Essays in empirical asset pricing and portfolio construction

Michael William Ashby

Faculty of Economics

This dissertation is submitted for the degree of
Doctor of Philosophy

Peterhouse August 2020

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the Economics Degree Committee of 60,000 words.

Michael William Ashby
August 2020

Summary

Essays in empirical asset pricing and portfolio construction

Michael William Ashby

The key thread running through this thesis is predictability and how it relates to asset pricing and portfolio construction.

Chapter 1, co-authored with Oliver Linton, tests for predictablity in asset pricing model residuals to check model specification. We estimate three consumption-based asset pricing models and derive ex-ante expected stock market returns from them. For each model, a suite of tests rejects the null that the model residual, the difference between the ex-ante expected market return and the actual return, is a martingale difference sequence. The ability of these models to explain the own-history predictability of the market return is therefore rejected. Further tests show that lagged returns have too much predictive power over current returns to be consistent with the state variables which explain the market return being the same as the state variables which explain the market return in any of the three models.

Chapter 2 focusses on a specific type of predictive information. I examine whether regulator-required public disclosures of large net short positions can be profitably used to build portfolios. These disclosures do not form the basis of a profitable trading strategy for UK stocks. Long-short portfolios based on these disclosures typically make a profit, but it is statistically insignificant. While certain long-only unit initial outlay portfolios can reliably significantly outperform the market, this outperformance is economically modest: about one percentage point a year in gross and risk-adjusted terms.

Finally, Chapter 3 considers how best to use predictive information. Using predictive information unconditionally optimally produces better portfolios than using the predictive information conditionally optimally. Unconditionally optimal portfolios have higher Sharpe ratios and certainty equivalents, plus lower turnover, leverage, losses and drawdowns than conditionally optimal portfolios. Moreover, the unconditionally optimal portfolios tend to stochastically dominate the conditionally optimal portfolios once transaction costs are accounted for. However, whether unconditionally optimal portfolios are preferred to minimum variance or $1/N$ portfolios depends on the asset universe.

Acknowledgements

I am greatly indebted to Oliver Linton, my Supervisor, and Mark Salmon, my Advisor, for their help and guidance in completing my research. I give special thanks to Oliver, as a co-author, and Mark, who has offered support well beyond the call of duty for an Advisor. Both Oliver and Mark have been enthusiastic and encouraging in receiving my ideas, and thoughtful and kind in how they have helped me shape them into papers. Oliver and Mark's excellent courses in empirical finance and asset management, respectively, sparked my interest in empirical asset pricing and portfolio construction as research topics in the first place.

I also thank Tiago Cavalcanti, Soenje Reiche and especially Donald Robertson and Melvyn Weeks for their extremely helpful comments on Chapter 1. In addition, I am very grateful to Yongmiao Hong and Yoon-Jin Lee for sharing the GAUSS code to implement their generalised spectral test for Chapter 1. Massoud Mussavian, Reto Tuffi, James Ng and especially Jacopo Capra provided very helpful comments for Chapter 2. Jacopo's insight into the practicalities and transaction costs of short selling was especially helpful. I also had helpful discussions regarding Chapter 2 with Marco Geraci and Pedro Saffi. I had helpful discussions with Hans-Peter Schrei about Chapter 3. I am grateful to the Econometrics Subject Group for their feedback on my work at the PhD workshops. Thanks also go to Silvana Dean and Louise Cross for their help navigating the bureaucracy of the PhD, and to Dan Wales for helpful conversations throughout.

Preface

Chapter 1 contains some material derived from the dissertation I submitted as part of my MPhil in Economic Research at the University of Cambridge. It draws on my MPhil dissertation only in the outline of the Bansal-Yaron and Campbell-Cochrane models and the estimation of their parameters.

Chapter 1 is co-authored with Oliver Linton. Professor Linton provided the theory for the weighted correlogram test and the bootstrap procedure used in the quantilogram and rescaled range tests.

Chapters 2 and 3 contain no co-authored work, nor work previously or due to be submitted for another qualification.

Table of contents

Chapter 1

Do consumption-based asset pricing models explain own-history predictability in stock market returns?

We show that three prominent consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the own-history predictability properties of stock market returns. We show this by estimating these models with GMM, deriving ex-ante expected returns from them and then testing whether the difference between realised and expected returns is a martingale difference sequence, which it is not. Furthermore, a semi-parametric test suggests that lagged returns have too much predictive power over current returns to be consistent with the state variables which explain market returns being the same as the state variables which explain market returns in any of the three models.

This Chapter is co-authored with Oliver Linton. Professor Linton provided the theory for the weighted correlogram test and the bootstrap procedure used in the quantilogram and rescaled range tests.

This Chapter contains some material derived from the dissertation I submitted as part of my MPhil in Economic Research at the University of Cambridge. It draws on my MPhil dissertation only in the outline of the Bansal-Yaron and Campbell-Cochrane models and the estimation of their parameters.

1.1 Introduction

Three prominent consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the own-history predictability structure of the US market return. The Bansal-Yaron and Campbell-Cochrane models are designed to explain the level of stock market returns, in particular to simultaneously resolve the equity premium and risk-free rate puzzles. Yet, whether these models can explain the degree of predictability in stock returns is of interest too, especially if investors want to time or beat the market. In this sense, the dynamics (second moment) of returns are important separately to their level (first moment). This is recognised by Cecchetti et al. (1990). The Cecchetti-Lam-Mark was developed specifically to explain return dynamics, rather than to price assets per se. Since own-history predictability is the most basic kind of predictability, this is what we consider.

Our tests of whether the three models can explain own-history predictability amount to testing whether the difference between the model-implied ex-ante expected market return and the realised market return - the residual - is a martingale difference sequence (MDS). Since the residuals are not MDS, there is some own-history predictability left over in realised returns not captured by the models. To construct the expected returns and residuals, we first estimate the models by GMM. Our testing procedures account for this estimation step.

We base our tests of the null that the residuals are MDS on serial correlation, quantile hits, the rescaled range and the generalised spectrum (Hong, 1999). The asymptotic distribution of the serial correlation and generalised spectrum-based tests accounts for the initial estimation step, while we use a bootstrap procedure to account for the estimation step in the quantile hits and rescaled range-based tests. We use a battery of tests since tests of the MDS null can suffer locally low power against certain alternatives (Poterba and Summers, 1988).

Our finding that none of the three models can explain the own-history predictability properties of the market return is robust to the empirical choices we make. It does not matter whether we use the optimal GMM weight matrix, or the identity matrix; whether we use size/book-to-market or industry portfolios to estimate the models; or whether we use quarterly, instead of annual, data. The only apparent hope comes from estimating the Cecchetti-Lam-Mark model using size/book-to-market portfolios and the identity GMM weight matrix at the quarterly frequency. However, using a quarterly sample gives a much larger number of observations and allows us to consider the robustness of our results over time by splitting the sample into two equal-length sub-samples. When we do this, we clearly reject the null that the Cecchetti-Lam-Mark residuals are MDS in both sub-samples.

In each of the robustness check cases, we consider only models that provide credible expected returns. Many of the robustness check specifications do not give plausible expected returns series. There is no point checking the second moment of a model that fits poorly in terms of the first moment, as one would not use it to price assets anyway. Moreover, the centred second moment (e.g. serial correlation coefficient) is a function of the first moment.

We also consider a semi-parametric test of whether the state variables of the three models can explain the own-history predictability properties of returns. Our test is an adaptation of the Huang and Zhou (2017) test. We test whether the R^2 from a predictive regression of returns on their lagged values exceeds a theoretical upper bound, $\bar{R}^2$. $\bar{R}^2$ depends on the state variables of the stochastic discount factor (i.e. the state variables which explain stock returns). Unlike the residual-based tests, this test does not depend on the functional form of the stochastic discount factor being correctly specified. It requires only that the state variables be correctly specified.

The Bansal-Yaron state variables cannot explain the own-history predictability of returns. We find find statistically significant excess predictability (excessively high R^2 significantly greater than $\bar{R}^2$) at four out of nine horizons using annual data and six out of nine horizons using quarterly data. While there is superficially more hope for the Campbell-Cochrane and Cecchetti-Lam-Mark model state variables, this turns out not to be robust. There is statistically significant excess predictability at only one of the nine horizons considered for the Campbell-Cochrane and Cecchetti-Lam-Mark models in our main results using annual data. However, this good performance does not survive switching to quarterly data. There are many R^2 bound exceedences for the Campbell-Cochrane state variable using quarterly data. There is only one significant R^2 bound violation for the Cecchetti-Lam-Mark state variable over the whole sample using quarterly data. But, again, there are many violations in each sub-sample when we split the sample into two equal-length sub-samples, and the ability of the Cecchetti-Lam-Mark state variable to explain return predictability is not robust over time.

Apart from the question of how well these models explain own-history predictability in asset returns being interesting in its own right, testing this property leads us naturally to residual-based testing. This is a standard time-series specification test, although not one that is commonly used in the context of consumption-based asset pricing models. In this setting, GMM estimation and an accompanying J-test is more common. The advantage of testing the residuals, in this case from the market return, is that it allows us to test models which are estimated in “stages” - i.e. where the estimation is not done in one single GMM implementation. Both the Campbell-Cochrane and Cecchetti-Lam-Mark models are estimated in stages in this way.

The Bansal-Yaron and Campbell-Cochrane models are two of the most prominent models designed to simultaneously explain the equity premium (Mehra and Prescott, 1985) and risk-free rate (Weil, 1989) puzzles. Assuming a standard endowment economy with a representative investor who has constant relative risk aversion (CRRA) preferences, the observed difference between stock returns and low-risk bond yields requires extremely high levels of risk aversion to explain. This is the equity premium puzzle. The risk-free rate puzzle compounds the equity premium puzzle. If CRRA investors are indeed as risk averse as they would need to be to justify the equity premium, low-risk bond yields are far too low. As a result, researchers such as Bansal and Yaron (2004) and Campbell and Cochrane (1999) have sought to modify the standard CRRA set-up in order to account for these puzzles. In terms of explaining the equity premium and risk-free rate puzzles simultaneously, these models do reasonably well. But they are yet to be examined in terms of their ability to capture the predictability of stock returns in any great detail.

Huang and Zhou (2017) is the main study of how well the Bansal-Yaron and Campbell-Cochrane models explain return predictability. They develop the R^2 bound test described above, but in the context of one-step-ahead predictability of the market return with respect to several well known predictors (the book-to-market ratio, term spread, CAY, investment-to-capital ratio, new-orders-to-shipments ratio, output gap and credit expansion).[1] Huang and Zhou use Constantinides and Ghosh's (2011) inversion of the Bansal-Yaron model which renders the state variables observable. For the Campbell-Cochrane model, the state variable is unobserved and Huang and Zhou extract it as per Campbell and Cochrane's (1999) calibration. They do not estimate the model first, but condition on the extracted state variable. Huang and Zhou show that the degree of predictability in the market return is greater than can be explained by the Bansal-Yaron and Campbell-Cochrane models' state variables.

Our residual-based approach is potentially more powerful, since it can detect situations where the asset pricing model suggests too little predictability. In addition, our residual-based tests have the advantage of accounting explicitly for any initial estimation of the model or its state variables. While the Bansal-Yaron model can be inverted so that its state variables are a function of observables, this inversion is not generally possible for other asset pricing models (e.g. the Campbell-Cochrane model).

There has been little recent work on explaining own-history stock return predictability in the context of consumption-based asset pricing models. Kandel and Stambaugh (1989) propose a model with a representative CRRA investor and where consumption growth is lognormally distributed with time-varying mean and variance. The mean and variance of consumption

[1] Our adaptation is to adapt the test for q-period-ahead predictability with respect to lagged returns.

growth follow a nine-state Markov-switching process and exhibit positive serial correlation. Kandel and Stambaugh's calibration exercise shows that the model produces the "U" shaped autocorrelation function observed in stock returns. However, the model is not able to replicate the observed pattern of small positive autocorrelations at short horizons followed by larger negative autocorrelations at longer horizons. Kandel and Stambaugh speculate that this is because their model is overly restrictive. In particular, current news only affects the conditional distribution of consumption one period in the future. Nonetheless, their model broadly matches the observed pattern of autocorrelations at horizons greater than 12 months.

Cecchetti et al. (1990) use a similar specification to Kandel and Stambaugh. Cecchetti et al. use a Markov-switching log endowment level and a more parsimonious two-state specification. They find that popular measures of serial correlation always lie within a 60% confidence interval of data simulated from the model. The Cecchetti et al. model has the same problem of not being able to generate negative autocorrelations at short horizons as the Kandel and Stambaugh model.

We update the Cecchetti et al. (1990) evidence in two ways. First, we formally estimate their model. This also allows for the development of asymptotic theory for the hypothesis tests used. Second, the Cecchetti et al. (1990) model rests on CRRA preferences. As discussed above, these have been much criticised on an empirical basis, in particular because of the equity premium and risk-free rate puzzles. We test more recent models that can potentially accommodate these two puzzles. However, we also include the Cecchetti-Lam-Mark model in our results as a benchmark, since it is a model explicitly designed to explain serial correlation in returns.

Other attempts have been made to explain own-history predictability in a risk-based framework. Kim et al. (2001) proxy risk by volatility and use a volatility feedback model (where an unexpected change in volatility has an immediate impact on stock prices) with volatility following a two-state Markov-switching process. Risk adjusting returns in this way accounts for the serial correlation observed in returns. We focus on consumption-based models, which micro-found their risk factors from the start, rather than more ad hoc risk adjustments.

More recently, Barroso et al. (2017) consider how conditional predictability of the short-run equity premium varies with economic and risk conditions.[2] They model the equity risk premium as a function of economic state variables. The extent to which these state variables forecast both the equity risk premium and consumption growth varies with time. When a state variable predicts consumption growth more strongly, it also contributes more to the equity premium. This is consistent with the intertemporal CAPM (Barroso et al., 2017). A consumption-based

[2] There are also non-risk based explanations for return predictability. These are beyond the scope of this paper.

asset pricing model is capable of explaining short-term conditional predictability, although no specific specification is tested.

This paper proceeds as follows. Section 1.2 outlines the three asset pricing models tested and their estimation. Section 1.3 discusses the tests we use and how we modify them to account for parameter estimation. Section 1.4 briefly describes the data and reports the estimation of the asset pricing models. Section 1.5 presents our empirical results regarding the predictability of the model residuals and Section 1.6 our robustness analysis. Section 1.7 concludes.

1.2 The models and their estimation

1.2.1 Bansal-Yaron model

The Bansal and Yaron (2004) model is as follows:

$$V_t = \left[(1-\delta) C_t^{1-\frac{1}{\psi}} + \delta \left(\mathrm{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{1.1}$$

$$x_{t+1} = \rho_x x_t + \psi_x \sigma_t \varepsilon_{t+1} \tag{1.2}$$

$$\Delta c_{t+1} = \mu_c + x_t + \sigma_t \eta_{t+1} \tag{1.3}$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \varphi \sigma_t u_{t+1} \tag{1.4}$$

$$\sigma_{t+1}^2 = \sigma^2 + \nu(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1} \tag{1.5}$$

$$\varepsilon_t, \eta_t, u_t, w_t \sim NID(0,1),$$

where V_t is the representative investor's value function, δ the subjective discount factor, $\gamma > 0$ the risk-aversion coefficient, $\psi > 0$ the elasticity of intertemporal substitution (EIS), C_t consumption, D_t dividends, E_t the expectation conditional on information at time t and lower-case variables denote logs of upper-case variables.

The model has three key ingredients. First, it has recursive preferences (1.1) à la Epstein and Zin (1989) and Weil (1989). These allow EIS and risk aversion to differ, unlike standard CRRA preferences. This is an advantage: risk aversion and intertemporal substitution are different concepts. EIS reflects the extent to which consumers are willing to smooth certain consumption through time, while risk aversion relates to the extent to which consumers are willing to smooth consumption across uncertain states of nature (Cochrane, 2008).

Second, consumption growth (1.3) has a small predictable component (the long-run risk, x_t). Consumption news in the present affects expectations of future consumption growth, increasing the impact of current consumption news on long-run consumption and therefore the difference between present discounted values (PDVs) of dividend streams which drives returns.

Third, there is time-varying economic volatility (1.5) in consumption growth. This reflects time-varying economic uncertainty and is a further source of investor uncertainty and risk.

In the Bansal and Yaron (2004) calibration, the model justifies the equity premium, risk-free rate and the volatilities of the market return, risk-free rate and price-dividend ratio.

When Constantinides and Ghosh (2011) estimate the Bansal-Yaron model by GMM, the results are mixed. Simulating through the model with the estimated parameter values, the model is able to justify the market return in all specifications considered. The mean risk-free rate can be a little high, although this too is justified when the model is estimated using the identity weight matrix. Meanwhile, the J-statistic p-value is less than 0.03 in all specifications considered. However, the estimated model still generates reasonable market returns in Constantinides and Ghosh's simulations and the model may therefore still be of interest from an asset pricing point-of-view.

To estimate the model, Constantinides and Ghosh (2011) show that the log-linearised version of the Bansal-Yaron model can be inverted, allowing the unobserved state variables to be written as a linear combination of observables as follows.

$$x_t = \alpha_0 + \alpha_1 r_{f,t} + \alpha_2 z_{m,t} \tag{1.6}$$

$$\sigma_t^2 = \beta_0 + \beta_1 r_{f,t} + \beta_2 z_{m,t} \tag{1.7}$$

where $\alpha_0, \ldots, \beta_2$ are functions of Bansal-Yaron model parameters, as detailed in Appendix A.1.1, and $r_{f,t}$ the (log) risk-free rate. This allows them to express the Bansal-Yaron Euler equation for a general asset as

$$\mathrm{E}_t \left[\exp \left\{ a_1 + a_2 \Delta c_{t+1} + a_3 \left(r_{f,t+1} - \frac{1}{\kappa_1} r_{f,t} \right) + a_4 \left(z_{m,t+1} - \frac{1}{\kappa_1} z_{m,t} \right) + r_{t+1} \right\} \right] - 1 = 0,$$

where r_t is the log asset return and $a_1, \ldots, a_4, \kappa_1$ are functions of the Bansal-Yaron model parameters, also given in Appendix A.1.1.

In addition, they derive eight unconditional moment restrictions for continuously compounded consumption and dividend growth, which are given in Appendix A.1.2. These moment conditions are derived from Bansal and Yaron's (2004) specification of consumption and dividend growth, the long-run risk and its conditional variance.

The model has 12 parameters to estimate and we use 15 moment conditions to allow for an overidentification test. Our set of moment conditions comprises an Euler equation for each of seven assets (the market index and six size and book-to-market double sorted portfolios, taken from Kenneth French's website), and the eight time-series restrictions.

Constantinides and Ghosh (2011) show that

$$\mathrm{E}_t\, r_{m,t+1} = B_0 + B_1 x_t + B_2 \sigma_t^2$$

where $r_{m,t}$ is the market return and $B_0, \ldots, B_2$ are non-linear combinations of the 12 model parameters provided in Appendix A.1.3. This yields a plug-in estimator of $\mathrm{E}_t\, r_{m,t+1}$, which we use as the ex-ante expected market return.

1.2.2 Campbell-Cochrane model

Campbell and Cochrane's (1999) model adds a slow-moving external habit to the standard power utility function. The representative agent's utility function is

$$U_t(C) = \mathrm{E}_t \sum_{s=0}^{\infty} \delta^s \frac{(C_{t+s} - H_{t+s})^{1-\gamma} - 1}{1-\gamma},$$

where δ is the subjective discount factor, γ the utility curvature and H_t the habit level of consumption. Defining $S_t \equiv (C_t - H_t)/C_t$ and $s_t \equiv \ln(S_t)$, the habit evolves according to

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t) v_{t+1}, \tag{1.8}$$

where $\bar{s}$ is the steady-state s, $\bar{S} = \sigma_v \sqrt{\gamma/(1-\phi)}$ and $\lambda(s_t)$ is a sensitivity function given by

$$\lambda(s_t) = \begin{cases} (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1, & \text{if } s_t \leq s_{\max} \\ 0, & \text{otherwise,} \end{cases} \tag{1.9}$$

with $s_{\max} \equiv \bar{s} + \frac{1}{2}(1 - \bar{S}^2)$. Campbell and Cochrane set ϕ to be equal to the first-order autocorrelation coefficient of the log market price-dividend ratio, $z_{m,t}$.

Consumption and dividends satisfy

$$\begin{aligned} \Delta c_t &= \bar{g} + v_t \\ \Delta d_t &= \bar{g} + w_t \end{aligned} \tag{1.10}$$

with Δ being the first difference operator and

$$\begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim NID\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_v^2 & \sigma_{vw} \\ \sigma_{vw} & \sigma_w^2 \end{pmatrix}\right), \tag{1.11}$$

where NID indicates normally and independently and identically distributed through time.

Campbell and Cochrane (1999) calibrate their model to match the annualised unconditional equity premium using monthly US data. When given actual data, the model replicates the main movements observed in stock prices. In simulations, the model is able to justify the means and standard deviations of excess returns and the price-dividend ratio, and the existence of a short-run and long-run equity premium. Moreover, this is achieved without a risk-free rate puzzle by construction: the habit is specified such that the risk-free rate remains constant and the model is calibrated such that the log risk-free rate is equal to its sample mean.[3]

In Garcia et al.'s (2004) GMM estimation of the Campbell-Cochrane model, the estimated γ is significantly greater than 0 and the δ significantly less than 1. The J-statistic p-value exceeds 0.2, although this does condition on earlier estimates of time-series parameters in the manner described below.

We estimate the Campbell-Cochrane model using a GMM procedure similar to Garcia et al. (2004). The procedure has three steps. First, we estimate the time-series parameters $\bar{g}$, σ_v^2 and σ_w^2 in (1.10) by GMM. Second, we estimate α and ϕ from the linear regression

$$z_{m,t+1} = \alpha + \phi z_{m,t} + e_{t+1}.$$

Based on these estimates, we generate the series s_t. We do so by initialising the series at $s_0 = \bar{s} = \ln(\sigma_v \sqrt{\gamma/(1-\phi)})$, using the estimates of the relevant time-series moments from above and assuming an initial γ of 2. This allows the series s_t to be generated as per (1.8) and (1.9).

We can then proceed to the third step: estimating the preference parameters δ and γ from the Euler equation

$$\mathrm{E}_t\left[\delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma}(1+R_t)\right] - 1 = 0, \tag{1.12}$$

using an Euler equation for each of our seven assets. We use this new estimate of γ to generate a new s_t series, and re-estimate (1.12) based on this new s_t series. We iterate this procedure

[3] Campbell and Cochrane (1999) argue this is realistic as the risk-free rate varies relatively little and does not vary cyclically.

until the estimates of δ and γ converge. The J-statistic p-values of Garcia et al. (2004) come from their final iteration of this third step, but do not account for the initial estimation steps.

We obtain $\mathrm{E}_t\, r_{m,t+1}$ from the Campbell-Cochrane model as follows. We use the fact that $1+R_t = (P_t + D_t)/P_{t-1}$, where P_t is the price of the asset and D_t its dividend. Iterating the Euler equation forwards, we have

$$P_t = \sum_{j=1}^{\infty} \delta^j \,\mathrm{E}_t \left[\left(\frac{S_{t+j}}{S_t} \frac{C_{t+j}}{C_t} \right)^{-\gamma} D_{t+j} \right] \tag{1.13}$$

when we impose the no-bubble condition

$$\lim_{j\to\infty} \delta^j \,\mathrm{E}_t \left[\left(\frac{S_{t+j}}{S_t} \frac{C_{t+j}}{C_t} \right)^{-\gamma} P_{t+j} \right] = 0.$$

Therefore,

$$\mathrm{E}_t(1+R_{t+1}) = \frac{\mathrm{E}_t \sum_{j=1}^{\infty} \delta^j \left(\frac{S_{t+1+j}}{S_{t+1}} \frac{C_{t+1+j}}{C_{t+1}} \right)^{-\gamma} D_{t+1+j}}{\mathrm{E}_t \sum_{j=1}^{\infty} \delta^j \left(\frac{S_{t+j}}{S_t} \frac{C_{t+j}}{C_t} \right)^{-\gamma} D_{t+j}}. \tag{1.14}$$

We estimate (1.14) for the market return by simulation. We simulate the series $v_{t+1}, v_{t+2}, v_{t+3}, \ldots$ and $w_{t+1}, w_{t+2}, w_{t+3}, \ldots$ according to (1.11). Based on these series, we compute the series $s_{t+1}, s_{t+2}, s_{t+3}, \ldots$, $c_{t+1}, c_{t+2}, c_{t+3}, \ldots$ and $d_{t+1}, d_{t+2}, d_{t+3}, \ldots$ conditional on s_t, c_t and d_t. We repeat this procedure 200 times, where each simulated v_{t+1} and w_{t+1} series is of length 100. We then compute the expectation on the right-hand side of (1.14) as the mean of the 200 simulated realisations of the fraction inside that expectation.

1.2.3 Cecchetti-Lam-Mark model

Cecchetti et al.'s (1990) model attempts to explain return autocorrelation in a rational framework. The model is an endowment economy where the representative consumer has CRRA preferences:

$$U_t(C) = \mathrm{E}_t \sum_{s=0}^{\infty} \delta^s \frac{C_{t+s}^{1-\gamma} - 1}{1-\gamma}.$$

Here, δ denotes the subjective discount factor and γ the coefficient of relative risk aversion. Taking (log) consumption as the appropriate endowment process,

$$\Delta c_{t+1} = \alpha_0 + \alpha_1 y_t + \varepsilon_{t+1}\,. \tag{1.15}$$

$y_t \in \{0,1\}$ is a first-order Markov process and $\varepsilon_t \sim NID(0,\sigma^2)$. $y_t = 1$ denotes a bad state, so α_1 is restricted to be less than zero.

Cecchetti et al. (1990) find that, using either risk-neutral ($\gamma = 0$) or risk-averse ($\gamma = 1.7$) preferences, serial correlation in the observed market return always lies within a 60% confidence interval of serial correlation in the market return generated by the model. The confidence intervals come from Monte Carlo distributions of the serial correlation statistics, obtained by simulating the model. The medians of the Monte Carlo distributions of the serial correlation statistics obtained using $\gamma = 1.7$ are closer to the observed serial correlation than the medians of the distributions using $\gamma = 0$, so Cecchetti et al. prefer the risk-averse specification. Cecchetti et al. measure serial correlation using variance ratios and Fama and French (1988) regression coefficients[4] using annual US/S&P data over 2-10 year horizons.

There is no guarantee that this model would simultaneously explain the equity premium and risk-free rate puzzles. Given the CRRA preferences, it probably would not. However, given the model's success in explaining market serial correlation, it is a useful benchmark for our analysis.

We use GMM to estimate δ and γ. The moment conditions comprise an Euler equation for each of our seven assets of the form

$$\mathrm{E}_t\left[\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}(1+R_t)\right]-1=0. \tag{1.16}$$

We estimate the Markov switching endowment process by maximum likelihood following Hamilton (1989). In a slight deviation from Cecchetti et al. (1990), we estimate a Markov-switching process where the consumption innovation $\varepsilon_{t+1}|y_t \sim N(0,\sigma^2_{y_t})$, since this is more numerically stable.

$\mathrm{E}_t\, r_{m,t+1} \approx \mathrm{E}_t\,[\ln(1+R_{m,t+1})]$ and Cecchetti et al. (1990) show that

$$\mathrm{E}_t\,[\ln(1+R_{m,t+1})] = \mathrm{E}_t\left[\ln\left(\frac{1+\kappa(y_{t+1})}{\kappa(y_t)}\right)+(\alpha_0+\alpha_1 y_t+\varepsilon_{t+1})\right] \tag{1.17}$$

where $\kappa(y_t)$ is a non-linear function of model parameters defined in Appendix A.2. Since y_t is a binary variable and the distribution the expectation in (1.17) is straightforward to compute.

[4] Fama and French (1988) regression coefficients are the slope coefficient from a regression of the q-period return from t to $t+q$ on the q-period return from $t-q$ to t.

1.3 Tests

To test whether the asset-pricing models discussed above capture the serial correlation structure of stock returns, we note that rational expectations imply

$$r_{m,t+1} = \mathrm{E}_t\, r_{m,t+1} + \xi_{t+1}, \tag{1.18}$$

where expectations are formed under the model in question and ξ_{t+1} is unforecastable at t. If the model accurately captures own-history predictability, ξ_t should be MDS. If not, there is clearly something in the own-history predictability structure of r_t not captured by $\mathrm{E}_{t-1}\, r_t$.

We denote by θ the parameters in the model in question and define $\mathrm{E}_t\, r_{m,t+1} = \mu_{t+1}(\theta)$, to make clear the dependence of the expected returns on θ. We estimate (1.18) using plug-in estimators, $\mu_t(\hat{\theta})$, of $\mathrm{E}_t\, r_{t+1}$. We base our tests on the resulting residual $\xi_t(\hat{\theta})$ and denote

$$\bar{\xi} = T^{-1}\sum_{t=1}^{T}\xi_t(\hat{\theta}), \qquad \hat{s}^2 = T^{-1}\sum_{t=1}^{T}(\xi_t(\hat{\theta}) - \bar{\xi})^2.$$

We consider tests of linear and non-linear predictability in $\xi_t(\hat{\theta})$, as well as a rescaled range test. In each case, we adapt the test to cope with the fact that $\mu_t(\theta) \equiv \mathrm{E}_{t-1}\, r_{m,t}$ is estimated and this estimate, $\mu_t(\hat{\theta})$, is a function of a parameter vector estimated by GMM. It is well known that this estimation can both affect the limiting distribution of the statistics considered and induce serial dependence in the estimated residuals not present in the population.

In light of Poterba and Summers's (1988) argument that tests of the MDS null can have locally low power against certain alternatives, we use a battery of tests. Different tests have different power properties against different (local) alternatives. It therefore seems prudent to cover all bases and consider several tests. This approach bears fruit. Throughout the results, there are examples where one test fails to reject while all the others reject. It is not the case that the same test keeps failing to reject.

1.3.1 Linear predictability

A natural place to start with testing whether or not the residuals are MDS is a test based on the residuals' autocorrelations. Since the MDS null implies that all autocorrelations are zero, it makes sense to use a test statistic that incorporates autocorrelations from more than one lag.

We use a weighted correlogram, of the form

$$C(q) = \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right) \rho(j), \tag{1.19}$$

where $\rho(j)$ is the j^{th} order serial correlation coefficient of ξ_t. $C(q)$ is a weighted sum of serial correlations. If $C(q) > 0$, positive autocorrelation predominates at horizon q. $C(q) < 0$ is evidence that negative autocorrelation predominates at horizon q. We consider $q \in \{2, 3, \ldots, 10\}$ years.

Within the class of tests based on multiple serial correlation coefficients, we use a test of the form in (1.19) for two reasons. First, it is a linear transformation of the variance ratio statistic. The variance ratio $VR(q)$ is the variance of the sum of q residuals divided by q times the variance of the residuals. That is $VR(q) = \mathrm{Var}(\xi_{t+1} + \xi_{t+2} + \ldots + \xi_{t+q})/q\,\mathrm{Var}(\xi_t)$. Since under the MDS null the residuals ξ_t and ξ_{t+j} $(j \neq 0)$ are uncorrelated, the variance ratio is equal to one under the null. Cochrane (1988) shows we can write $VR(q) = 1 + 2C(q)$, hence the connection between (1.19) and $VR(q)$.

Second, Poterba and Summers (1988) and Lo and MacKinlay (1989) show variance ratio tests are generally more powerful tests of the martingale difference hypothesis than unit root and autoregressive tests. We are unable to compute $VR(q)$ directly in a way that accounts for the estimation of $\hat{\theta}$, since the Delgado and Velasco (2011) formulae extend only to serial correlations, not variances. Given that we can only work with serial correlations, there is then little point multiplying $C(q)$ by two and adding one. In addition, we prefer a test of the form in (1.19) over a Box-Pierce type test because there is information in the sign of the test regarding whether positive or negative serial correlation prevails.

In terms of estimating $C(q)$, we cannot simply treat the estimated residuals $\xi_t(\hat{\theta})$ as if they are the population residuals $\xi_t(\theta)$. The estimation of $\hat{\theta}$ affects the limiting distribution of $\hat{\rho}(j)$ under the MDS null (Delgado and Velasco, 2011). We therefore use Delgado and Velasco's (2011) transformation of the residual sample serial correlations. We denote the transformed autocorrelations by $\bar{\rho}(j)$. Delgado and Velasco start by standardising the autocorrelations so that they have a unit variance. To do this, they define the matrix A^m such that

$$(A^m)^{-1/2} \hat{\rho}^m \sim \mathrm{N}(0, I_m),$$

with $\hat{\rho}^m = [\hat{\rho}(1), \ldots, \hat{\rho}(m)]$. To make the transformation feasible, Delgado and Velasco (2011) use Lobato et al.'s (2002) estimate of A^m

$$\hat{A}^m = \frac{1}{T\hat{s}^4}\left[g^m(0) + \sum_{j=1}^{\ell-1}\left(1-\frac{j}{\ell}\right)\{g^m(j)+g^m(j)'\}\right]$$

where $g^m(j) = T^{-1}\sum_{t=1+j}^{T} w_t^m w_{t-j}^{m\prime}$, $w_t^m = (w_{1,t}, \ldots, w_{m,t})'$, $w_{k,t} = \left(\hat{\xi}_t(\hat{\theta}) - \bar{\xi}\right)\left(\hat{\xi}_{t-j}(\hat{\theta}) - \bar{\xi}\right)$ and ℓ is a bandwidth parameter. We use $\ell = \left\lceil T^{1/3}\right\rceil$.

Delgado and Velasco (2011) rid the estimated serial correlations collected in $\hat{\rho}^m$ of their dependence on $\hat{\theta}$ by projecting them onto the derivatives of $\hat{\xi}_t(\hat{\theta})$. First, define

$$\hat{\zeta}^m = \left[\hat{\zeta}(1)', \ldots, \hat{\zeta}(m)'\right]'$$
$$\hat{\zeta}(j) = \frac{1}{T\hat{s}^2}\sum_{t=j+1}^{T}\dot{\xi}_t(\hat{\theta})\left(\hat{\xi}_{t-j}(\hat{\theta}) - \bar{\xi}\right) + \frac{1}{T\hat{s}^2}\sum_{t=j+1}^{T}\dot{\xi}_{t-j}(\hat{\theta})\left(\hat{\xi}_t(\hat{\theta}) - \bar{\xi}\right)$$
$$\dot{\xi}_t(\theta) = \frac{\partial}{\partial\theta}\xi_t(\theta)$$

Then, let $\tilde{\xi}^m = \left(\hat{A}^m\right)^{-1/2}\hat{\zeta}^m$ and $\tilde{\rho}^m = \left(\hat{A}^m\right)^{-1/2}\hat{\rho}^m$. Finally, let

$$\bar{\rho}^m(j) = \frac{\breve{\rho}^m(j)}{\breve{s}^m(j)}$$
$$\breve{\rho}^m(j) = \tilde{\rho}^m(j) - \tilde{\zeta}(j)'\left(\sum_{k=j+1}^{m}\tilde{\zeta}(k)\tilde{\zeta}(k)'\right)^{-1}\sum_{k=j+1}^{m}\tilde{\zeta}(k)\tilde{\rho}^m(j)$$
$$\breve{s}^m(j)^2 = 1 + \tilde{\zeta}(j)'\left(\sum_{k=j+1}^{m}\tilde{\zeta}(k)\tilde{\zeta}(k)'\right)^{-1}\sum_{k=j+1}^{m}\tilde{\zeta}(k)$$

Delgado and Velasco (2011) show that

$$\bar{\rho}^m \xrightarrow{d} N(0, I_{m-d}) \tag{1.20}$$

where $d = \dim(\theta)$, $\bar{\rho}^m = (\bar{\rho}(1), \ldots, \bar{\rho}(m-d))'$ and $\xrightarrow{d}$ denotes convergence in distribution. Notice that the projections sacrifice d degrees of freedom, so that only the first $m-d$ can be transformed.

Based on (1.20), we estimate the weighted correlogram in (1.19) by

$$\bar{C}(q) = \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right) \bar{\rho}^{q-1+d}(j).$$

Because of the degrees of freedom sacrificed in the projections, we must estimate $q-1+d$ autocorrelations in order to transform the first $q-1$ autocorrelations. It follows from (1.20) that

$$\bar{C}(q) \xrightarrow{d} N\left(0, \left[\sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right)^2\right]\right)$$

under the MDS null.

1.3.2 Non-linear predictability

The weighted correlogram statistic is a function of the sample autocorrelations of $\hat{\xi}_t = \xi_t(\hat{\theta})$ and therefore does not exploit the full hypothesised MDS structure of $\xi_t = \xi_t(\theta)$. In particular it neglects non-linear predictability. We test for non-linear predictability using Linton and Whang's (2007) quantilogram, which is based on the correlation of quantile hits. If ξ_t is MDS, the probability ξ_{t+k} is in the α quantile given ξ_t is in the α quantile should remain α. The quantile hits are uncorrelated. The quantilogram is a more general version of Wright's (2000) sign tests, which focus on whichever quantile zero is in.

In our test statistic, we weight the quantilogram estimates analogously to the variance ratios. This gives

$$\widehat{W}_\alpha(q) = \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right) \hat{\rho}_\alpha(j), \tag{1.21}$$

where

$$\begin{aligned}
\hat{\rho}_\alpha(j) &= \frac{\sum_{t=1}^{T-j} \psi_\alpha(\hat{\xi}_t - \hat{\mu}_\alpha)\psi_\alpha(\hat{\xi}_{t+j} - \hat{\mu}_\alpha)}{\sqrt{\sum_{t=1}^{T-j} \psi_\alpha^2(\hat{\xi}_t - \hat{\mu}_\alpha)}\sqrt{\sum_{t=1}^{T-j} \psi_\alpha^2(\hat{\xi}_{t+j} - \hat{\mu}_\alpha)}} \\
\psi_\alpha(\cdot) &= \alpha - 1(\cdot < 0) \\
\hat{\mu}_\alpha &= \operatorname*{argmin}_{m\in\mathbb{R}} \sum_{t=1}^{T} (\hat{\xi}_t - m) \times \psi_\alpha(\hat{\xi}_t - m).
\end{aligned}$$

and $1(\cdot)$ is the indicator function. We evaluate (1.21) over the same q as in the correlograms and over a range of both extreme and moderate quantiles, namely $\alpha \in \{0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.99\}$.

We use a wild bootstrap for inference. This allows us to account for the estimation step involved in constructing $\hat{\xi}_t$. $\hat{\xi}_t$ is pre-multiplied by ι_t^* at each t, where $\mathrm{E}(\iota_t^*) = 0$ and $\mathrm{Var}(\iota_t^*) = 1$. We use Mammen's (1993) two-point distribution for ι_t^*.[5] Then, we use the bootstrapped residuals to extract a pseudo-sample of returns $r_{m,t}^*$ by the relationship

$$r_{m,t}^* = \mu_t(\hat{\theta}) + \iota_t^* \hat{\xi}_t.$$

We use $r_{m,t}^*$ to generate a new series for the market value and therefore obtain the pseudo-sample of the log price-dividend ratio, $z_{m,t}^*$. We then re-estimate the asset pricing model parameters using the modified data, generating a pseudo-sample of expected returns and thus a (new) pseudo-sample of residuals.

The empirical distribution of the weighted quantilograms thus obtained is used for inference and the bootstrap procedure is repeated 200 times.[6] Notice that our procedure conditions on consumption and dividends.

1.3.3 Hong-Lee generalised spectral test

The Hong and Lee (2005) generalised spectral test can detect both linear and non-linear predictability. We add it to our battery of MDS tests because the known low power problems of MDS tests (Poterba and Summers, 1988) mean it is useful to have additional tests. The test is based on the Hong (1999) generalised spectrum, corrected for the estimation of the parameters of the residual series in a way that yields a test statistic which has a nuisance parameter-free limiting distribution.

