88, 365–411.

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

Lehman, B. N., 1990, "Fads, Martingales, and Market Efficiency," *Quarterly Journal of Economics*, 105, 1–28.

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

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.

Mech, T. S., 1993, "Portfolio Return Autocorrelation," *Journal of Financial Economics*, 34, 307–344.

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

Merton, R. C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Merton, R. C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information," *Journal of Finance*, 42, 483–510.

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.

Peng, L., and W. Xiong, 2006, "Investor Attention, Overconfidence, and Category Learning," *Journal of Financial Economics*, 80, 563–602.

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

Ramnath, S., 2002, "Investor and Analyst Reactions to Earnings Announcements of Related Firms: An Empirical Analysis," *Journal of Accounting Research*, 40, 1351–1376.

Rouwenhorst, K. G., 1998, "International Momentum Strategies," *Journal of Finance*, 53, 267–84.

Scharfstein, D. S., and J. C. Stein, 1990, "Herd Behavior and Investment," *American Economic Review*, 80, 465–79.

Schwarz, K., 1978, "Estimating the dimension of a model," *Annals of Statistics*, 6, 461–464.

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

Zhang, X. F., 2006, "Information Uncertainty and Stock Returns," *Journal of Finance*, 61, 105–136.