The test statistic is

$$\widehat{G}(q) = \frac{\sum_{j=1}^{q-1}\left(1-\frac{j}{q}\right)^2 (T-j)\int_{-3}^{3} |\hat{\varsigma}_j^{(1,0)}(0,v)|^2 dW(v) - \widehat{D}(q)}{\sqrt{\widehat{E}(q)}}$$

[5] ι_t^* is *iid* through time and has probability mass function

$$f_I(\iota_t^*) = \begin{cases} \frac{\sqrt{5}+1}{2\sqrt{5}}, & \iota_t^* = \frac{1-\sqrt{5}}{2} \\ \frac{\sqrt{5}-1}{2\sqrt{5}}, & \iota_t^* = \frac{1+\sqrt{5}}{2} \end{cases}$$

[6] While 200 repetitions is a fairly low number, we are constrained by computational power in our ability to do more since the simulations for the Campbell-Cochrane expected returns each involve 200 repetitions themselves at each point in time in each bootstrap repetition.

where

$$\widehat{D}(q) = \sum_{j=1}^{q-1} \left(1 - \frac{j}{q}\right)^2 \frac{1}{T-j} \sum_{t=j+1}^{T-1} \hat{\xi}_t^2 \int_{-3}^{3} |\hat{\pi}_{t-j}(v)|^2 dW(v)$$

$$\widehat{E}(q) = 2 \sum_{j=1}^{T-2} \sum_{k=1}^{T-2} \left(1 - \frac{j}{q}\right)^2 \left(1 - \frac{k}{q}\right)^2 \int_{-3}^{3} \int_{-3}^{3} \left| \frac{1}{T - \max\{j,k\}} \right.$$
$$\left. \times \sum_{t=\max\{j,k\}+1}^{T} \hat{\xi}_t^2 \hat{\pi}_{t-j}(v) \hat{\pi}_{t-k}(v') \right|^2 dW(v) dW(v')$$

$W(\cdot)$ is the standard Normal distribution truncated on the interval $[-3,3]$, $\hat{\pi}(v) = e^{iv\hat{\xi}_t} - T^{-1}\sum_{t=1}^{T} e^{iv\hat{\xi}_t}$, $i = \sqrt{-1}$, and

$$\hat{\varsigma}_j^{(1,0)}(0,v) = \frac{\partial}{\partial u} \hat{\varsigma}_j(u,v)|_{u=0}$$
$$\hat{\varsigma}_j(u,v) = \hat{\varpi}_j(u,v) - \hat{\varpi}_j(u,0)\hat{\varpi}_j(0,v)$$
$$\hat{\varpi}_j(u,v) = \frac{1}{T-|j|} \sum_{t=|j|+1}^{T} e^{iu\hat{\xi}_t + iv\hat{\xi}_{t-|j|}}.$$

Under the MDS null and the technical conditions laid out in Hong and Lee (2005, p.p. 509-510), Hong and Lee show that

$$\widehat{G}(q) \xrightarrow{d} N(0,1).$$

1.3.4 Rescaled range

We also consider a rescaled range test. We do so as the rescaled range can be more powerful than other MDS tests in the presence of long-range dependence (Lo, 1991). The rescaled range is

$$\widehat{Q} = \frac{1}{\hat{s}\sqrt{T}} \left[\max_{k \leq j \leq T} \sum_{t=k}^{j} (\xi_t(\hat{\theta}) - \bar{\xi}) - \min_{k \leq j \leq T} \sum_{t=1}^{j} (\xi_t(\hat{\theta}) - \bar{\xi}) \right].$$

$\hat{s}^2$ is a consistent estimator of $\mathrm{Var}(\xi_t(\theta))$. Given the issue of the estimation of $\hat{\theta}$ distorting the limiting distribution of the statistic, we conduct inference using the same wild bootstrap procedure as for the quantilogram.

1.3.5 Maximal predictability

Huang and Zhou (2017) develop a Wald test of whether the predictability of excess market returns, $\tilde{r}_{m,t+1} = r_{m,t+1} - r_{f,t+1}$, is too large. Predictability is measured with respect to a forecasting variable, f_t. "Too large" is defined as too large to be consistent with $\widetilde{M}_t$, the stochastic discount factor (SDF) normalised such that $\mathrm{E}\widetilde{M}_{t+1} = 1$, being a function of a given set of state variables ω_t.[7] The test is semi-parametric in that the functional form of the SDF need not be known. The Wald statistic tests whether theoretical upper bound on R^2 implied by the state variables is exceeded by the empirical R^2 from the univariate one-step-ahead predictive regression of $\tilde{r}_{t+1}$ on f_t.

It is straightforward to verify that this test applies almost directly to the q-step-ahead predictive regression

$$\tilde{r}_{t+q} = \alpha + \beta f_t + \varepsilon_{t+q} .$$

In this context, when bounding R^2 with $\mathrm{SR}(r_m)$, the market Sharpe ratio, the bound becomes

$$R^2 \leq \bar{R}^2 = \phi^2_{\omega,rf} h^2 \mathrm{SR}^2(r_m),$$

where

$$\phi^2_{\omega,rf} = \rho^2_{\omega,rf} \frac{\mathrm{Var}[\tilde{r}_{t+q}(\tilde{r}_t - \mu_f)]}{\mathrm{Var}(\tilde{r}_{t+q})\,\mathrm{Var}(f_t)}$$

$$\rho^2_{\omega,rf} = \frac{\mathrm{Cov}[\omega_{t+q}, \tilde{r}_{t+q}(f_t - \mu_f)]'\,\mathrm{Var}^{-1}(\omega_{t+q})\,\mathrm{Cov}[\omega_{t+q}, \tilde{r}_{t+q}(f_t - \mu_f)]}{\mathrm{Var}[\tilde{r}_{t+q}(f_t - \mu_f)]},$$

and $\mu_f = \mathrm{E}(f_t)$. h is a parameter chosen by the marginal investor. We follow Cochrane and Saá-Requejo (2000) in using $h = 2$. This bound requires ω to have an elliptical distribution, which it does in all models.[8]

Huang and Zhou's (2017) test exploits the asymptotic normality of standard estimators of the mean and covariance matrix of $(r_{t+q}, f_t, r_{t+q} f_t, \omega'_{t+q})'$. These means and covariances, which comprise θ_{SR}, are all that is required to calculate the empirical R^2 and its bound. We follow Huang and Zhou and estimate θ_{SR} by GMM.

Testing whether R^2 exceeds $\bar{R}^2$ is equivalent to a one-sided test of the null $f(\theta_{SR}) \equiv R^2 - \bar{R}^2 = 0$ against the alternative that $f(\theta_{SR}) \equiv R^2 - \bar{R}^2 > 0$ (Huang and Zhou, 2017). The

[7] Our other tests relate to actual, not excess returns.

[8] The state variables for the Bansal-Yaron and Campbell-Cochrane models are conditionally lognormal, and the Cecchetti-Lam-Mark state variable has a binomial distribution.

Wald statistic for this test is

$$W_{RA} = Tf(\hat{\theta}_{SR})\left[\frac{df}{d\theta_{SR}}\mathrm{Var}(\hat{\theta}_{SR})\frac{df}{d\theta_{SR}}\right]^{-1} f(\hat{\theta}_{SR}) \xrightarrow{d} \chi^2(1).$$

This procedure can then be applied to the predictive regression Fama and French (1988) use to test for serial correlation in the market return

$$\tilde{r}_{m,t+q}(q) = \alpha_q + \beta_q \tilde{r}_{t,m}(q) + \varepsilon_{t+q}, \tag{1.22}$$

albeit, with the regression specified in terms of excess, rather than actual, returns.

For the Campbell-Cochrane and Cecchetti-Lam-Mark models, this test requires us to condition on our estimated state variables. The state variable for the Campbell-Cochrane model is s_t, which we extract as explained in Section 1.2.2. The state variable for the Cecchetti-Lam-Mark model is y_t, which we extract by estimating the Markov-switching model for consumption and taking $y_t = 1$ if the estimated smoothed probability $\Pr(y_t = 1|\mathscr{F}_{t+1}) \geq \frac{1}{2}$, where $\mathscr{F}_t$ is information available at t. The state variables for the Bansal-Yaron model are Δc_t, x_t and σ_t^2. Since we extract x_t and σ_t^2 as a linear function of $r_{f,t}$ and $z_{m,t}$, we take Δc_t, $r_{f,t}$ and $z_{m,t}$ to be the three Bansal-Yaron state variables, so that the results are not dependent on the estimation of the model.

1.4 Data

Data for our main results are from the US from 1930 to 2016. The time period is annual and, as is standard in the asset pricing literature, the agent's decision interval is assumed to be the time horizon considered. We consider results are robust to using quarterly data and a quarterly decision interval instead as a robustness check (see Section 1.6.3).

The market index is the value-weighted CRSP index, obtained from WRDS. The risk-free rate is the US one-month Treasury bill, from Ibbotson Associates via French's website. The set of assets used to estimate the asset pricing models also includes the six double-sorted size/book-to-market portfolios from Ken French's website. In our robustness checks, we consider replacing the six double-sorted size/book-to-market portfolios with the five industry portfolios, also from Ken French's website, in the estimation of the models (see Section 1.6.2).

Consumption is seasonally adjusted per-capita non-durables and services personal consumption expenditures from the BEA. We deflate nominal data by the BEA's consumption deflator. Table 1.1 summarises the data.

Table 1.1 Data summary statistics

	Mean	Median	Std dev	SC(1)
r_m	0.063	0.105	0.194	-0.024
r_f	0.005	0.008	0.037	0.762
Δc	0.020	0.023	0.022	0.466
Δd	0.017	0.023	0.111	0.192
z_m	3.409	3.393	0.455	0.885

Descriptive statistics for our key variables at the annual frequency over the period 1930-2016. r_m denotes the log market return, r_f the quarterly log risk-free rate (the rolled over 1 month US T-bill), Δc log consumption growth, Δd log dividend growth and z_m the log price-dividend ratio.

Table 1.2 Bansal-Yaron model estimates

μ_c	μ_d	ϕ	φ	ρ_x	ψ_x	σ	ν	σ_w	δ	ψ	γ
0.020	0.035	3.499	5.697	1.320	0.810	0.008	0.207	-0.006	0.929	2.310	7.755
(0.002)	(0.008)	(0.348)	(1.702)	(0.161)	(0.289)	(0.002)	(0.503)	(0.005)	(0.000)	(0.324)	(0.065)
0.000	0.009	0.000	0.000	0.000	0.014	0.000	0.511	0.357	0.000	0.000	0.000
J-stat	121.7	p-value	0.000								

Estimates of the Bansal-Yaron model parameters using annual US data 1930-2016. Point estimates are displayed in the first row, standard errors (in parentheses) in the second and p-values in the third. All p-values are asymptotic.

1.4.1 Model estimation

Our main results relate to when the asset pricing models are estimated at the annual frequency where the set of assets used to estimate the Euler equations comprises the market return, the risk-free rate and the size double-sorted size/book-to-market portfolios and we use the optimal weight matrix in GMM estimation. This is the specification that gives the most reasonable expected returns series across the board (Section 1.6 gives details of the residuals for other specifications; because actual returns are the sum of the expected return and the residual, only models with reasonable residual series will have reasonable expected returns).

Many of the other specifications do not give reasonable expected returns series. We look only at specifications where the expected returns are plausible. As much as our focus is on the dynamics of returns, rather than the levels, the first and second moments are related. Serial correlation (a centred second moment) depends on the first moment. But, even if we only used uncentred second moments, there is no reason to think that a model that fails to fit the first moment would fit the second. Even if it did, it would be of little practical relevance for pricing assets. While we focus on the specification that generally gives the most reasonable expected returns, our results are robust to considering other specifications giving reasonable expected returns.

Table 1.3 Campbell-Cochrane model estimates

$\bar{g}$	$\text{Var}(\Delta c)$	$\text{Var}(\Delta d)$	$\text{Cov}(\Delta c, \Delta d)$	α	ϕ	δ	γ
0.021	4.07×10^{-4}	0.012	0.001	0.424	0.879	0.926	10^{-7}
(0.003)	(1.68×10^{-4})	(0.004)	(6.93×10^{-4})	(0.173)	(0.050)	(0.016)	(0.322)
0.000	0.015	0.001	0.070	0.017	0.000	0.000	1.000
J-stat	0.068			R^2	0.783	J-stat	38.37
p-value	0.795					p-value	3.18×10^{-7}

Estimates of the Campbell-Cochrane model parameters using annual US data 1930-2016. Each panel (set of columns) refers to a separate estimation. The estimates of δ and γ, and the associated p-values, condition on the estimates in the first two panels. Point estimates are displayed in the first row, standard errors (in parentheses) in the second and p-values in the third. All p-values are asymptotic.

Table 1.4 Cecchetti-Lam-Mark model estimates

(a) Consumption model

α_0	α_1	p	q	σ_0^2	σ_1^2
0.023	-0.016	0.956	0.876	0.012	0.040

(b) Preference parameters

δ	γ
0.966	2.431
(0.290)	(15.38)
0.001	0.874
J-stat	37.18
p-value	6×10^{-7}

Estimates of the Cecchetti-Lam-Mark model parameters, estimated using annual US data 1930-2016. Panel (a) presents point estimates only. In panel (b), point estimates are displayed in the first row, standard errors (in parentheses) in the second and p-values in the third. All p-values are asymptotic.

Table 1.2 suggests the Bansal-Yaron model is mis-specified. The J-statistic has a vanishingly small p-value. Worryingly, the long-run risk is estimated to be non-stationary ($\hat{\rho}_x > 1$), although this could simply be a function of more general model mis-specification. It is unsurprising that the resulting expected returns do not form a plausible financial time series. By extension, the model residuals do not form a plausible financial time series, either, as shown in Table 1.5. It is nonetheless interesting to test whether the Bansal-Yaron residuals are predictable. Since the model is so clearly mis-specified, it is worth checking that the residual-based tests do in fact reject it.

As we see in Table 1.3, there was some difficulty in estimating the Campbell-Cochrane Euler equations. In order to generate an s_t series, we constrain the estimate of γ to be no less

Table 1.5 Properties of $\hat{\xi}_t$

Model	Mean	Median	Std dev	SC(1)
Bansal-Yaron	-1248	-1248	1.694	0.816
Campbell-Cochrane	0.007	-0.012	0.210	-0.115
Cecchetti-Lam-Mark	-0.017	0.014	0.191	-0.080

Summary statistics for the model-implied ex-ante residuals. "Std dev" denotes standard deviation and "SC(1)" first-order serial correlation. The models are estimated and residuals computed using annual US data over the period 1930-2016.

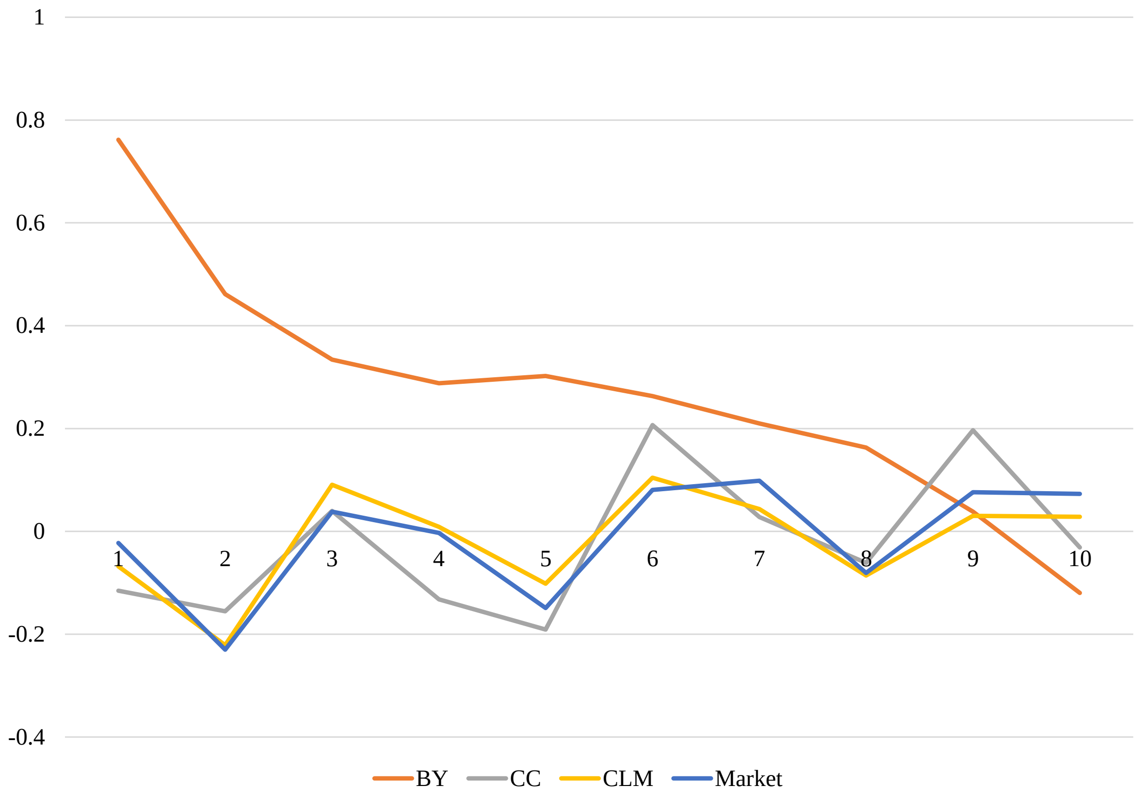
than 10^{-7} and this constraint binds. Not imposing this constraint gives $\hat{\gamma} = -0.078$ with a standard error of 0.753, so the estimates are not very different relative to their standard errors. The subjective discount factor is significantly less than one. The J-test rejects the model's Euler equations. Nonetheless, this is only indicative of how well specified the Euler equations are. The Euler equation estimation conditions on earlier estimates of time-series parameters ($\bar{g}$, $\mathrm{Var}(\Delta c)$, $\mathrm{Var}(\Delta d)$, $\mathrm{Cov}(\Delta c, \Delta d)$, α and ϕ), yet the over-identification test in the third panel of Table 1.3 does not account for this estimation. We cannot firmly reject the model on this basis. Table 1.5 shows that the mean residual is close to zero, just 0.7%. The Campbell-Cochrane model therefore seems to give reasonable expected returns, despite the issue of the estimation constraint binding.

Table 1.4 shows that the Cecchetti-Lam-Mark model preference parameter estimates are also generally reasonable. The subjective discount factor is less than one and the utility curvature greater than zero. The Euler equations are rejected by the J-test, but this test does not enforce the Markov-switching structure on consumption growth. Enforcing this structure may still yield reasonable expected returns. Table 1.5 suggests this is indeed the case. The mean residual for the Cecchetti-Lam-Mark model is fairly low at around -1.7% a year.

Figure 1.1 shows the autocorrelation functions of the observed market return and the model-implied ex-ante expected returns. This graph is only indicative. We must be mindful of the distortions in the model-implied autocorrelation functions induced by parameter estimation. In the graph, the Bansal-Yaron is a long way from matching the market autocorrelation function. The Campbell-Cochrane and Cecchetti-Lam-Mark model expected return autocorrelations are fairly close to the observed market autocorrelations.

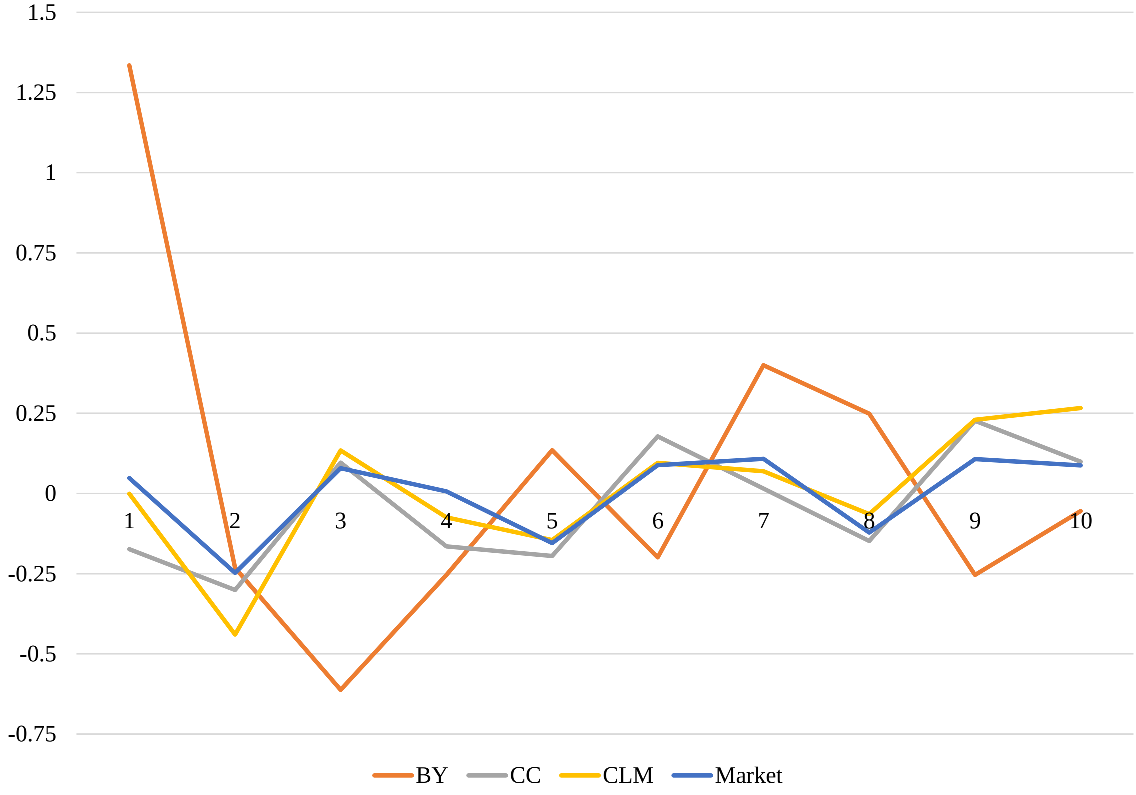
To remove the effect of estimation in the autocorrelations of the expected returns, we can apply the Delgado and Velasco (2011) procedure to them. Note that the Delgado and Velasco procedure transforms the standardised autocorrelations $\tilde{\rho}^m = (\hat{A}^m)^{1/2}$. It transforms the autocorrelations divided by their standard errors. So, in order to see the effect of the transformation, we need to consider the (untransformed) standardised autocorrelations and the

Fig. 1.1 Market and model autocorrelation functions

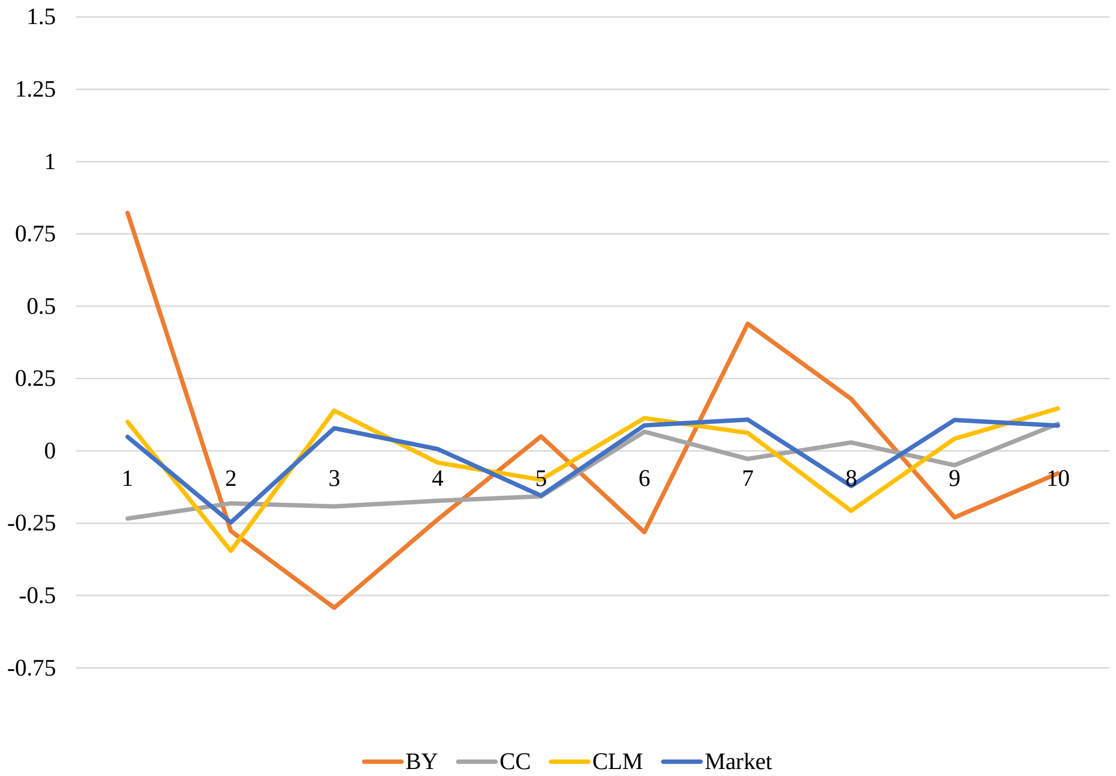


Autocorrelation functions for the market return and the model-implied ex-ante expected returns. Serial correlation is computed up to lag 10. The models are estimated and expected returns computed over 1930-2016. These estimates of the model-implied autocorrelation functions are biased due to the estimation of the parameters of the expected returns and it is therefore difficult to draw many firm conclusions from this figure, which is provided for illustrative purposes only.

Fig. 1.2 Market and model standardised autocorrelation functions



(a) Unadjusted standardised autocorrelation function ($\tilde{\rho}$)



(b) Adjusted standardised autocorrelation function ($\bar{\rho}$)

Transformed and untransformed standardised autocorrelation for the model-implied ex-ante expected returns compared to (untransformed) standardised autocorrelation for the market. Serial correlation is computed up to lag 10. The models are estimated and expected returns computed over 1930-2016.

transformed standardised autocorrelations. These are shown in panel (a) of Figure 1.2, where $m = 10 + d$ for each model. Panel (b) shows the transformed standardised autocorrelations, $\bar{\rho}^m$. The market autocorrelation in panel (b) remains $\tilde{\rho}^m$, since there is no adjustment needed.

We can see that, both with and without the Delgado and Velasco (2011) adjustment, the Bansal-Yaron model's standardised autocorrelations are the furthest from the market's, which is not a great surprise given the mis-specification of the model. Oddly, the Campbell-Cochrane standardised autocorrelations appear to be closer to that of the market before applying the adjustment. This would imply that the bias in the autocorrelation function of the Campbell-Cochrane expected returns arising from the estimation of the model parameters was making the Campbell-Cochrane autocorrelations artificially close to the market's autocorrelations. The adjustment does not appear to impact how close the Cecchetti-Lam-Mark autocorrelations are to the market autocorrelations: they seem to be close in both cases.

1.5 Serial dependence in the model residuals

Our results for the Bansal-Yaron model are in Table 1.6. Unsurprisingly but reassuringly, given how poorly the model performs in terms of the levels of returns, the quantilogram, Hong-Lee and rescaled range tests resoundingly reject the null that the residuals are MDS. Not only can the Bansal-Yaron model not explain mean returns (the first moment), it cannot explain return dynamics (the second moment) either. Curiously, the correlogram does not reject the MDS null at any lag.

Moreover, the maximal predictability results suggest that the Bansal-Yaron state variables do not explain observed predictability, either. Changing the functional form of the SDF would not enable a model based on the Bansal-Yaron state variables to explain the dynamics of returns. There are extremely significant exceedences of the R^2 bound, $\bar{R}^2$, at four horizons: four, five, six and seven years.

However, we express some caution regarding these results for two reasons. First, $\bar{R}^2$ is, for the Bansal-Yaron model, almost always either less than zero or greater than one for the holding periods considered. So either any degree of predictability is consistent with consumption growth, the long-run risk and time-varying economic volatility being risk factors in the stochastic discount factor or no predictability is consistent with these risk factors. Second, the parameters of R^2 and $\bar{R}^2$ are jointly estimated using GMM. The R^2 does not come directly from a regression themselves. The methods ought to be equivalent but it is not computationally possible to satisfy the moment conditions exactly here, despite the system being exactly identified. Therefore the methods are not equivalent in a finite sample. Because of this, the

Table 1.6 Bansal-Yaron model results

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	-0.024	-0.041	-0.039	-0.037	0.001	-0.063	-0.100	-0.124	-0.099
(Std Err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.660	0.606	0.698	0.751	0.996	0.666	0.532	0.468	0.586

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.013	-0.016	-0.02	-0.024	-0.028	-0.032
	0.59	0.59	0.59	0.59	0.60	0.61	0.63	0.63	0.63
0.05	0.219	0.33	0.375	0.456	0.557	0.644	0.704	0.745	0.772
	0.12	0.16	0.22	0.20	0.20	0.18	0.17	0.17	0.17
0.1	0.247	0.425	0.587	0.782	1.005	1.205	1.374	1.508	1.631
	0.01	0.04	0.03	0.03	0.02	0.01	0.01	0.00	0.00
0.25	0.366	0.678	0.945	1.180	1.391	1.590	1.771	1.935	2.079
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.5	0.375	0.725	1.043	1.323	1.574	1.801	1.994	2.160	2.293
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.75	0.354	0.611	0.823	1.033	1.235	1.416	1.574	1.717	1.844
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.9	0.363	0.648	0.887	1.081	1.234	1.345	1.414	1.455	1.477
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	0.282	0.402	0.453	0.476	0.485	0.487	0.483	0.476	0.466
	0.00	0.00	0.01	0.04	0.04	0.11	0.11	0.11	0.11
0.99	-0.009	-0.018	-0.024	-0.030	-0.036	-0.041	-0.047	-0.052	-0.057
	0.06	0.33	0.42	0.42	0.42	0.42	0.42	0.42	0.37

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	85.14	85.15	85.13	85.09	85.03	84.95	84.87	84.78	84.68
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

(d) Rescaled range

$\widehat{Q}$	2.926
p-value	0.00

Panels (a)-(d) report tests of the MDS null for the Bansal-Yaron residuals, over the period 1930-2016. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 1.6 Bansal-Yaron model results
(e) Maximal predictability

q	2	3	4	5	6	7	8	9	10
R^2	0.100	0.044	0.462	0.002	0.038	0.043	0.125	0.028	0.000
$\bar{R}^2$	21.68	0.122	-12.29	-59.18	-0.725	-2.738	1993	67.57	1.058
Wald stat	-	-	42.19	54.06	50.13	5.156	-	-	-
p-value	-	-	0.000	0.000	0.000	0.023	-	-	-

Panel (e) reports tests of the null that the market return is no more predictable than implied by the Bansal-Yaron model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1930-2016. The Wald statistic and its asymptotic p-value are reported.

reported R^2 for the predictive regression for a given horizon is not the same for the Bansal-Yaron model as it is for the Campbell-Cochrane and Cecchetti-Lam-Mark models, even though it should be. These discrepancies highlight the numerical challenges of the GMM estimation undertaken to compute the tests. However, these numerical issues do not affect the maximal predictability tests for the Campbell-Cochrane or Cecchetti-Lam-Mark models so may simply be a further reflection of the mis-specification of the Bansal-Yaron state variables. Overall, the best available evidence is that the state variables of the Bansal-Yaron model cannot explain the predictability of market returns.

Our main results regarding the Campbell-Cochrane model are in Table 1.7. We reject the null that the Campbell-Cochrane residuals are MDS: the correlogram rejects the MDS null at all lags considered. However, the Hong-Lee test and rescaled range provide no rejections and only three of the 81 quantilograms reject the MDS null at the 10% level. This shows the benefits of using a battery of test statistics: the correlogram lacked power against the specific alternative characterising the Bansal-Yaron residuals but has it against the alternative characterising the the Campbell-Cochrane residuals.

Turning to our maximal predictability tests, there are three exceedences of the R^2 bound, only one of which is significant. On the basis of annual data, it therefore appears possible that a model based on the surplus consumption state variable but with a different functional form of the SDF could explain the dynamics of returns. This conclusion, however, is not robust to using quarterly data (see Section 1.6.3).

Table 1.8 shows the results for the Cecchetti-Lam-Mark model. The residuals are clearly not MDS. The correlogram rejects the MDS null from $q = 5$ onwards and the rescaled range also rejects the MDS null. Both rejections suggest negative serial dependence: that higher values are followed by lower ones. Neither the quantilogram nor the Hong-Lee tests provide

Table 1.7 Campbell-Cochrane model results

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	-0.117	-0.170	-0.196	-0.239	-0.304	-0.341	-0.430	-0.442	-0.452
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.030	0.035	0.052	0.043	0.022	0.020	0.007	0.010	0.013

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.013	-0.016	-0.020	-0.024	-0.028	0.004
	0.02	0.09	0.21	0.25	0.45	0.45	0.51	0.52	0.47
0.05	-0.014	-0.028	-0.043	-0.058	-0.073	-0.042	-0.001	0.029	0.052
	0.47	0.83	0.88	0.92	0.97	1.00	0.96	0.98	0.7
0.1	0.017	-0.012	-0.052	-0.070	-0.101	-0.099	-0.092	-0.098	-0.113
	0.67	0.81	0.74	0.76	0.82	0.86	0.81	0.82	0.97
0.25	0.026	-0.020	-0.086	-0.140	-0.183	-0.217	-0.234	-0.230	-0.212
	0.67	0.56	0.61	0.67	0.76	0.78	0.78	0.76	0.95
0.5	-0.006	-0.080	-0.114	-0.150	-0.179	-0.186	-0.190	-0.201	-0.198
	0.80	0.79	0.85	0.85	0.87	0.92	0.92	0.94	0.92
0.75	-0.059	-0.120	-0.132	-0.141	-0.157	-0.150	-0.141	-0.143	-0.142
	0.77	0.64	0.76	0.81	0.86	0.87	0.93	0.90	0.98
0.9	-0.064	-0.088	-0.129	-0.137	-0.163	-0.160	-0.173	-0.197	-0.222
	0.82	0.89	0.93	0.90	0.90	0.88	0.87	0.86	0.99
0.95	-0.030	0.024	0.038	0.037	0.029	0.015	-0.001	-0.020	-0.040
	0.42	0.97	0.98	0.99	0.95	0.88	0.91	0.91	0.85
0.99	-0.009	-0.018	-0.026	-0.035	-0.044	-0.053	-0.062	-0.071	-0.080
	0.01	0.23	0.28	0.31	0.35	0.44	0.48	0.53	0.48

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.381	0.386	0.388	0.386	0.383	0.379	0.373	0.365	0.345
p-value	0.703	0.699	0.698	0.699	0.702	0.705	0.709	0.715	0.730

(d) Rescaled range

$\widehat{Q}$	0.911
p-value	0.28

Panels (a)-(d) report tests of the MDS null for the Campbell-Cochrane residuals, over the period 1930-2016. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 1.7 Campbell-Cochrane model results
(e) Maximal predictability

q	2	3	4	5	6	7	8	9	10
R^2	0.057	0.022	0.012	0.000	0.029	0.035	0.039	0.020	0.002
$\bar{R}^2$	0.092	0.000	0.004	0.074	0.079	0.009	0.057	0.085	0.076
Wald stat	-	27.35	0.604	-	-	2.194	-	-	-
p-value	-	0.000	0.437	-	-	0.139	-	-	-

Panel (e) reports tests of the null that the market return is no more predictable than implied by the Campbell-Cochrane model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1930-2016. The Wald statistic and its asymptotic p-value are reported.

any rejections of the MDS null. This serves to further illustrate the power issues of MDS tests and justify our approach of considering multiple different tests.

Considering the maximal predictability tests, we see only one significant exceedence of the R^2 bound, which is at $q = 2$. Annual data therefore suggests that an SDF based on the good/bad state indicator variable but with a different functional form may be able to explain the dynamics of the market return. Again, though, this conclusion is not robust to using quarterly data (see Section 1.6.3).

1.6 Robustness

We consider the robustness of our results to (i) using the identity weight matrix in GMM estimation rather than the optimal weight matrix, (ii) using the five Fama-French industry portfolios in place of the six Fama-French size/value portfolios when estimating the asset pricing models and (iii) using quarterly data instead of annual data. Overall, we find that, even where the models produce reasonable residual and expected returns series, they cannot explain return dynamics.

In terms of whether the state variables can explain return dynamics, in the sense that the R^2 of the predictive regressions does not exceed its theoretical upper bound, the finding that the Bansal-Yaron state variable cannot explain the own-history predictability of returns is robust to using quarterly data. The finding that the Campbell-Cochrane state variable may be able to explain the predictability of returns is not robust to using quarterly data. The finding that the Cecchetti-Lam-Mark model may be able to explain the predictability of returns survives switching to quarterly data in the whole sample, but this finding is not robust over time. When we split the sample period into two equal-length sub-samples, we get many more significant R^2 bound exceedences in both sub-samples than in the whole sample.

Table 1.8 Cecchetti-Lam-Mark model results

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	0.043	-0.087	-0.201	-0.275	-0.342	-0.361	-0.428	-0.455	-0.526
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.425	0.278	0.047	0.020	0.010	0.014	0.007	0.008	0.004

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.009	0.018	0.080	0.165	0.262	0.359	0.448	0.526	0.587
	0.39	0.77	0.98	0.78	0.74	0.67	0.67	0.66	0.71
0.05	-0.006	-0.017	-0.036	-0.053	-0.070	-0.090	-0.104	-0.116	-0.126
	0.72	0.78	0.72	0.66	0.61	0.57	0.57	0.58	0.47
0.1	-0.004	-0.014	-0.030	-0.047	-0.063	-0.084	-0.099	-0.116	-0.132
	0.76	0.67	0.62	0.60	0.53	0.51	0.47	0.45	0.39
0.25	0.015	0.004	-0.005	-0.016	-0.026	-0.039	-0.053	-0.072	-0.094
	0.86	0.75	0.70	0.67	0.60	0.57	0.53	0.48	0.42
0.5	0.031	0.028	0.032	0.035	0.029	0.016	0.000	-0.021	-0.045
	0.99	0.95	0.87	0.85	0.78	0.73	0.72	0.70	0.59
0.75	0.089	0.115	0.147	0.170	0.179	0.178	0.167	0.145	0.120
	0.97	0.87	0.87	0.85	0.81	0.77	0.77	0.73	0.66
0.9	0.069	0.081	0.082	0.082	0.076	0.063	0.049	0.030	0.009
	0.98	0.92	0.88	0.86	0.86	0.80	0.76	0.73	0.60
0.95	0.009	0.009	-0.004	-0.018	-0.035	-0.055	-0.075	-0.098	-0.119
	0.91	0.84	0.75	0.74	0.69	0.67	0.64	0.63	0.49
0.99	-0.009	-0.025	-0.051	-0.100	-0.170	-0.259	-0.367	-0.495	-0.650
	0.41	0.45	0.47	0.49	0.52	0.51	0.52	0.53	0.48

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.884	0.897	0.904	0.907	0.907	0.903	0.894	0.882	0.867
p-value	0.377	0.370	0.366	0.364	0.364	0.367	0.371	0.378	0.386

(d) Rescaled range

$\widehat{Q}$	0.698
p-value	0.02

Panels (a)-(d) report tests of the MDS null for the Cecchetti-Lam-Mark model residuals, estimated over the period 1930-2016. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 1.8 Cecchetti-Lam-Mark model results
(e) Maximal predictability

q	2	3	4	5	6	7	8	9	10
R^2	0.057	0.022	0.012	0.000	0.029	0.035	0.039	0.020	0.002
$\bar{R}^2$	0.006	0.026	0.033	0.160	0.194	0.026	0.050	0.362	0.346
Wald stat	134.1	-	-	-	-	0.189	-	-	-
p-value	0.000	-	-	-	-	0.664	-	-	-

Panel (e) reports tests of the null that the market return is no more predictable than implied by the Cecchetti-Lam-Mark model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1930-2016. The Wald statistic and its asymptotic p-value are reported.

Table 1.9 Properties of $\hat{\xi}_t$ - Identity matrix

Model	Mean	Median	Std dev	SC(1)
Bansal-Yaron	-13905871	-13905873	31.21	0.762
Campbell-Cochrane	-0.204	-0.228	0.206	-0.126
Cecchetti-Lam-Mark	-0.038	-0.040	0.191	-0.088

Summary statistics for the model-implied ex-ante residuals. “Std dev” denotes standard deviation and “SC(1)” first-order serial correlation. The models are estimated and residuals computed using annual US data over the period 1930-2016.

We consider the robustness of the residual-based tests (i.e. the correlogram, quantilogram, Hong-Lee tests and rescaled range) only in scenarios where the model provides credible residuals, and therefore credible expected returns. There is no point checking the second moment of a model that fits poorly in terms of the first moment, as one would not use it to price assets anyway. Moreover, the centred second moment (e.g. serial correlation coefficient) is a function of the first moment.

For the robustness of the maximal predictability test, note that the state variables in the Bansal-Yaron and Cecchetti-Lam-Mark are independent of the asset sets or GMM weighting matrices used. As such, the maximal predictability results for these models depend only on the data frequency and sample period. The extraction of the the Campbell-Cochrane state variable depends on, amongst other things, the estimated utility curvature. Therefore, the (estimated) state variable does depend on the asset set and GMM weighting matrix. As a result, consider the robustness of the Campbell-Cochrane maximal predictability tests in each of the scenarios set out above.

1.6.1 Identity weight matrix

Table 1.9 shows that estimating the Bansal-Yaron model using the identity weight matrix produces an even less credible time-series of expected returns that when estimating it with the optimal weight matrix. The Campbell-Cochrane model's average residual of -20.4% coupled with the mean market return of 6.3% implies a mean expected market return of almost 30% a year under the Campbell-Cochrane model. This is almost five times the actual value. Again, this is not really a credible time series of expected returns. Only the Cecchetti-Lam-Mark model gives rise to a credible expected returns series: the mean residual of 3.8% implies a mean expected market return of 10% a year.

The results of the MDS tests for the Cecchetti-Lam-Mark residuals when the model is estimated with the identity weight matrix are shown in Table 1.10. They paint a similar picture to the results with the optimal weight matrix: the correlograms reject the MDS null (at the 5% level) from $q = 5$ onwards and the rescaled range rejects the MDS null too. Again, both tests imply anti-persistence in the residuals, while the quantilogram and Hong-Lee tests do not reject the null.

Notice that the choice of weight matrix does not affect the extraction of the Bansal-Yaron or Cecchetti-Lam-Mark state variables, so these maximal predictability results are unchanged. The GMM estimation for R^2 and $\bar{R}^2$ using the extracted Campbell-Cochrane state variable did not converge, so maximal predictability results are not available. This may be a reflection of the more general mis-specification of the Campbell-Cochrane model in this case.

1.6.2 Industry portfolios

Table 1.11 shows summary statistics of the residuals where we replace the six Fama-French size/value portfolios with the five Fama-French industry portfolios in the set of assets used to estimate the asset pricing models. Only the Campbell-Cochrane model estimated with the identity weight matrix produces a credible residual, and therefore expected return, series. With a mean residual of -11.5% and a mean market return of 6.3%, the mean expected market return is 17.8%. Even this may be stretching the bounds of credibility. But there is little harm in considering the robustness of the residual-based tests in this scenario in any case.

The Campbell-Cochrane model results when estimating the model using the industry portfolios and the identity weight matrix are shown in Table 1.12. We resoundingly reject the null that the residuals are MDS. The correlogram test produces two rejections at the 5% level, at the two shortest horizons considered. There are 72 rejections of the MDS null out of 81 quantilogram tests. The 99th percentile is the only one where we do not reject the MDS null.

Table 1.10 Cecchetti-Lam-Mark model results - Identity matrix

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	0.026	-0.122	-0.190	-0.260	-0.300	-0.311	-0.351	-0.381	-0.390
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.633	0.129	0.059	0.028	0.025	0.034	0.028	0.026	0.032

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.013	-0.016	-0.020	-0.024	-0.028	0.004
	0.40	0.75	0.99	0.83	0.80	0.70	0.70	0.69	0.76
0.05	-0.014	-0.028	-0.043	-0.058	-0.073	-0.042	-0.022	-0.007	0.004
	0.74	0.80	0.78	0.70	0.69	0.65	0.61	0.59	0.47
0.1	0.017	0.011	-0.018	-0.003	0.012	0.047	0.060	0.067	0.063
	0.68	0.66	0.54	0.55	0.51	0.49	0.46	0.45	0.38
0.25	0.026	-0.053	-0.111	-0.111	-0.136	-0.157	-0.154	-0.146	-0.147
	0.74	0.70	0.60	0.54	0.53	0.49	0.47	0.41	0.37
0.5	-0.030	-0.136	-0.162	-0.192	-0.260	-0.301	-0.318	-0.337	-0.350
	0.70	0.66	0.60	0.48	0.46	0.44	0.35	0.33	0.25
0.75	-0.059	-0.100	-0.140	-0.134	-0.140	-0.119	-0.126	-0.130	-0.137
	0.64	0.62	0.56	0.50	0.46	0.47	0.44	0.39	0.30
0.9	-0.003	0.034	0.084	0.139	0.199	0.227	0.251	0.275	0.310
	0.76	0.79	0.65	0.57	0.49	0.44	0.42	0.36	0.26
0.95	-0.030	0.024	0.070	0.137	0.197	0.234	0.258	0.272	0.280
	0.79	0.68	0.58	0.56	0.49	0.45	0.40	0.38	0.33
0.99	-0.009	-0.018	-0.024	-0.030	-0.036	-0.041	-0.047	-0.052	-0.057
	0.31	0.38	0.40	0.35	0.37	0.39	0.38	0.40	0.34

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.782	0.795	0.801	0.803	0.803	0.800	0.793	0.783	0.770
p-value	0.434	0.427	0.423	0.422	0.422	0.424	0.428	0.434	0.441

(d) Rescaled range

$\widehat{Q}$	0.694
p-value	0.01

Panels (a)-(d) report tests of the MDS null for the Cecchetti-Lam-Mark model residuals, estimated over the period 1930-2016. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 1.11 Properties of $\hat{\xi}_t$ - Industry portfolios

Model	Mean	Median	Std dev	SC(1)
Optimal weight matrix				
Bansal-Yaron	0.311	0.299	0.371	0.558
Campbell-Cochrane	-0.250	-0.279	0.232	-0.045
Cecchetti-Lam-Mark	-0.185	-0.158	0.194	-0.041
Identity weight matrix				
Bansal-Yaron	2.140	2.075	0.537	0.652
Campbell-Cochrane	-0.115	-0.130	0.257	-0.126
Cecchetti-Lam-Mark	-0.242	-0.098	0.376	0.567

Summary statistics for the model-implied ex-ante residuals. “Std dev” denotes standard deviation and “SC(1)” first-order serial correlation. The models are estimated and residuals computed using annual US data over the period 1930-2016.

While the Hong-Lee test produces no rejections, the rescaled range test also rejects the MDS null. Whether or not one considers the residuals to be a plausible financial time series, they are not MDS and the model is again rejected.

Turning to the maximal predictability tests, note again that the Bansal-Yaron and (extracted) Cecchetti-Lam-Mark state variables are unaffected by the change in the assets set, as well as the change in weight matrix. The Campbell-Cochrane state variable is, however, affected. The GMM estimation of R^2 and $\bar{R}^2$ does not converge for the Campbell-Cochrane state variable extracted based on parameter estimates using the optimal weight matrix to estimate the model. The estimation does converge, though, when the identity weight matrix is used in the Campbell-Cochrane model estimation.

These maximal predictability results are in Table 1.13. There is more evidence here that the Campbell-Cochrane state variable is unable to explain the own-history predictability of returns than in our main results from earlier. There are two significant exceedences of the R^2 bound at the three and six-year horizons.

1.6.3 Quarterly data

Returning to using the six size/value portfolios in the set of assets for estimating the models, rather the five industry portfolios, we consider the robustness of our results when estimating the models at the quarterly frequency. Quarterly data is only available from 1947Q1 and our sample period becomes 1947Q1-2017Q1. In this case, the summary statistics for our data are altered, as shown in Table 1.14 (note that none of the figures presented in this subsection are

Table 1.12 Campbell-Cochrane model results - Industry portfolios and identity matrix

(a) Correlogram

q	2	3	4	5	6	7	8	9	10
$\bar{C}(q)$	-0.142	-0.220	0.028	-0.040	0.039	-0.149	0.179	0.303	0.318
(Std err)	(0.054)	(0.080)	(0.101)	(0.118)	(0.133)	(0.147)	(0.159)	(0.171)	(0.182)
p-value	0.009	0.006	0.782	0.733	0.768	0.310	0.261	0.076	0.080

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	2	3	4	5	6	7	8	9	10
0.01	-0.003	-0.006	-0.009	-0.009	-0.006	-0.003	0.001	0.006	0.010
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.05	-0.014	-0.028	0.078	0.138	0.176	0.224	0.259	0.285	0.305
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.1	0.017	0.011	0.083	0.135	0.155	0.194	0.246	0.278	0.311
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.25	-0.022	-0.037	-0.027	-0.058	-0.082	-0.069	-0.065	-0.057	-0.064
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.5	-0.065	-0.111	-0.113	-0.158	-0.195	-0.210	-0.207	-0.208	-0.212
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.75	-0.029	-0.090	-0.121	-0.151	-0.154	-0.133	-0.113	-0.115	-0.116
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.9	-0.034	-0.047	-0.083	-0.099	-0.130	-0.148	-0.160	-0.183	-0.199
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.95	-0.030	-0.059	-0.086	-0.080	-0.083	-0.092	-0.106	-0.121	-0.139
	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.99	-0.009	-0.018	-0.026	-0.041	-0.059	-0.079	-0.099	-0.121	-0.144
	0.62	0.47	0.47	0.47	0.44	0.42	0.42	0.42	0.42

(c) Hong-Lee tests

q	2	3	4	5	6	7	8	9	10
$\widehat{G}(q)$	0.491	0.511	0.527	0.539	0.548	0.556	0.562	0.568	0.575
p-value	0.624	0.609	0.598	0.590	0.584	0.579	0.574	0.570	0.566

(d) Rescaled range

$\widehat{Q}$	0.946
p-value	0.00

Panels (a)-(d) report tests of the MDS null for the Campbell-Cochrane model residuals, estimated over the period 1930-2016. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In Panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. $\widehat{Q}$ denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 1.13 Campbell-Cochrane maximal predictability - Industry portfolios and identity weight matrix

q	2	3	4	5	6	7	8	9	10
R^2	0.057	0.022	0.012	0.000	0.029	0.035	0.039	0.020	0.002
$\bar{R}^2$	0.130	0.000	0.050	0.000	0.007	0.046	0.010	0.028	0.029
Wald stat	-	25.34	-	-	6.015	-	3.572	-	-
p-value	-	4.8×10^{-7}	-	-	0.014	-	0.059	-	-

Tests of the null that the market return is no more predictable than implied by the Campbell-Cochrane model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1930-2016. The Wald statistic and its asymptotic p-value are reported.

Table 1.14 Quarterly data summary statistics

	Mean	Median	Std dev	SC(1)
r_m	0.018	0.029	0.081	0.077
r_f	0.002	0.003	0.007	0.745
Δc	0.005	0.006	0.005	0.279
Δd	0.007	0.001	0.148	0.584
z_m	4.871	4.851	0.426	0.937

Descriptive statistics for our key variables at the quarterly frequency over the period 1947Q1-2017Q1. r_m denotes the log market return, r_f the quarterly log risk-free rate (the rolled over 1 month US T-bill), Δc log consumption growth, Δd log dividend growth and z_m the log price-dividend ratio.

Table 1.15 Properties of $\hat{\xi}_t$ - Quarterly

Model	Mean	Median	Std dev	SC(1)
Optimal weight matrix				
Bansal-Yaron	-670.6	-670.6	0.361	0.912
Campbell-Cochrane	-0.351	-0.341	0.091	0.264
Cecchetti-Lam-Mark	-0.389	-0.886	0.726	0.871
Identity weight matrix				
Bansal-Yaron	-395.9	-395.9	0.118	0.569
Campbell-Cochrane	-0.203	-0.186	0.094	0.318
Cecchetti-Lam-Mark	-0.008	0.003	0.081	0.074

Summary statistics for the model-implied ex-ante residuals. “Std dev” denotes standard deviation and “SC(1)” first-order serial correlation. The models are estimated and residuals computed using quarterly US data over the period 1947Q1-2017Q1.

annualised). In particular, the mean market return is slightly higher, at around 1.8% per quarter (or 7.2% a year).

We estimate the models using both the optimal and identity weight matrices. Summary statistics for the residuals are shown in Table 1.15. Note that these are quarterly figures (one could annualise them by multiplying them by four). As we can see in Table 1.15, only the Cecchetti-Lam-Mark model estimated with the identity matrix provides a credible residual series and therefore a credible expected return series, with a mean residual of -0.8% per quarter. The Bansal-Yaron model certainly does not provide credible residual series: it has mean quarterly residuals of -67000% per quarter with the optimal weight matrix and -40000% per quarter with the identity weight matrix! The Campbell-Cochrane model generates mean residuals of -35% per quarter with the optimal weight matrix and -20% per quarter with the identity weight matrix.

The MDS results for the Cecchetti-Lam-Mark model estimated at the quarterly frequency with the identity weight matrix are in Table 1.16. We also include the maximal predictability results in Table 1.16, since the Cecchetti-Lam-Mark state variable is affected by the change of data frequency. Note that q indicates the horizon in quarters. The choice of $q = 8, 12, 16, 20, 24, 28, 32, 36, 40$ quarters aligns with the earlier choice of $q = 2, 3, 4, 5, 6, 7, 8, 9, 10$ years. There are no rejections of the MDS null for the residuals, which would suggest the model does explain the dynamics of returns. In addition, the maximal predictability results only show one significant exceedence of the R^2 bound. Note, however, that the R^2 bound exceeds one on three occasions, which may be a symptom of numerical issues in computing the bounds.

Table 1.16 Cecchetti-Lam-Mark model - quarterly results with identity weight matrix

(a) Correlogram

q	8	12	16	20	24	28	32	36	40
$\bar{C}(q)$	-0.029	-0.095	-0.068	-0.007	-0.017	-0.073	-0.038	-0.040	-0.055
(Std err)	(0.089)	(0.112)	(0.132)	(0.149)	(0.164)	(0.178)	(0.191)	(0.203)	(0.215)
p-value	0.739	0.396	0.607	0.965	0.916	0.682	0.840	0.843	0.796

(b) Quantilogram

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	0.007	0.003	-0.003	-0.008	-0.014	-0.020	-0.025	-0.030	-0.036
	0.57	0.78	0.85	0.95	0.78	0.66	0.48	0.39	0.35
0.05	0.044	0.054	0.054	0.049	0.038	0.024	0.013	0.006	0.000
	0.37	0.37	0.35	0.37	0.38	0.33	0.30	0.29	0.27
0.1	0.054	0.061	0.059	0.054	0.041	0.023	0.010	0.001	-0.006
	0.46	0.40	0.33	0.33	0.36	0.36	0.35	0.35	0.32
0.25	0.043	0.048	0.045	0.043	0.035	0.018	0.003	-0.006	-0.014
	0.42	0.41	0.41	0.39	0.36	0.32	0.34	0.36	0.33
0.5	0.008	0.009	0.005	0.001	-0.004	-0.008	-0.014	-0.020	-0.025
	0.62	0.56	0.59	0.60	0.51	0.48	0.45	0.41	0.38
0.75	0.033	0.034	0.032	0.026	0.014	-0.003	-0.019	-0.029	-0.036
	0.49	0.46	0.41	0.39	0.41	0.43	0.40	0.37	0.30
0.9	0.042	0.041	0.034	0.024	0.009	-0.011	-0.028	-0.041	-0.052
	0.43	0.45	0.44	0.46	0.53	0.53	0.55	0.49	0.43
0.95	0.045	0.049	0.044	0.034	0.020	0.004	-0.009	-0.018	-0.027
	0.43	0.36	0.38	0.37	0.37	0.39	0.41	0.37	0.33
0.99	0.007	0.007	0.004	0.000	-0.008	-0.018	-0.025	-0.030	-0.035
	0.77	0.99	0.90	0.73	0.66	0.62	0.53	0.46	0.39

(c) Hong-Lee tests

q	8	12	16	20	24	28	32	36	40
$\widehat{G}(q)$	0.002	-0.211	-0.367	-0.468	-0.542	-0.603	-0.682	-0.731	-0.769
p-value	0.998	0.833	0.714	0.640	0.588	0.547	0.495	0.465	0.442

(d) Rescaled range

$\widehat{Q}$	1.076
p-value	0.87

Panels (a)-(d) report tests of the MDS null for the Cecchetti-Lam-Mark model residuals when the model is estimated with the identity matrix, over the period 1947Q1-2017Q1. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath. In panel (b), the estimated weighted quantilogram is given in larger font for the appropriate (α,q) combination. Its bootstrapped p-value is given underneath in smaller font. $\widehat{G}(q)$ in panel (c) denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. $\widehat{Q}$ in panel (d) denotes the estimated rescaled range. Its bootstrapped p-value is given beneath.

Table 1.16 Cecchetti-Lam-Mark model - quarterly results with identity weight matrix
(e) Maximal predictability

q	8	12	16	20	24	28	32	36	40
R^2	0.367	0.011	0.105	0.005	0.001	0.099	0.012	0.023	0.072
$\bar{R}^2$	1.298	0.347	0.411	3.858	5.6×10^{-7}	0.499	3.071	3.196	0.861
Wald stat	-	-	-	-	10.95	-	-	-	-
p-value	-	-	-	-	0.001	-	-	-	-

Panel (e) reports tests of the null that the market return is no more predictable than implied by the Cecchetti-Lam-Mark model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1947Q1-2017Q1. The Wald statistic and its asymptotic p-value are reported.

The findings that the Cecchetti-Lam-Mark model and its state variable can explain return dynamics, however, are not themselves robust. Having a larger sample allows us to look at performance in sub-samples. We divide our sample in two with the break in the middle of the sample, so that our sub-samples are 1947Q1-1982Q1 and 1982Q2-2017Q1. Dividing the sample into two in this way ensures a sample size in excess of 120 (i.e. $3 \times \max\{q\}$) in each sub-sample, which helps ensure the accuracy of the long-horizon serial correlation estimates.

In addition, we can examine robustness to dealing with look-ahead bias in the second sub-sample. In the above results, the parameters of the ex-ante $(t-1)$ expectations are estimated over future data, which could induce a finite-sample bias in the test statistics even when the test statistics are asymptotically valid. Note that these concerns apply only to the correlogram and Hong-Lee tests. The quantilogram and rescaled range bootstrap procedures explicitly account for the estimation method and the finite sample. The maximum predictability test conditions on the parameter estimates in any case. We evaluate the robustness of our correlogram and Hong-Lee results to using past data only to estimate the parameters of the model residuals. We compute residuals for the second sub-sample which are formed using parameters estimated over an expanding window. The expanding window begins at the first observation in the whole sample (1947Q1) and ends at the $(t-1)$th observation when computing the $t-1$ expectations of returns at t. We compare these results to those obtained for the second sub-sample above to evaluate the effect of restricting the data sample to past data only.

Looking at the Cecchetti-Lam-Mark residuals estimated with the identity matrix in the sub-samples in this way, we see that the MDS null is rejected in both sub-samples and when we account for look-ahead bias. The MDS null is clearly rejected by the quantilograms in the first sub-sample (Table 1.17a): 37 of the 81 weighted quantilograms are significant at the 10% level and 25 of those are significant at the 5% level. Untabulated results show that this is the only test to reject the null in the first sub-sample, re-iterating why it is important to consider

Table 1.17 Cecchetti-Lam-Mark model quarterly sub-sample results using identity weight matrix

(a) Quantilogram - sub-sample 1: 1947Q1-1982Q1

$\alpha \downarrow / q \rightarrow$	8	12	16	20	24	28	32	36	40
0.01	0.050	0.076	0.100	0.122	0.142	0.163	0.193	0.337	0.589
	0.02	0.02	0.02	0.06	0.07	0.09	0.09	0.65	0.62
0.05	0.092	0.093	0.075	0.076	0.065	0.042	0.010	-0.009	-0.036
	0.82	0.97	0.86	0.75	0.62	0.48	0.33	0.24	0.19
0.1	0.008	-0.002	-0.017	0.002	0.004	-0.006	-0.012	-0.007	-0.018
	0.52	0.41	0.40	0.49	0.45	0.42	0.35	0.35	0.29
0.25	-0.095	-0.147	-0.181	-0.189	-0.198	-0.212	-0.226	-0.239	-0.244
	0.06	0.04	0.03	0.03	0.03	0.03	0.02	0.02	0.01
0.5	-0.052	-0.133	-0.191	-0.207	-0.200	-0.203	-0.221	-0.236	-0.232
	0.30	0.10	0.07	0.05	0.04	0.04	0.03	0.02	0.02
0.75	0.036	-0.049	-0.111	-0.136	-0.161	-0.195	-0.227	-0.234	-0.225
	0.84	0.26	0.10	0.07	0.04	0.03	0.01	0.01	0.01
0.9	0.063	0.060	0.043	0.013	-0.039	-0.095	-0.140	-0.167	-0.186
	0.97	0.75	0.59	0.38	0.23	0.16	0.11	0.09	0.08
0.95	0.008	0.011	-0.002	-0.025	-0.066	-0.107	-0.142	-0.166	-0.183
	0.57	0.49	0.41	0.33	0.22	0.19	0.17	0.12	0.09
0.99	-0.029	-0.046	-0.059	-0.073	-0.091	-0.108	-0.128	-0.153	-0.184
	0.16	0.17	0.17	0.18	0.17	0.07	0.01	0.00	0.00

(b) Hong-Lee tests - sub-sample 2: 1982Q2-2017Q1

q	8	12	16	20	24	28	32	36	40
$\widehat{G}(q)$	20.29	19.83	19.52	19.30	19.08	18.78	18.46	18.15	17.87
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

(c) Correlogram - sub-sample 2: 1982Q2-2017Q1, accounting for look-ahead bias

q	8	12	16	20	24	28	32	36	40
$\bar{C}(q)$	-3.548	-11.02	-26.13	-35.71	196.3	-39.16	-38.37	-13.32	21.03
(Std err)	0.251	0.318	0.373	0.422	0.465	0.504	0.541	0.575	0.608
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Panel (a) reports the quantilogram tests of the MDS null for the Cecchetti-Lam-Mark model residuals estimated with the identity weight matrix, over the first sub-sample 1947Q1-1982Q1. The estimated weighted quantilogram is given in larger font for the appropriate (α, q) combination. Its bootstrapped p-value is given underneath in smaller font. Panel (b) gives the Hong-Lee tests for the residuals from the second sub-sample 1982Q2-2017Q1. $\widehat{G}(q)$ denotes the Hong-Lee generalised spectral statistic. Its asymptotic p-value is given beneath. Panel (c) reports the weighted correlogram tests for the second sub-sample where estimation uses the identity weight matrix but also accounts for possible look-ahead bias. $\bar{C}(q)$ denotes the estimated transformed weighted correlogram statistic, $\sum_{j=1}^{q-1}(1-j/q)\bar{\rho}(q)$. Its standard error and asymptotic p-value are given underneath.

Table 1.17 Cecchetti-Lam-Mark model quarterly sub-sample results using identity weight matrix

(d) Maximal predictability

q	8	12	16	20	24	28	32	36	40
Sub-sample 1: 1947Q1-1982Q1									
R^2	0.072	0.072	0.137	0.228	0.106	0.153	0.276	0.630	0.416
$\bar{R}^2$	0.467	0.107	1.658	0.666	0.009	0.044	0.136	14.81	2.064
Wald stat	-	-	-	-	494.3	806.1	553.0	-	-
p-value	-	-	-	-	0.000	0.000	0.000	-	-
Sub-sample 2: 1982Q2-2017Q1									
R^2	0.023	0.156	0.262	0.161	0.071	0.076	0.546	0.792	0.519
$\bar{R}^2$	0.084	0.126	0.002	0.010	0.409	0.009	17.462	2.665	0.204
Wald stat	-	0.771	442.1	167.2	-	924.5	-	-	13078
p-value	-	0.380	0.000	0.000	-	0.000	-	-	0.000

Panel (d) reports tests of the null that the market return is no more predictable than implied by the Cecchetti-Lam-Mark model state variables (i.e. $R^2 \leq \bar{R}^2$) in each of the two sub-samples. The Wald statistic and its asymptotic p-value are reported.

a battery of test statistics. Looking at the second sub-sample (Table 1.17b), the MDS null is easily rejected by the Hong-Lee tests. When accounting for look-ahead bias in the estimation (Table 1.17c), the MDS null remains strongly rejected, this time by the weighted correlograms.

Moreover, there are now three significant exceedences of the R^2 bound in each sub-sample, although not necessarily at the same horizons. The R^2 bound is significantly exceeded at $q = 28$ in both sub-samples, but not the whole sample. The ability of the Cecchetti-Lam-Mark model state variable to explain the dynamics of returns also appears not to be robust.

We lastly consider the robustness of the Bansal-Yaron and Campbell-Cochrane maximal predictability results to using quarterly data. Note that the Bansal-Yaron state variables do not depend on whether we estimate the Bansal-Yaron model using the identity or optimal weight matrix, but the Campbell-Cochrane state variables do depend on the weight matrix used.

Table 1.18 shows the results of these robustness checks. For the Bansal-Yaron model we see similar results to when using the annual data: the model's state variables cannot explain the own-history predictability of returns. For the Campbell-Cochrane model, things look a little more hopeful. There are two significant exceedences of the R^2 bound using the optimal weight matrix and three using the identity weight matrix. However, untabulated results show further rejections at horizons $q < 8$. Using the optimal weight matrix, the R^2 bound is exceeded for $q = 3, 4, 5$ and 6 quarters and these exceedences are significant at the 1% level. Using the

Table 1.18 Quarterly maximal predictability results

q	8	12	16	20	24	28	32	36	40
Bansal-Yaron model									
R^2	0.057	0.018	10^{-8}	0.069	0.025	0.012	0.119	0.011	0.159
$\bar{R}^2$	-0.217	-11.357	5.508	-3.740	-0.282	0.303	-10.089	-6.736	6.579
Wald stat	216041	148018	-	267587	416715	-	593143	1307207	-
p-value	0.000	0.000	-	0.000	0.000	-	0.000	0.000	-
Campbell-Cochrane model - optimal weight matrix									
R^2	0.029	0.041	0.059	1.2×10^{-4}	0.001	0.005	1.4×10^{-4}	0.006	0.082
$\bar{R}^2$	0.003	0.018	0.421	0.039	0.006	0.556	0.352	1.505	0.689
Wald stat	462.1	7.954	-	-	-	-	-	-	-
p-value	0.000	0.005	-	-	-	-	-	-	-
Campbell-Cochrane model - identity weight matrix									
R^2	-30.42	0.022	0.034	0.445	0.082	0.085	1.5×10^{-6}	0.025	1.000
$\bar{R}^2$	1225	0.622	4.4×10^{-4}	0.300	0.903	7.0×10^{-5}	0.831	0.335	9.772
Wald stat	-	-	25.26	8.190	-	197.0	-	-	-
p-value	-	-	5.0×10^{-7}	0.004	-	0.000	-	-	-

Tests of the null that the market return is no more predictable than implied by the Bansal-Yaron/Campbell-Cochrane model state variables (i.e. $R^2 \leq \bar{R}^2$), estimated over the period 1947Q1-2017Q1. The Wald statistic and its asymptotic p-value are reported.

identity weight matrix, there are exceedences for $q = 1$ and 6 quarters. Overall, it does not seem as if the Campbell-Cochrane state variable can explain own-history predictability of returns when using quarterly data.

We take these maximal predictability results with a little caution, however. Table 1.18 shows that there are numerical difficulties in estimating the R^2 and $\bar{R}^2$ parameters. These are estimated jointly by GMM (no regression is run to obtain R^2). As a result, even though the R^2 for the predictive regressions should be the same for both models and whether the optimal or identity weight matrix is used to estimate the model, this is not the case. Moreover, we see some R^2 and $\bar{R}^2$ which are either greater than one or less than zero. These numerical issues may be a function of the mis-specification of the state variables in terms of being able to explain own-history predictability of returns. Or they may reflect more general numerical issues.

1.7 Conclusion

We show that three consumption-based asset pricing models - the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark models - cannot explain the own-history predictability structure of the US market return. We focus on how well the three models explain stock return predictability because, from an investor's point of view, it is a key characteristic of returns. It has received relatively little attention in the context of the Bansal-Yaron and Campbell-Cochrane models, two of the leading models in explaining the equity premium puzzle. Within predictability, we focus on own-history predictability as it is the most basic form of predictability.

In order to test whether the three models can explain the own-history predictability properties of the US market return, we first estimate the models' parameters by GMM before computing model implied ex-ante expected returns. If the model can capture the own-history predictability of the market, the difference between the realised market return and the model implied ex-ante expected return will be MDS due to rational expectations. We test whether these residuals are MDS, ensuring that our tests account for the initial estimation step. In this sense, our tests can be interpreted as a time-series specification test of the models. However, unlike a J-test, our procedure allows us to test models which are not estimated in single GMM implementation, such as the Campbell-Cochrane and Cecchetti-Lam-Mark models here.

We find that the Bansal-Yaron, Campbell-Cochrane and Cecchetti-Lam-Mark model residuals are not MDS. This finding is robust to the choice of GMM weight matrix, using quarterly in instead of annual data and using industry instead of size/book-to-market portfolios to estimate the models. There appears to be some hope, in that we cannot reject the null that the Cecchetti-Lam-Mark residuals are MDS using quarterly data, the identity weight matrix and size/book-to-market portfolios to estimate the model. However, this non-rejection of the MDS null is not robust over time. When we divide the sample period into two equal-length sub-samples, we clearly reject the MDS null in both sub-samples.

Moreover, our tests of maximal predictability suggest that returns are more predictable with respect to their past history than is consistent with the state variables which truly explain the market return being the state variables of any of the three models considered. While the annual data suggested that the Campbell-Cochrane and Cecchetti-Lam-Mark state variables could explain the own-history predictability of the market return, this finding was not robust to using quarterly data. The Bansal-Yaron state variables were not able to explain the own-history predictability of returns at either the annual or quarterly horizon.

The failure of the models considered to capture the own-history predictability of stock returns has several different interpretations. The first is that perhaps some auxiliary assumption in the models has failed. For example, the assumed joint normality of consumption and dividend growth in the Campbell-Cochrane model (used to derive expected returns) or the assumed joint normality of consumption growth, dividend growth, the long-run risk and economic volatility in the Bansal-Yaron model (used by Constantinides and Ghosh (2011) to invert the model and derive the moment conditions to estimate it). Note that these normality assumptions are used when backing out the state variables for the maximal predictability tests too, so both the residual-based and maximal predictability tests would be affected in this scenario. In this interpretation, the models are basically correct, but the auxiliary assumptions need to be relaxed in future empirical work.

A second interpretation in which the models are basically correct is to say that the models presented are equilibrium models, but that financial markets are often out of equilibrium. Therefore, to model market dynamics, it is necessary to consider a framework in which markets adjust to a (possibly time-varying) equilibrium. Adam et al. (2016) present such a model. They have an agent with CRRA preferences who knows the risk-adjusted stock price is a random walk (a result due to Samuelson, 1965) but who observes the risk-adjusted price plus mean-zero noise. Optimal updating of beliefs under subjective expected utility maximisation produces a feedback loop: expectations affect prices, as in the classical model, but prices also affect expectations, due to updating. This feedback imparts serial correlation and excess volatility upon the returns, even when the estimated prior uncertainty (noise variance) is small. In general, this model is able to match many facts about asset prices, including the long-horizon predictability of excess returns with respect to the price-dividend ratio. However, rather like the standard CRRA model, it cannot account for the equity premium and risk-free rate puzzles. Nonetheless, it is possible that by applying this framework to, say, the Campbell-Cochrane model would account for these puzzles.

Finally, it may simply be that the model state variables are mis-specified: that more state variables need to be considered or some of those considered need to be dropped. Or, given that the models here are strictly rational models of investor behaviour, it may be that an “outright” behavioural model (going beyond, say, rational learning) is required.

Chapter 2

Is regulatory short sale data a profitable predictor of UK stock returns?

Regulator-required public disclosures of net short positions do not provide a profitable investment signal for UK stocks. While long-short (zero initial outlay) portfolios based on this signal usually make a profit on average, it is rarely statistically significant in either gross or risk-adjusted terms. The issue is that the short sides of the portfolios make substantial losses. This is true even when using information in the trend in disclosures to form portfolios, rather than using the most recent disclosures, which is a more standard procedure. Unit initial outlay portfolios based on the disclosures that are allowed to take short positions do not reliably significantly outperform the market. Certain long-only unit initial outlay portfolios based on the disclosures do reliably significantly outperform the market. However, this outperformance is economically modest: about 1 percentage point a year in gross and risk-adjusted terms.

JEL classification: G11, G14
Keywords: short sales, Short Selling Regulation, net short position disclosure, investment signal, anomaly

2.1 Introduction

In November 2012, European Union (EU) Regulation 236/2012 brought into force disclosure requirements that give rise to a freely available database of all large net short positions held in stocks traded on EU markets. The database is made available to investors with only a short lag of up to a couple of days. It gives access to daily data on short holdings in specific stocks, something which has previously only been available for a fee. Conversations with practitioners indicate that these fees are high and that uptake is subsequently limited. It is therefore of great practical interest to examine whether the new, freely available information on other investors' short positions can be used profitably.

I examine this question from the point of view of the UK stock market - the largest and most liquid in the EU.[1] All in all, there appears to be little profit to be gained from using this information to form portfolios.

First, I look at standard long-short portfolios, of the sort commonly used to study new potential investment signals. I consider equal-, value- and net short positions-weighted long-short portfolios with zero initial outlay. These portfolios go long in stocks with a low level of total (aggregated across investors) declared net short positions and short in stocks with a high level of total declared net short positions. Since the rules exempt market making and hedging trades from notifications, the rationale is that short sellers are revealing their private information. As short sellers are likely to be more sophisticated investors than the average investor, mimicking their positions could therefore be profitable. This has proved to be the case in previous studies (e.g. Boehmer et al., 2008; Diether et al., 2009).

While these long-short portfolios are profitable on average, only the equal-weighted portfolio has a mean gross return significantly different from zero at the (5% level). Its average risk-adjusted returns are not significant. Scaling positions by volatility to reduce portfolio volatility does not improve the statistical significance of any profitability. The long sides of these long-short portfolios tend to make substantial gains, but the short sides tend to make substantial losses. Going beyond the standard approach of forming portfolios based on the most recent declarations does not solve this problem. I find similar results using information about the trend in short positions, which captures signal strength. Rebalancing the portfolios less frequently to make them less responsive to signal noise does not remedy the lack of long-short portfolio profitability, either. This is the case even when losses in the less frequently rebalanced portfolios are controlled with stop-loss rules. Nonetheless, I find that the long sides of the

[1] My entire sample period is prior to the UK's exit from the EU.

standard equal and value-weighted long-short portfolios are highly profitable. Moreover, they produce significant risk-adjusted returns, as measured against a variety of risk factors.

I therefore construct fully invested (unit initial investment) portfolios based on the net short position disclosures too. Using portfolios of all stocks as a benchmark shows how much information is in the short disclosures, above and beyond a strategy that simply buys every stock listed on the FTSE350. When these fully invested portfolios based on declared net short positions are allowed to take short positions, they do not generally significantly outperform portfolios of all stocks. In fact, the value-weighted versions of these fully invested portfolios underperform the value-weighted portfolio of all stocks. I also look at fully invested portfolios based on the disclosures which are long-only. Both the equal-weighted long-only portfolio based on the most recent disclosures and the long-only portfolio weighted by the strength of the trend in disclosures significantly outperform the comparable portfolios of all stocks. However, the gain is relatively modest in both cases: one percentage point per year in both gross and risk-adjusted terms.

This paper relates to a literature that began with studies of US short sale data. Early studies of whether short sale data could be used profitably in US stock markets were positive. Boehmer et al. (2008) use proprietary NYSE order data between January 2000 and April 2004 to show that long-short strategies of the kind I use here can yield a substantial mean return of 3.8% per month, or an annualised average three-factor alpha of 16%. Given the rarity of recalls and the relatively low direct costs of shorting most stocks, Boehmer et al. argue their strategies' profitability would survive accounting for all short selling costs.

Similarly, Diether et al. (2009) study NYSE, AMEX and Nasdaq stocks in 2005 and find that long-short strategies based on daily short volume are very profitable. They construct their short sale measures from SEC-required disclosure data.[2] While Diether et al. find positive and significant abnormal returns to short sale activity strategies, these returns are less extreme than in Boehmer et al. (2008). Moreover, Diether et al. (2009) are more cautious about the impact of costs. They argue that when one considers the costs of shorting smaller stocks, which are more heavily represented in the short basket than in the market as a whole, it is quite possible that costs would annul any long-short strategy profits.

UK studies of strategies based on short sale data have been less optimistic than those in the US. Au et al. (2009) analyse equal- and value-weighted long-short portfolios which go long in stocks with low short interest and short in stocks with high short interest. Their short interest measures are based on subscription-access short sale data for FTSE350 constituents between

[2] This data was only available for a trial period between January and December 2005. Such disclosures have subsequently become mandatory again, but the exchanges are permitted to charge for the data. The fees are high.

September 2003 and September 2006. Only the equal-weighted long-short portfolio makes a statistically significant average profit. For both the equal- and value-weighted portfolios, the short side makes a loss on average. Likewise, Andrikopoulos et al. (2012) find that profits to equal- and value-weighted long-short portfolios formed on short interest are generally insignificant. Again, the short legs cost these portfolios money. Andrikopoulos et al. use a broader and longer sample of 1645 UK stocks between August 2004 and February 2012, again with subscription-access data.

The data made available as a result of the EU short selling regulations have started to be studied, although not, as far as I am aware, as the basis for an investment strategy. The closest work to doing this is Jank and Smajlbegovic (2017), who consider all EU stock markets. They analyse how actual short positions have performed, rather than hypothetical positions taken based on information about others' positions. The difference is due to a difference in aims. I am testing potential investment strategies, Jank and Smajlbegovic test how actual investments performed. They find that the average short trade is profitable in terms of mean and excess returns, although the excess returns are not significant and the mean is only significant at the 5% level on a value-weighted basis. My finding that hypothetical short trades based on past information about others' short positions are not profitable is consistent with Jank and Smajlbegovic. By basing my strategies on past short selling activity, I am considering an investor who is relatively late to the short selling party. It is natural that this position is less profitable than that of the average investor already shorting the stock.

The paper proceeds as follows. Section 2.2 explains the short disclosures rules and the practicalities of making a disclosure. It also sets outlines trends in the dataset and discusses portfolios how I evaluate portfolios. Section 2.3 deals with the construction and performance of the long-short (zero initial outlay) portfolios. Section 2.4 does the same for the fully invested (unit initial outlay) portfolios. Section 2.5 shows that the results described above are robust to the empirical choices which had to be made throughout the analysis, while Section 2.6 shows that these results are even robust to pretending that the lag between positions being taken and the declaration being published does not exist. Section 2.7 concludes.

2.2 Data and portfolio evaluation

2.2.1 Net short position disclosures

As of 1 November 2012, an investor in any EU-listed stock with a sufficiently large net short position must notify the national regulator. For the UK, this is the Financial Conduct Authority

(FCA). I work with public notifications, which must be made once the net short position is 0.5% or greater of the issued share capital of a given company. Further notifications must be made at each 0.1% increment or decrement and a notification must also be made once the position drops below 0.5%.[3] These notifications are published on the FCA's website by the close of business on the day the notification is made. Notifications are made the trading day after the position crosses the relevant threshold. Therefore the data is made easily available to investors in a timely manner.

The disclosures ought to be informative about short sellers' private information and beliefs. Market making and liquidity providing trades are exempt from the notification rules. Moreover, net short positions are calculated net of delta-adjusted derivative positions. A short position created to hedge a derivative position does not count towards the net total. Because market making, liquidity providing and hedging trades are unrelated to anticipated price changes, they do not reveal a short seller's private information.

In addition, synthetic short positions - holdings of derivatives such as options that perfectly replicate the payoff of a short position - must also be reported, again on a delta-adjusted basis. Synthetic short trades can attract lower transaction costs than conventional short trades (Daske et al., 2005). They may therefore be an important source of information about investors' expectations of price movements.

Note that the rules apply to any investor, no matter where they are domiciled. In fact, Jank et al. (2016) find that a sizeable proportion of disclosures for EU-listed stocks come from the US.

Net short position declarations required by the regulations outlined above must be made by 3.30pm the trading day after the position was established/modified. The FCA publishes the public notifications by the close of business that same day. A notification triggered by a trade on day t must be made by 3.30pm on $t+1$ and that information is published by the FCA by close of business on $t+1$. However, since the FCA's close of business is 6pm and FTSE trading closes at 4.30pm, it is unclear whether the FCA disclosures will be available to trade on before the close of trading on $t+1$, or if investors would have to wait until the market opens on $t+2$. I err on the side of caution and assume that investors have to wait until the market opens at $t+2$. However, assuming that investors can instead trade at the close of day $t+1$ makes little difference to the results (see Section 2.5.1).

Given the truncated nature of the data, it is necessary to make some assumptions to construct a net short positions measure from the public disclosures. Note that disclosures are made by

[3] In addition, there are private notifications, which the regulator keeps confidential. These must be made once the position crosses a threshold of 0.2% of share capital outstanding, and at each 0.1% increment/decrement in the position.

investor by stock. I assume that open positions with no new disclosures on a given day are unchanged. If there is a new declaration of a position of 0.51% of issued share capital on Monday and no new declaration on Tuesday, I assume the position remains 0.51% on Tuesday. I also assume that positions that are below the 0.5% reporting threshold are 0%. If Wednesday's position falls to 0.49%, then I take the position to be 0% from Wednesday onwards since no further tracking of the position is possible. This gives a daily position series for each investor for each stock. To get total declared net short positions, I sum declarations across investors for each stock on each day. To evaluate the robustness of the results to these assumptions I consider a measure of net short positions that requires no such assumptions: the number of distinct declarations of positions greater than 0.5% per stock per day. The results are robust to such a change (see Section 2.5.2).

2.2.2 Sample and data

My sample is the constituents of the FTSE350 index. These are all large and liquid stocks. I obtain the return and characteristic data needed to form and evaluate portfolios from Thomson Eikon. I adjust returns for dividends since short sellers must pay any dividends distributed to the stock owner, and this allows for the reinvestment of dividends on the long side - a key source of growth.

Short disclosure data is available for positions taken as of 31 October 2012[4] and therefore the first position can be taken as of 2 November 2012, assuming the FCA's public disclosures of positions on day t reach traders between market close on $t+1$ and market open on $t+2$. The first return in the back test return series is realised on 1 November 2013, since some portfolio formation schemes use up to 252 days of net short positions information in their formation. The last date in my sample is 13 December 2018. The back test return series contain 1295 observations.

2.2.3 Net short position disclosures across the sample

I look at disclosures for stocks in the FTSE350 index. There is a total of 30357 disclosures (position openings, updates and closures) across the sample period. 101 stocks have no disclosures associated with them. There are 114 disclosures made when the disclosure regulations initially enter into force.[5] Table 2.1 shows summary statistics for the number of daily FTSE350

[4] Despite the regulation entering into force on 1 November, there are some disclosures for 31 October.

[5] There are three on 31 October 2012 and 111 on 1 November 2012. None of the 1 November declarations are updates to the 31 October ones.

Table 2.1 Summary statistics for the number of public net short position disclosures per day

Mean	18.9
Median	16.0
Standard deviation	13.1
Maximum	68.0
Minimum	0.00

Summary statistics for the number of public net short position disclosures in FTSE350 listed stocks per day over the sample period (2 November 2012 - 13 December 2018). The first day of the regulations is excluded as the declaration of existing positions opened before the regulations took effect is likely to distort the summary statistics. Disclosures include both declarations of positions that have just crossed the 0.5% reporting threshold and updates to positions which have already been declared as being above the threshold. The standard deviation is computed using the unbiased estimator.

Table 2.2 Summary statistics for the number of public net short position disclosures per stock

Mean	29.7
Median	19.0
Standard deviation	30.4
Maximum	149
Minimum	0.00

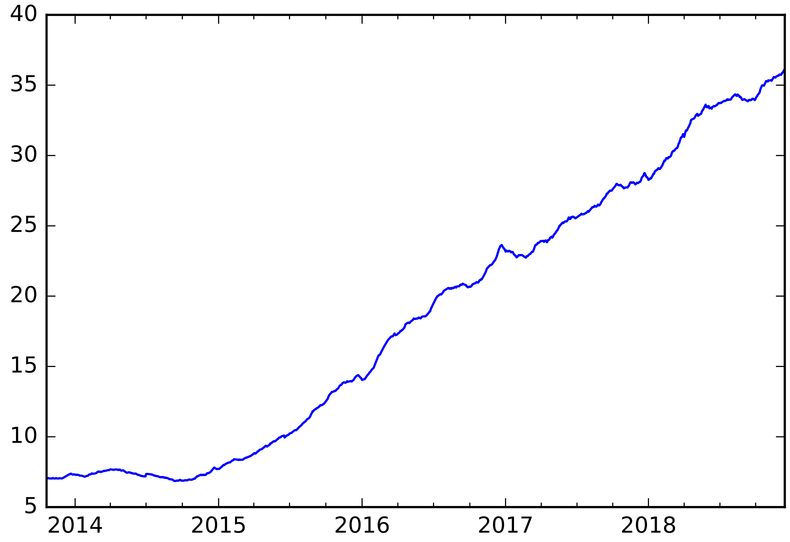
Summary statistics for the number of public net short position disclosures in FTSE350 listed stocks per stock made over the entire sample period (1 November 2012 - 13 December 2018). Disclosures include both declarations of positions that have just crossed the 0.5% reporting threshold and updates to positions which have already been declared as being above the threshold. The standard deviation is computed using the unbiased estimator.

disclosures over time, from the day after the regulation enters force (i.e. from 2 November 2012). This is to prevent the initial rush of declarations of existing positions from distorting the averages of the number of fresh daily disclosures. We see that there are approximately 19 disclosures per day on average (16 on the median day), although this number ranges between zero and 68. Figure 2.1 shows a one year rolling average of the daily number of disclosures. It is clearly increasing over time.

We can also consider the number of public disclosures per stock over the whole sample, as in Table 2.2. The mean stock has a total of 30 disclosures over the sample, while the median stock has 19. The maximum number of disclosures is 149.

Table 2.3 shows the size of positions, where these are not aggregated across investors, both as a percentage of outstanding share capital and in monetary value. The mean publicly disclosed declared net short position for an individual investor is 0.94% of share capital outstanding, while the median is 0.71%. The highest net short position for an individual investor is 8.03% of share capital outstanding. In monetary terms, the mean size of a declaration is £30.5 million

Fig. 2.1 One year rolling mean of the number of public net short position disclosures per day



Graph shows the 252 trading day rolling mean of the number of public disclosures of net short positions in FTSE350 listed stocks from 2 November 2012 until 13 December 2018. The first day of the regulations is excluded as the declaration of existing positions opened before the regulations took effect is likely to distort the summary statistics. The date on the horizontal axis represents the end of the rolling window.

and the median is £19.3 million. There is a large range in the monetary value of declared positions, with (non-zero) declarations ranging between £71,000 and £1.4 billion.[6]

Figure 2.2 shows the (cross-sectional) mean *open* (i.e. $\geq 0.5\%$) net short position size over time, where size is measured both in terms of the percentage of outstanding share capital and monetary value. The open short positions have been aggregated across investors. Stocks with no declared net short positions are included in the cross-sectional averages. There is a clear upward trend in the total size (aggregated across investors) of declared short positions in the average stock over time on both measures, albeit with greater volatility in the monetary values, and potentially some levelling off or even a fall at the end of the sample.

Finally, the mean declared net short position is open (i.e. $\geq 0.5\%$) for 45 trading days. Table 2.4 summarises average length of declared net short position duration by stock. The mean stock has an average net short disclosure duration of 49 days, the median stock 36 days with the maximum average duration is 529 days.

[6]Note that the largest position in monetary terms and the largest position in terms of outstanding share capital are two different positions. The 8.03% position is a position in Melrose Industries PLC held by Guevoura Fund Ltd and was declared on 19 August 2016. The £1.4 billion position is a position in British American Tobacco PLC held by Millennium International Management LP and was declared on 24 July 2017. The position amounted to 1.43% of the outstanding share capital.

Table 2.3 Summary statistics for the size of public net short position disclosures (% outstanding share capital)

	% outstanding share capital	£GBP
Mean	0.94%	£30.5 million
Median	0.71%	£19.3 million
Standard deviation	0.68%	£44.5 million
Maximum	8.03%	£1443 million
Minimum	0.00%	£0.07 million

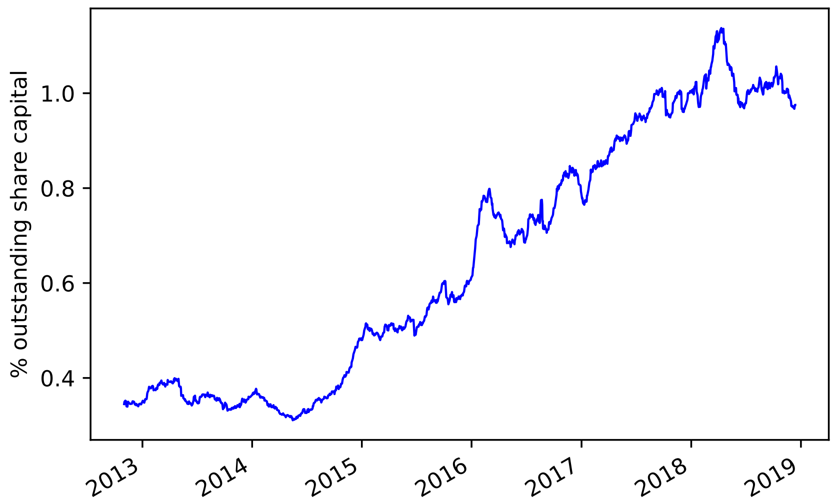
Summary statistics for the size (as a percentage of outstanding share capital and in monetary value) of public net short position disclosures in FTSE350 over the sample period (1 November 2012 - 13 December 2018). Disclosures for a given stock are not aggregated across investors. Disclosures include both declarations of positions that have just crossed the 0.5% reporting threshold and updates to positions which have already been declared as being above the threshold. The standard deviation is computed using the unbiased estimator.

Table 2.4 Cross-sectional summary statistics for the mean duration of public net short positions (in days)

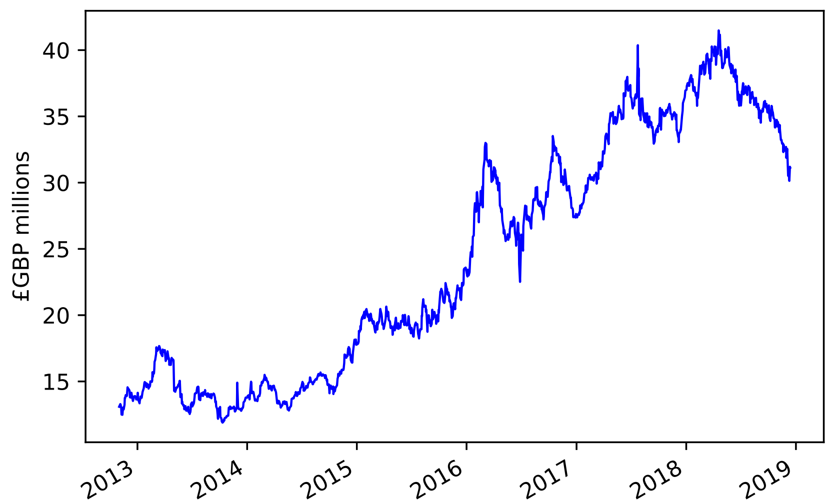
Mean	49.2
Median	35.8
Standard deviation	49.5
Maximum	529
Minimum	1.00

Cross-sectional (across stocks) summary statistics for the mean duration of public net short position disclosures in FTSE350 listed stocks over the sample period (1 November 2012 - 13 December 2018). The standard deviation is computed using the unbiased estimator.

Fig. 2.2 Cross-sectional mean size of aggregate disclosed net short positions in a given stock over time



(a) % outstanding share capital



(b) £GBP (millions)

Graph shows the evolution of the cross-sectional (across stocks) mean size (% outstanding share capital and £GBP billions) of open disclosed net short positions, aggregated across investors, over the sample period 1 November 2012 - 13 December 2018.

2.2.4 Portfolio evaluation

I evaluate portfolios based on their mean, Sharpe ratio and alphas from Fama-French type regressions. The alphas are effectively average risk-adjusted returns, where the risk adjustment is with respect to the risk factors included in the Fama-French style regressions. The Sharpe ratios and alphas are calculated with respect to returns in excess of the risk-free rate, where the risk-free rate is the SONIA overnight rate (data obtained from the Bank of England). I compute HAC p-values for the mean and the alphas. All return series are daily and I scale the means and alphas by 252 to approximately annualise them. The p-values are based on the daily returns, however. I annualise the Sharpe ratios as per Lo (2002), taking 252 trading days to be a year.[7]

I compute the alphas using three different sets of factors to ensure and evaluate robustness to the factors used. I obtain all factor data from AQR's online data library. The first set of factors is the almost canonical Fama and French (1993) three factors: market, size and value. I denote the resulting alphas as α^{FF3}. The second set of factors adds momentum to the Fama-French three factors, as per Carhart (1997). This could be an important factor. If short sellers correctly anticipate price falls and price falls are persistent, my strategies will inevitably contain momentum exposure. I denote the alphas from this four factor model α^{FF4}. The final set of factors adds quality minus junk (Asness et al., 2019) to the four-factor model. The resulting alphas are termed α^{QMJ}. The quality minus junk factor encompasses profitability, growth and safety (Asness et al., 2019). It is clear that each of these may be related to shorting activity. All else equal, investors are likely to be more willing to trade in safe firms. However, one would expect those with low/negative growth and profitability to be the main candidates for short trades.

2.3 The failure of long-short portfolios

2.3.1 Portfolios using the most recent declared net short positions

First, I construct standard long-short arbitrage portfolios. This is the standard method of constructing and testing a possibly profitable investment signal in the literature. Long-short

[7] The estimated Sharpe ratio for strategy i is given as $\widehat{SR}_i = \overline{r_i - r_f}/\mathrm{sd}(r_i - r_f)$, where $\overline{r_i - r_f}$ is the mean daily return to strategy i in excess of the risk free rate and $\mathrm{sd}(\cdot)$ the sample standard deviation. Lo (2002) then shows that the q-day Sharpe ratio is given by

$$\widehat{SR}_i(q) = \frac{q}{\sqrt{q + 2\sum_{k=1}^{q-1}(q-k)\hat{\rho}_i(k)}}\widehat{SR}_i,$$

where $\hat{\rho}_i(k)$ is the kth order autocorrelation coefficient of the excess returns to strategy i.

portfolios are arbitrage portfolios and involve a zero initial outlay. They take long positions totalling £1 in a basket of stocks (the "long side" or "long basket") and short positions totalling £1 in a different basket of stocks. The returns to the long-short portfolio are therefore the returns to the long basket less the returns to the short basket.

The broad idea of the strategy in this paper is that short sellers are informed investors who reveal their private information and beliefs through their net short positions. I therefore assign stocks in the top quintile of declared net short positions on day t to the short basket on day $t+2$ (using day $t+2$ due to the timing convention in Section 2.2.1). Likewise, I assign stocks in the bottom quintile of declared net short positions to the long basket. In practice, the 20th percentile of declared net short positions is always zero, so all stocks with zero declared net short positions go into the long basket. On a typical day, around 70% of stocks of no declared short positions. The long basket therefore typically contains far more than 20% of stocks.[8] Occasionally, the 80th percentile of declared net short positions is also zero. In this case, I assign stocks with zero declared net short positions to the long basket and those with non-zero declared net short positions to the short basket.

I first consider three weighting schemes: equal weighting, value weighting and net short positions weighting. For the value-weighted portfolios, I weight the short side by one over the market capitalisation at $t+1$. The reason for this is that, with short-term momentum in stock prices, market value weights assign ever increasing weights to losing positions, harming the portfolio. Empirically, this does turn out to be the case in my sample: a short basket formed with inverse market capitalisation weights performs better than a short basket with market capitalisation weights. To see the problem, consider a short position taken on day s. Suppose the stock rises in price between s and $s+1$. The short position has lost money, but the rise in price implies a rise in market value and so an increase in the market value weight. Short-term price momentum means that the stock price is likely to rise again between $s+1$ and $s+2$, thus the standard value weighting has just assigned a higher weight to a position likely to loose money. Compared to using standard value weights in the short basket, we see that the median weight using the inverse value weighting scheme is slightly higher (1.04% versus 0.96%) but the inverse value weights have a lower standard deviation (1.4% compared to 2.1%).

[8] There are many more than the target number of stocks in the long basket. The number of stocks in the long basket could be reduced, e.g. by looking at stocks with the lowest net declared short positions over the period $t-k$ to t, for some arbitrary k rather than just focussing on positions above the threshold on day t. I do not use this approach for two reasons. First, because it is not really in the spirit of assigning stocks with the lowest quintile of short interest to the long basket (the lowest quintile is zero, so all stocks with zero declared short interest should go to the long basket). Second, such an approach would give a significant degree of freedom to the researcher and may tempt data mining.

Table 2.5 Standard long-short portfolio performance

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Value-weighted									
Long	0.075	0.062	0.797	0.073	0.021	0.069	0.037	0.082	0.017
Short	0.053	0.401	0.377	0.055	0.275	0.065	0.193	0.070	0.184
L-S	0.022	0.491	0.415	0.018	0.519	0.004	0.873	0.012	0.658
Equal-weighted									
Long	0.119	0.007	1.170	0.118	0.001	0.114	0.002	0.127	0.001
Short	0.075	0.161	0.679	0.077	0.076	0.084	0.053	0.096	0.035
L-S	0.044	0.042	1.127	0.041	0.046	0.029	0.133	0.031	0.121
Net short position-weighted									
Long	0.119	0.007	1.170	0.118	0.001	0.114	0.002	0.127	0.001
Short	0.092	0.121	0.777	0.094	0.059	0.104	0.036	0.113	0.029
L-S	0.027	0.449	0.360	0.024	0.479	0.010	0.772	0.013	0.692

Performance evaluation measures for daily rebalanced declared net short positions portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

In the equal-weighted portfolios, the long and short sides are the portfolio are both separately equally weighted. For the net short positions-weighted portfolios, the long side is equal-weighted, since all stocks in the long basket have zero declared net short positions. The short side weights are proportional the level of declared net short positions, which are aggregated across investors for each stock on each day. I normalise the sum of the short basket weights to be one. Unlike the equal- and value-weighted portfolios, the net short positions-weighted portfolio uses information in how much the declared net short positions exceed the threshold by.

In Table 2.5, we see that the long-short ("L-S") value- and net short positions-weighted portfolios are profitable on average, but their mean returns are not significantly different from zero. The picture is similar for their alphas, too. While the long legs of these portfolios make healthy gross and risk-adjusted returns, the short sides lose a substantial amount of money.

The equal-weighted long-short portfolio makes positive returns that are significantly different from zero on average. However, its risk-adjusted returns (alphas) are not generally significant. Even taking the significant profit for the equal-weighted long-short portfolio at face value, the short leg loses 7.5% per year. If we conceive the short leg loss as the cost of borrowing to invest in the long leg, there are surely cheaper means of financing the investments in the long basket. These investments in the long basket return a strong 11.9% per year on average.

The problem of the short side making substantial losses persists throughout the sample. Figure 2.3 shows the annualised (scaled by 252) one-year rolling mean daily return to the short basket of stocks under all three weighting schemes. In all three cases, this return is positive for the great majority of the sample. Since short sellers profit from stock prices falling, a positive return to stocks in the short basket means that the short basket is losing the portfolio money. It is not the case that one or two bad patches are distorting the profitability of the short side.

Neither is it the case that there are any great periods where the long side significantly outperforms the short side. Figure 2.4 shows the rolling t-statistics for α^{QMJ}, although the results are essentially the same for α^{FF3} and α^{FF4}. The rolling window Fama-French alphas are hardly ever both positive and significantly different from zero at the 5% level for any of the three long-short portfolios.

The finding that long-short portfolios have empirically positive means and alphas which are not significantly different from zero may be driven by excessive volatility in the portfolios. Controlling the volatility in the portfolios may make their profits more stable and reliable, and so their means and alphas more likely to be significant. Of course, changing the weights may also reduce, or increase, the means and alphas.

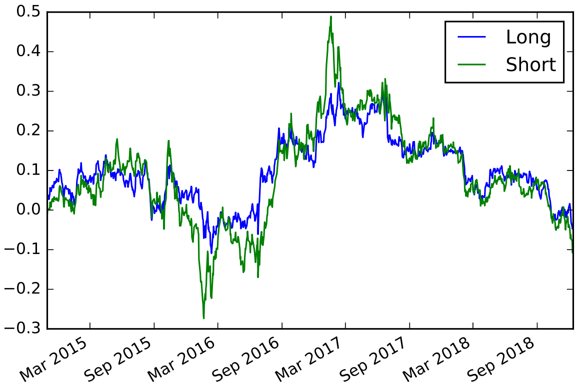
To control portfolio volatility, I divide stock i's weight in each basket by its volatility. I then re-normalise the weights to sum to one in each basket again. All else equal, low volatility stocks get a higher weight and high volatility stocks a lower weight. I follow Elaut and Erdos (2019) in estimating stock i's volatility as the square root of its exponentially weighted moving variance (EWMV), where the EWMV has a 60 day centre of mass.

Untabulated results show that volatility scaling does not really affect the portfolios' returns. The average gross and risk-adjusted returns to the long-short portfolios remain positive but insignificant overall. The mean return to the equal-weighted long-short portfolio remains positive and significantly different from zero. However, its alphas remain insignificant.

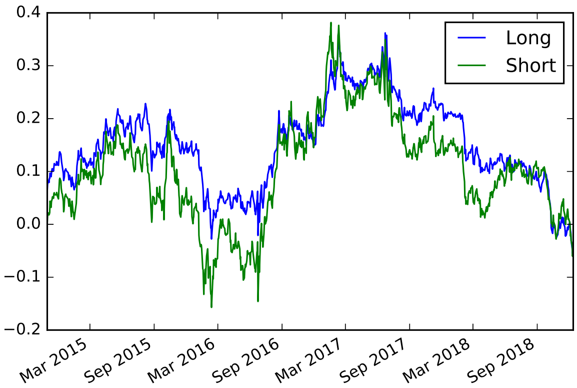
Daily rebalancing is another possible impediment to the long-short portfolios described above. Daily updating and rebalancing of the portfolios may be excessively frequent. Daily rebalancing ensures the portfolios rapidly respond to spikes in relative declared net short positions, but can also induce excessive volatility. Less frequent rebalancing could smooth these responses out.

To analyse the effect of rebalancing frequency, I form long and short (and therefore long-short) portfolios of which $1/q$ of the portfolio is rebalanced each day, where q days is the rebalancing frequency. This provides a daily time series of over-lapping q-day rebalanced

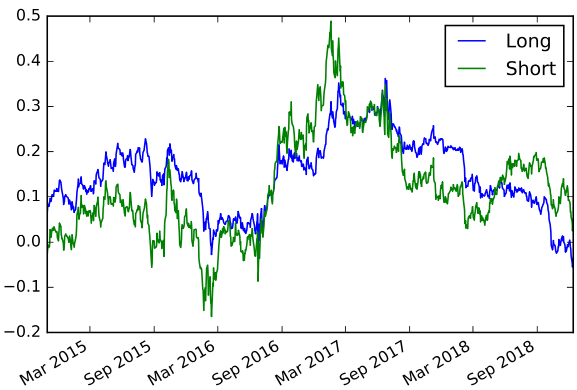
Fig. 2.3 Rolling window mean returns to long and short side of standard long-short portfolios



(a) Value-weighted



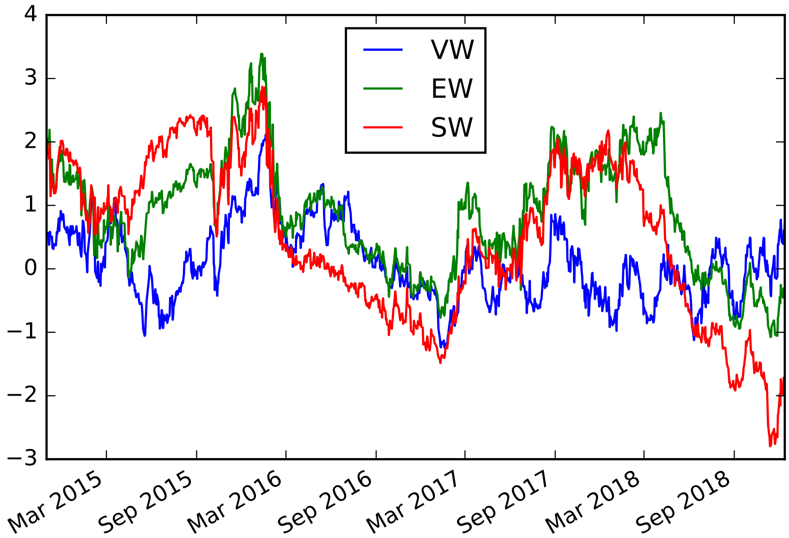
(b) Equal-weighted



(c) Net short position-weighted

252 day rolling window mean return for both the long and short sides of the declared net short positions portfolios with: value weights ("VW"), equal weights ("EW") and net short positions weights ("SW"). The mean return is scaled by 252 to make it an approximately annual figure. The date on the horizontal axis shows the rolling window end. The first rolling window begins on 1 November 2013 and ends on 30 October 2014. The last rolling window ends on 13 December 2018.

Fig. 2.4 Rolling window α^{QMJ} t-statistic for standard long-short portfolios



HAC 252 day rolling window t-statistic on α^{QMJ} for the long-short declared net short positions portfolios with: value weights ("VW"), equal weights ("EW") and net short positions weights ("SW"). The date on the horizontal axis shows the rolling window end. The alphas are scaled by 252 to make them approximately annual figures. The first rolling window begins on 1 November 2013 and ends on 30 October 2014. The last rolling window ends on 13 December 2018.

portfolio returns. Boehmer et al. (2008) use this approach. I consider monthly ($q = 21$) and annual ($q = 252$) rebalancing.

The results are essentially the same for the equal- and net short positions-weighted portfolios when rebalancing less frequently. The value-weighted portfolio performs worse with either monthly or annual rebalancing: its long side profits fall and its short side losses rise.

I also consider adding stop-loss rules to the less frequently rebalanced portfolios. These rules exit positions after a maximum loss of 1% for a monthly rebalanced portfolio and 10% for an annually rebalanced portfolio. The stop-loss rules do not prevent the short sides of the value-, equal- or net short positions-weighted portfolios from continuing to lose a considerable amount of money. The rules do not, therefore, alter the results very much.

I now consider combinations of the volatility scaling and less frequent rebalancing fixes described above. Rebalancing volatility scaled portfolios less frequently has little impact on their returns.

Adding stop-loss rules to monthly and annually rebalanced portfolios does improve the results, however, as Table 2.6 shows. The equal-weighted long-short portfolio performs the best in this set-up, with a mean annual return of 7.8%, annualised alphas of 5.8%-7.5% and an annual

Sharpe ratio of 1.1. While the mean return and three-factor alpha are significantly different from zero at the 5% level, the four-factor and QMJ alphas are not. Moreover, the t-statistics on the mean and three-factor alpha are some way from Harvey et al.'s (2016) recommended enhanced threshold of 3.0.[9] The net short position-weighted long-short portfolio provides greater profits than with daily rebalancing. However, its mean return and alphas are not significantly different from zero at any conventional level. The value-weighted long-short portfolio now makes lower losses, thanks to lower losses on the short side.

2.3.2 Portfolios using multiple days' declarations

In Section 2.3.1, I implement less frequent rebalancing by rebalancing $1/q$ of the portfolio every q days. This rebalancing scheme is closely related to a portfolio formed using an average of the past q days' net short position declarations.[10] Using the average of these declarations incorporates the persistence of high/low declared net short positions into the portfolio weighting scheme.

An alternative means of incorporating the persistence of high/low net short position declarations is to consider signals based on net short positions averaged over different time periods. My approach follows Elaut and Erdos (2019). I compute the mean of declared net short positions over a set of horizons, $\mathscr{H}$. For each $h \in \mathscr{H}$, and remembering that net short position declarations on trading day t only become available to traders on trading day $t+2$, we have

$$\overline{SI}^{h}_{i,t} = \frac{1}{h}\sum_{j=0}^{h-1} SI_{i,t-j-2}.$$

I then set $S^{h}_{i,t} = +1$ if $\overline{SI}^{h}_{i,t}$ is less than or equal to the 20th cross-sectional percentile of $\overline{SI}^{h}_{i,t}$ at time t and $S^{h}_{i,t} = -1$ if $\overline{SI}^{h}_{i,t}$ is greater than or equal to the 80th percentile. Finally, I compute

$$\bar{S}^{\mathscr{H}}_{i,t} = \frac{1}{\dim(\mathscr{H})}\sum_{h\in\mathscr{H}} S^{h}_{i,t}. \tag{2.1}$$

[9] The enhanced threshold is due to multiple testing concerns: a large amount of research on possible investment signals is focussed on a small number of underlying datasets. The mean returns and alphas for portfolios based on new potential signals must therefore exceed a higher t-statistic threshold to be deemed profitable.

[10] When rebalanced once every q days, the long and short sides of the equal-weighted long-short portfolio have weights

$$w^{b}_{i,t}(q) = \frac{1}{q}\sum_{s=0}^{q-1}\frac{1}{N^{b}_{t-s}}Q^{b}_{i,t-s},$$

where N^{b}_{t} is the number of stocks in basket b at time t and $Q^{b}_{i,t}$ indicates if stock i is assigned to basket b at time t.

Table 2.6 Volatility scaled standard long-short portfolios with stop-loss rules

(a) Monthly rebalancing

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Value-weighted									
Long	0.051	0.140	0.647	0.048	0.078	0.039	0.178	0.050	0.093
Short	0.103	0.085	0.777	0.105	0.036	0.115	0.019	0.122	0.022
L-S	-0.053	0.166	-0.722	-0.057	0.093	-0.076	0.016	-0.072	0.034
Equal-weighted									
Long	0.106	0.004	1.251	0.104	0.000	0.098	0.002	0.110	0.001
Short	0.028	0.607	0.209	0.029	0.518	0.040	0.359	0.049	0.298
L-S	0.078	0.030	1.009	0.075	0.026	0.058	0.062	0.061	0.062
Net short position-weighted									
Long	0.106	0.004	1.251	0.104	0.000	0.098	0.002	0.110	0.001
Short	0.030	0.641	0.199	0.032	0.564	0.046	0.391	0.052	0.359
L-S	0.076	0.154	0.674	0.072	0.142	0.053	0.260	0.058	0.215

(b) Annual rebalancing

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Value-weighted									
Long	0.047	0.187	0.580	0.045	0.114	0.042	0.160	0.054	0.084
Short	0.113	0.040	0.977	0.114	0.013	0.118	0.010	0.127	0.009
L-S	-0.066	0.029	-1.245	-0.068	0.014	-0.076	0.005	-0.072	0.009
Equal-weighted									
Long	0.104	0.006	1.225	0.102	0.001	0.098	0.002	0.111	0.001
Short	0.045	0.396	0.382	0.046	0.287	0.052	0.231	0.062	0.171
L-S	0.059	0.032	1.087	0.057	0.020	0.047	0.044	0.048	0.039
Net short position-weighted									
Long	0.104	0.006	1.225	0.102	0.001	0.098	0.002	0.111	0.001
Short	0.044	0.455	0.337	0.046	0.367	0.054	0.280	0.064	0.217
L-S	0.060	0.143	0.697	0.057	0.131	0.044	0.229	0.046	0.201

Performance evaluation measures for declared net short positions based portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

I assign stocks with $\bar{S}^{\mathscr{H}}_{i,t} > 0$ at time t to the long basket with weights proportional to $\bar{S}^{\mathscr{H}}_{i,t}$. I normalise the weights in the long basket to sum to one. Likewise, I assign stocks with $\bar{S}^{\mathscr{H}}_{i,t} < 0$ at time t to the short basket. Again, the weights are proportional to $\bar{S}^{\mathscr{H}}_{i,t}$ and normalised to sum to one. I call this the multiple signals approach.

An alternative is to use a regression-based approach, analogous to Han et al. (2016). Here, in each time period, I run the cross-sectional regression

$$R_{i,t+1} = \gamma_{0,t+1} + \sum_{h\in\mathscr{H}} \gamma_{h,t+1} S^h_{i,t} + \upsilon_{i,t+1}, \tag{2.2}$$

where $\upsilon_{i,t+1}$ is an error term.[11] Letting $\hat{\gamma}_t$ denote the OLS estimates from (2.2), I generate expected returns $\hat{R}_{i,t+1}$ as

$$\begin{aligned}\hat{R}_{i,t+1} &= \tilde{\gamma}_{0,t+1} + \sum_{h\in\mathscr{H}} \tilde{\gamma}_{h,t+1} S^h_{i,t}\\ \tilde{\gamma}_{t+1} &= \frac{1}{P}\sum_{s=0}^{P-1} \hat{\gamma}_{t-s}.\end{aligned}$$

Note that $\tilde{\gamma}_{t+1}$ does not contain $\hat{\gamma}_{t+1}$. I take an average of past $\hat{\gamma}_t$ as the expectation for γ_{t+1}, given information at t.

I assign stocks to the short and long baskets at time t based on their expected returns $\hat{R}_{i,t+1}$. Stocks with expected returns in the top cross-sectional quintile of $\hat{R}_{i,t+1}$ are assigned to the long basket. Those with expected returns in the bottom quintile are assigned to the short basket.

I consider three different weighting schemes for the long-short portfolios. First, I consider value (market capitalisation) weighting within both the long and the short basket, again using the inverse market capitalisation for the short basket weights. Second, equal weighting within each basket. And third, using weights proportional to the stock's expected return. As before, I normalise the long and the short basket weights separately to sum to one in each case. I term this method of forming portfolios the regression-based approach.

Table 2.7 shows the results of the following implementations of the multiple signals and regression-based approaches. These are representative of other implementations (see Section 2.5.3). For the multiple signals approach, I use $\mathscr{H} = \{1,2,5,10,21,42,63,126,189,252\}$ days, corresponding to one-day, two-day, one-week, one-month, two-month, three-month, six-month, nine-month and one-year horizons. For the regression-based approach, I use $\mathscr{H} = \{1,5,21,63,126,189\}$. This reduced set of horizons is to reduce problems of multicollinearity. I set $P = 63$, so use three months of regressions to compute the coefficients.

[11] I cannot use $\overline{SI}^h_{i,t}$ in place of $S^h_{i,t}$ in (2.2) due to collinearity issues.

Table 2.7 Performance of long-short portfolios based on multiple days' disclosures

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Multiple signals approach									
Long	0.107	0.008	1.133	0.105	0.001	0.101	0.002	0.113	0.001
Short	0.077	0.151	0.692	0.078	0.068	0.084	0.052	0.096	0.033
L-S	0.030	0.169	0.832	0.027	0.183	0.017	0.382	0.017	0.382
Regression-based approach: Value weights									
Long	0.099	0.025	0.957	0.099	0.007	0.098	0.009	0.110	0.004
Short	0.057	0.285	0.516	0.058	0.161	0.057	0.162	0.065	0.129
L-S	0.042	0.172	0.966	0.041	0.192	0.041	0.196	0.045	0.148
Regression-based approach: Equal weights									
Long	0.122	0.008	1.110	0.122	0.001	0.121	0.001	0.135	0.001
Short	0.067	0.179	0.660	0.067	0.084	0.065	0.103	0.078	0.058
L-S	0.056	0.017	1.118	0.055	0.015	0.056	0.010	0.057	0.011
Regression-based approach: Expected return weights									
Long	0.165	0.052	0.760	0.166	0.027	0.165	0.039	0.171	0.037
Short	0.066	0.342	0.451	0.065	0.264	0.067	0.260	0.076	0.187
L-S	0.099	0.262	0.500	0.101	0.255	0.098	0.276	0.095	0.297

Performance evaluation measures for daily rebalanced declared net short positions portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

Like many of the portfolios using only the most recent net short position declarations in Section 2.3.1, the long-short multiple signals portfolio is profitable on average but not significantly so. The alphas are also positive but insignificant. The long side makes a strong 10.7% a year, with alphas very close to this. These are all very significant at conventional levels. However, the short side continues to lose the long-short portfolio a considerable amount of money: 7.7% per year. Untabulated results show this is an issue throughout the sample, similar to the portfolios in Section 2.3.1.

Turning to the regression-based portfolios, the returns to the expected return-weighted long-short portfolio have an annualised mean of 9.9%, which is high. Its alphas are similarly high. However, the p-values on the mean return and alphas are large at around 0.25-0.3. The problem is that the portfolio is very volatile: its Sharpe ratio is just 0.48, despite its high mean return. The returns are too volatile for the profit to this strategy to be statistically reliable.

The equal-weighted long-short portfolio produces a positive mean return (5.6% a year) that is significantly different from zero at the 5% level. Its alphas are of a similar magnitude and

Table 2.8 Performance of volatility scaled long-short portfolios based on multiple days' declarations

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Multiple signals approach									
Long	0.113	0.002	1.376	0.111	0.000	0.106	0.000	0.117	0.000
Short	0.068	0.148	0.691	0.068	0.067	0.071	0.059	0.084	0.030
L-S	0.045	0.003	1.946	0.043	0.003	0.035	0.012	0.033	0.017
Regression-based approach: Value weights									
Long	0.081	0.037	0.899	0.080	0.011	0.077	0.019	0.089	0.008
Short	0.059	0.259	0.564	0.058	0.155	0.053	0.192	0.062	0.139
L-S	0.022	0.524	0.398	0.022	0.546	0.024	0.505	0.027	0.455
Regression-based approach: Equal weights									
Long	0.108	0.006	1.184	0.107	0.000	0.106	0.001	0.119	0.000
Short	0.062	0.157	0.687	0.062	0.070	0.057	0.109	0.069	0.052
L-S	0.046	0.005	1.617	0.045	0.004	0.049	0.002	0.050	0.001
Regression-based approach: Expected return weights									
Long	0.142	0.072	0.707	0.142	0.036	0.140	0.055	0.147	0.046
Short	0.047	0.470	0.317	0.045	0.395	0.046	0.405	0.057	0.269
L-S	0.095	0.202	0.608	0.097	0.195	0.094	0.223	0.090	0.248

Performance evaluation measures for daily rebalanced declared net short positions portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

also significant at conventional levels. The alphas' t-statistics do not exceed Harvey et al.'s (2016) enhanced threshold of 3.0, though. Moreover, the short side of the portfolio makes considerable losses: 6.7% a year. The short side losses can be seen as the cost of financing investments in the long side of the portfolio. There must surely be a more cost-effective means of financing these investments.

The value-weighted regression-based portfolios behave similarly to the expected return-weighted portfolios. The long-short portfolio makes a positive but statistically insignificant mean return. Its alphas are positive and insignificant, too. The short side loses 5.7% per year.

To examine whether the lack of significance in the portfolios' profitability is a result of excessive portfolio volatility, I volatility scale the multiple signals and regression-based portfolios as in Section 2.3.1.

Table 2.8 shows that volatility scaling these portfolios does affect the results. The returns to both the long and the short sides of the long-short multiple signals portfolio improve. As a result, the long-short portfolio's mean return improves to be 4.5% per year. Moreover, the long-short portfolio's volatility falls. The portfolio's mean return is significantly different from zero at the 5% level and has a t-statistic in excess of 3.0. Its annualised alphas are 3.3%-4.3% and are all significant at the 5% level. Only the three-factor alpha has a t-statistic greater than 3.0, though. In addition, the short side continues to lose 6.8% per year, leading one to believe that there would be cheaper ways of financing the investments in the long basket.

Looking at the equal-weighted regression-based portfolios, the returns to the long basket fall. The losses to the short basket fall too, although by less. The net effect is that the returns to the long-short portfolio fall. Nonetheless, the long-short portfolio's volatility falls substantially. As a result, the mean and all of the alphas are significant at the 1% level, and the four-factor and QMJ alphas have t-statistics exceeding 3.0. Like with the multiple signals long-short portfolio, the short side still makes considerable losses: this time of 6.2% a year.

Volatility scaling the value-weighted regression-based portfolios harms long-short performance. Returns to the long basket fall, while losses to the short basket remain broadly unchanged. Volatility scaling the expected return-weighted regression-based portfolios has no net effect on the long-short portfolio. The fall in returns to the long basket approximately offsets the fall in losses to the short basket.

Another option for smoothing portfolio response to signals is to rebalance less frequently. I implement this in a similar way to in Section 2.3.1, except I consider one-month ($q = 21$) and six-month ($q = 126$) rebalancing. Untabulated results show that rebalancing less frequently has very little impact on the multiple signals portfolios. This lack of impact is not surprising given the portfolios are already a function of a smoothed signal. Rebalancing the regression-based portfolios less frequently hurts their returns substantially by reducing signal exposure. The returns to the long basket fall and losses to the short side rise. Volatility scaling these less frequently rebalanced portfolios makes little difference to their performance.

I also consider adding stop-loss rules to the less frequently rebalanced portfolios. These allow a maximum position loss of 1% for monthly rebalanced portfolios and 5.5% for six-monthly rebalanced portfolios. The only portfolios to benefit from these rules are the monthly rebalanced multiple signals portfolios. Table 2.9a shows that the long-short portfolio now makes an impressive 8.3% a year, which is significantly different from zero at the 5% level. The annualised alphas range from 6.3-7.5%, although only the three-factor alpha is significant. The short side loss falls to 2.9% a year, too. Table 2.9b shows that volatility scaling the monthly rebalanced multiple signals portfolios with stop loss rules improves things even further. The long-short portfolio now makes 8.8% a year with annualised alphas of 7.2%-8.2%. The mean and three-factor alpha now have t-statistics greater than 3.0.

Table 2.9 Multiple signal portfolio performance with monthly rebalancing and stop-loss rules

(a) No volatility scaling

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Multiple signals approach									
Long	0.112	0.011	1.200	0.115	0.001	0.112	0.001	0.122	0.001
Short	0.029	0.670	0.186	0.040	0.466	0.050	0.364	0.055	0.329
L-S	0.083	0.041	0.998	0.075	0.046	0.063	0.088	0.067	0.071

(b) With volatility scaling

	Mean	p-value	Sharpe	α^{FF3}	p-value	α^{FF4}	p-value	α^{QMJ}	p-value
Multiple signals approach									
Long	0.117	0.004	1.364	0.120	0.000	0.116	0.001	0.126	0.000
Short	0.029	0.607	0.224	0.038	0.398	0.044	0.324	0.052	0.259
L-S	0.088	0.002	1.400	0.082	0.003	0.072	0.010	0.074	0.009

Performance evaluation measures for declared net short positions portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

2.4 Fully invested portfolios

I now consider fully invested - unit (£1) initial outlay - portfolios. I allow these portfolios to take long and short positions, and term them unconstrained fully invested portfolios. The key advantage to these portfolios over the long-short portfolios already discussed is that the size of the short side of the portfolio relative to the long side can vary endogenously over time. We have already seen that the short-only portfolios lose money. If the relative size of the short basket were fixed, it would simply act as a drain on returns. In order to have a chance of making money, the proportion of capital allocated to the short basket must be timed.

I consider not only the risk-adjusted returns of the fully invested portfolios, but also how they compare to equal- and value-weighted portfolios of all stocks. The all stock portfolios can be thought of as naive portfolios that simply buy a bit of everything. By comparing the fully invested portfolios to these naive portfolios, I can test how informative, if at all, the net short position disclosures are. Moreover, portfolios of all stocks are a relevant benchmark for investors. After all, if a strategy cannot even beat the market, investors are unlikely to find it attractive.

I also consider long-only fully invested portfolios based on net short position disclosures. These portfolios are similar to the unconstrained portfolios, except they have a lower bound on weights of zero. Given the strong performance of the long sides of the long-short portfolios considered in Section 2.3, they are likely to perform well. Comparing these to the portfolios of all stocks will reveal to what extent there is information in the absence of net short position disclosures. Moreover, comparing these to the unconstrained short disclosure based positions (which can go long and short), we can evaluate the benefit of allowing the fully invested portfolios to take short positions.

Since all these comparisons are a question of relative performance, it is important to account for transaction costs. I assume these to be 10bps each way in what follows.[12,13]

I consider portfolios formed on both the most recent day's disclosures, as well as multiple days' disclosures. I stick to the timing convention that the disclosures relating to net short positions taken or adjusted on day t are made public overnight between $t+1$ and $t+2$.

For the unconstrained portfolios based on the most recent day's disclosures, the portfolio weights $w_{i,t}$ are

$$w_{i,t} \propto 1(SI_{i,t-2} = 0) - 1(SI_{i,t-2} > 0), \tag{2.3}$$

where $1(\cdot)$ is the indicator function, $SI_{i,t-2}$ the total size of declared net short positions open in stock i at time $t-2$ across all investors. I normalise $w_{i,t}$ such that $\sum_i w_{i,t} = 1$, where i indexes all stocks in the FTSE350. A short position is taken in stock i at t when $SI_{i,t-2} > 0$, and a long position is taken when $SI_{i,t-2} = 0$. I use a constant threshold of zero for $SI_{i,t-2}$ to form the portfolios so that the relative size of the short basket to varies endogenously through time in a consistent manner. Zero is a natural value for that threshold, as it means that at least one investor has a large net short position in i at $t-2$. The long-only version of this portfolio has weights

$$w_{i,t} \propto 1(SI_{i,t-2} = 0). \tag{2.4}$$

Again, I normalise these weights to sum to one.

Both (2.3) and (2.4) are equal-weighted in the sense that all long positions in stocks are of the same size and all short positions in stock are also the same size. The natural comparison portfolio of all stocks is therefore the equal-weighted portfolio of all stocks. This is the comparison I use for (2.3) and (2.4), which I term the equal-weighted unconstrained and equal-weighted long-only portfolios, respectively. In any case, the equal-weighted portfolio of all stocks transpires to be a tougher comparison portfolio than the value-weighted portfolio of all stocks.

[12] I thank Jacopo Capra of Cantab Capital for our discussions of this assumption.

[13] Since the earlier long-short portfolios do not generally make a significant profit without transaction costs, they will obviously not make one with transaction costs. So transaction costs are unimportant in Section 2.3.

I also consider a value-weighted versions of (2.3) and (2.4), with $w_{i,t} \propto [1(SI_{i,t-2} = 0) - 1(SI_{i,t-2} > 0)] \times \text{market cap}_{i,t}$ and $w_{i,t} \propto 1(SI_{i,t-2} = 0) \times \text{market cap}_{i,t}$, respectively. I compare these value-weighted portfolios to the value-weighted portfolio of all stocks.

The unconstrained fully invested portfolio based on multiple days' declarations has weights

$$w_{i,t} \propto \bar{S}_{i,t}^{\mathscr{H}},$$

with $\bar{S}_{i,t}^{\mathscr{H}}$ defined as in (2.1) and $\mathscr{H} = \{1,2,5,10,21,42,63,126,189,252\}$. Again, I normalise the weights to sum to one. These weights also capture the persistence of the net short disclosure signal, incorporating extra information. It is not clear whether to compare this portfolio to the equal- or value-weighted portfolio of all stocks. I use equal-weighted, since it is the more exacting comparison. I term these weights the unconstrained multiple signals weights. In addition, I consider a long-only portfolio based on multiple days' declarations with weights

$$w_{i,t} \propto 1(\bar{S}_{i,t}^{\mathscr{H}} > 0) \times \bar{S}_{i,t}^{\mathscr{H}}.$$

These weights are the long-only multiple signals weights. Finally, I consider value-weighted versions of these multiple signals portfolios (weights proportional to $\bar{S}_{i,t}^{\mathscr{H}} \times \text{market cap}_{i,t}$ and $1(\bar{S}_{i,t}^{\mathscr{H}} > 0) \times \bar{S}_{i,t}^{\mathscr{H}} \times \text{market cap}_{i,t}$, respectively), and compare these to the value-weighted portfolio of all stocks.

To compare the portfolios, I compute gross and risk-adjusted average returns, as well as the differences between them. I also compute turnover, for an indication of costs, and maximum drawdown, as an indicator of tail risk. A drawdown is a loss from local peak to local trough. The maximum drawdown therefore gives the return for the investor who times his entry into and exit from the portfolio perfectly badly. I further examine tail risk through the 99%, 95% and 90% expected shortfalls: the expected loss given that returns are in the worst 1%, 5% and 10% of the distribution, respectively. Computing tail risks allows us to make risk comparisons beyond the Fama-French style factors. I give all losses as positive numbers, so that a higher number is worse.

2.4.1 Portfolios using the most recent day's declarations

Table 2.10 shows that, with daily rebalancing, the equal-weighted unconstrained (UC, can take both long and short positions) and long-only (LO) portfolios perform very well. They both have mean gross and risk-adjusted returns in the region of 12% per year. However, the turnover of the unconstrained portfolio is rather high at 7% per day and it has a slightly higher maximum

Table 2.10 Fully invested portfolios using the most recent day's declarations - daily rebalancing

(a) Fully invested portfolios

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val	Turn	MDD
Equal-weighted											
UC	0.129	0.006	1.085	0.126	0.001	0.114	0.006	0.126	0.003	0.070	0.160
LO	0.117	0.009	1.151	0.116	0.001	0.112	0.002	0.125	0.001	0.009	0.127
AS	0.108	0.019	1.059	-	-	-	-	-	-	0.000	0.139
Value-weighted											
UC	0.054	0.129	0.665	0.051	0.072	0.042	0.159	0.055	0.072	0.032	0.139
LO	0.072	0.075	0.760	0.070	0.027	0.065	0.048	0.079	0.022	0.013	0.155
AS	0.079	0.068	0.783	0.078	0.023	0.076	0.034	0.089	0.016	0.009	0.167

(b) Differences between portfolios

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val
Equal-weighted									
UC-AS	0.022	0.358	0.525	0.019	0.461	0.009	0.713	0.008	0.749
LO-AS	0.010	0.038	1.266	0.009	0.062	0.007	0.160	0.007	0.157
FI-LO	0.012	0.536	0.349	0.010	0.644	0.003	0.903	0.001	0.953
Value-weighted									
UC-AS	-0.025	0.159	-0.749	-0.027	0.104	-0.034	0.049	-0.034	0.045
LO-AS	-0.008	0.191	-0.744	-0.008	0.126	-0.011	0.057	-0.011	0.052
FI-LO	-0.018	0.147	-0.744	-0.019	0.097	-0.024	0.047	-0.024	0.043

(c) Expected shortfalls (daily returns)

	ES_{99}	ES_{95}	ES_{90}
Equal-weighted			
UC	0.029	0.016	0.012
LO	0.028	0.016	0.012
AS	0.030	0.017	0.013
Value-weighted			
UC	0.031	0.018	0.014
LO	0.030	0.018	0.014
AS	0.032	0.019	0.014

Performance evaluation measures for fully invested portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002). "Turn" denotes daily turnover and "MDD" maximum drawdown. Panel (a) shows the mean gross and risk-adjusted returns to the unconstrained (UC), long-only (LO) and all stock (AS) portfolios in isolation. Alphas are not reported for the all share portfolios since the risk models all contain a market factor. In panel (b), row "X-Y" shows the difference in returns between X and Y. The Sharpe ratio is the Sharpe ratio of a portfolio which is £1 long in X and £1 short in Y. Panel (c) shows the expected shortfall at the 99% (ES_{99}), 95% (ES_{95}) and 90% (ES_{90}) level.

drawdown than the other two portfolios. All three portfolios have near-identical tail risk in terms of 90%, 95% and 99% expected shortfall, and relatively similar maximum drawdowns.

The unconstrained and long-only mean returns compare favourably with the almost 11% per year average return to the equal-weighted portfolio of all stocks (AS) in the FTSE350. Nonetheless, the difference in means is not significant for the unconstrained portfolio. It also does not survive transaction costs if these are 50bps each way.[14,15] While the long-only portfolio outperforms the all share portfolio significantly in terms of mean returns, the significance does not survive risk adjustment. The unconstrained portfolio does not significantly outperform the long-only one. There is no discernible profitability in using the net short position disclosures to form a fully invested equal-weighted portfolio, or in allowing such a portfolio to go short.

For the value-weighted portfolios, the picture is even gloomier. The unconstrained and long-only portfolios underperform versus the all share portfolio. Depending on the factor model used, this underperformance can be statistically significant. Moreover, the long-only portfolio outperforms the unconstrained one. The expected shortfalls and maximum drawdowns remain similar between the three portfolios. The net short disclosures do not appear to bring any profitable information on a value-weighted basis, either.

The high turnover of the equal-weighted unconstrained portfolio highlights that daily rebalancing may expose the fully invested portfolios to excessive noise. I therefore consider rebalancing the portfolios monthly and six-monthly, following the scheme in Sections 2.3.1 and 2.3.2. I leave these results untabulated.

Rebalancing at the monthly frequency improves both gross and risk-adjusted returns to the equal-weighted unconstrained portfolio. Rebalancing less frequently also cuts the equal-weighted unconstrained portfolio turnover to 2.4% per day and its maximum drawdown falls, too. The equal-weighted long-only and all share portfolio returns change little. The unconstrained portfolio's advantage over the equal-weighted long-only and all share portfolios in terms of average returns increases. However, this outperformance remains statistically insignificant. The long-only portfolio continues to significantly outperform the all share portfolio in terms of gross mean returns but not risk-adjusted returns. Moving to six-monthly rebalancing, the equal-weighted unconstrained portfolio does significantly outperform the long-only portfolio, although this significance does not survive 50bps each-way transaction costs.

For the value-weighted portfolios, the long-only and all share portfolios' returns fall somewhat when rebalancing less frequently. The unconstrained portfolio's returns fall a little.

[14] 50bps each way is a common assumed level of transaction costs in the literature. However, my conversations with practitioners suggest this is somewhat higher than the level investors typically face.

[15] I leave results using 50bps each-way transaction costs untabulated throughout in the interests of space.

The net effect is that all three portfolios produce very similar returns to each-other, both in gross and risk-adjusted returns.

Adding the stop-loss rules described in Sections 2.3.1 and 2.3.2 to these less frequently rebalanced portfolios has almost no impact on their gross and risk-adjusted returns.

I also consider volatility scaling the fully invested portfolios, as in Sections 2.3.1 and 2.3.2. Volatility scaling the daily rebalanced fully invested portfolios reduces all the equal-weighted returns, while having only a minor impact on the value-weighted returns. The differences in performance among the equal-weighted portfolios remain roughly the same as in the non-volatility scaled case. The performance of the value-weighted unconstrained and long-only portfolios remains poor: they continue to underperform the value-weighted all stock portfolio.

The one time the equal-weighted unconstrained and long-only portfolios do significantly outperform the equal-weighted all share portfolio is when the portfolios are rebalanced less frequently (monthly or six-monthly) and are volatility scaled. This statistically significant outperformance occurs both in terms of gross and risk-adjusted returns. With 50bps each-way transaction costs, the equal-weighted unconstrained portfolio significantly outperforms the long-only and all share portfolios only in terms of mean returns, not alphas. Volatility scaling causes little change to the returns to the less frequently rebalanced value-weighted portfolios.

Combining volatility scaling and stop-loss rules in the less frequently rebalanced portfolios makes very little difference to the gross and risk-adjusted returns and return differences.

2.4.2 Portfolios using multiple days' declarations

Table 2.11 shows that the unconstrained multiple signals portfolio insignificantly outperforms the long-only multiple signals portfolio and the equal-weighted all share portfolio. The long-only multiple signals portfolio itself outperforms the equal-weighted all share portfolio. However, the value-weighted all share portfolio insignificantly outperforms both the unconstrained and long-only multiple signals portfolios.

The tail risks are similar within each comparison set. The long-only and unconstrained multiple signals portfolios both have similar expected shortfalls. These are also similar to the equal-weighted all share portfolio's expected shortfalls. The maximum drawdowns of these three portfolios are similar, too. Likewise, the value-weighted unconstrained and long-only multiple signals portfolios have similar expected shortfalls. These are similar to the value-weighted all share portfolio's expected shortfalls. And the three value-weighted portfolios also have similar maximum drawdowns.

Table 2.11 Fully invested portfolios using multiple days' declarations - daily rebalancing

(a) Fully invested portfolios

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val	Turn	MDD
Multiple signals approach											
UC	0.123	0.004	1.276	0.118	0.003	0.104	0.010	0.105	0.014	0.102	0.256
LO	0.107	0.008	1.135	0.105	0.001	0.101	0.002	0.112	0.001	0.005	0.119
AS	0.108	0.019	1.059	-	-	-	-	-	-	0.000	0.139
Value-weighted multiple signals approach											
UC	0.052	0.125	0.696	0.049	0.083	0.039	0.191	0.051	0.087	0.031	0.136
LO	0.073	0.069	0.771	0.071	0.026	0.066	0.046	0.079	0.022	0.011	0.155
AS	0.079	0.068	0.783	-	-	-	-	-	-	0.009	0.167

(b) Differences between portfolios

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val
Multiple signals approach									
UC-AS	0.016	0.688	1.035	0.011	0.788	-0.001	0.987	-0.013	0.732
LO-AS	-0.001	0.895	-0.085	-0.002	0.772	-0.004	0.555	-0.005	0.449
UC-LO	0.017	0.618	1.874	0.013	0.710	0.004	0.915	-0.007	0.815
Value-weighted multiple signals approach									
UC-AS	-0.027	0.185	-0.689	-0.030	0.115	-0.037	0.064	-0.038	0.053
LO-AS	-0.006	0.296	-0.618	-0.007	0.206	-0.009	0.110	-0.010	0.084
UC-LO	-0.021	0.159	-0.690	-0.022	0.097	-0.028	0.058	-0.028	0.050

(c) Expected shortfalls (daily returns)

	ES_{99}	ES_{95}	ES_{90}
Multiple signals			
UC	0.026	0.014	0.011
LO	0.038	0.021	0.015
AS	0.030	0.017	0.013
VW multiple signals			
UC	0.031	0.018	0.014
LO	0.030	0.018	0.015
AS	0.032	0.019	0.014

Performance evaluation measures for fully invested portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002). "Turn" denotes daily turnover and "MDD" maximum drawdown. Panel (a) shows the mean gross and risk-adjusted returns to the unconstrained (UC), long-only (LO) and all stock (AS) portfolios in isolation. Alphas are not reported for the all share portfolios since the risk models all contain a market factor. In panel (b), row "X-Y" shows the difference in returns between X and Y. The Sharpe ratio is the Sharpe ratio of a portfolio which is £1 long in X and £1 short in Y. Panel (c) shows the expected shortfall at the 99% (ES_{99}), 95% (ES_{95}) and 90% (ES_{90}) level. "VW" stands for "value-weighed".

Rebalancing the portfolios less frequently does not see the fully invested portfolios based on declared net short positions significantly outperform the all share portfolios. Nor does adding the stop-loss rules. The tail risks remain similar in each comparison set when rebalancing less frequently, with and without stop-loss rules.

Volatility scaling the daily rebalanced portfolios does, however, improve the performance of the unconstrained and long-only portfolios relative to the all share portfolios. I show this in Table 2.12. The unconstrained multiple signals portfolio now outperforms the equal-weighted all share portfolio significantly in gross terms, although not significantly in risk-adjusted terms. The difference in mean returns is now 3.2% per year. The long-only multiple signals portfolio outperforms the all shares portfolio by a more modest 1.2% a year. This difference is, however, statistically significant for gross and risk-adjusted returns and this significance survives 50bps each way transaction costs. The unconstrained multiple signals portfolio continues to outperform the long-only portfolio statistically insignificantly.

For the value-weighted portfolios, the outperformance of the all share portfolio compared to the other two is reduced. The underperformance of the value-weighted unconstrained multiple signals portfolio against the long-only portfolio is also reduced. The tail risks in each comparison group continue to be similar, hence I suppress them.

Untabulated results show that rebalancing the volatility scaled portfolios monthly or six-monthly produces qualitatively similar outcomes. The unconstrained portfolios benefit most from less frequent rebalancing, then the long-only portfolios and then the all share portfolios. Therefore, both the standard and value-weighted unconstrained multiple signals portfolios perform better relative to their long-only and all share counterparts than with daily rebalancing. Likewise, the standard and value-weighted long-only multiple signals portfolios now perform better relative to their all share counterparts. In fact, the value-weighted unconstrained portfolio marginally outperforms the long-only and all share portfolios with monthly or six-monthly rebalancing. Similarly, the value-weighted long-only portfolio marginally outperforms the all share portfolio. Adding stop-loss rules to the less frequently rebalanced volatility scaled unconstrained and long-only portfolios harms their performance.

2.5 Robustness

The key results - that standard long-short portfolios based on public net short position declarations are not profitable in the UK and using these declarations to form fully invested portfolios gives no great advantage, either - are robust to the various portfolio formation and data choices made in the preceding Sections.

Table 2.12 Fully invested portfolios using multiple days' declarations - daily rebalancing and volatility scaling

(a) Fully invested portfolios

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val	Turn	MDD
Multiple signals approach											
UC	0.129	0.000	1.613	0.125	0.000	0.117	0.000	0.124	0.000	0.066	0.090
LO	0.109	0.003	1.321	0.107	0.000	0.103	0.001	0.113	0.000	0.019	0.104
AS	0.097	0.013	1.111	-	-	-	-	-	-	0.015	0.118
Value-weighted multiple signals approach											
UC	0.065	0.061	0.819	0.062	0.030	0.053	0.070	0.063	0.035	0.034	0.126
LO	0.072	0.052	0.846	0.069	0.019	0.064	0.038	0.075	0.017	0.016	0.142
AS	0.076	0.084	0.738	-	-	-	-	-	-	0.009	0.169

(b) Differences between portfolios

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val
Multiple signals approach									
UC-AS	0.032	0.029	1.350	0.030	0.050	0.024	0.118	0.019	0.190
LO-AS	0.012	0.005	1.566	0.012	0.005	0.010	0.015	0.009	0.032
UC-LO	0.020	0.065	1.196	0.018	0.111	0.014	0.229	0.010	0.338
Value-weighted multiple signals approach									
UC-AS	-0.010	0.528	-0.365	-0.013	0.384	-0.019	0.212	-0.022	0.138
LO-AS	-0.004	0.686	-0.202	-0.005	0.523	-0.008	0.361	-0.010	0.261
UC-LO	-0.006	0.402	-0.600	-0.007	0.322	-0.011	0.137	-0.012	0.088

Performance evaluation measures for volatility scaled fully invested portfolios over the sample period 1 November 2013 - 13 December 2018. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002). "Turn" denotes daily turnover and "MDD" maximum drawdown. Panel (a) shows the mean gross and risk-adjusted returns to the unconstrained (UC), long-only (LO) and all stock (AS) portfolios in isolation. Alphas are not reported for the all share portfolios since the risk models all contain a market factor. In panel (b), row "X-Y" shows the difference in returns between X and Y. The Sharpe ratio is the Sharpe ratio of a portfolio which is £1 long in X and £1 short in Y.

2.5.1 Closing prices

Perhaps the biggest empirical choice made in the earlier sections is to use the $t+2$ day opening prices for back-testing, where t is the day the short position was taken/adjusted. The alternative is to use the $t+1$ closing price, and assume the FCA information is published and analysed by traders before the market close. Using $t+1$ closing prices instead of $t+2$ opening prices has relatively little impact on the results.

Table 2.13 shows some improvement for certain value-weighted long-short portfolios when using $t+1$ closing prices. Looking at the daily rebalanced value-weighted portfolio based on the most recent declarations, the loss on the short side drops to 1.2% per year. The profit to the long-short portfolio becomes 7%. The Sharpe ratio also improves to 1.1. The mean return is significantly different from zero at the 5% level. However, the alphas range from 2% (four-factor model) to 5.4% (three factors) and these are not close to being significant at any conventional level.

There is a similar improvement for the value-weighted regression-based long-short portfolios. These portfolios are based on multiple days' declarations. The long basket return increases to 12.4% per year and the short basket loss falls to 4.1% per year, leaving a net long-short portfolio profit of 8.3% per year. Unlike the $t+2$ opening prices case, this is significantly different from zero at the 5% level, as are the alphas. Only the QMJ alpha has a t-statistic in excess of 3.0. However, the short side of the portfolio still loses 4% a year.

The improvements in these value-weighted portfolios' performance remain with volatility scaling, although the non-scaled portfolios perform better in terms of means and alphas. However, the improvements do not survive less frequent rebalancing. Stop-loss rules does not help the less frequently rebalanced value-weighted portfolios regain their strong daily rebalanced performance.

For the equal- and net short positions-weighted long-short portfolios, however, everything remains more or less the same as when using $t+2$ opening prices. There is a slight deterioration in long-short performance on average, but it is small. Moreover, there is little change in the performance of the multiple signals and non-value weighted regression-based portfolios. If anything, these perform slightly worse overall. This lack of difference to the results in Sections 2.3.1 and 2.3.2 remains with volatility scaling, less frequent rebalancing and stop-loss rules.

The results for the fully invested portfolios are very similar whether I use $t+1$ closing prices or $t+2$ opening prices.

Table 2.13 Value-weighted long-short portfolios using $t+1$ closing prices

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val
Value-weighted portfolio based on most recent day's declaration									
Long	0.082	0.040	0.882	0.068	0.049	0.056	0.131	0.082	0.027
Short	0.012	0.855	0.059	0.014	0.743	0.036	0.417	0.047	0.271
L-S	0.070	0.048	1.110	0.054	0.152	0.020	0.579	0.035	0.361
Value-weighted regression-based portfolio									
Long	0.124	0.008	1.176	0.112	0.001	0.110	0.003	0.133	0.000
Short	0.041	0.454	0.359	0.037	0.364	0.041	0.328	0.055	0.187
L-S	0.083	0.009	1.729	0.076	0.004	0.069	0.005	0.078	0.002

Daily rebalanced portfolios over the sample period 1 November 2013 - 13 December 2018. Performance evaluation measures for declared net short positions portfolios. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

2.5.2 Portfolio formation

Short position measure

One way to circumvent many of the assumptions made to turn the net short position disclosures into a continuous series for each stock is to use the number of distinct investors with net short positions above the declaration threshold as the short position measure instead. Doing so has little impact on the results. In some cases the long-short portfolios improve in performance a little. In other cases they worsen. There are no clear differences overall. Notice that stocks with zero declared net short positions in aggregate also have zero investors with declared net short positions. Therefore the fully invested portfolio results are entirely unchanged.

High/low total net short position threshold

Returning to forming portfolios based on aggregate declared net short positions per stock, the results are robust to the high net short position threshold. I consider using the 70th and 90th percentiles of declared total net short positions as the cut-off for being placed in the short basket. The long legs of the portfolios (and therefore the fully invested portfolio results) are unchanged, since far more than 30% of stocks have zero declared net short positions on any given day. Generally, the short side losses decrease when using the 90th percentile as the threshold. However, they remain positive and economically substantial and this improvement is not uniform across the various portfolio formation methodologies. Likewise, the short side losses increase slightly when the 70th percentile is the cut-off. Again, these differences are economically fairly small. The high net short position threshold does not affect how I form the fully invested portfolios, and so these results remain unchanged.

Forming portfolios based on changes in declared net short positions

It is possible that there is information in the change in declared net short positions above and beyond what is encapsulated in the level. The strategies in this paper are rebalanced up to the daily frequency and changes in declared net short positions may have greater predictive power over short-run trends. In fact, Boehmer et al. (2008) and Diether et al. (2009) use (scaled) changes in short interest for their main results. (Au et al., 2009, use the level of short positions in their UK study.) However, since positions must only be reported once their size crosses certain thresholds, most daily changes in declared total net short positions are zero. Even the 95th cross sectional percentile of the change in declared total net short positions is zero on 94% of trading days in my sample, while the 80th and 90th percentiles are always zero. Using daily changes in total declared net short positions would not be a feasible strategy here.

The weekly change in total declared net short positions suffers a similar issue: its 80th percentile is always zero and its 90th percentile is zero on 66% of trading days. Even monthly changes in total declared net short positions suffer an issue of lack of information. The 80th percentile of this series is zero on 87% of days and the 90th is zero on 36% of days. There would not be enough days with stocks in the short basket even using the 90th percentile of the monthly change in total declared net short positions for meaningful back tests.

2.5.3 Sensitivity of multiple signals and regression-based portfolios to set of horizons

Long-short portfolios

For the multiple signals portfolios, I consider a reduced set of horizons of up to three months ($\mathscr{H} = \{1,2,5,10,21,42,63\}$ days) and an increased set of horizons of up to two years ($\mathscr{H} = \{1,2,5,10,21,42,63,126,189,252,378,504\}$ days). The main findings are robust to the horizon choice. The long-short portfolios do not generally make gross or risk-adjusted returns significantly different from zero. Volatility scaling, rebalancing less frequently and using stop-loss rules do not change this.

When it comes to the regression-based portfolios, I use an extended set of horizons of up to nine months ($\mathscr{H} = \{1,5,21,63,126,189,252,378\}$ days) and a shortened set of up to three months ($\mathscr{H} = \{1,5,21,63\}$ days). The above conclusions are robust to using either the reduced or extended set of horizons. In fact, the returns to the daily rebalanced long-short portfolios are markedly lower for the reduced set of horizons. Otherwise, like in Section 2.3.2, volatility scaling has little impact on the portfolios, rebalancing less frequently harms portfolio performance and stop-loss rules do not reverse this.

Fully invested portfolios

For the fully invested portfolios based on multiple days' declarations, the daily rebalanced portfolio returns are very similar when using $\mathscr{H} = \{1,2,5,10,21,42,63\}$ days. However, rebalancing less frequently gives the unconstrained and equal-weighted long-only portfolios less of a boost, and any advantage to the short disclosure-based portfolios becomes statistically insignificant. Volatility scaling the portfolio weights gives returns extremely similar to when volatility scaling the weights based on the baseline $\mathscr{H}$ ($\mathscr{H} = \{1,2,5,10,21,42,63,126,189, 252\}$ days).

Extending the set of horizons considered to $\mathscr{H} = \{1,2,5,10,21,42,63,126,189,252,378, 504\}$ days leads to rather high portfolio turnover and more extreme versions of the results presented for the baseline $\mathscr{H}$. Rebalancing less frequently ameliorates, but does not cure, this turnover issue. Applying volatility scaling to the weights resolves the problem, and returns results very similar to those when volatility scaling the weights based on the baseline $\mathscr{H}$.

2.6 Day-t strategies

I have so far focussed on strategies which trade on disclosures either at the market opening two days after the disclosure is made, or at the close of trading the day after the disclosure is made. This is because strategies trading earlier than this are not feasible, as the disclosures are published with a lag (see Section 2.2.1). However, we have seen that feasible long-short strategies are not profitable, mainly because of heavy losses to the short leg, and feasible fully invested portfolios do not reliably outperform the market. These findings are very robust. To investigate whether the publication lag causes this disappointing performance, I now suspend reality and assume that a disclosure becomes available on the day it was required (i.e. the day the position in question was taken, or day t in Section 2.2.1's terminology).

I start with the daily rebalanced long-short portfolios which use a single day's declarations. The losses to the short baskets, and subsequent non-profitability of these long-short portfolios, are not caused by the publication lag. The results are very similar to those using the $t+1$ closing prices (see Section 2.5.1). Compared the the baseline $t+2$ opening prices case in

Table 2.14 Value-weighted long-short portfolios using day-t closing prices and day-t declarations

	Mean	p-val	Sharpe	α^{FF3}	p-val	α^{FF4}	p-val	α^{QMJ}	p-val	
Non-volatility scaled										
Long	0.082	0.040	0.880	0.171	0.068	0.049	0.056	0.132	0.082	0.027
Short	-0.001	0.986	-0.034	0.953	0.003	0.955	0.034	0.535	0.041	0.449
L-S	0.084	0.119	0.896	0.282	0.065	0.226	0.022	0.656	0.042	0.432
Volatility scaled										
Long	0.081	0.038	0.905	0.162	0.065	0.069	0.050	0.183	0.072	0.053
Short	-0.006	0.932	-0.071	0.890	-0.007	0.888	0.013	0.796	0.022	0.651
L-S	0.087	0.017	1.291	0.111	0.072	0.102	0.037	0.375	0.050	0.250

Daily rebalanced portfolios over the sample period 1 November 2013 - 13 December 2018. Performance evaluation measures for declared net short positions portfolios. Means and alphas are scaled by 252 days to make them approximately annual figures. Mean and alpha p-values are based on daily returns and are HAC. The Sharpe ratio is the annualised Sharpe ratio, allowing for serial correlation, computed as per Lo (2002).

Section 2.3.1, the returns to the equal- and net short positions-weighted long and short baskets remain similar. The returns to the value-weighted long basket increase to 8.2% a year, while the losses to the short basket fall to 1.2%. The average annual return of the value-weighted long-short portfolio is 7.0% and is significantly different from zero. However, the alphas remain insignificant. Volatility scaling the portfolios does not enable any of the long baskets to make a profit, nor does rebalancing the portfolios less frequently. This latter finding is unsurprising, since the gain to more timely access to information is diluted as the portfolio is rebalanced less frequently.

Returning to daily rebalancing, Table 2.14 shows that using the 90th percentile of declared open net short positions as the cut-off for forming the long and short baskets does allow the value-weighted short basket to make a small profit. This does not happen when using the $t+1$ closing prices or $t+2$ opening prices. At 0.1% per year, however, the profit is minuscule. Volatility scaling the portfolio increases the profit to 0.6% a year. However, the profits are not large enough for either the volatility scaled or non-volatility scaled long-short portfolio to have an alpha significantly different from zero. The volatility scaled long-short portfolio does, however, have a mean return which is significant. Whether volatility scaled or not, the equal- and net short positions-weighted short baskets continue to make substantial losses. Rebalancing the portfolios less than daily sees all the short baskets (including the value-weighted short basket) make substantial losses, too.

These stubborn losses to the short baskets may seem to suggest that the actual (declared) net short positions taken were not profitable on average. This is not necessarily the case, however, as the strategies discussed here do not perfectly track the declared positions - the entry (when the position first rises above the 0.5% threshold) and exit (when the position falls below the threshold) times are not necessarily the same. Thus, the actual positions may be timed better than the hypothetical positions considered here. Jank and Smajlbegovic (2017) do track the entry and exit times of the actual positions and find them to be profitable on average. However, they consider positions taken in all EU-listed stocks rather than positions taken in only UK-listed stocks. Moreover, the profitability of the actual positions is not typically significant.

In common with the less frequently rebalanced long-short portfolios using a single day's declarations, the short baskets in the long-short strategies based on multiple days' declarations continue to make substantial losses. Moreover, the fully invested portfolios - whether equal- or value-weighted, whether formed using a single day's declarations or multiple days' declarations, whether allowed to take short positions or not - continue to underperform the market. This is hardly a surprise. Changing the assumed disclosure availability has a relatively small impact on the composition of the long basket, which makes up the majority the fully invested portfolios which can take short positions (and, of course, 100% of the long-only portfolios).

2.7 Conclusion

I examine whether public net short position disclosures can be used as the basis for a profitable investment strategy in the UK. New rules introduced in 2012 mean that all net short positions above 0.5% of issued share capital must be publicly disclosed, making freely available for the first time information about large net short positions. There is a clear practical interest in evaluating the profitability of this new information.

In general, regulatory net short position disclosures do not form the basis of profitable long-short portfolios, in either gross or risk-adjusted terms. Even where there are statistically significant average profits to these strategies, the short sides of the portfolios lose a considerable amount of money. If an investor wants a zero initial outlay portfolio, she would surely be better off financing the investments in the long side of the portfolio by borrowing, rather than taking the short positions considered here.

The regulatory net short position disclosures do not form the basis of fully invested (unit initial outlay) portfolios that substantially outperform comparable portfolios of all stocks, either. When such fully invested portfolios are allowed to take short positions, they do not, in general, significantly outperform comparable portfolios of all stocks. Certain long-only portfolios formed on the basis of the disclosures do tend to significantly outperform comparable portfolios of all stocks. However, such outperformance is typically economically modest: around one percentage point per year.

The practical implications of this research are clear. For the UK at least, the regulatory net short position disclosures do not seem to form the basis of a profitable short selling signal or a profitable buy signal. The lack of profitability of the short selling signal is clear from the fact that the short side of the long-short portfolios loses money in every formulation considered. For the buying signal, while the long sides of the long-only portfolios do well when considered in isolation, the long-only portfolios only modestly outperform portfolios of all stocks. There is little economic gain to buying stocks on the basis of this buy signal, as opposed to simply buying all stocks.

The short selling signal may not be profitable because, by the time the net short positions exceed the disclosure threshold, there is already as much short interest in the stock as the market can profitably handle.That is, the strategy described in this paper is simply too late to the party. After all, we know from Jank and Smajlbegovic (2017) that the disclosed short positions are profitable on average and that positions taken earlier perform better overall.

It is also possible that the bull market which coincided with the sample period made it an unusually difficult time to profit from short positions, making the losses in the short baskets of the strategies considered potentially anomalous. Of course, this possibility cannot be ruled out with the data available. However, while the market rose strongly overall over the whole sample period, on any given day there were typically many individual stocks which did fall in value. The short baskets could still have made money by identifying these stocks. On the median day, 145 of the 351 stocks in the sample fell in value (152 stocks fell on the mean day). The lower quartile of the number of stocks to fall in value on a given day was 101, while the number ranged from 8 to 338. So the short baskets did have the opportunity to make money, had the signal been more informative as to which stocks were likely to fall in value on a given day.

As for why the buy signal is not too informative, about 70% of stocks have a buy signal on the average day. It is therefore not necessarily surprising that this portfolio should not perform much differently to a portfolio of all stocks. The portfolios are rather similar. In any case, the information generating the buy signal is public and easy to act on - certainly easier to act on than a short selling signal - and so perhaps any information in the signal has already been exploited by the time the marginal investor (what I study here) arrives.

Chapter 3

The value of using predictive information optimally

For mean-variance investors, using predictive information unconditionally optimally produces better portfolios than using predictive information conditionally optimally. The latter is more usually done in practice. Empirically, the unconditionally optimal portfolios have higher Sharpe ratios and certainty equivalents than the conditionally optimal portfolios. They also have lower turnover, leverage, losses and drawdowns. Moreover, measures of the whole distribution tend to prefer the unconditionally optimal portfolios, especially once transaction costs are accounted for. With transaction costs, the unconditionally optimal portfolios often second-order stochastically dominate the conditionally optimal portfolios. The unconditionally optimal portfolios are also preferred in terms of Sharpe ratio, certainty equivalent, costs, losses, drawdowns and stochastic dominance to mean-variance optimal portfolios that do not use predictive information. However, whether unconditionally optimal portfolios are preferred to minimum variance or $1/N$ portfolios depends on the asset universe.

JEL classification: G11, G14, G17
Keywords: conditional efficiency, unconditional efficiency, signal, predictive information, prediction, risk-return trade-off, mean-variance

3.1 Introduction

A mean-variance optimiser is better off using predictive information unconditionally optimally than conditionally optimally. Using the information unconditionally optimally produces portfolios with a better risk-to-reward profile in terms of Sharpe ratios and certainty equivalents. Moreover, the unconditionally optimal portfolios have lower turnover and leverage and therefore lower transaction costs. They also produce lower expected shortfalls (extreme losses) and drawdowns (runs of losses). In addition, the utopia index - a measure of almost stochastic dominance (the entire return distribution) - tends to prefer unconditionally optimal portfolios, too. Especially when transaction costs are considered, the unconditionally optimal portfolios often second-order stochastically dominate the conditionally optimal ones.

The question of how using predictive information unconditionally, rather than conditionally, optimally affects portfolio performance is an important one. Most studies that consider the gain to a mean-variance optimiser of using predictive information consider conditionally optimal portfolios (e.g. Allen et al., 2019). However, in the plausible scenario where an uninformed investor, with no access to predictive information, delegates to an informed manager, with access to predictive information, the investor will assess the manager in terms of the unconditional mean and variance of the portfolio. Furthermore, in practice, studies such as Allen et al. (2019) tend to evaluate conditionally optimal portfolios in terms of statistics that are a function of unconditional moments (e.g. a Sharpe ratio depending on the unconditional mean and variance of the portfolio's excess return). Yet, Hansen and Richard (1987) show that the conditionally optimal portfolio is not necessarily unconditionally optimal. From an unconditional perspective, the conditionally optimal portfolio may not use predictive information optimally, and the gain to using the information will be understated.

In addition, similar to Ferson and Siegel (2001) and Abhyankar et al. (2012), I show that unconditionally optimal weights have a conservative response property. The unconditionally optimal weights are substantially less sensitive to changes in the predicted returns than the conditionally optimal weights are over the most likely values of the predicted returns. This conservative response can reduce turnover and provide a degree of robustness to estimation error which is useful in practice, beyond considerations of whether the investor is motivated by the unconditional or conditional moments of their portfolio.

Empirically, I show that the unconditionally optimal portfolios are also much preferred to the mean-variance optimal portfolios which ignore the conditioning information (the no-information optimal portfolios). This is in terms of Sharpe ratio, certainty equivalent, utopia index, costs, drawdowns and losses. Subject to fees, the investor is always better off delegating to an informed manager or using predictive information.

How the unconditionally optimal portfolio compares to non-optimal benchmarks depends on the asset universe. The non-optimal benchmarks I use are the $1/N$ portfolio, the minimum conditional variance portfolio and then no-information minimum variance portfolio (i.e. the minimum unconditional variance portfolio). It is important to compare to these benchmarks as DeMiguel et al. (2009) show that mean-variance optimal portfolios can be subject to severe estimation error, meaning non-optimal portfolios can perform better in practice.

In the event, the unconditionally optimal portfolios outperform the non-optimal benchmarks in universes of size/book-to-market double-sorted portfolios but not universes of industry portfolios. This outperformance survives transaction costs of 50bps each-way in the smaller size/book-to-market universe and 10bps each-way in the larger size/book-to-market universe.

The asset means and variances are less dispersed in the industry universes, meaning the non-optimal allocations are closer to the true optimal allocations (Kirby and Ostdiek, 2012a). This lack of dispersion could therefore explain why the non-optimal portfolios outperform the unconditionally optimal portfolio, even with zero costs, in the industry but not size/book-to-market universes.

In the empirical work, I use four asset universes: six size/book-to-market double-sorted portfolios and 25 size/book-to-market portfolios, as well as 10 industry portfolios and 30 industry portfolios. The time period is January 1990 to December 2019. My predictive information is the lagged market return, which I use in a univariate linear model. This predictor gives the best information coefficient in all universes. I reoptimise and rebalance the portfolios each month, and use the standard conditional variance estimator, which does not account for possible conditional heteroscedasticity.

The results are robust to using alternative predictors with positive information coefficients. These are reversal and one/12-month trend change, which I use in univariate predictive models, and a machine learning predictor which combines information from 10 variables using an elastic net-targeted random forest.

I compute the information coefficients in one-step-ahead rolling predictive regressions over the same sample as I run the portfolio strategies and focus only on cases where the information coefficient is positive. I do so since the question of the benefits of using predictive information unconditionally optimally relies on having valid predictive information in the first place. This approach may introduce look-ahead bias into the comparisons between portfolios using the predictive information and those not using it. These comparisons effectively condition on the predictor used having a positive information coefficient in the sample and remain interesting for two reasons. First, it is not a given that using a even predictor with a "guaranteed" positive information coefficient will lead to better portfolio performance if estimating the predictive model introduces additional estimation error. Second, we see that portfolios using the predictive information typically struggle to outperform the minimum variance portfolio that does not use this information. This finding is particularly damning given that the predictor has been chosen on the basis of in-sample predictive performance. Moreover, the failure of the machine learning predictor to improve portfolio performance is all the more striking given that the machine learning predictor was chosen in the basis of in-sample predictive performance.

Reoptimising and rebalancing the portfolios each quarter, rather than each month, improves Sharpe ratios and certainty equivalents across the board, but leaves the relative preferences for the portfolios largely unchanged. Since return volatility falls with quarterly rebalancing, the p-values on the various comparisons tend to fall as well.

Using an asymmetric dynamic conditional correlation model to account for possible conditional heteroscedasticity actually harms portfolio performance, but does not affect the relative ranking of the portfolios.

The literature regarding the unconditionally optimal use of predictive information dates back to Hansen and Richard (1987). They show that the portfolio that uses predictive information unconditionally optimally is different to the portfolio which uses this information conditionally optimally. The solution to the conditional mean-variance problem is not necessarily the solution to the unconditional mean-variance problem. Ferson and Siegel (2001) derive the unconditionally optimal weights explicitly.

Abhyankar et al. (2012) are the first to analyse the benefit of using predictive information unconditionally optimally, rather than conditionally optimally. They find that using predictive information unconditionally, rather than conditionally, optimally portfolio improves performance both in- and out-of-sample. They compare the maximum Sharpe ratio attainable using the predictive information unconditionally optimally to that using the information conditionally optimally. While this provides a convenient measure of the expansion of the mean-variance frontier, the two maximum Sharpe ratio portfolios may have different mean targets and therefore aggressiveness. In this sense, the maximum Sharpe ratio approach may not be comparing like with like, which Kirby and Ostdiek (2012a) show can lead to misleading comparisons. For this reason, I use portfolios with the same mean target.

Chiang (2015) considers a variation of the Ferson and Siegel (2001) problem. Rather than considering returns in excess of the risk-free rate, Chiang (2015) considers returns in excess of a (potentially time-varying) benchmark. This extension is useful since many fund managers' performance is benchmarked in practice. Like Abhyankar et al. (2012), Chiang (2015) finds in- and out-of-sample information ratio gains to using predictive information unconditionally, rather than conditionally, optimally. The information ratio is the analogue of the Sharpe ratio in the presence of a benchmark.

I go beyond comparing Sharpe/information ratios by comparing the entire distributions of returns through the utopia index (Anderson et al., 2019). The utopia index comparisons turn out to be statistically more powerful than those based on Sharpe ratios in this paper. They are also more general than considering Sharpe/information ratios, which are specific functions of means and variances. I also go beyond linear prediction methods by considering very recent machine learning methods. Ironically, these do not help portfolio performance.

Other work considers the effectiveness of using predictive information unconditionally optimally. Kirby and Ostdiek (2012b) consider the optimal use of predictive information from an unconditional perspective as part of a broader framework that accounts for estimation risk, specification error and transaction costs directly in the (unconditional) optimisation. Kirby and Ostdiek (2012b) find that portfolios using predictive information optimally in this setting outperform the $1/N$ portfolio and the S&P500 index. They do not consider other benchmarks. Zhou (2008) formulates a new version of the fundamental law of active management that maximises the unconditional value added, in the spirit of Ferson and Siegel (2001). This new law performs better in simulations than one based on conditional value-added.

The rest of this paper proceeds as follows. Section 3.2 discusses the portfolios I consider. Section 3.3 outlines how I compare portfolios. Section 3.4 describes the sample data and predictors used. Section 3.5 gives the main results and Section 3.6 verifies these are robust. Section 3.7 concludes.

3.2 Portfolios

3.2.1 Portfolio construction with predictive information

Consider a mean-variance investor. At time t, she chooses the vector of portfolio weights $w_t = (w_{1,t}, \ldots, w_{N,t})'$ which minimises her portfolio variance subject to the mean target μ_P. Note that I consider only single-period problems: the t subscript on the weights indicates that they may depend on predictive information available at time t.

Suppose further that the excess returns (returns in excess of the risk-free rate) $R_t = (R_{1,t}, \ldots, R_{N,t})'$ are generated by

$$R_{t+1} = \mu(S_t) + \varepsilon_{t+1}, \tag{3.1}$$

where $\mu(S_t) = \mathrm{E}(R_{t+1}|S_t)$ is the time t conditional mean of R_{t+1}, $\mu(\cdot)$ is a function, S_t is a vector, matrix, or array of predictors, $\mathrm{E}(\varepsilon_{t+1}|S_t) = 0$ and $\mathrm{Var}(\varepsilon_{t+1}|S_t) = \Sigma_\varepsilon(S_t)$. The portfolio construction results below allow for both homoscedastic and conditionally heteroscedastic ε_{t+1}, so $\mathrm{Var}(\varepsilon_{t+1}|S_t)$ can depend on S_t.

So that the allocation to the risk-free asset does not distort my results, I consider an investor investing only in risky assets. My investor therefore faces the budget constraint that her portfolio weights must sum to one: $\iota' w_t = 1$, where ι is an N-length vector of ones. I work in terms of excess returns as the investor allocates only over risky assets.

Ferson and Siegel (2001) Theorem 3 shows that the portfolio weights using the predictive information S_t that minimise the unconditional portfolio variance $\mathrm{Var}(w_t' R_{t+1})$ subject to the unconditional mean target $\mathrm{E}(w_t' R_{t+1}) \geq \mu_P$ and the budget constraint $\iota' w_t = 1$ are

$$w_t^{UO} = \frac{\Lambda(S_t)\iota}{\iota'\Lambda(S_t)\iota} + \frac{\mu_P - \alpha_2}{\alpha_3}\left(\Lambda(S_t) - \frac{\Lambda(S_t)\iota\iota'\Lambda(S_t)}{\iota'\Lambda(S_t)\iota}\right)\mu(S_t). \tag{3.2}$$

where

$$\begin{aligned}
\alpha_2 &= \mathrm{E}\left(\frac{\iota'\Lambda(S_t)\mu(S_t)}{\iota'\Lambda(S_t)\iota}\right) \\
\alpha_3 &= \mathrm{E}\left[\mu(S_t)'\left(\Lambda(S_t) - \frac{\Lambda(S_t)\iota\iota'\Lambda(S_t)}{\iota'\Lambda(S_t)\iota}\right)\mu(S_t)\right] \\
\Lambda(S_t) &= \left[\mathrm{E}(R_{t+1}R_{t+1}'|S_t)\right]^{-1}.
\end{aligned}$$

I term these weights the unconditionally optimal (UO) weights.

These UO weights are in contrast the optimal weights for an investor who minimises the *conditional* portfolio variance $\mathrm{Var}(w_t' R_{t+1}|S_t)$ subject to the *conditional* mean target $\mathrm{E}(w_t' R_{t+1}|S_t) \geq \mu_P$ and the budget constraint $\iota' w_t = 1$. In this case, Kirby and Ostdiek (2012a) show the optimal weights are

$$w_t^{CO} = \left(\frac{\mu_P - \mu_{CMIN,t}}{\mu_{CTP,t} - \mu_{CMIN,t}}\right) w_t^{CTP} + \left(1 - \frac{\mu_P - \mu_{CMIN,t}}{\mu_{CTP,t} - \mu_{CMIN,t}}\right) w_t^{CMIN}. \tag{3.3}$$

I term these weights the conditionally optimal (CO) weights. w_t^{CTP} gives the time t conditional tangency portfolio (CTP) weights, and w_t^{CMIN} the global minimum conditional variance (CMIN) weights. The CTP maximises the conditional Sharpe ratio $\mathrm{E}(w_t' R_t|S_t)/\sqrt{\mathrm{Var}(w_t' R_t|S_t)}$ (recall that R_t denotes excess returns). $\mu_{CMIN,t} = \mathrm{E}[(w_t^{CMIN})' R_{t+1}|S_t]$ and $\mu_{CTP,t} = \mathrm{E}[(w_t^{CTP})' R_{t+1}|S_t]$.

The CMIN and CTP portfolios have weights:

$$w_t^{CTP} = \frac{\Sigma_\varepsilon(S_t)^{-1}\mu(S_t)}{\iota'\Sigma_\varepsilon(S_t)^{-1}\mu(S_t)} \qquad w_t^{CMIN} = \frac{\Sigma_\varepsilon(S_t)^{-1}\iota}{\iota'\Sigma_\varepsilon(S_t)^{-1}\iota}. \tag{3.4}$$

In general, the CO weights differ from the UO weights: the CO weights are not necessarily UO (Hansen and Richard, 1987). In practice, predictive information is often used *conditionally* optimally, rather than unconditionally optimally (e.g. Allen et al., 2019, consider conditionally optimal strategies using predictive information). Yet there are good reasons to consider the UO weights.

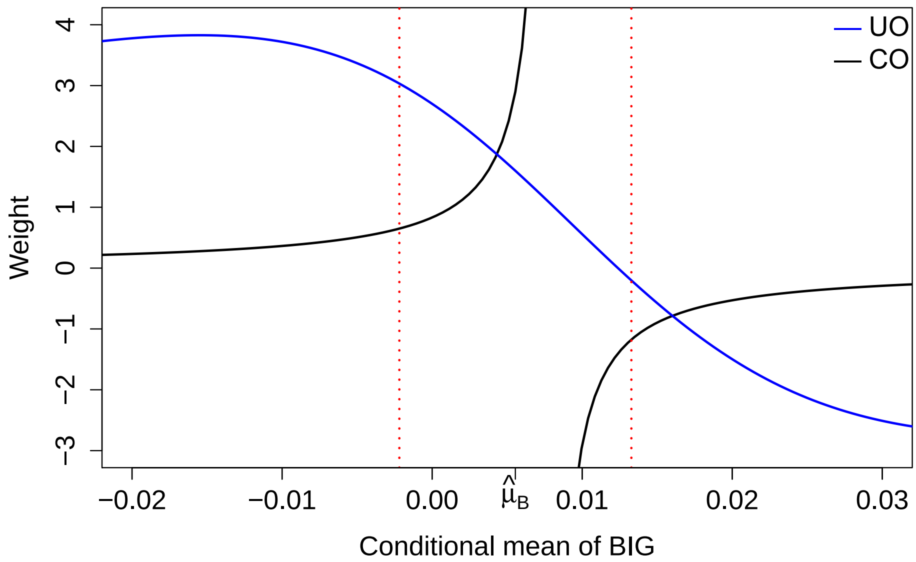
First, the benefit of using predictive information to form conditionally optimal portfolios is often assessed using statistics which are functions of *unconditional* portfolio means and variances (e.g. in Allen et al., 2019). Since we would expect the unconditionally optimal portfolios to perform better on unconditional performance measures, such studies likely understate the usefulness of predictive information.

Second, using predictive information to optimise an objective function which depends on unconditional moments reflects the situation facing an investment manager with access to predictive information investing on behalf of an uninformed client, who does not have access to this information (Ferson and Siegel, 2001). This is a situation of practical interest for fund managers. It is reasonable to assume that one reason that clients are willing to pay investment managers' fees is because the investment manager is better informed. If the client does not have access to the predictive information, the client will assess the manager in terms of unconditional performance. Therefore, the manager should use the predictive information unconditionally optimally.

Finally, the UO weights also have a conservative response property (Abhyankar et al., 2012; Ferson and Siegel, 2001). This property can provide a useful defence against estimation error coming from estimating $\mu(S_t)$. If the portfolio weights are very sensitive to even small changes in the conditional mean, they are also very sensitive to estimation error. A small amount of estimation error would lead to weights a long way from the true optimum. In addition, the conservative response property helps to control turnover and therefore transaction costs.

I illustrate the UO portfolio's conservative response property in Figure 3.1. In this example, I allocate between the BIG and SMALL portfolios: the large-cap and small-cap sides of the Fama-French *SMB* factor, respectively. I predict the returns to the BIG and SMALL portfolios using the lagged market return as the predictive variable in 120-month rolling window regressions and assuming homoscedastic $\varepsilon_{i,t}$. I use data from January 1990-December 2019. My mean target is 0.13% per month, which is the mean of the optimal portfolio for a mean-variance

Fig. 3.1 Unconditionally and conditionally optimal portfolio weights as a function of the conditional mean for the BIG portfolio of large stocks



The black line ($-1/x$-like shape) shows the conditionally optimal (CO) weight for BIG, while the blue line (sigmoid-like shape) shows the unconditionally optimal (UO) weight. The CO curve has a vertical asymptote at the sample mean of SMALL $\hat{\mu}_S = 0.0078$. $\hat{\mu}_B = 0.0055$ is the sample mean of BIG. The red dotted lines are one standard deviation of the conditional mean either side of $\hat{\mu}_B$.

optimiser without access to predictive information and a risk aversion of 5.[1] The sample mean return to the BIG portfolio is 0.55% per month, marked $\hat{\mu}_B$ in Figure 3.1.

We can see that the CO portfolio weights respond considerably to small changes in the conditional mean of BIG around $\hat{\mu}_B$, which the conditional mean is centred on. In fact, the CO response function has a vertical asymptote at $\hat{\mu}_S = 0.78\%$, which is very close to $\hat{\mu}_B$. Around this asymptote, a small change in the conditional mean of BIG can see BIG go from having an extreme positive weight to an extreme negative weight, or vice-versa. By contrast, the UO weights are much less sensitive to changes in the conditional mean of BIG around $\hat{\mu}_B$. We see this over almost all of the range between the red dotted lines, which show one standard deviation of the BIG conditional mean either side of $\hat{\mu}_B$. The UO weights also have no asymptote: the UO weights are bounded from above by 3.8 and below by -2.7.

[1]This mean-variance optimiser's problem is

$$\max_w \left\{ \mathrm{E}(w'R_t) - \frac{5}{2}\mathrm{Var}(w'R_t) \right\}.$$

Away from $\hat{\mu}_B$, in the region of the more extreme and less likely values of the conditional mean of BIG, the CO weights do respond a bit more conservatively to changes in the conditional mean than the UO weights. However, the difference in the strength of response between the CO and UO weights is smaller than where the UO weights respond more conservatively. Moreover, the region where the CO weights respond more conservatively corresponds to less likely realisations of the conditional mean. Therefore, while it is empirically possible that CO weights overall respond more conservatively to changes in the conditional mean, and so suffer less from estimation error and have lower turnover, it is unlikely.

3.2.2 No-information benchmark

The UO weights use predictive information to minimise (with respect to w) the unconditional portfolio variance $\mathrm{Var}(w'R_t)$, subject to the unconditional mean target $\mathrm{E}(w'R_t) \geq \mu_P$ and budget constraint $\iota' w = 1$. I consider as a benchmark the weights which minimise the same unconditional portfolio variance, subject to the same unconditional mean target and the same budget constraint, but which do not use the predictive information. I term these weights the no-information optimal (NIO) weights.

The NIO weights depend only on the unconditional means $\mathrm{E}(R_t) = \mu$ and variances $\mathrm{Var}(R_t) = \Sigma$ of asset returns, and are (Kirby and Ostdiek, 2012a):

$$w^{NIO} = \left(\frac{\mu_P - \mu_{NIMIN}}{\mu_{NITP} - \mu_{NIMIN}}\right)\left(\frac{\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\mu}\right) + \left(1 - \frac{\mu_P - \mu_{NIMIN}}{\mu_{TP} - \mu_{NIMIN}}\right)\left(\frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}\right).$$

In practice, I estimate μ and Σ with rolling window estimators and update the estimates of μ and Σ, and therefore the weights, each period.

NIMIN denotes the no-information global minimum variance portfolio and NITP the no-information tangency portfolio. $\mu_{NIMIN} = \mathrm{E}[(w^{NIMIN})'R_t]$ denotes and $\mu_{NITP} = \mathrm{E}[(w^{NITP})'R_t]$. The NITP maximises the unconditional Sharpe ratio $\mathrm{E}(w'R_t)/\sqrt{\mathrm{Var}(w'R_t)}$ and the NIMIN portfolio is the global minimum unconditional variance portfolio (it minimises $\mathrm{Var}(w'R_t)$). The NITP and NIMIN portfolios have weights:

$$w^{NITP} = \frac{\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\mu} \qquad\qquad w^{NIMIN} = \frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota} \tag{3.5}$$

I consider the NIO benchmark since the NIO portfolio represents the optimal portfolio for an uninformed investor who does not delegate to an informed manager. Therefore, the difference in performance between the UO and NIO portfolios is the value of delegating to the informed manager. More generally, this difference measures the value of having predictive information versus not having it, while keeping the objective function and constraints the same in both cases.

3.2.3 Mean target

The UO, CO and NIO portfolios above depend on the mean target μ_P. I ensure all three portfolios have the same mean target, in order to control their aggressiveness. Portfolios with higher mean targets are more aggressive. Kirby and Ostdiek (2012a) show that more aggressive portfolios are more exposed to estimation error and it is therefore important to compare "like with like", in terms of mean targets.

I derive the mean target from the mean-variance utility optimisation problem of an investor without access to predictive information, so that the mean target is

$$\mu_P = \frac{\iota'\Sigma^{-1}\mu}{\gamma}\left(\frac{\mu'\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\mu}\right) + \left(1 - \frac{\iota'\Sigma^{-1}\mu}{\gamma}\right)\left(\frac{\iota'\Sigma^{-1}\mu}{\iota'\Sigma^{-1}\iota}\right).$$

μ_P is the mean return to the portfolio that maximises, with respect to w, the certainty equivalent

$$CEQ_\gamma(w) = w'\mu - \frac{\gamma}{2}w\Sigma w' \tag{3.6}$$

where γ represents the investor's risk aversion.

I use the problem for an investor without access to predictive information to represent the situation of an uninformed investor delegating to an informed manager. Here, the investor gives the manager a target rate of return and the manager's task is to achieve that return for the lowest possible risk.

Note that, in this case, neither the UO nor CO portfolios will be utility-optimal for the investor (uninformed principal) with risk aversion γ. The UO portfolio will, however, give the lowest unconditional variance given the unconditional mean target. Likewise, the CO portfolio will yield the lowest conditional variance given the conditional mean target. In this sense, the UO and CO portfolios are mean-variance optimal.

The non-utility optimality of the UO and CO portfolios comes from the fact that the usual Lagrangian duality results that mean that the utility maximisation problem has the same solution as an appropriately constrained variance minimisation problem do not hold when the information sets underlying the two problems are different. For similar reasons, the NIO portfolio *is* utility-optimal for the (uninformed) investor with risk aversion γ studied here.

That the UO and CO portfolios are not utility-optimal for our uninformed investor information is not a concern here. The purpose of giving each portfolio the same mean target is to ensure comparable aggressiveness. The reason for deriving that mean target from a utility optimisation problem is to ensure that it is economically grounded.

3.2.4 Other benchmarks

I consider three non-optimal benchmarks for the UO portfolio: the $1/N$ portfolio, the CMIN portfolio in (3.4) and the NIMIN portfolio in (3.5). DeMiguel et al. (2009) show that estimation error can be a severe problem for mean-variance optimal portfolios, leading to non-optimal allocations, such as a simple $1/N$ portfolio, outperforming the theoretically optimal allocations empirically. Minimum variance portfolios can be seen as an aggressive form of shrinkage, which assumes that all assets in a universe have the same mean. Since DeMiguel et al. (2009) find that estimation error in the mean is the main issue for mean-variance allocations, using minimum variance portfolios could provide a challenging benchmark, too.

I also consider CTP and NITP, defined in (3.4) and (3.5) respectively, as mean-variance optimal benchmarks for the UO portfolio. I consider these portfolios as benchmarks as, without estimation error, the CTP and NITP maximise the conditional and unconditional Sharpe ratio, respectively. However, the CTP and NITP tend to have extreme mean targets and so are very sensitive to estimation error, meaning they do not necessarily have the highest Sharpe ratios in practice (Kirby and Ostdiek, 2012a).

3.3 Portfolio performance measures

I consider three sets of performance measures: standard mean-variance performance measures, measures of tail risk and a measure of almost stochastic dominance, which compares the entire distributions of returns.

The mean-variance performance measures are the Sharpe ratio and the certainty equivalent, which is calculated as the sample analogue of (3.6). I compute HAC p-values for the differences between the UO Sharpe ratio and the CO and (no-information and other) benchmark Sharpe ratios in bivariate comparisons. I also compute HAC p-values for the differences between the UO certainty equivalent and the CO and benchmark certainty equivalents in bivariate comparisons. The null in each case is that the difference is zero and the p-values are for two-tailed tests. My method for computing these p-values is in Appendix B.1.

Mean-variance measures are most natural in this setting, since the portfolios are designed to minimise (conditional or unconditional) variance given a (conditional or unconditional) mean target. The Sharpe ratio has the advantage of being independent of risk aversion and shows the risk-to-return ratio of the portfolio. For a given mean, a mean-variance investor prefers the portfolio with a higher Sharpe ratio. The certainty equivalent does depend on risk aversion. It is a risk-adjusted return. The difference in certainty equivalent returns between two portfolios can be interpreted as the maximum management fee (as a percentage of assets) that an investor would pay to switch from one portfolio to the other.

To illustrate the practicalities of the portfolios, I also report their leverage and turnover. Leverage shows us the percentage of the investor's assets that must be borrowed to build the portfolio. In practice, borrowing capacity is finite and portfolios with very high leverage may be infeasible. Turnover will give an indication of transaction costs. A standard transaction cost assumption is 50bps each-way, proportional to turnover (e.g. DeMiguel et al., 2009). However, 10bps each-way may be more reasonable, as discussed in Chapter 2.

Beyond mean-variance measures of risk, I also consider tail risk. To do this, I look at the profit and loss distribution, computing the empirical 99%, 95% and 90% expected shortfalls. These are the expected losses given that returns are in the worst 1%, 5% and 10% of their empirical distributions, respectively.

In addition, I consider drawdowns. These measure runs of losses, computed from peak to trough. I compute the empirical 90%, 95% and 99% percentiles of portfolio drawdowns and the maximum drawdowns. The maximum drawdown shows the loss suffered by an investor who times their investment perfectly badly: the worst loss suffered by buying at a peak and selling at a trough.

It is important to consider tail risk. While mean-variance investing is widely used in practice, the risk that investors face in practice is fundamentally asymmetric. Once an investor loses 100% of their capital, they can invest no further and therefore cannot benefit from future upturns. Using variance as a risk measure does not capture this effect since variance is symmetric about the mean.

Losses represent a form of tail risk. Drawdowns are another and perhaps of even more practical interest, since they consider runs of losses. In practice, investors do not hold a portfolio for one period only, but for successive periods. Drawdowns, then, capture the tail risk associated with timing entry and exit poorly: suffering runs of losses by buying high and selling low.

I do not confine my examination of non-variance risk to tail risk only. I also consider the whole return distribution. I compute the utopia index (Anderson et al., 2019), a measure of almost stochastic dominance.[2] I use the utopia index to compare the UO portfolio to the CO and benchmark portfolios in bivariate comparisons.[3]

Portfolio p second-order stochastically dominates (SOSD) q if its second-order integrated cumulative distribution function (CDF) lies below that of q. If p SOSD q, any risk-averse investor prefers to invest in p rather than q. Almost stochastic dominance refers to how much of the domain of the second-order integrated CDF of p is below that of q. This is normalised to be a percentage in the utopia index, so that the utopia index of p and the utopia index of q sum to one. Specifically, the utopia index of portfolio p is defined as

$$\begin{aligned}
\mathscr{I}_p &= 1 - \frac{\mathscr{A}_p}{\mathscr{T}} \\
\mathscr{A}_p &= \int_{x \in \mathscr{X}} \left(G_p(x) - \underline{\mathscr{G}}(x)\right) dx \\
\mathscr{T} &= \int_{x \in \mathscr{X}} \left(\bar{\mathscr{G}}(x) - \underline{\mathscr{G}}(x)\right) dx
\end{aligned} \tag{3.7}$$

$G_p(x)$ is the second-order integrated CDF for p, $\mathscr{X}$ the range of all possible values of returns, $\bar{\mathscr{G}}(x)$ the upper envelope of the second-order integrated CDFs for all portfolios being compared and $\underline{\mathscr{G}}(x)$ the lower envelope. Since it is clear from (3.7) that $\mathscr{A}_p \leq \mathscr{T}$, it is also clear that $0 \leq \mathscr{I}_p \leq 1$. The best case scenario (p SOSD q) has $G_p(x) = \underline{\mathscr{G}}(x)$ for all $x \in \mathscr{X}$. Therefore a higher $\mathscr{I}_p$ indicates a better portfolio.

In practice, the utopia index must be computed over empirical integrated CDFs. This makes it a statistic with a distribution and the empirical utopia index is subject to estimation error. For inference, I use the sub-sampling procedure described in Anderson et al. (2019) for weakly dependent data. I use the same tuning parameter values as Anderson et al. (2019).

Using the utopia index allows for more general comparisons than the Sharpe ratio or certainty equivalent, since SOSD relates to preferences for all risk-averse investors. Moreover, the utopia index turns out to provide more statistically powerful comparisons than the Sharpe ratio or certainty equivalent in this paper.

[2] What I refer to as simply the utopia index, Anderson et al. (2019) term the "second-degree utopia index". I use the utopia index only in the context of second-order stochastic dominance, which is what the second-degree utopia index tests for in Anderson et al. (2019).

[3] The utopia index can handle comparisons between more than two alternatives, but I focus on bivariate comparisons. While the set of alternatives makes no difference to portfolio utopia index rankings in the population, it does in finite samples. In particular, including one portfolio which has a much worse utopia index than all of the others in the set skews the comparisons and makes it difficult to rank the alternatives which are not clearly awful. Solving this problem requires iterative deletion of portfolios which typically results in bi- or trivariate comparisons. In any case, there is a natural portfolio to compare against in all cases in this paper: the UO portfolio.

3.4 Data

I consider four asset universes. They all come from Kenneth French's website and are: the six and 25 value-weighted size/book-to-market double-sorted portfolios, and the 10 and 30 value-weighted industry portfolios. Since I allocate over risky assets only, I work with excess returns and subtract the risk-free rate (the one-month Treasury bill rate, computed by Ibbotson Associates and obtained via Kenneth French's website) from the portfolio returns.

My sample for returns is January 1990-December 2019 (360 months). Since I lag all predictors by one period, the sample for predictors is December 1989-November 2019. I consider portfolios that re-optimise and rebalance each month. All statistics (means, Sharpe ratios, etc.) presented are monthly statistics.

3.4.1 Predictors

I consider a suite of ten predictive variables. Note that the predictive models used to generate the results in the rest of this paper do not necessarily use all of the predictive variables. Which variables are used in each case is specified with the relevant results. Following Abhyankar et al. (2012), I consider the following six economic variables, each with a one-month lag: the return to the market index (from CRSP), the dividend yield on the market index (computed from CRSP data), the one-month Treasury bill rate (from Ibbotson Associates, via Kenneth French's website), CPI inflation (from FRED), the term spread (10-year Treasury yield minus one-year Treasury yield, computed using FRED data) and the credit spread (10-year BAA corporate bond spread minus 10-year Treasury yield, computed using FRED data).[4]

I also consider four technical predictors, following Neely et al. (2014). These measures are: reversal (previous month's return), momentum (previous 12 months' cumulative return), one/12-month trend change ($MA_{i,t}^{1,12}$) and three/12-month trend change ($MA_{i,t}^{3,12}$). For asset i,

$$MA_{i,t}^{s,12} = \frac{1}{s}\sum_{k=1}^{s} R_{i,t-k} - \frac{1}{12}\sum_{k=1}^{12} R_{i,t-k}. \tag{3.8}$$

$MA_{i,t}^{s,12}$ is the difference between the s-month moving average and the 12 month moving average. These moving average differences can be interpreted as capturing changes in trends, since the shorter moving average is "more sensitive to recent price movements" than the longer one (Neely et al., 2014).

[4]Due to missing data over the period of the financial crisis in the longer horizon (e.g. 20- or 30-year) FRED Treasury yield series, it is not possible to use exactly the same credit spread measure as Abhyankar et al. (2012), nor to compute Abhyankar et al.'s convexity of the term structure measure.

It is important to consider technical predictors as well as economic predictors. Neely et al. (2014) show that technical predictors provide an important complement to economic predictors when it comes to predicting stock returns. Moreover, Neely et al. find that technical indicators are, in fact, stronger predictors of stock returns than economic predictors.

3.5 Results

I evaluate the effectiveness of unconditionally optimal (UO) portfolio compared to the conditionally (CO) and no-information (NIO) optimal portfolios, the tangency portfolios and the other benchmarks in an out-of-sample exercise. This removes concerns regarding look-ahead bias. I estimate the parameters of $\mu(S_t)$ and $\Sigma_\varepsilon(S_t)$, as well as μ and Σ, using a rolling window approach. That is, I estimate μ and Σ and the parameters of $\mu(S_t)$ and $\Sigma_\varepsilon(S_t)$ using data from $t-b$ to $t-1$. The results I present below use a window of $b = 120$, although the results are robust to using a 60-month look-back period ($b = 60$) or an expanding window ($b = t-1$) - see Section 3.6.2. In all three cases, I reserve the first 120 observations for the initial estimation window. Therefore, all the returns series for the portfolios run from January 2000 to December 2019 and each returns series is 240 observations in length.

In the main results below, I use a linear, univariate conditional mean function where the sole predictive variable is the (lagged) market return. I choose this specification since, in each universe, it has the best information coefficient (correlation between predicted and realised returns) of all the predictive models I consider. The results below are robust to using other predictors with a positive information coefficient - see Sections 3.5.3 and 3.6.3. (Details of the predictive models considered and their information coefficients are in Appendix B.2.)

I compute the information coefficients based on one-step-ahead out-of-sample predictions, produced using rolling windows of length 120, using the same sample as for the portfolio returns. This means that the first returns I forecast are in January 2000 (where I estimate the predictive model using data up to December 1999) and the final returns I forecast are in December 2019. The reason for computing the information coefficients over the same period I run the portfolio strategies is that studying the value of using predictive information, and comparing different methods of using that information, implicitly assumes, and really relies on, the information being used having predictive power, i.e. a positive information coefficient. As discussed in Section 3.1, this may introduce a degree of look-ahead bias into the comparisons between portfolios using the predictive information and those not using it. These comparisons look at the value of using predictive information compared to not using it *conditional on that* *information having predictive power*. There is unlikely to be much value in using "predictors" which do not have predictive power!

I treat ε_t as homoscedastic so that $\widehat{\Sigma}_\varepsilon(S_t) = b^{-1}\sum_{s=1}^{b}(R_{t+1-s} - \hat{\mu}(S_{t-s}))(R_{t+1-s} - \hat{\mu}(S_{t-s}))'$, where $\hat{\mu}(S_t)$ is the estimated conditional mean of R_{t+1} given S_t. Section 3.5.2 shows that allowing for conditional heteroscedasticity in ε_t does not change the key conclusions of the results.

At each t, I estimate μ and Σ with the standard estimators $\hat{\mu} = b^{-1}\sum_{s=1}^{b} R_{t-s}$ and $\widehat{\Sigma} = b^{-1}\sum_{s=1}^{b}(R_{t-s} - \hat{\mu})(R_{t-s} - \hat{\mu})'$, respectively. Note that the estimates of μ and Σ change each period as new information arrives.

I re-estimate the weights and rebalance each month and use a risk aversion of $\gamma = 5$ for the UO, CO and NIO portfolios. Section 3.5.1 shows moving to quarterly rebalancing does not change the overall pattern of the results. Section 3.6.1 shows that the results are robust to using a risk aversion of $\gamma = 1$ or $\gamma = 10$.

The results I present are calculated in the absence of transaction costs. I do, however, consider the impact of transactions costs throughout the discussion of the results. I leave the results calculated with transaction costs untabulated in the interests of space. The transaction costs I consider are proportional to turnover at both the standard level of 50bps each-way and also the lower, but potentially more realistic, level of 10bps each-way (see Chapter 2 for a greater discussion of a realistic level of each-way transaction costs). Where I do not explicitly mention a cost level, the discussion takes the costs to be zero.

Tables 3.1-3.4 show the portfolio performance across the four universes (six and 25 size/book-to-market portfolios, and 10 and 30 industry portfolios) using the univariate (lagged) market return predictor. Overall, we clearly see that the UO portfolios have the best Sharpe ratios and certainty equivalents of all the mean-variance optimal portfolios, including the tangency portfolios. They also have the lowest turnover and leverage, so will have the lowest costs. The preference for the UO portfolios over other mean-variance optimal portfolios is increasing in transaction costs.

In the six size/book-to-market universe, the UO portfolio has a higher Sharpe ratio and certainty equivalent than both the CO and NIO portfolios, while having lower turnover and leverage than both of these portfolios, too. None of these differences is significant, even when 10bps of 50bps each-way transactions costs are accounted for. However, the difference in Sharpe ratio between UO and NIO without costs is economically meaningful: nearly 10 percentage points. Further, the UO portfolio has less severe extreme losses and drawdowns than either CO or NIO.

Table 3.1 Six size/book-to-market portfolios

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.019	0.052	0.370	0.013	0.437	3.257
CO	0.022	0.064	0.348	0.012	0.636	4.659
p-value			0.524	0.820		
NIO	0.020	0.074	0.273	0.007	0.941	6.064
p-value			0.176	0.220		
$1/N$	0.007	0.049	0.137	0.001	0.053	0.000
p-value			0.026	0.028		
CMIN	0.011	0.037	0.302	0.008	0.190	1.446
p-value			0.369	0.171		
NIMIN	0.011	0.037	0.301	0.008	0.190	1.450
p-value			0.366	0.168		
CTP	0.002	0.249	0.007	-0.153	4.084	8.952
p-value			0.002	0.287		
NITP	0.016	0.057	0.276	0.008	0.621	4.322
p-value			0.161	0.175		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.100	0.075	0.043	0.230	0.149	0.111	0.043
CO	0.118	0.089	0.053	0.303	0.186	0.165	0.101
NIO	0.140	0.107	0.069	0.371	0.243	0.206	0.134
$1/N$	0.111	0.090	0.058	0.558	0.313	0.260	0.116
CMIN	0.079	0.061	0.037	0.342	0.177	0.115	0.041
NIMIN	0.079	0.061	0.037	0.335	0.171	0.115	0.041
CTP	0.633	0.360	0.166	1.064	1.039	0.984	0.974
NITP	0.113	0.086	0.054	0.307	0.178	0.143	0.072

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	0.462	0.538	0.000	0.000	0.950
NIO	0.943	0.057	0.083	0.000	0.000
$1/N$	1.000	0.000	1.000	0.000	0.000
CMIN	0.867	0.133	0.177	0.000	0.077
NIMIN	0.866	0.134	0.171	0.000	0.077
CTP	1.000	0.000	1.000	0.000	0.000
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.2 25 size/book-to-market portfolios

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.047	0.128	0.369	0.006	2.920	14.255
CO	0.048	0.142	0.336	-0.003	3.776	17.050
p-value			0.218	0.177		
NIO	0.057	0.212	0.267	-0.056	7.239	22.193
p-value			0.112	0.002		
$1/N$	0.007	0.051	0.145	0.001	0.057	0.000
p-value			0.063	0.655		
CMIN	0.011	0.037	0.295	0.008	0.337	2.934
p-value			0.491	0.916		
NIMIN	0.011	0.037	0.299	0.008	0.328	2.861
p-value			0.517	0.907		
CTP	0.109	0.666	0.164	-0.999	4.229	15.891
p-value			0.008	0.310		
NITP	0.026	0.076	0.338	0.011	1.168	7.764
p-value			0.611	0.568		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.214	0.165	0.096	0.468	0.413	0.342	0.198
CO	0.265	0.201	0.118	0.640	0.570	0.535	0.338
NIO	0.392	0.302	0.183	0.782	0.714	0.669	0.460
$1/N$	0.115	0.093	0.060	0.560	0.305	0.249	0.117
CMIN	0.076	0.060	0.038	0.341	0.236	0.137	0.050
NIMIN	0.076	0.060	0.037	0.345	0.230	0.132	0.056
CTP	0.236	0.164	0.090	0.673	0.410	0.229	0.143
NITP	0.128	0.100	0.062	0.268	0.217	0.183	0.118

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	0.888	0.112	0.425	0.000	0.088
NIO	0.712	0.288	0.028	0.000	0.182
$1/N$	0.926	0.074	0.387	0.000	0.000
CMIN	0.880	0.120	0.370	0.000	0.061
NIMIN	0.879	0.121	0.370	0.000	0.061
CTP	0.007	0.993	0.000	0.519	0.033
NITP	0.844	0.156	0.320	0.000	0.017

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.3 10 industry portfolios

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.007	0.045	0.153	0.002	0.193	1.248
CO	0.006	0.052	0.107	-0.001	0.250	1.687
p-value			0.266	0.162		
NIO	0.003	0.067	0.046	-0.008	0.360	2.163
p-value			0.154	0.055		
$1/N$	0.006	0.041	0.140	0.002	0.048	0.000
p-value			0.884	0.943		
CMIN	0.007	0.033	0.208	0.004	0.089	0.531
p-value			0.463	0.449		
NIMIN	0.007	0.033	0.209	0.004	0.090	0.536
p-value			0.459	0.445		
CTP	0.024	0.277	0.087	-0.167	1.294	4.527
p-value			0.547	0.181		
NITP	0.004	0.059	0.074	-0.004	0.283	1.492
p-value			0.415	0.352		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.103	0.083	0.051	0.381	0.295	0.220	0.154
CO	0.125	0.099	0.062	0.432	0.374	0.324	0.252
NIO	0.159	0.127	0.084	0.718	0.691	0.660	0.596
$1/N$	0.093	0.076	0.048	0.496	0.334	0.246	0.118
CMIN	0.076	0.060	0.037	0.326	0.235	0.183	0.072
NIMIN	0.076	0.060	0.037	0.319	0.227	0.177	0.076
CTP	0.434	0.264	0.134	1.568	1.274	1.225	1.187
NITP	0.139	0.104	0.068	0.623	0.572	0.530	0.400

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	1.000	0.000	0.000
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.421	0.579	0.000	0.000	1.000
CMIN	0.000	1.000	0.000	1.000	0.000
NIMIN	0.000	1.000	0.000	1.000	0.000
CTP	0.394	0.606	0.000	0.000	0.724
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.4 30 industry portfolios

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.013	0.097	0.131	-0.011	1.030	5.618
CO	0.011	0.111	0.100	-0.020	1.377	6.993
p-value			0.356	0.072		
NIO	0.003	0.161	0.019	-0.061	2.654	9.266
p-value			0.122	0.002		
$1/N$	0.007	0.046	0.143	0.001	0.057	0.000
p-value			0.923	0.205		
CMIN	0.006	0.034	0.167	0.003	0.168	1.210
p-value			0.761	0.124		
NIMIN	0.006	0.034	0.167	0.003	0.168	1.204
p-value			0.765	0.126		
CTP	-0.039	1.482	-0.026	-5.527	17.315	19.854
p-value			0.173	0.233		
NITP	-0.066	1.244	-0.053	-3.933	13.999	14.247
p-value			0.078	0.285		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.180	0.151	0.101	0.683	0.645	0.633	0.550
CO	0.216	0.185	0.125	0.841	0.813	0.795	0.749
NIO	0.379	0.293	0.193	0.990	0.987	0.982	0.978
$1/N$	0.107	0.083	0.052	0.544	0.267	0.211	0.092
CMIN	0.075	0.062	0.039	0.308	0.209	0.167	0.095
NIMIN	0.075	0.063	0.039	0.318	0.209	0.165	0.098
CTP	1.000	1.000	0.706	37.238	15.148	1.893	1.217
NITP	1.000	1.000	0.638	23.598	7.136	0.998	0.928

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	0.569	0.000	0.000
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.336	0.664	0.000	0.000	0.652
CMIN	0.348	0.652	0.000	0.000	0.541
NIMIN	0.348	0.652	0.000	0.000	0.541
CTP	1.000	0.000	1.000	0.000	0.000
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Considering the whole distribution, it has hard to rank UO and CO. Both estimated utopia indices are close to 0.5 and not significantly different from each other. We reject the null that UO stochastically dominates CO and vice-versa. With costs of 50bps, though, UO does appear to stochastically dominate CO. We do not reject the null that UO stochastically dominates CO, and UO has a significantly better utopia index.

In addition, there is evidence that UO stochastically dominates NIO, even without costs. The utopia indices are significantly different from each other with UO having the higher utopia index. Moreover, we do not (quite) reject the null that UO stochastically dominates NIO at the 5% level. With transaction costs, we do not reject the null that UO stochastically dominates NIO at any conventional level. The UO portfolio also has a significantly better utopia index than NIO for all levels of transaction costs.

We clearly prefer the UO portfolio to either of the tangency portfolios. The UO portfolio has a somewhat better Sharpe ratio and certainty equivalent than either CTP or NITP, along with lower costs. UO also has lower extreme losses and drawdowns and, even without costs, clearly second order stochastically dominates both portfolios.

We also clearly prefer UO to $1/N$: it has a significantly higher Sharpe ratio and certainty equivalent, lower losses and drawdowns and stochastically dominates $1/N$. We also prefer the UO portfolio to the minimum variance portfolios in terms of Sharpe ratio, certainty equivalent and utopia index, albeit insignificantly. These preferences remain even with 50bps each-way costs. With 50bps each-way costs, however, the UO portfolio's utopia index is no longer significantly greater than that of $1/N$. Being more defensive, the minimum variance portfolios do have lower losses and drawdowns, even before costs.

The results for the 25 size/book-to-market universe are fairly similar. The UO portfolio continues to have the best Sharpe ratio of any portfolio, although not the highest certainty equivalent. The minimum variance portfolios now have the highest certainty equivalents.

UO continues to be preferred to CO and NIO in terms of Sharpe ratio, certainty equivalent and extreme losses and drawdowns, while having lower turnover and leverage. Now, though, the UO portfolio also has a significantly better certainty equivalent than NIO without costs, and a significantly better Sharpe ratio with 50bps each-way costs. In addition, the estimated utopia index now favours UO over CO and NIO, although the differences in utopia index are not significant. With 10bps each-way costs, however, UO has a significantly better utopia index and stochastically dominates both CO and NIO.

While the UO portfolio has an insignificantly better Sharpe ratio than NITP, it has an inferior certainty equivalent. Nonetheless, UO has lower costs and the utopia index shows it stochastically dominates NITP. NITP does, however, have lower losses and drawdowns than UO.

Unusually, CTP has a significantly better utopia index than UO, despite having a significantly worse Sharpe ratio, even when accounting for the CTP's much higher transaction costs. The UO portfolio does, though, have lower drawdowns and similar losses to CTP before costs.

Again, we prefer UO to $1/N$ and the minimum variance portfolios in terms of Sharpe ratio and utopia index. These preferences survive 10bps each-way transaction costs, but are reversed with 50bps each-way costs. Only the without-cost utopia index difference between UO and $1/N$ is significant at the 5% level. The $1/N$ and minimum variance portfolios are more defensive, so have lower losses and drawdowns.

Shifting to the industry universes, we again find that the UO portfolio is the best mean-variance portfolio. In both industry universes, it has a better Sharpe ratio, certainty equivalent and utopia index than CO, NIO, CTP or NITP and lower costs. It also produces lower extreme losses and drawdowns. While the Sharpe ratio and certainty equivalent differences are insignificant, the UO portfolio does have a significantly higher utopia index than, and stochastically dominate, CO, NIO and NITP in both industry universes before costs. The UO and CTP utopia indices are not significantly different in the 10 industry universe without costs. However, with 10bps each-way costs, the UO portfolio has a significantly higher utopia index than and stochastically dominates the CTP.

The big difference between the size/book-to-market and industry universes is that, in the industry universes, the minimum variance portfolios outperform the UO portfolio in terms of Sharpe ratio and utopia index without costs, while also having lower turnovers and leverages. The $1/N$ and minimum variance portfolios also have lower drawdowns and losses.

Mean-variance optimisation does not perform well in the industry universes. This is in-line with Kirby and Ostdiek (2012a), who argue that mean-variance optimisation works better in size/book-to-market universes because there is greater dispersion in the means and variances of the size/book-to-market portfolios. Therefore, the minimum variance and $1/N$ allocations are further from the optimum and more estimation error is required for these non-optimal allocations to outperform mean-variance optimisation.

3.5.1 Quarterly rebalancing

It is possible that re-optimising and rebalancing each month is unnecessary and simply adding noise to the portfolios. However, there is a tension between not wanting to add noise to the portfolios by re-optimising and rebalancing too frequently and wanting to update the weights

Table 3.5 Six size/book-to-market portfolios - quarterly rebalancing

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.020	0.028	0.701	0.018	0.411	3.052
CO	0.022	0.036	0.617	0.019	0.616	4.388
p-value			0.085	0.277		
NIO	0.020	0.043	0.479	0.016	0.945	6.069
p-value			0.007	0.507		
$1/N$	0.007	0.029	0.223	0.004	0.053	0.000
p-value			0.000	0.000		
CMIN	0.011	0.021	0.545	0.010	0.189	1.448
p-value			0.135	0.000		
NIMIN	0.011	0.021	0.548	0.010	0.189	1.452
p-value			0.138	0.000		
CTP	0.002	0.127	0.014	-0.038	18.899	7.962
p-value			0.000	0.107		
NITP	0.016	0.031	0.510	0.013	0.621	4.336
p-value			0.017	0.033		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.039	0.029	0.013	0.194	0.107	0.062	0.014
CO	0.047	0.037	0.021	0.219	0.130	0.101	0.044
NIO	0.067	0.051	0.029	0.267	0.195	0.144	0.074
$1/N$	0.073	0.056	0.032	0.499	0.268	0.223	0.089
CMIN	0.042	0.030	0.016	0.295	0.156	0.097	0.013
NIMIN	0.042	0.030	0.016	0.287	0.151	0.095	0.013
CTP	0.393	0.231	0.105	1.051	1.001	1.001	1.000
NITP	0.054	0.038	0.022	0.188	0.133	0.096	0.037

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	0.228	0.772	0.000	0.709	0.704
NIO	0.921	0.079	0.067	0.000	0.000
$1/N$	1.000	0.000	1.000	0.000	0.000
CMIN	1.000	0.000	1.000	0.000	0.028
NIMIN	1.000	0.000	1.000	0.000	0.028
CTP	1.000	0.000	1.000	0.000	0.000
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with quarterly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.6 25 size/book-to-market portfolios - quarterly rebalancing

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.046	0.078	0.595	0.031	2.414	12.078
CO	0.046	0.084	0.549	0.029	3.356	14.905
p-value			0.136	0.278		
NIO	0.056	0.129	0.430	0.014	7.451	22.073
p-value			0.015	0.040		
$1/N$	0.007	0.031	0.235	0.005	0.057	0.000
p-value			0.005	0.000		
CMIN	0.011	0.021	0.530	0.010	0.335	2.923
p-value			0.573	0.001		
NIMIN	0.011	0.021	0.535	0.010	0.326	2.849
p-value			0.610	0.001		
CTP	0.109	0.398	0.273	-0.287	5.628	13.245
p-value			0.000	0.066		
NITP	0.025	0.046	0.557	0.020	1.159	7.733
p-value			0.546	0.011		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.096	0.070	0.037	0.360	0.314	0.266	0.060
CO	0.130	0.094	0.050	0.512	0.452	0.381	0.191
NIO	0.179	0.147	0.090	0.686	0.619	0.416	0.296
$1/N$	0.074	0.057	0.033	0.502	0.261	0.203	0.090
CMIN	0.040	0.029	0.016	0.301	0.200	0.110	0.020
NIMIN	0.040	0.030	0.016	0.303	0.195	0.100	0.020
CTP	0.094	0.070	0.038	0.519	0.191	0.165	0.061
NITP	0.064	0.048	0.026	0.203	0.166	0.130	0.053

(c) Stochastic dominance tests

			p-value	p-value	p-value
	UI^{UO}	UI^n	$UI^{UO}=1$	$UI^{UO}=0$	$UI^{UO}=UI^n$
CO	0.866	0.134	0.413	0.000	0.302
NIO	0.539	0.461	0.000	0.000	0.860
$1/N$	0.990	0.010	0.453	0.000	0.000
CMIN	0.965	0.035	0.385	0.000	0.000
NIMIN	0.965	0.035	0.385	0.000	0.000
CTP	0.009	0.991	0.000	0.542	0.212
NITP	0.958	0.042	0.358	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with quarterly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.7 10 industry portfolios - quarterly rebalancing

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.007	0.027	0.280	0.006	0.169	1.090
CO	0.006	0.031	0.198	0.004	0.221	1.453
p-value			0.040	0.121		
NIO	0.003	0.040	0.075	-0.001	0.356	2.146
p-value			0.007	0.024		
$1/N$	0.006	0.024	0.234	0.004	0.048	0.000
p-value			0.626	0.530		
CMIN	0.007	0.019	0.352	0.006	0.088	0.531
p-value			0.326	0.882		
NIMIN	0.007	0.020	0.354	0.006	0.089	0.537
p-value			0.314	0.857		
CTP	0.023	0.157	0.149	-0.038	2.024	4.036
p-value			0.239	0.021		
NITP	0.004	0.035	0.121	0.001	0.298	1.485
p-value			0.125	0.204		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.057	0.045	0.027	0.336	0.257	0.188	0.115
CO	0.069	0.053	0.035	0.370	0.306	0.264	0.188
NIO	0.103	0.082	0.049	0.659	0.632	0.590	0.493
$1/N$	0.060	0.046	0.026	0.442	0.303	0.221	0.099
CMIN	0.046	0.033	0.019	0.286	0.197	0.150	0.055
NIMIN	0.046	0.033	0.019	0.280	0.191	0.147	0.051
CTP	0.238	0.162	0.082	0.843	0.824	0.811	0.781
NITP	0.093	0.069	0.039	0.505	0.477	0.439	0.338

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	1.000	0.000	0.000
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.866	0.134	0.413	0.000	0.385
CMIN	0.076	0.924	0.000	0.508	0.374
NIMIN	0.065	0.935	0.000	0.503	0.369
CTP	0.302	0.698	0.000	0.436	0.492
NITP	1.000	0.000	1.000	0.000	0.168

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with quarterly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.8 30 industry portfolios - quarterly rebalancing

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.013	0.057	0.229	0.005	0.828	4.690
CO	0.012	0.067	0.171	0.000	1.185	6.079
p-value			0.091	0.049		
NIO	0.004	0.097	0.044	-0.019	2.653	9.194
p-value			0.004	0.001		
$1/N$	0.007	0.028	0.239	0.005	0.057	0.000
p-value			0.930	0.970		
CMIN	0.006	0.020	0.286	0.005	0.165	1.205
p-value			0.611	0.956		
NIMIN	0.006	0.020	0.287	0.005	0.165	1.199
p-value			0.611	0.960		
CTP	-0.039	0.706	-0.056	-1.286	21.099	17.475
p-value			0.020	0.038		
NITP	-0.066	0.699	-0.095	-1.288	19.456	13.678
p-value			0.001	0.060		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.102	0.083	0.054	0.633	0.607	0.590	0.478
CO	0.118	0.095	0.069	0.802	0.769	0.753	0.688
NIO3	0.225	0.178	0.115	0.955	0.947	0.943	0.926
$1/N$	0.068	0.050	0.029	0.485	0.239	0.168	0.071
CMIN	0.040	0.031	0.020	0.282	0.181	0.132	0.076
NIMIN	0.040	0.031	0.020	0.289	0.191	0.137	0.077
CTP	1.000	1.000	0.529	8.838	4.667	0.999	0.999
NITP	1.000	1.000	0.510	37.461	3.697	0.853	0.794

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	1.000	0.000	0.000
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.575	0.425	0.000	0.000	0.855
CMIN	0.575	0.425	0.000	0.000	0.793
NIMIN	0.574	0.426	0.000	0.000	0.799
CTP	1.000	0.000	1.000	0.000	0.045
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with quarterly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

to account for new information. I evaluate this tension by moving to quarterly rebalancing, following the Boehmer et al. (2008) scheme of rebalancing one-third of the portfolio each month.

Tables 3.5-3.8 show that rebalancing each quarter is very effective at reducing portfolio noise. In all four universes, the Sharpe ratios of all portfolios depending on estimated parameters increase substantially through reduced portfolio variances. Turnover and leverage fall in general, too. How the UO portfolio ranks in terms of performance is generally unchanged from the monthly rebalancing case. However, the statistical significance of its advantage over other portfolios generally improves as the differences become larger and standard errors fall. The UO portfolios tend to perform better relative to $1/N$, CMIN and NIMIN in terms of utopia indices with quarterly, rather than monthly, rebalancing. Losses and drawdowns generally fall for all portfolios but little changes in how the portfolios rank.

How the UO portfolio compares to the other portfolios changes little in the six size/book-to-market portfolio universe. If anything, the preference for the UO portfolio is stronger with quarterly rebalancing. It now has the lowest extreme drawdowns of any portfolio for all cost levels, and the lowest expected shortfalls for zero and 10bps each-way costs (the minimum variance portfolios have lower expected shortfalls with 50bps each-way costs). Moreover, the UO portfolio has a significantly better certainty equivalent and utopia index than the minimum variance portfolios with zero and 10bps each-way costs. UO's certainty equivalent advantage over CMIN and NIMIN remains significant with 50bps each-way costs, too. In contrast to the monthly rebalanced case, UO continues to have a significantly higher utopia index than and stochastically dominate $1/N$ even with 50bps each-way costs.

A similar pattern emerges in the 25 size/book-to-market universe. The UO portfolio now has a significantly better Sharpe ratio than the NIO portfolio and $1/N$, and significantly better certainty equivalents than NIO, $1/N$, CMIN, NIMIN and NITP. UO also now has a significantly better utopia index than CMIN and NIMIN. All of these preferences remain significant with 10bps each-way costs, but none is significant with 50bps each-way costs.

Again, in the 10 industry universe, little changes in how the UO portfolio compares to the others. The minor changes are that the UO portfolio's Sharpe ratio advantage over the CO and NIO portfolios is now significant without costs. Likewise, the UO portfolio's zero-cost certainty equivalent advantage over NIO and CTP is now significant, too. UO also now has an insignificantly better utopia index than $1/N$ for zero and 10bps, but not 50bps, each-way transaction costs.

Similarly, the changes in the 30 industry universe are very minor. The UO Sharpe ratio becomes significantly higher than the NIO, CTP and NITP Sharpe ratios. The UO certainty equivalent becomes significantly higher than those of CO and CTP and remains significantly higher than that of NIO. While the utopia index now indicates an insignificant preference for UO over $1/N$, CMIN and NIMIN without costs, that preference reverses with 10bps each-way costs. $1/N$, CMIN and NIMIN all have a significantly higher utopia index than UO with 50bps each-way costs.

3.5.2 Conditional heteroscedasticity

Reverting to monthly rebalancing, another possible factor holding back the UO portfolio is the assumption that the error variance in (3.1) is homoscedastic. By not allowing for conditional heteroscedasticity and not forecasting the variance, I may be depriving the unconditionally optimal portfolio of useful information and the chance to engage in volatility timing.[5] To rectify this, I assume that ε_t follows an asymmetric dynamic conditional correlation model (aDCC), whose specification is in Appendix B.3, while continuing to use the standard estimator of the covariance matrix for the no-information portfolios.[6]

Tables 3.9-3.12 show the aDCC results. Using the aDCC specification is actually unhelpful: it results in lower Sharpe ratios and certainty equivalents for the portfolios using conditional means. This may be the result of specification error (the aDCC specification being further from the "true" ε_t variance specification than the homoscedastic specification) or increased estimation error (the aDCC model requires estimation of more parameters: there are $4N+3+(N-1)N/2$ in the aDCC model compared to $(N+1)N/2$ in a standard covariance matrix). The turnover and leverage of portfolios using predictive information also increase. Ultimately, though, when moving from homoscedastic to aDCC ε_t, little changes in terms of the relative merits of the UO portfolio compared to the other mean-variance optimal portfolios.

The main change is in the six size/book-to-market universe. The UO portfolio is now second-best - behind the CO portfolio - in terms of Sharpe ratio, certainty equivalent and utopia index. Moreover, the CO portfolio has slightly better extreme drawdowns, although not extreme losses, for all cost levels. This is broadly true for all cost levels, although with 10bps and 50bps each-way costs, UO has slightly lower extreme losses and, with 50bps each-way costs, UO has a marginally higher certainty equivalent. The only change in the 25 size/book-to-market universe is that the minimum variance portfolios also now have higher certainty equivalents than the UO portfolio without costs.

[5] Note that the homoscedastic specification does allow the covariance matrix of ε_t to vary over time in practice, since it is estimated by rolling window.

[6] Assuming a standard (symmetric) DCC specification for ε_t makes little difference to the results. This is not overly surprising since accounting for asymmetry only requires the estimation of one additional parameter.

Table 3.9 Six size/book-to-market portfolios - aDCC

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.018	0.054	0.344	0.011	0.511	3.675
CO	0.021	0.059	0.357	0.012	0.642	4.688
p-value			0.769	0.630		
NIO	0.020	0.074	0.273	0.007	0.941	6.064
p-value			0.433	0.417		
$1/N$	0.007	0.049	0.137	0.001	0.053	0.000
p-value			0.059	0.062		
CMIN	0.008	0.039	0.212	0.004	0.182	1.272
p-value			0.125	0.099		
NIMIN	0.011	0.037	0.301	0.008	0.190	1.450
p-value			0.628	0.395		
CTP	0.018	1.288	0.014	-4.126	5.616	28.968
p-value			0.014	0.240		
NITP	0.016	0.057	0.276	0.008	0.621	4.322
p-value			0.434	0.435		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.103	0.080	0.048	0.343	0.193	0.127	0.057
CO	0.106	0.083	0.051	0.326	0.165	0.119	0.063
NIO	0.140	0.107	0.069	0.371	0.243	0.206	0.134
$1/N$	0.111	0.090	0.058	0.558	0.313	0.260	0.116
CMIN	0.089	0.069	0.042	0.431	0.295	0.215	0.066
NIMIN	0.079	0.061	0.037	0.335	0.171	0.115	0.041
CTP	1.000	1.000	0.492	52.413	8.483	7.095	1.124
NITP	0.113	0.086	0.054	0.307	0.178	0.143	0.072

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	0.161	0.839	0.000	0.425	0.287
NIO	0.828	0.172	0.033	0.000	0.000
$1/N$	0.989	0.011	0.602	0.000	0.022
CMIN	0.888	0.112	0.249	0.000	0.017
NIMIN	0.739	0.261	0.265	0.000	0.320
CTP	1.000	0.000	1.000	0.000	0.000
NITP	0.972	0.028	0.492	0.000	0.017

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as following an aDCC process. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.10 25 size/book-to-market portfolios - aDCC

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.040	0.126	0.320	0.001	3.380	15.843
CO	0.041	0.144	0.283	-0.011	3.846	16.662
p-value			0.323	0.194		
NIO	0.057	0.212	0.267	-0.056	7.239	22.193
p-value			0.536	0.016		
$1/N$	0.007	0.051	0.145	0.001	0.057	0.000
p-value			0.144	0.989		
CMIN	0.011	0.039	0.280	0.007	0.281	2.454
p-value			0.719	0.585		
NIMIN	0.011	0.037	0.299	0.008	0.328	2.861
p-value			0.851	0.557		
CTP	0.019	0.149	0.129	-0.036	10.238	10.728
p-value			0.073	0.199		
NITP	0.026	0.076	0.338	0.011	1.168	7.764
p-value			0.828	0.309		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.230	0.185	0.112	0.588	0.488	0.458	0.270
CO	0.284	0.217	0.127	0.769	0.644	0.582	0.389
NIO	0.392	0.302	0.183	0.782	0.714	0.669	0.460
$1/N$	0.115	0.093	0.060	0.560	0.305	0.249	0.117
CMIN	0.085	0.065	0.038	0.360	0.157	0.110	0.054
NIMIN	0.076	0.060	0.037	0.345	0.230	0.132	0.056
CTP	0.347	0.219	0.114	1.006	1.005	1.003	1.002
NITP	0.128	0.100	0.062	0.268	0.217	0.183	0.118

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	0.980	0.020	0.464	0.000	0.155
NIO	0.473	0.527	0.000	0.000	0.912
$1/N$	0.821	0.179	0.243	0.000	0.094
CMIN	0.747	0.253	0.144	0.000	0.149
NIMIN	0.741	0.259	0.171	0.000	0.204
CTP	1.000	0.000	0.564	0.000	0.000
NITP	0.598	0.402	0.000	0.000	0.470

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as following an aDCC process. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.11 10 industry portfolios - aDCC

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.006	0.045	0.132	0.001	0.204	1.341
CO	0.004	0.052	0.073	-0.003	0.239	1.672
p-value			0.109	0.053		
NIO	0.003	0.067	0.046	-0.008	0.360	2.163
p-value			0.294	0.098		
$1/N$	0.006	0.041	0.140	0.002	0.048	0.000
p-value			0.938	0.870		
CMIN	0.007	0.033	0.214	0.004	0.089	0.514
p-value			0.304	0.302		
NIMIN	0.007	0.033	0.209	0.004	0.090	0.536
p-value			0.346	0.336		
CTP	0.003	0.560	0.006	-0.782	3.327	9.950
p-value			0.318	0.281		
NITP	0.004	0.059	0.074	-0.004	0.283	1.492
p-value			0.562	0.441		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.112	0.088	0.055	0.318	0.236	0.199	0.120
CO	0.128	0.105	0.065	0.342	0.301	0.272	0.193
NIO	0.157	0.126	0.083	0.712	0.684	0.651	0.578
$1/N$	0.093	0.076	0.048	0.495	0.332	0.243	0.117
CMIN	0.077	0.063	0.037	0.264	0.203	0.150	0.087
NIMIN	0.076	0.060	0.037	0.316	0.223	0.172	0.071
CTP	0.910	0.496	0.218	7.580	6.111	5.341	1.361
NITP	0.138	0.103	0.067	0.604	0.551	0.525	0.393

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	1.000	0.000	0.000
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.099	0.901	0.000	0.597	0.503
CMIN	0.000	1.000	0.000	1.000	0.000
NIMIN	0.000	1.000	0.000	1.000	0.000
CTP	1.000	0.000	1.000	0.000	0.000
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as following an aDCC process. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.12 30 industry portfolios - aDCC

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.014	0.096	0.151	-0.009	1.146	6.218
CO	0.014	0.105	0.138	-0.013	1.260	6.635
p-value			0.667	0.244		
NIO	0.003	0.161	0.019	-0.061	2.654	9.266
p-value			0.119	0.003		
$1/N$	0.007	0.046	0.143	0.001	0.057	0.000
p-value			0.948	0.316		
CMIN	0.009	0.035	0.246	0.006	0.157	1.121
p-value			0.418	0.131		
NIMIN	0.006	0.034	0.167	0.003	0.168	1.204
p-value			0.897	0.235		
CTP	0.145	1.799	0.081	-7.945	2.137	11.149
p-value			0.455	0.306		
NITP	-0.066	1.244	-0.053	-3.933	13.999	14.247
p-value			0.058	0.285		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.198	0.158	0.104	0.626	0.565	0.537	0.462
CO	0.208	0.173	0.116	0.769	0.734	0.708	0.641
NIO	0.379	0.293	0.193	0.990	0.987	0.982	0.978
$1/N$	0.107	0.083	0.052	0.544	0.267	0.211	0.092
CMIN	0.074	0.059	0.038	0.328	0.185	0.117	0.060
NIMIN	0.075	0.063	0.039	0.318	0.209	0.165	0.098
CTP	0.468	0.288	0.147	1.127	1.058	1.048	1.028
NITP	1.000	1.000	0.638	23.598	7.136	0.998	0.928

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	0.997	0.003	0.707	0.000	0.088
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.401	0.599	0.000	0.000	0.514
CMIN	0.257	0.743	0.000	0.000	0.000
NIMIN	0.397	0.603	0.000	0.000	0.343
CTP	0.006	0.994	0.000	0.757	0.144
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using the (lagged) market return in a univariate linear prediction model and treat the prediction error ε_t as following an aDCC process. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

In the industry universes, the only differences compared to the homoscedastic ε_t case are some minor changes in the p-values of the utopia indices. But the utopia indices' conclusions about preference and stochastic dominance broadly remain. The sole exception is in the 30 industry portfolio universe, where the preference for the UO portfolio over the CO portfolio becomes insignificant at the 5% level ($p = 0.088$), despite $\widehat{UI}^{UO} = 0.997$ and the null that $UI^{UO} = 1$ not being rejected at any conventional level.

3.5.3 Machine learning predictor

Returning to treating ε_t as conditionally homoscedastic, another possibility is that the portfolios using predictive information are being held back by using a fairly basic univariate prediction equation. It transpires that more modern methods do not produce a predictor with a better information coefficient than the univariate model with the market return. Appendix B.2 provides the full details of the machine learning predictors considered and their information coefficients. The machine learning predictor giving rise to the highest information coefficients in all universes is a targeted random forest with an elastic net targeting step. Tables 3.13-3.16 show that using this predictor does not necessarily help the UO portfolio.

In the six book-to-market portfolio universe, the Sharpe ratio and certainty equivalent of the UO portfolio fall below that of the NIO portfolio. Using the predictive information does not now lead to better portfolios. The UO portfolio also falls behind CMIN, NIMIN and NITP in terms of Sharpe ratio and certainty equivalent. NIO and NITP now have insignificantly higher utopia indices than UO, too. CMIN and NIMIN's utopia index advantages over UO are now significant. Since the UO portfolio has lower costs than NIO or NITP, it does have an insignificantly better Sharpe ratio, certainty equivalent and utopia index than both NIO and NITP with 50bps each-way transaction costs. The UO portfolio no longer has the lowest expected shortfalls and extreme drawdowns of all portfolios, although it does still have the lowest expected shortfalls of all the mean-variance optimal portfolios. With 50bps each-way transaction costs, the UO portfolio also has the lowest extreme drawdowns of all the mean-variance optimal portfolios.

The 25 size/book-to-market portfolio universe results are much less affected by the change of predictor. How the UO portfolio ranks compared to the other portfolios remains very similar. In fact, the UO portfolio now has a significantly higher Sharpe ratio and certainty equivalent than $1/N$ with both zero and 10bps, but not 50bps, each-way costs. The UO utopia index is also significantly higher the CO utopia index. Moreover, the UO portfolio now stochastically dominates CMIN and NIMIN with zero and 10bps each-way costs. It also has an insignificantly better utopia index than CMIN and NIMIN with 50bps each-way costs. While NIO and NITP

Table 3.13 Six size/book-to-market portfolios - targeted random forest

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.009	0.042	0.222	0.005	0.203	1.602
CO	0.010	0.047	0.218	0.005	0.292	2.371
p-value			0.903	0.907		
NIO	0.020	0.074	0.273	0.007	0.941	6.064
p-value			0.573	0.779		
$1/N$	0.007	0.049	0.137	0.001	0.053	0.000
p-value			0.227	0.207		
CMIN	0.009	0.039	0.228	0.005	0.155	1.143
p-value			0.857	0.897		
NIMIN	0.011	0.037	0.301	0.008	0.190	1.450
p-value			0.096	0.123		
CTP	-0.032	0.718	-0.045	-1.320	4.312	13.723
p-value			0.012	0.443		
NITP	0.016	0.057	0.276	0.008	0.621	4.322
p-value			0.469	0.483		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.091	0.074	0.045	0.388	0.220	0.158	0.064
CO	0.104	0.079	0.048	0.288	0.215	0.172	0.085
NIO	0.140	0.107	0.069	0.371	0.243	0.206	0.134
$1/N$	0.111	0.090	0.058	0.558	0.313	0.260	0.116
CMIN	0.084	0.067	0.042	0.380	0.227	0.163	0.058
NIMIN	0.079	0.061	0.037	0.335	0.171	0.115	0.041
CTP	1.000	0.766	0.351	7.073	5.151	3.623	3.146
NITP	0.113	0.086	0.054	0.307	0.178	0.143	0.072

(c) Stochastic dominance tests

			p-value	p-value	p-value
	UI^{UO}	UI^n	$UI^{UO}=1$	$UI^{UO}=0$	$UI^{UO}=UI^n$
CO	0.609	0.391	0.000	0.000	0.762
NIO	0.272	0.728	0.000	0.331	0.403
$1/N$	1.000	0.000	1.000	0.000	0.000
CMIN	0.136	0.864	0.000	0.420	0.044
NIMIN	0.000	1.000	0.000	1.000	0.000
CTP	1.000	0.000	1.000	0.000	0.000
NITP	0.273	0.727	0.000	0.315	0.392

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using an elastic net-targeted random forest and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.14 25 size/book-to-market portfolios - targeted random forest

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.019	0.050	0.384	0.013	0.639	5.125
CO	0.019	0.052	0.370	0.012	0.681	5.521
p-value			0.218	0.417		
NIO	0.057	0.212	0.267	-0.056	7.239	22.193
p-value			0.266	0.006		
$1/N$	0.007	0.051	0.145	0.001	0.057	0.000
p-value			0.020	0.024		
CMIN	0.011	0.035	0.315	0.008	0.290	2.475
p-value			0.322	0.112		
NIMIN	0.011	0.037	0.299	0.008	0.328	2.861
p-value			0.240	0.103		
CTP	0.338	2.994	0.113	-22.078	79.179	115.786
p-value			0.005	0.328		
NITP	0.026	0.076	0.338	0.011	1.168	7.764
p-value			0.588	0.755		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.090	0.069	0.042	0.227	0.120	0.101	0.047
CO	0.093	0.073	0.044	0.220	0.123	0.103	0.050
NIO	0.392	0.302	0.183	0.782	0.714	0.669	0.460
$1/N$	0.115	0.093	0.060	0.560	0.305	0.249	0.117
CMIN	0.070	0.055	0.034	0.262	0.126	0.092	0.039
NIMIN	0.076	0.060	0.037	0.345	0.230	0.132	0.056
CTP	1.000	0.909	0.455	5.111	2.274	1.753	1.342
NITP	0.128	0.100	0.062	0.268	0.217	0.183	0.118

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	0.343	0.000	0.000
NIO	0.285	0.715	0.000	0.298	0.343
$1/N$	1.000	0.000	1.000	0.000	0.000
CMIN	0.856	0.144	0.193	0.000	0.000
NIMIN	0.903	0.097	0.381	0.000	0.006
CTP	0.017	0.983	0.000	0.635	0.144
NITP	0.257	0.743	0.000	0.420	0.420

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using an elastic net-targeted random forest and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.15 10 industry portfolios - targeted random forest

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.007	0.034	0.209	0.004	0.094	0.588
CO	0.007	0.036	0.192	0.004	0.111	0.735
p-value			0.482	0.528		
NIO	0.003	0.067	0.046	-0.008	0.360	2.163
p-value			0.085	0.034		
$1/N$	0.006	0.041	0.140	0.002	0.048	0.000
p-value			0.328	0.328		
CMIN	0.007	0.032	0.209	0.004	0.082	0.464
p-value			0.984	0.966		
NIMIN	0.007	0.033	0.209	0.004	0.090	0.536
p-value			0.996	0.984		
CTP	0.022	0.616	0.036	-0.925	4.154	12.297
p-value			0.161	0.130		
NITP	0.004	0.059	0.074	-0.004	0.283	1.492
p-value			0.166	0.172		

(b) Loss distribution

	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.075	0.062	0.038	0.273	0.182	0.138	0.075
CO	0.080	0.065	0.041	0.244	0.197	0.178	0.083
NIO	0.157	0.126	0.083	0.712	0.684	0.651	0.578
$1/N$	0.093	0.076	0.048	0.495	0.332	0.243	0.117
CMIN	0.074	0.059	0.036	0.313	0.229	0.178	0.073
NIMIN	0.076	0.060	0.037	0.316	0.223	0.172	0.071
CTP	1.000	0.702	0.335	5.417	0.980	0.925	0.642
NITP	0.138	0.103	0.067	0.604	0.551	0.525	0.393

(c) Stochastic dominance tests

	UI^{UO}	UI^n	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^n$
CO	1.000	0.000	1.000	0.000	0.000
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	1.000	0.000	1.000	0.000	0.000
CMIN	0.216	0.784	0.000	0.492	0.470
NIMIN	0.237	0.763	0.000	0.613	0.602
CTP	0.858	0.142	0.845	0.000	0.464
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using an elastic net-targeted random forest and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^n that for the portfolio indicated in the left-most column.

Table 3.16 30 industry portfolios - targeted random forest

(a) Summary performance statistics

	Mean	Std. dev	Sharpe ratio	CEQ_5	Turnover	Leverage
UO	0.005	0.043	0.120	0.001	0.257	1.931
CO	0.005	0.044	0.121	0.000	0.262	1.984
p-value			0.884	0.957		
NIO	0.003	0.161	0.019	-0.061	2.654	9.266
p-value			0.351	0.001		
$1/N$	0.007	0.046	0.143	0.001	0.057	0.000
p-value			0.834	0.875		
CMIN	0.006	0.033	0.174	0.003	0.154	1.086
p-value			0.463	0.410		
NIMIN	0.006	0.034	0.167	0.003	0.168	1.204
p-value			0.541	0.475		
CTP	-0.468	6.009	-0.078	-90.736	511.803	60.624
p-value			0.043	0.410		
NITP	-0.066	1.244	-0.053	-3.933	13.999	14.247
p-value			0.142	0.284		

(b) Loss distribution

Sortino	$ES_{95\%}$	$ES_{90\%}$	$ES_{75\%}$	MaxDD	$DD_{95\%}$	$DD_{90\%}$	$DD_{75\%}$
UO	0.105	0.082	0.053	0.397	0.300	0.265	0.181
CO	0.106	0.083	0.054	0.398	0.304	0.265	0.179
NIO	0.379	0.293	0.193	0.990	0.987	0.982	0.978
$1/N$	0.107	0.083	0.052	0.544	0.267	0.211	0.092
CMIN	0.074	0.060	0.038	0.339	0.239	0.188	0.081
NIMIN	0.075	0.063	0.039	0.318	0.209	0.165	0.098
CTP	1.000	1.000	1.000	247.47	5.179	2.513	1.272
NITP	1.000	1.000	0.638	23.598	7.136	0.998	0.928

(c) Stochastic dominance tests

	UI^{UO}	UI^{n}	p-value $UI^{UO}=1$	p-value $UI^{UO}=0$	p-value $UI^{UO}=UI^{n}$
CO	0.763	0.237	0.287	0.000	0.265
NIO	1.000	0.000	1.000	0.000	0.000
$1/N$	0.158	0.842	0.000	0.547	0.199
CMIN	0.000	1.000	0.000	1.000	0.000
NIMIN	0.000	1.000	0.000	1.000	0.000
CTP	1.000	0.000	1.000	0.000	0.000
NITP	1.000	0.000	1.000	0.000	0.000

Out-of-sample portfolio performance computed over a 120-month rolling window for the unconditionally optimal (UO), conditionally optimal (CO) and benchmark portfolios (no information optimal [NIO], $1/N$, conditional minimum variance [CMIN], no-information minimum variance [NIMIN], conditional tangency portfolio [CTP] and unconditional tangency portfolio [NITP]) with monthly rebalancing. I compute $\hat{\mu}(S_t)$ using an elastic net-targeted random forest and treat the prediction error ε_t as homoscedastic. Risk aversion is $\gamma = 5$ in the UO, CO and NIO portfolios. The p-values beneath the Sharpe ratios in panel (a) are those for the difference between the Sharpe ratio immediately above the p-value and the UO Sharpe ratio. Likewise for the certainty equivalent with risk aversion 5 (CEQ_5). $ES_{95\%}$ denotes the 95% expected shortfall in panel (b), MaxDD the maximum drawdown and $DD_{95\%}$ the 95th percentile of the drawdowns. Positive numbers in panel (b) indicate a loss. In panel (c), UI^{UO} is the unconditionally optimal utopia index, UI^{n} that for the portfolio indicated in the left-most column.

have insignificantly better utopia indices than the UO portfolio with zero and 10bps costs, this is reversed with 50bps costs. The UO portfolio actually stochastically dominates the NIO portfolio with 50bps each-way costs.

The change of predictor changes very little in the 10 industry universe. The UO certainty equivalent is now significantly higher than the the NIO certainty equivalent. The UO portfolio stochastically dominates and has a significantly higher utopia index than $1/N$ for all cost levels and now has a higher utopia index than CTP.

Slightly more changes in the 30 industry universe, where the UO portfolio's performance is worse than with the univariate market predictor. The CO portfolio now has a marginally higher Sharpe ratio, although the UO portfolio retains a marginally higher certainty equivalent. While the UO portfolio still has a higher utopia index than CO, this difference is no longer significant at any cost level. Moreover, the minimum variance portfolios now clearly stochastically dominate the UO portfolio, even without costs.

3.6 Robustness

The main results are robust to the empirical choices I make regarding risk aversion, look-back window length for computing the parameters the weights are based on, and the predictor I use.

3.6.1 Risk aversion

The UO, CO and NIO portfolio weights depend on the level of risk aversion. In the above results, I use $\gamma = 5$. Reducing risk aversion to $\gamma = 1$ or increasing it to $\gamma = 10$ makes little difference to the overall results. The risk aversion drives the aggressiveness of the UO, CO and NIO portfolios. In common with Kirby and Ostdiek (2012a), the more aggressive portfolios (here $\gamma = 1$) do not perform as well in terms of risk/return trade-off as less aggressive portfolios ($\gamma = 5$ or 10). This does not much change the ranking of the portfolios in terms of Sharpe ratio, certainty equivalent or utopia index, but rather the distance between them. As expected, turnover and leverage are decreasing in γ (increasing in aggression).

In general, whatever γ is, the UO portfolio has the best Sharpe ratio and certainty equivalent, the lowest turnover and leverage, and the lowest extreme losses and drawdowns of the mean-variance optimal portfolios. It is also usually preferred by the utopia index. The UO portfolio does not, however, necessarily outperform the non-optimal benchmarks. In particular, the UO portfolio does not generally outperform the non-optimal benchmarks in the industry universes.

3.6.2 Look-back window

Reverting to a risk-aversion of $\gamma = 5$, we see that changing the look-back window in the estimation makes almost no difference to the results. The above results use a 120-month rolling window. When switching to a shorter rolling window of 60 months, the main difference is that all portfolios perform worse as they are subject to greater estimation error (except $1/N$, which is unchanged). However, the ranking of the portfolios barely changes. The results are quantitatively very similar when I move from using a 120-month rolling window to using an expanding window.

3.6.3 Choice of predictor

I now return to a 120-month rolling window and risk aversion of $\gamma = 5$. I continue to restrict attention to predictors with a positive information coefficient. I therefore consider the reversal and one/12-month trend change ($MA^{1,12}$) univariate predictors in the size/book-to-market universes and the targeted random forest with a LASSO targeting step for the industry universes. There is almost no difference in how the portfolios rank in terms of Sharpe ratio, certainty equivalent, utopia index, losses, drawdowns, turnover or leverage in any of these cases.

3.7 Conclusion

Using data for four asset universes (six and 25 size/book-to-market portfolios and 10 and 30 industry portfolios) from January 1990 to December 2019, I show that portfolios using predictive information unconditionally optimally are preferred to portfolios using such information conditionally optimally. The UO portfolios have higher Sharpe ratios and certainty equivalents, and score higher on a measure of almost stochastic dominance. This superior performance is achieved with lower turnover, leverage and tail risk, where tail risk is measured in terms of expected shortfalls and extreme drawdowns. We also prefer portfolios using predictive information unconditionally optimally to mean-variance optimal portfolios not using predictive information (no-information optimal portfolios) on these same metrics. However, one must note the possible look-ahead bias in these latter comparisons, as the predictors are chosen based on their predictive performance over the same time period that the portfolio strategies are run.

How well the UO portfolios compare to non-optimal benchmarks ($1/N$, CMIN and NIMIN) depends on the asset universe. The UO portfolios outperform the non-optimal benchmarks in the size/book-to-market universes, but not the industry universes. The fact that the UO portfolios do not routinely outperform $1/N$ and NIMIN is all the more concerning when one remembers that the predictors used for the UO portfolios are chosen based on in-sample predictive performance.

These results are robust to the predictive information used, the investor's risk aversion, allowing the prediction error to be conditionally heteroscedastic and re-optimising and rebalancing the portfolios quarterly rather than monthly.

There are several implications of these results. When judging portfolio performance on the basis of measures which are functions of unconditional moments, it is better to use information unconditionally optimally than conditionally optimally. This is true, for example, for an empirical study looking at the value of using predictive information in portfolio construction where the Sharpe ratios depend on unconditional means and variances. It is also true for an informed manager making investments on behalf of an uninformed client. The client is better served if the manager uses the information unconditionally optimally. Moreover, the lower turnover and leverage of the UO portfolios suggest that they have lower costs and are more practical for investors than CO portfolios. However, the failure of the UO portfolios to outperform the non-optimal benchmarks in the industry universes suggests better predictive information or predictive techniques need to be found to truly assert the dominance of unconditionally optimal mean-variance portfolios.

Appendix A

Appendices to Chapter 1

A.1 Bansal-Yaron model estimation

A.1.1 Inversion and stochastic discount factor coefficients

Constantinides and Ghosh (2011) show that

$$
\begin{aligned}
x_t &= \alpha_0 + \alpha_1 r_{f,t} + \alpha_2 z_{m,t} \\
\sigma_t^2 &= \beta_0 + \beta_1 r_{f,t} + \beta_2 z_{m,t},
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha_0 &= \frac{A_{2,m}A_{0,f} - A_{0,m}A_{2,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\
\alpha_1 &= \frac{-A_{2,m}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\
\alpha_2 &= \frac{A_{2,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\
\beta_0 &= \frac{A_{0,m}A_{1,f} - A_{1,m}A_{0,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\
\beta_1 &= \frac{A_{1,m}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}} \\
\beta_2 &= \frac{-A_{1,f}}{A_{1,m}A_{2,f} - A_{2,m}A_{1,f}}.
\end{aligned}
$$

The expressions for the $A_0, \ldots, A_{2,f}$ coefficients are given by

$$A_1 = \frac{1-\frac{1}{\psi}}{1-\kappa_1\rho_x}$$

$$A_2 = \frac{\frac{1}{2}\left[\left(-\frac{\theta}{\psi}+\theta\right)^2 + (\theta\kappa_1 A_1\psi_x)^2\right]}{\theta(1-\kappa_1\nu)}$$

$$A_0 = \frac{\ln(\delta) + \left(1-\frac{1}{\psi}\right)\mu_c + \kappa_0 + \kappa_1 A_2\sigma^2(1-\nu) + \frac{1}{2}\theta\kappa_1^2 A_2^2\sigma_w^2}{1-\kappa_1}$$

$$A_{1,m} = \frac{\phi-\frac{1}{\psi}}{1-\kappa_{1,m}\rho_x}$$

$$A_{2,m} = \frac{(1-\theta)A_2(1-\kappa_1\nu) + \frac{1}{2}\left[\gamma^2+\varphi^2+\left((\theta-1)\kappa_1 A_1 + \kappa_{1,m}A_{1,m}\right)^2\psi_x^2\right]}{1-\kappa_{1,m}\nu}$$

$$A_{0,m} = \frac{\theta\ln(\delta) + \left(-\frac{\theta}{\psi}+\theta-1\right)\mu_c + (\theta-1)\kappa_0 + (\theta-1)(\kappa_1-1)A_0 + (\theta-1)\kappa_1 A_2\sigma^2(1-\nu)}{1-\kappa_{1,m}}$$
$$+ \frac{\kappa_{0,m} + \mu_d + \kappa_{1,m}A_{2,m}\sigma^2(1-\nu) + \frac{1}{2}\left[(\theta-1)\kappa_1 A_2 + \kappa_{1,m}A_{2,m}\right]^2\sigma_w^2}{1-\kappa_{1,m}}$$

$$A_{0,f} = \theta\ln(\delta) - \left(-\frac{\theta}{\psi}+\theta-1\right)\mu_c - (\theta-1)\kappa_0 - (\theta-1)(\kappa_1-1)A_0 - (\theta-1)\kappa_1 A_2(1-\nu)\sigma^2$$
$$-0.5(\theta-1)^2\kappa_1^2 A_2^2\sigma_w^2$$

$$A_{1,f} = -\left[\left(\frac{\theta}{\psi}+\theta-1\right) + (\theta-1)(\kappa_1\rho_x-1)A_1\right]$$

$$A_{2,f} = -\left[(\theta-1)(\kappa_1\nu-1)A_2 + \frac{1}{2}\left(\left(-\frac{\theta}{\psi}+\theta-1\right)^2 + (\theta-1)^2\kappa_1^2 + A_1^2\psi_x^2\right)\right].$$

In the stochastic discount factor

$$\exp\left\{a_1 + a_2\Delta c_{t+1} + a_3\left(r_{f,t+1} - \frac{1}{\kappa_1}r_{f,t}\right) + a_4\left(z_{m,t+1} - \frac{1}{\kappa_1}z_{m,t}\right)\right\} = 0,$$

we have:

$$
\begin{aligned}
a_1 &= \theta \ln(\delta) + (\theta - 1)[\kappa_0 + (\kappa_1 - 1)(A_0 + A_1\alpha_0 + A_2\beta_0)] \\
a_2 &= -\frac{\theta}{\psi} + (\theta - 1) \\
a_3 &= (\theta - 1)\kappa_1[A_1\alpha_1 + A_2\beta_1] \\
a_4 &= (\theta - 1)\kappa_1[A_1\alpha_2 + A_2\beta_2].
\end{aligned}
$$

The linearisation constants κ_0 and κ_1 derive from applying the Campbell and Shiller (1988) log-linearisation procedure to the returns to the consumption claim and market portfolio (Bansal and Yaron, 2004). These constants satisfy

$$
\begin{aligned}
\kappa_1 &= \frac{\exp\{\bar{z}\}}{1 + \exp\{\bar{z}\}} \\
\kappa_0 &= \ln(1 + \exp\{\bar{z}\}) - \kappa_1 \bar{z},
\end{aligned}
$$

where z_t is the log price/consumption ratio of an asset whose dividend stream is identical to consumption. Similar expressions are obtained for $\kappa_{0,m}$ and $\kappa_{1,m}$ when z is replaced by z_m. These are identified under the assumption that $\bar{z}$ and $\bar{z}_m$ are equal to the unconditional expectation of z_t and $z_{m,t}$ respectively.

A.1.2 Time-series moment conditions

The nine time-series moment conditions derived by Constantinides and Ghosh (2011) are:

$$
\begin{aligned}
\mathrm{E}(\Delta c_t) &= \mu_c \\
\mathrm{Cov}(\Delta c_t, \Delta c_{t+1}) &= \rho_x \frac{\varphi_x^2 \sigma^2}{1-\rho_x^2} \\
\mathrm{E}(\Delta d_t) &= \mu_d \\
\mathrm{Var}(\Delta d_t) &= \phi^2 \frac{\psi_x^2 \sigma^2}{1-\rho_x^2} + \sigma^2 \varphi_u^2 \\
\mathrm{Cov}(\Delta d_t, \Delta d_{t+1}) &= \phi^2 \rho_x \frac{\varphi_x^2 \sigma^2}{1-\rho_x^2} \\
\mathrm{Cov}(\Delta c_t, \Delta d_t) &= \phi \frac{\varphi_x^2 \sigma^2}{1-\rho_x^2} \\
\mathrm{Var}\left[(\Delta c_t)^2\right] &= \frac{3\varphi_x^4 \sigma_w^2 (1+\pi\rho_x^2)}{(1-\rho_x^4)(1-\pi^2)(1-\pi\rho_x^2)} + \frac{1}{1-\rho_x^4}\left[2\sigma^4 + \frac{4\rho_x^2 \varphi_x^4 \sigma^4}{1-\rho_x^2}\right] + 2\sigma^4 \\
&\quad + \frac{3\sigma_w^2}{1-\pi^2} + 4\mu_c^2 \frac{\varphi_x^2 \sigma^2}{1-\rho_x^2} + \frac{6\varphi_x^2 \sigma_w^2 \pi}{(1-\pi^2)(1-\pi\rho_x^2)} + \frac{4\varphi_x^2 \sigma^4}{1-\rho_x^2} + 4\mu_c^2 \sigma^2 \\
\mathrm{Var}\left[(\Delta d_t)^2\right] &= \phi^4 \left[\frac{3\varphi_x^4 \sigma_w^2 (1+\pi\rho_x^2)}{(1-\rho_x^4)(1-\pi^2)(1-\pi\rho_x^2)} + \frac{2\sigma^4}{1-\rho_x^4} + \frac{4\rho_x^2 \varphi_x^4 \sigma^4}{(1-\rho_x^4)(1-\rho_x^2)}\right] \\
&\quad + \frac{3\sigma_w^2 \varphi_u^4}{1-\pi^2} + 4\mu_c^2 \frac{\varphi_x^2 \sigma^2}{1-\rho_x^2} \phi^2 + \frac{6\varphi_x^2 \sigma_w^2 \pi \phi^2 \varphi_u^2}{(1-\pi^2)(1-\pi\rho_x^2)} + \frac{4\varphi_x^2 \sigma^4}{1-\rho_x^2} \phi^2 \varphi_u^2 \\
&\quad + 2\sigma^4 \varphi_u^4 + 4\mu_d^2 \varphi_u^2 \sigma^2.
\end{aligned}
$$

A.1.3 Expected return coefficients

The expected market return in the Bansal-Yaron model is

$$\mathrm{E}_t\, r_{m,t+1} = B_0 + B_1 x_t + B_2 \sigma_t^2,$$

where

$$
\begin{aligned}
B_0 &= \kappa_{0,m} + (\kappa_{1,m} - 1) A_{0,m} + \mu_d + \kappa_{1,m} A_{2,m} (1-\nu) \sigma^2 - 3\kappa_{1,m} \\
B_1 &= A_{1,m} (\kappa_{1,m} \rho_x - 1) + \phi \\
B_2 &= A_{2,m} (\kappa_{1,m} \nu - 1).
\end{aligned}
$$

A.2 Cecchetti-Lam-Mark $\kappa(y_t)$

$$\kappa(y_t) = \begin{cases} \tilde{\delta}(1 - \tilde{\delta}\tilde{\alpha}_1(p+q-1))/\Delta & ,y_t = 0 \\ \tilde{\delta}\tilde{\alpha}_1(1 - \tilde{\delta}(p+q-1))/\Delta & ,y_t = 1, \end{cases}$$

where

$$\begin{aligned} \tilde{\delta} &= \delta \exp\{\alpha_0(1-\gamma) + (1-\gamma)^2 \sigma_{y_t}^2/2\} \\ \tilde{\alpha}_1 &= \exp\{\alpha_1(1-\gamma)\} \\ \Delta &= 1 - \tilde{\delta}(p\tilde{\alpha}_1 + q) + \tilde{\delta}^2 \tilde{\alpha}_1 (p+q-1). \end{aligned}$$

Appendix B

Appendices to Chapter 3

B.1 HAC p-values for the difference in Sharpe ratios and certainty equivalents

The following extends the analysis of Lo (2002), which considers the asymptotic distribution of a single Sharpe ratio. Here, I derive the asymptotic distribution of the difference of two Sharpe ratios.

Consider the Sharpe ratio for portfolio p: $SR_p = \mu_p/\sigma_p$. Suppose we wish to consider the null that $SR_p = SR_q$. This requires the estimation of four parameters, collected in the vector $\theta = (\mu_p, \sigma_p^2, \mu_q, \sigma_q^2)'$. Our test statistic is $\Delta^S \equiv SR_p - SR_q$ and the null hypothesis is that $\Delta^S = 0$. Note that we can write Δ^S as a function of θ: $\Delta^S(\theta) = \mu_p/\sigma_p - \mu_q/\sigma_q$. I estimate θ by considering the moment conditions

$$h(R_t, \theta) = \begin{bmatrix} R_{p,t} - \mu_p \\ (R_{p,t} - \mu_p)^2 - \sigma_p^2 \\ R_{q,t} - \mu_q \\ (R_{q,t} - \mu_q)^2 - \sigma_q^2 \end{bmatrix}.$$

Based on these moment conditions, the GMM estimate of θ, $\hat{\theta}$, is given by the standard maximum likelihood estimators of each of the four parameters (Lo, 2002). Further, following Lo (2002) and by Hansen (1982),

$$\sqrt{T}(\hat{\theta} - \theta_0) \overset{a}{\sim} N(0, V_{GMM})$$

where θ_0 is the population θ and $V_{GMM} = V_B^{-1} V_M (V_B^{-1})^T$, with

$$V_B \equiv \lim_{T\to\infty} \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T}\frac{D}{D\theta}h(R_t,\theta_0)\right]$$
$$V_M \equiv \lim_{T\to\infty} \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}h(R_t,\theta_0)h(R_s,\theta_0)'\right].$$

Since

$$\frac{D}{D\theta}h(R_t,\theta_0) = \begin{bmatrix} -1 & 0 & 0 & 0 \\ -2(R_{p,t}-\mu_p) & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -2(R_{q,t}-\mu_q) & -1 \end{bmatrix},$$

it is clear that $V_B = -I$ and $V_{GMM} = V_M$.

I follow Lo (2002) in using the Newey-West estimator for V_M

$$\hat{V}_M = \hat{V}_{M,0} + \sum_{j=1}^{m-1}\left(1-\frac{j}{m}\right)(\hat{V}_{M,j}+\hat{V}'_{M,j})$$
$$\hat{V}_{M,j} = \frac{1}{T}\sum_{t=j+1}^{T} h(R_t,\hat{\theta})h(R_{t-j},\hat{\theta})',$$

and follow Delgado and Velasco (2011) in setting $m = 2(T/100)^{1/3}$.

From here, it is straightforward to apply the delta method, which, under the null that $\Delta^S = 0$, yields

$$\sqrt{T}\hat{\Delta}^S \overset{a}{\sim} N\left(0, \frac{\partial\Delta(\theta)}{\partial\theta}V_M\frac{\partial\Delta(\theta)}{\partial\theta^T}\right)$$

where

$$\frac{\partial\Delta^S(\theta)}{\partial\theta} = \left(\frac{1}{\sigma_p},\quad -\frac{1}{2}\frac{\mu_p}{\sigma_p^3},\quad -\frac{1}{\sigma_q},\quad \frac{1}{2}\frac{\mu_q}{\sigma_q^3}\right)'.$$

Similar logic works for the difference in certainty equivalents, $\Delta^C(\theta) \equiv CEQ_p - CEQ_q$, where $CEQ_p = \mu_p - \frac{\gamma}{2}\sigma_p^2$. In this case, V_M is as before, the asymptotic distribution is analogous to that above and

$$\frac{\partial\Delta^C(\theta)}{\partial\theta} = \left(1, -\frac{\gamma}{2}, -1, \frac{\gamma}{2}\right)'.$$

B.2 Predictive models

Suppose that S_t is a $T \times N \times K$ array of predictive information, where K is the number of predictors. For each asset $i = 1, \ldots, N$, $S_{i,t} = (s^i_{1,t}, \ldots s^i_{K,t})'$. $\mathrm{E}(R_{i,t+1}|S_t) = \mu_i(S_{i,t})$.

I first consider a univariate linear function for $\mu_i(S_{i,t})$:

$$\mu_i(S_{i,t}) = \kappa_{0,i} + \kappa_{1,i} s^i_{k,t}$$

for some $k \in 1, \ldots, K$. In this case, I consider K different sets of conditional mean functions and see how well using each individual predictor works in turn.

Table B.1 shows the out-of-sample information coefficients for one-step-ahead univariate predictive regressions computed over 120-month rolling windows, using each variable in turn. I compute the information coefficients over the same sample as I run the portfolio strategies in Section 3.5, i.e. the first returns I forecast are in January 2000 (where I estimate the predictive model using data up to December 1999) and the final returns I forecast are in December 2019. The (lagged) market return is the only predictor with a positive information coefficient in all four universes. It also has the highest information coefficient in each universe. Since the question of how to use predictive information optimally implicitly assumes the information used has some predictive ability, it makes sense to focus on the results with the market return as the predictor. The reversal and $MA^{1,12}$ predictors also have positive information coefficients in the size/book-to-market universes. Hence the use of these two predictors for robustness checks.

Table B.1 Information coefficients for univariate predictors

	6 size/BTM	25 size/BTM	10 industry	30 industry
Market return	0.061	0.068	0.028	0.047
Dividend yield	-0.044	-0.043	-0.048	-0.039
One-month Treasury	-0.059	-0.056	-0.037	-0.019
Treasury spread	-0.087	-0.092	-0.083	-0.052
Credit spread	-0.036	-0.052	-0.041	-0.036
CPI inflation	-0.039	-0.030	-0.072	-0.033
Reversal	0.022	0.023	-0.042	-0.021
Momentum	-0.055	-0.069	-0.056	-0.055
$MA^{1,12}$	0.027	0.032	-0.061	-0.024
$MA^{3,12}$	-0.092	-0.082	-0.101	-0.029

Out-of-sample information coefficients for the univariate predictors computed using a 120-month rolling window. $MA^{1,12}$ is the difference between the (lagged) one-month and twelve-month moving average.

The focus on univariate predictors may seem overly simplistic. I turn now to multivariate methods. I use machine learning methods, in order to reduce the collinearity problems of using a multivariate regression for prediction, as illustrated so strongly in Welch and Goyal (2008).

B.2.1 Variable selection

Variable selection methods such as the elastic net can improve forecasts of asset returns (Li and Tsiakas, 2016). In the linear conditional mean model

$$R_{i,t+1} = \kappa_{0,i} + \sum_{k=1}^{K} \kappa_{k,i} s^{i}_{k,t} + \varepsilon_{i,t}, \tag{B.1}$$

the elastic net estimator $\tilde{\kappa}_i = (\tilde{\kappa}_{0,i}, \ldots, \tilde{\kappa}_{K,i})'$ solves

$$\min_{\kappa_i} \left\{ \sum_{t=1}^{T} \left(R_{i,t} - \kappa_{0,i} - \sum_{k=1}^{K} \kappa_{k,i} s^{i}_{k,t} \right)^2 + \lambda_{L,i} \sum_{k=1}^{K} |\kappa_{k,i}| + \lambda_{R,i} \sum_{k=1}^{K} \kappa^2_{k,i} \right\}. \tag{B.2}$$

This is a penalised least squares estimator, where the penalty comprises a LASSO term ($\lambda_{L,i} \sum_k |\kappa_{k,i}|$) and a ridge term ($\lambda_{R,i} \sum_k \kappa^2_{k,i}$). The pure LASSO estimator has $\lambda_{R,i} = 0$. In this case, I fix $\lambda_{R,i} = 1$ in the elastic net estimation. In both the LASSO and elastic net estimation, $\lambda_{L,i}$ is chosen to minimise in-sample prediction error in 10-fold cross-validation, while K is chosen to minimise Mallows's C_p. $\lambda_{L,i}$ is chosen first for a given K and then the C_p computed for each K, using the optimal $\lambda_{L,i}$ for that K.

The LASSO term helps with variable selection, by forcing estimated coefficients that are small in magnitude to be zero. By dropping irrelevant variables, LASSO can help reduce estimation error. However, pure LASSO will not necessarily select two highly correlated predictors with non-zero population coefficients, even asymptotically. The ridge term helps with this grouped selection problem.

Variable selection techniques may also reduce specification error. Welch and Goyal (2008) show that the predictive ability of individual variables over the equity risk premium varies over time. Using a rolling or expanding window LASSO or elastic net model allows different predictors to drop in and out of the model at different times, depending on the economic state.

Relaxed methods

The elastic net and LASSO estimators are both biased. This can be resolved by using relaxed elastic net or LASSO. The relaxed LASSO/elastic net is a two step process. First, LASSO/elastic net is run over all predictive variables. Then, a standard multivariate OLS regression is run over only the variables with non-zero coefficients in the initial LASSO/elastic net step. This produces an unbiased forecast. Biased predictions may, however, be preferred by an investors (Kirby and Ostdiek, 2012b). I therefore consider both relaxed and non-relaxed LASSO and elastic net.

B.2.2 Machine learning

Random forests

By repeatedly partitioning the regression surface, regression trees can provide powerful, non-parametric prediction. This can help overcome functional form mis-specification. Random forests randomly grow many deep (with several splits) regression trees and average over them to reduce variance. Gu et al. (2020) show that random forests can produce substantial gains in terms of out-of-sample forecasting of stock returns compared to OLS. Here, my random forests grow 500 trees, where splits are based on variance. Each tree samples the original data with replacement (to form a sample of the same length as the original sample), has a minimum node size of five, unlimited maximum depth (subject to the minimum node size of five), and splits over up to three variables at each node.

Gradient boosting

Gradient boosting is also based on regression trees. However, instead of growing many deep trees and averaging over them, a series of shallow trees are grown and summed up. Each new tree is fitted to the gradient of the loss function (which is the residual with a squared loss function). Each time a new tree is fitted, its aim is to reduce the forecast error. The trees have to be kept shallow and we must be careful of fitting too many trees for fear of over-fitting.

Gradient boosting too can enjoy good forecast performance. With a Huber loss function Gu et al. (2020) show that gradient boosting can produce substantial gains in terms of out-of-sample forecasting of stock returns compared to OLS. Its forecasting performance is similar to random forests in Gu et al. (2020). Here, each boosted regression grows 100 trees with maximum depth one, where splits minimise a squared loss function.

B.2.3 Combining variable selection and machine learning

Borup et al. (2020) find that the performance of random forests for forecasting can be improved by including a variable selection step, similar to relaxed LASSO or elastic net. Borup et al.

Table B.2 Information coefficients for machine learning predictors

	6 size/BTM	25 size/BTM	10 industry	30 industry
Random forest	-0.054	-0.043	-0.076	-0.025
Gradient boost	-0.066	-0.052	-0.043	0.010
LASSO	-0.034	-0.045	-0.040	-0.021
Rel. LASSO	-0.035	-0.038	-0.031	-0.015
Targ. rand. forest (LASSO)	-0.013	0.014	0.006	0.010
Targ. grad. boost (LASSO)	-0.057	-0.037	-0.021	-0.015
Elastic net	-0.057	-0.038	-0.031	-0.011
Rel. el. net	-0.026	-0.022	-0.026	-0.011
Targ. rand. forest (el. net)	0.016	0.025	0.022	0.005
Targ. grad. boost (el. net)	-0.007	-0.014	-0.004	-0.026

Out-of-sample information coefficients for the machine learning predictors computed using a 120-month rolling window. The variables used are those listed in Table B.1. Targ. rand. forest (LASSO) refers to a targeted random forest with a LASSO targeting step.

(2020) call this a “targeted random forest”. LASSO or elastic net is run in the first stage and then, in the second stage, random forest is run over the variables with non-zero coefficients in the first stage. This selection step can help to remove the weakest predictors and subsequently improve forecast performance. I therefore consider both targeted random forest and targeted gradient boosting (gradient boosting with an initial variable selection step).

B.2.4 Information coefficients

The variable selection and machine learning methods do not necessarily give rise to better information coefficients than the univariate methods. Table B.2 shows the out-of-sample information coefficients, where the one-step-ahead prediction models are estimated with a 120-month rolling window, again over the same sample as I run the portfolio strategies in Section 3.5. The set of variables used in each prediction method is as in Table B.1.

The targeted random forest, using the elastic net in its targeting step, is the only prediction method which produces a positive information coefficient in all four universes. Hence, it is the machine learning predictor considered in Section 3.5.3.

The LASSO targeted random forest is the only other predictor to have a positive information coefficient in both industry universes. I therefore consider using it as a robustness check in Section 3.6.3

No predictor other than the targeted random forest with elastic net targeting step has a positive information coefficient in both size/book-to-market universes.

B.3 Conditional heteroscedasticity specification

To account for possible conditional heteroscedasticity in the conditional mean/prediction error ε_t in (3.1), I use an asymmetric dynamic conditional correlation (aDCC) model (Cappiello et al., 2006). The advantage to using an aDCC model, rather than a standard dynamic conditional correlation model (DCC), is that the aDCC model allows for correlations between assets to be different when the prediction error is below its sample mean compared to when it is above. The aDCC model contains only one additional parameter to estimate compared to the DCC model.

aDCC models the ex-ante variance of ε_t as

$$\mathrm{Var}(\varepsilon_t|\mathscr{F}_{t-1}) = D_t^{1/2} P_t D_t^{1/2},$$

where P_t is the matrix of conditional correlations of $\varepsilon_{i,t}$, D_t is the diagonal matrix of conditional variances of $\varepsilon_{i,t}$ and $\mathscr{F}_t$ the information available at t. Element i of D_t's diagonal is $\mathrm{Var}(\varepsilon_{i,t}|\mathscr{F}_{t-1})$. To obtain $\mathrm{Var}(\varepsilon_{i,t}|\mathscr{F}_{t-1})$, I specify each time series of errors $\varepsilon_{i,t}$ to follow a GARCH(1,1) process.

I use the `rmgarch` package of Ghalanos (2019b) to estimate the aDCC model in `R`. Ghalanos (2019a) explains that P_t is estimated by the transformation

$$P_t = \mathrm{diag}(Q_t)^{-1/2} Q_t \mathrm{diag}(Q_t)^{-1/2}.$$

Q_t evolves over time as

$$Q_t = (1-a-b)\bar{Q} - g\bar{Q}^{-} + a z_{t-1} z_{t-1}' + b Q_{t-1} + g z_t^{-}(z_t^{-})'$$

where z_t are the standardised errors with $z_t = D_t^{-1}\varepsilon_t$, $z_t^{-} = 1(z_t < 0) z_t$, $1(\cdot)$ the indicator function, $\bar{Q}$ and $\bar{Q}^{-}$ the unconditional variance of z_t and z_t^{-} respectively and a, b and g scalar parameters.

References

Abhyankar, A., Basu, D., and Stremme, A. (2012). The optimal use of return predictability: An empirical study. *Journal of Financial and Quantitative Analysis*, 47(5):973–1001.

Adam, K., Marcet, A., and Nicolini, J. P. (2016). Stock market volatility and learning. *Journal of Finance*, 71(1):33–82.

Allen, D., Lizieri, C., and Satchell, S. (2019). In defense of portfolio optimisation: What if we can forecast? *Financial Analysts Journal*, 75(3):20–38.

Anderson, G., Post, T., and Whang, Y.-J. (2019). Somewhere between utopia and dystopia: Choosing from multiple incomparable prospects. *Journal of Business and Economic Statistics*, In press.

Andrikopoulos, P., Clunie, J., and Siganos, A. (2012). UK short selling activity and firm performance. *Journal of Business Finance and Accounting*, 39(9-10):1403–1417.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, 24(1):34–112.

Au, A. S., Doukas, J. A., and Onayev, Z. (2009). Daily short interest, idiosyncratic risk, and stock returns. *Journal of Financial Markets*, 12(2):290–316.

Bansal, R. and Yaron, A. (2004). Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509.

Barroso, P., Boons, M., and Karehnke, P. (2017). Time-varying predictability of consumption growth, macro-uncertainty, and risk premiums. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2933449.

Boehmer, E., Jones, C. M., and Zhang, X. (2008). Which shorts are informed? *Journal of Finance*, 63(2):491–527.

Borup, D., Christensen, B. J., Mühlbach, N. N., and Nielsen, M. S. (2020). Targeting predictors in random forest regression. Mimeo, Department of Economics and Business Economics, Aarhus University.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: a consumption-based explanation of aggregate stock market behaviour. *Journal of Political Economy*, 107(2):205–251.

Campbell, J. Y. and Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, 1(3):195–227.

Cappiello, L., Engle, R. F., and Sheppard, K. (2006). Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics*, 4(4):537–572.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82.

Cecchetti, S. G., Lam, P.-S., and Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review*, 80(3):398–418.

Chiang, I.-H. E. (2015). Modern portfolio management with conditioning information. *Journal of Empirical Finance*, 33:114–134.

Cochrane, J. H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, 96(5):893–920.

Cochrane, J. H. (2008). Financial markets and the real economy. In Mehra, R., editor, *Handbook of the Equity Risk Premium*, pages 237–330. Elsevier Science, Amsterdam, Netherlands, 1st edition.

Cochrane, J. H. and Saá-Requejo, J. (2000). Beykond arbitrage: good deal asset price bounds in incomplete markets. *Journal of Finance*, 108(1):79–119.

Constantinides, G. M. and Ghosh, A. (2011). Asset pricing tests with long-run risks in consumption growth. *Review of Asset Pricing Studies*, 1(1):96–136.

Daske, H., Richardson, S. A., and Tuna, I. (2005). Do short sale transactions precede bad news events? Available at SSRN: https://ssrn.com/abstract=722242.

Delgado, M. A. and Velasco, C. (2011). An asymptotically pivotal transform of the residuals sample autocorrelations with application to model checking. *Journal of the American Statistical Association*, 106(495):946–958.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953.

Diether, K. B., Lee, K.-H., and Werner, I. M. (2009). Short-sale strategies and return predictability. *Review of Financial Studies*, 22(2):575–607.

Elaut, G. and Erdos, P. (2019). Trends' signal strength and the performance of CTAs. *Financial Analysts Journal*, 74(1):64–83.

Epstein, L. G. and Zin, S. E. (1989). Substitution, risk aversion, and the temporal behaviour of consumption and asset returns: a theoretical framework. *Econometrica*, 57(4):937–969.

Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246–273.

Fama, E. F. and French, K. R. (1993). Common risk factors in stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Ferson, W. E. and Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *Journal of Finance*, 41(3):967–982.

Garcia, R., Renault, E., and Semenov, A. (2004). A consumption CAPM with a reference level. Université de Montréal, Québec, Canada; available online at http://epge.fgv.br/files/2056.pdf.

Ghalanos, A. (2019a). *The rmgarch models: Background and properties.* https://cran.r-project.org/web/packages/rmgarch/vignettes/The_rmgarch_models.pdf.

Ghalanos, A. (2019b). rmgarch: Multivariate garch models. https://cran.r-project.org/web/packages/rmgarch/index.html.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, 33(5):2223–2273.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Han, Y., Zhou, G., and Zhu, Y. (2016). A trend factor: Any economic gains from using information over investment horizons? *Journal of Financial Economics*, 122(2):352–375.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054.

Hansen, L. P. and Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica*, 55(3):587–613.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Hong, Y. (1999). Hypothesis tests in time series via the empirical characteristic function: a generalised spectral density approach. *Journal of the American Statistical Association*, 94(448).

Hong, Y. and Lee, Y.-J. (2005). Generalised spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form. *Review of Economic Studies*, 72(2).

Huang, D. and Zhou, G. (2017). Upper bounds on return predictability. *Journal of Financial and Quantitative Analysis*, 52(2):401–425.

Jank, S., Roling, C., and Smajlbegovic, E. (2016). Flying under the radar: The effects of short-sale disclosure rules on investor behavior and stock prices. *Bundesbank Discussion Paper*, (25/2016). Available at SSRN: https://ssrn.com/abstract=2809498.

Jank, S. and Smajlbegovic, E. (2017). Dissecting short-sale performance: Evidence from large position disclosures. *SSRN*. Available at https://ssrn.com/abstract=2631266.

Kandel, S. and Stambaugh, R. F. (1989). Modeling expected stock returns for long and short horizons. Department of Finance, Wharton School, University of Pennsylvania.

Kim, C.-J., Morley, J. C., and Nelson, C. R. (2001). Does an intertemporal tradeoff between risk and return explain mean reversion in stock prices? *Journal of Empirical Finance*, 8(4):403–26.

Kirby, C. and Ostdiek, B. (2012a). It's all in the timing: Simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis*, 47(2):437–467.

Kirby, C. and Ostdiek, B. (2012b). Optimising the performance of sample mean-variance efficient portfolios. Available at SSRN: https://ssrn.com/abstract=1821284.

Li, J. and Tsiakas, I. (2016). Equity premium prediction: The role of economic and statistical constraints. Mimeo, Rimini Centre for Economic Analysis.

Linton, O. B. and Whang, Y.-J. (2007). The quantilogram: With an application to evaluating directional predictability. *Journal of Econometrics*, 141(1):250–282.

Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica*, 59(5):1279–1313.

Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 48(4):36–52.

Lo, A. W. and MacKinlay, A. C. (1989). The size and power of the variance ratio test in finite samples: a Monte Carlo investigation. *Journal of Econometrics*, 40(2):203–238.

Lobato, I. N., Nankervis, J. C., and Savin, N. E. (2002). Testing for zero autocorrelation in the presence of statistical dependence. *Econometric Theory*, 18(3):730–743.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, 21(1):255–285.

Mehra, R. and Prescott, E. C. (1985). The equity premium: a puzzle. *Journal of Monetary Economics*, 15(2):145 – 161.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, 60(7):1772–1791.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics*, 22(1):27–59.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–49.

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24(3):401 – 421.

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

Wright, J. H. (2000). Alternative variance-ratio tests using ranks and signs. *Journal of Business and Economics Statistics*, 18(1):1–9.

Zhou, G. (2008). On the fundamental law of active portfolio management: How to make conditional investments unconditionally optimal. *Journal of Portfolio Management*, 35(1):12–21.